COLLECTION *Musée Claude Bernard*

SINGULAR SELVES

Historical Issues and Contemporary Debates in Immunology

Editors : A. M. Moulin and A. Cambrosio

ELSEVIER

SINGULAR SELVES
HISTORICAL ISSUES AND CONTEMPORARY DEBATES IN IMMUNOLOGY

Musée Claude Bernard

Collection Musée Claude Bernard, dirigée par Annick Opinel

SINGULAR SELVES
HISTORICAL ISSUES AND CONTEMPORARY DEBATES IN IMMUNOLOGY

DIALOGUES ENTRE SOI
QUESTIONS HISTORIQUES ET DÉBATSCONTEMPORAINS EN IMMUNOLOGIE

Edited by
Anne-Marie Moulin, Alberto Cambrosio

ELSEVIER

Amsterdam, New York, Oxford, Paris, Shannon, Tokyo

A member of Elsevier Science

Front cover: *Carnival*, Lena Subota, 2000.

Born in southern Ukraine, Lena Subota is a talented painter who also works with engravings, photography and graphic techniques. She currently lives and works in St. Petersburg and has taken part in numerous exhibitions both in Russia and abroad. *Carnival* foregrounds a modern individual - a "Singular Self" - against a fairy backdrop reminiscent of a Russian tale. The other human beings are distant, masked, some turning their backs to the viewer. The silhouette in the front acts as a boundary, separating the right side of the painting, dominated by human artifacts (a tower, a clock), from the left side of the painting, featuring a luscious nature with exotic birds. The eyes of the central figure are focused on some distant scene, taking place in a yet undefined setting. Subota's painting suits a book that started as a dialogue between individuals from various disciplines, looking for ways of accounting for the elusive immunological Self and its interface with many other social Selves

23, rue Linois, 75724 Paris cedex 15, France.
www.elsevier.fr

Imprimé en France par Stedi, 75018 Paris *Printed in France*
Dépôt légal: 6790 – Février 2001 – ISBN 2-84299-210-5

Anne-Marie Moulin and Alberto Cambrosio, co-editors, would like to thank the Mérieux Foundation, and more particularly, Claude Lardy, Secretary General, Mérieux Foundation, and Annick Opinel, Curator, Claude-Bernard Museum, for their generous help in organizing the conference entitled "Immunology: Historical Issues and Contemporary Debates," and publishing this book. Though mainly based on the lectures that were given during the meeting, this book also includes contributions from authors who, while not having attended the conference, were eager to participate in this dialog about issues and debates that have over time helped to build today's cutting-edge immunology, in which various disciplines and thereby "Singular Selves" interact, in search of new avenues that may only be opened through comprehensive understanding of the past.

Created in the house where Claude Bernard was born and lived in Saint-Julien-en-Beaujolais (France), the Claude Bernard Museum has been managed and financially supported by the Mérieux Foundation since 1965. Its location, recollection and memories of Claude Bernard, and permanent collections account for its uniqueness.

Not only devoted to history, the Claude-Bernard Museum organizes, with the partnership of major universities, meetings that focus on the history of medicine, philosophy of sciences and epistemology, and publishes their related proceedings.

Issued from a conference entitled "Immunology: Historical Issues and Contemporary Debates" that was held on 4–6 June 1998 in Saint-Julien-en-Beaujolais (France), "Singular Selves. Historical Issues and Contemporary Debates in Immunology" gathers the conference proceedings updated by their respective author(s) and independent contributions from other authors. This book therefore represents the Claude-Bernard Museum continuing effort, as it belongs to a recently launched collection entitled "Collection Musée Claude Bernard."

Annick Opinel,
Curator
Claude-Bernard Museum

Contents

PUBLIC IMAGES OF IMMUNOLOGY

THE HISTORIOGRAPHY OF IMMUNOLOGY

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Introduction
Historical issues and contemporary debates

Anne-Marie Moulin, Alberto Cambrosio

Most of the chapters included in this book were originally presented at a conference entitled "Immunology: Historical Issues and Contemporary Debates" held in June 1998 at the Musée Claude Bernard in Saint-Julien-en-Beaujolais (France), under the sponsorship of the Mérieux Foundation, with additional support from the Wellcome Institute (UK). The conference was planned as a workshop, to which a carefully selected mix of historians, social scientists and immunologists were invited. Two previous, similar international workshops had acted as the immediate stimulus for the St. Julien conference [1, 2], but the latter can be placed in a broader context.

One of the very first international meetings specifically devoted to the history of immunology had been convened in Toronto in 1985 by Pauline Mazumdar. In spite of the diversified contents of the resulting book [3], the Toronto meeting centered on a key figure of pre-World War II immunology, Felix Haurowitz, who was to die shortly after the meeting, and who used that occasion as a last attempt to defend his, by then, largely superseded template theory of antibody production. Although bracketed at the beginning and at the end by two historians, most of the Toronto presenters were immunologists. Following the Toronto meeting, on the joint initiative of Anne-Marie Moulin and Arthur Silverstein, a section on the history of immunology was embedded in the large triennial venture known as the International Congress of Immunology (Berlin 1988, Budapest 1991). A few veteran immunologists, and a handful of young ones attended the presentations, but without reaching the critical threshold that would have allowed the transformation of those occasional events into permanent ones. A summer school on the history of immunology, to which a few historians but especially veteran immunologists had been convened, was held in Ischia in 1992 as part of the International School of the History of Biological Sciences under the title "From Immunity to Cellular and Molecular Immunology" [4]. The event was characterized by a palpable tension between at least some of the immunologists and those among the historians who were no longer willing "to listen to the siren calls of the senior statesmen of the field," and were instead interested in producing "a history of recent immunology that [did] not degenerate into partisanship or advocacy of the present intellectual preferences and aesthetic ideals of the discipline" [5]. Finally, it is worth mentioning that in 1993 a Witness seminar series meeting covering a major event in the recent history of immunology, namely the discovery of monoclonal antibodies, was organized by the Wellcome Institute for the History of Medicine in London and attended by

both scientists and historians of contemporary medicine [6]. In contrast with the preceding conferences, the two workshops that we introduced as the immediate predecessors of the St. Julien conference, namely the 1993 Boston meeting on "Conceptual Issues in Immunology: Experimental and Clinical Foundations" [1] and the 1995 meeting held in Italy on "Conceptual Issues at the Interface between Immunology and Epidemiology" [2], were characterized by the attempt to define an agenda that would be firmly grounded in historical and epistemological considerations.

In spite of their undeniable differences, these various activities, taken together, bear witness to the willingness shown by historians, philosophers, sociologists and scientists to cross the "two-cultures" divide in order to confront their analyses and interpretations of a discipline – immunology – that has become a prime example of the challenges and opportunities facing contemporary biomedicine. Yet, this endeavor is not without difficulties, for, as shown by the mixed results of the above-mentioned meetings, goodwill is not enough to overcome the epistemological and methodological challenges facing those attempting to walk down this path. All the previously mentioned meetings have raised, implicitly or explicitly, the thorny question of who should write the history of immunology, according to which criteria, and for what purpose. Readers looking to the present collection for easy answers to these questions will be disappointed. As a matter of fact, we have decided to stress, rather than hide, differences, for, in our opinion, identifying and confronting differences is the surest way to progress.

The very fact that the history of immunology has become a contentious field can be accounted for in terms of the growth of the discipline itself. Immunology has indeed undergone remarkable developments during the last few decades, becoming a focal point of biomedical research and practice. Historians, philosophers, sociologists and anthropologists of science and medicine who had heretofore neglected immunology, have become interested in the discipline, as shown by the increasing number of books (not to mention the articles) published on the topic [7-21]. Prompted by the often dramatic events that have punctuated the rapid development of their discipline, practicing immunologists have also felt the need to share in print their historical recollections, sometime adopting a scholarly format [22, 23] and sometime a more popular one [24]. The time is past when one needed to justify his or her somewhat esoteric interest in the field. Yet, now that the interest and legitimacy of the endeavor has been established, other questions have surfaced. For instance, a recent review of several books on the history of immunology [25] came to the conclusion that while the number of scholars and publications had reached a critical mass, the field still lacked maturity, for its agenda was in large part still being dictated by scientific, rather than historical considerations. The tension between these two agendas is probably best epitomized by the different role assigned to historical, social and technical contingencies. Whereas social scientists see contingencies as the major focus of their endeavor, many (although far from all) scientists perceive them as a threat to the ideology of truth they use as a rallying cry. So, for instance, while immunologists, in surveying debates about any given key issues of immunology – say, the Self-NonSelf distinction – will mainly worry about the truth status of the latter, historians will be busy trying to understand how that same distinction has functioned as a metaphorical tool in the production of immunological practices and how the latter have, in turn, "performed" that distinction [26]. The difficulty, thus, lies not so much in the existence of different interpretations

of a given topic, but in the very definition of what should count as a relevant topic. Meaningful exchanges can only take place if people agree on a common ground for discussion.

A prototypical example of these sort of difficulties can be found in a exchange of messages between one of the scientists who attended the St. Julien meeting and one of the editors of the present collection. The discussion turned on whether notions such as right or wrong, or rationality and irrationality, could act as guiding principles of historical inquiry. The exchange began with the scientist's remark that if Melville in Moby Dick had interrupted the story at Chapter 32 (entitled "Cetology") with a detailed discussion of the biology and behavior of whales, he had most certainly done so on the assumption that readers ignoring everything about whaling and the habits of whales would not be able to appreciate Captain Ahab's actions. Similarly, in order to be successful in producing accounts of interest to practicing immunologists, historians of immunology would have to understand their subject matter and ground their accounts in such an understanding, the obvious implication being that this was not presently the case.

The reply to this argument was that while some degree of practical and theoretical understanding of the field was obviously important, the historian's task was precisely to question what the scientist took for granted, namely, that there was one and only one "right" understanding of, say, the structure and functions of the immune system and that the history of immunology should consist in comparing how notions and theories of the past compared to the understanding provided by hindsight. Arguing that phenomena such as the Self-NonSelf discrimination were being defined by the immune system and not by the experimenter, the philosopher or the historian, the scientist was asking the historian to accept as self-evident the very practices the latter had set out to explore and that resulted in the production and stabilization of notions such as that of an immune system capable of discriminating between Self and NonSelf [27]. Confronted, at any given time, with immunologists who differed on key issues (ranging from overarching theoretical frameworks to the interpretation of a given piece of experimental evidence) historians had the choice between focusing on who was right or wrong by giving a set of (usually extra-scientific) reasons for why some scientists subscribed to the wrong claims, or of rejecting this asymmetrical approach by replacing it with an account of the material culture and the experimental systems that sustained various immunological approaches. A symmetrical account would thus focus on the dynamics of these various immunological practices, not on the adequacy between a given piece of immunological knowledge and the biological reality allegedly underlying it.

The scientist rebuttal consisted in arguing that he was trying to distinguish rational from irrational, as distinct from erroneous or mistaken approaches. In other words, partisanship in the answer was less important than understanding the underpinnings of the subject and being able to formulate the questions that the scientists who advanced the subject were asking, "however unclearly." This latter proviso – "however unclearly" – was an important one, since it required the imposition of an external rationality on the events. In short, the move from right/wrong to rational/irrational, did not seem to radically alter the scientist's stance that some sort of normative intervention should define the analysis of historical events. Moreover, only a restricted number of events qualified as relevant in the production of such an account. "Personal struggles and so on" – in short, the "incarnate" aspects of laboratory life – did not count, for all that really mattered was the immunologist's interac-

tion with an anonymous body of knowledge, some of which would be lost as wrong and some of which would be passed on as right. The discussion had obviously reached an impasse: while the historian was interested in the ethnographic reconstruction of practices, the scientist was asking him to adopt an a priori conceptual framework in order to reconstruct and categorize the practices under analysis. In short, history as meta-science.

There are thus many criteria – differences of methods, epistemology, and subject matter – according to which a distinction between a scientific and a historical agenda can be established. One needs not mobilize Georges Canguilhem's powerful argument [28] that the object of science and the object of the history of science should be carefully distinguished – the latter entertaining a meta-relation to the former – in order to maintain that, say, a posteriori logical reconstructions of the development of immunological notions and theories correspond to a profoundly different endeavor than, say, a historical inquiry into the institutional entrenchment of the discipline. For, even if we agree to define "immunology" as corresponding to the activities of card-carrying immunologists, one can disagree precisely on what immunologists actually do, or, in other words, on what, among the heterogeneous things they do, should count as constitutive of immunology: to develop theories and concepts, to perform experiments, to build a rich material culture of instruments and techniques, to engage in institution or network building, to carve out biographical trajectories combining personal and scientific issues... or all of the above? A choice between these various alternatives will shape answers to subsequent questions, such as how one should analyze the relationship between the laboratory and the clinic that has characterized immunology from its beginnings [29], or define the interaction between the various biomedical fields contributing to or being affected by immunology.

A more general way of phrasing these questions is to ask what the point of a history of immunology is, a question one cannot ask without simultaneously defining the audiences targeted by such a history. A strong declaration of independence on the part of historians – according to which the history of immunology is a discipline whose accomplishments and shortcomings should be judged according to the relevant standards as defined by historians without undue interference by practitioners of other disciplines (namely, natural scientists) – is certainly a possible answer, one to which the editors of this book are quite sympathetic. For, in the same way as it would be absurd for historians to dictate to immunologists how they should go about practicing immunology, historians are entitled to a high degree of professional autonomy in the choice of the relevant themes and assessment criteria. However, this solution does not quite solve all problems, for both theoretical and practical reasons.

First, the relationship between historians and scientists is not a symmetrical one. Historians of immunology write about what immunologists do, but not vice versa. Now, it is often claimed that one of the golden rules of historical scholarship is that the history of a given scientific episode should be written by adopting the point of view of the actors of that time, i.e., without indulging in anachronism and/or hindsight. However, while this is certainly sound methodological advice, one should not read it as a form of naïve historical empiricism, according to which one could reach a comprehensive, factual and final understanding of "what really happened." Rather, one should recognize that events of the past are irremediably transformed by our present understanding and that, in this sense, history has to be continuously re-written [30, esp. pp. 176-86]. Insofar as scientists, by

their ongoing practices, decisively contribute to a recasting of the past, a rigid dichotomy between historians and scientists is, at best, simplistic. Second, 'in practice', historians of immunology, especially those working on contemporary topics, do entertain, at different degrees, a working relationship with practicing immunologists and this requires some sort of give-and-take. Why should scientists bother about history and historians? Possible, but far from satisfactory, answers have ranged from the pragmatic argument that knowledge of past debates will reduce the occurrence of the "reinventing the wheel" syndrome, thus contributing to a better understanding of contemporary debates [31], to the more detached one that history, in addition to being a source of insight, can be a source of wonderment [23]. But, maybe, these are exactly the kinds of questions that are not worth asking from a general, abstract point of view, and that should rather be formulated in ad hoc, concrete terms, every time that, for whatever specific reason, one feels the need to ask them.

Given this general state of affairs, a reader, confronted with the present collection, might wonder about its purpose. Is this a book on the history/epistemology of immunology? Or is it a book about the debates surrounding the history/epistemology of immunology? A glance at the table of contents shows that both aspects are present, and that the various contributions have been organized on a spectrum ranging from historical to meta-historical issues. Moreover, and in spite of the neat separation between scientists and historians displayed in the title of the book's first two sections, no unique historical or meta-historical approach or principle is to be found within a given section that could be easily mapped onto the professional, or, for that matter, any other general qualification of that section's authors.

The first three authors of section I are immunologists best known for their contributions to the development of the field. Michael Potter and Leslie Brent describe and analyze events in which they have been personally involved, thus exemplifying, in their different ways, a "protagonist's approach" to history. Potter's contribution focuses on myeloma proteins, a topic that points to the existence of an intimate link between research materials and conceptual developments, while Brent's chapter examines the issue of tolerance, that, in a sense, and as shown by the recent debates surrounding the issue of Self-NonSelf discrimination, has never lost its status as a constitutive, yet controversial tenet of modern immunology. A comparison between Brent's chapter and the following one by Melvin Cohn is certainly instructive, insofar as their subject matter overlaps. Cohn adopts a militant stance, openly endorsing a particular theoretical approach to Self-NonSelf discrimination which he then uses as a yardstick to review the historical contributions by other immunologists. Kenneth Schaffner is a philosopher, not an immunologist, but his chapter has been included in Section I because it illustrates an attempt to intervene, with the help of the Internet, into contemporary debates over Self-NonSelf discrimination and thus represents, so to speak, a philosophy 'in', as opposed to a philosophy 'of' immunology. In this sense, it can be used to contextualize the preceding contributions. Section I closes with a chapter by Patrick Triadou that marks the transition from a more scientific to a historical and epistemological agenda, not only because of the author's personal qualifications as both a practicing scientist and a historian, but, more importantly, because his chapter raises the issue of the relationship between the normal and the pathological and of the models used to tackle it (see also Peter Keating's chapter in section IV). It does so by exploring the relationship of immunology to hematology, from a "mere" producer of tools for

diagnostic practices to a "method-theory package" [32] for redefining our understanding of pathological processes.

By choosing to open section II with a chapter by Arthur Silverstein, we have opted for an even smoother transition between the first two sections. As in Triadou's case, this is so not only because of the dual scientific and historical qualifications of the author, but also because of the author's particular focus on conceptual issues, namely the early development of Paul Ehrlich's notion of a receptor that, according to Silverstein, has conditioned all subsequent work by the influential German immunologist. Covering pretty much the same historical period, the chapter by Eileen Crist and Alfred Tauber also tackles conceptual issues, but by espousing a more explicit epistemological agenda. Crist and Tauber reject linear accounts of conceptual developments and focus on the controversy between humoral and cellular approaches, showing how the latter can be accounted for in terms of competing visions of biology and how the traditional picture of the controversy in terms of winner and losers should be replaced by a more subtle understanding of how elements from both approaches were quietly incorporated into our present understanding of immunological mechanisms. The next three chapters mark a shift toward a different historical approach, by refusing to privilege the role of concepts in the history of immunology. This does not mean that concepts are evacuated but, rather, that they are analyzed as part of a complex network of practices. In keeping with the recent interest in the material culture of scientific practices, the chapter by Angela Creager shows how particular objects and materials (for instance, antibodies as material substances obtained through peculiar production practices) have helped to stabilize immunology as a clinical and scientific field of knowledge. The chapters by Mark Jackson and Jennifer Stanton qualify even more directly as social and institutional history. Jackson accounts for the controversies about allergen immunotherapy in the United Kingdom by arguing that they have not been uniquely determined either by concerns about technical issues (such as protocols, testing, standardization) or by their outcomes as perceived by patients and doctors. Rather, in order to understand the therapy's "checkered history," one has to factor in professional and socio-economic concerns. Stanton, examining hepatitis B vaccination, similarly claims that the use of vaccines depended on a complex of social, economic and political factors. Thus, a notion such as "vaccine viability" is to be understood in social, economic and political terms rather than biological, scientific and medical ones.

The first two chapters in section III, by Ed Cohen and Philippe Menut, could have been accommodated within section II, but we decided to include them in a separate section with the chapters by Dominique Frommel, Daniel Jacobi and Anne-Marie Moulin insofar as these five contributions share a common focus on public images of immunology. As shown by the different approach and subject matter of the five chapters, the term public image is to be taken in a broad sense, sometimes metaphorically and sometimes quite literally. Cohen's chapter covers a broad historical spectrum and tracks the migration of a metaphor – immunity – from the political to the biological and medical realms. Thus, rather than focusing on the social and political consequences of an already constituted biomedical field, the chapter focuses on the co-construction of nature (immune bodies) and society (political bodies), and examines the bio-political consequences of such a mutual production. In contrast to Cohen's chapter, Menut's contribution focuses on a limited, yet (in)famous event, namely the trial that followed the death of 77 children

following a 1930 BCG vaccination in Lübeck. Menut shows how the juridical system became a hybrid forum for debates over the existence of bacterial species, and how the legal outcome of the Lübeck tragedy was predicated upon the temporarily successful attempt to untangle scientific and political connotations that are still with us in present-day controversies over vaccination. Frommel's chapter examines, in a sense, the equivalent for developing countries of the dynamics explored by Menut in the case of the Lübeck trial. The argument is derived from the author's personal experience with immunization campaigns in developing countries: on this basis, the author examines the gap, and the resulting clashes, between the allegedly neutral textual and visual imagery conveyed by Western immunological documents and the connotations evoked by that same imagery in non-Western technical and lay practitioners. Jacobi's chapter introduces a much needed semiotic dimension, by looking at the imagery used in the advertisements produced by pharmaceutical companies to promote their immunotherapeutic substances. Jacobi's analysis is to be read as part of the renewed interest by sociologists and historians of science in the visual dimension of scientific practices. By focusing on the end-product of a long chain of visual productions, Jacobi also highlights the mutually-reinforcing links between the representational practices that characterize the esoteric world of laboratories and those displayed in the public sphere of biomedical communication. The last chapter in Section III is yet again a transitional text, as it combines the public image of a central immunological notion with historiographic considerations, the topic of the next section. Moulin's analysis of the "multiple versions of the immune system" shows how these different representations move back and forth not only between what Fleck [33] has termed the esoteric and exoteric circles – namely the specialized world of the laboratory and the general public – but also Western and non-Western cultures. A philosopher and a practicing doctor, Anne-Marie Moulin shows that common reveries underlie immunological concepts, and that the metaphysical character of the immune system turns it into an anthropological tool available for the interpretation of various cultural patterns. Thus, Moulin's contribution addresses not simply the issue of scientific popularization, but also the historiographic issue of how a sensitivity to, inter alia, anthropological elements should inform the writing of a discipline that is constantly in the process of being deconstructed and reconstructed.

The chapters in the fourth and final section openly address historiographic issues. The chapter by Alfred Tauber echoes both his joint contribution with Eileen Crist in section II and Anne-Marie Moulin's chapter in the previous section by pleading for the return, under new modalities, of a historical genre taking into proper account the metaphysical assumptions that explicitly and implicitly guide scientific practices. Peter Keating's chapter closely examines one of the central thesis of French historical epistemology, namely George Canguilhem's groundbreaking, yet often misunderstood analysis of the relationship between the normal and the pathological, and shows how it can be used to analyze more recent biomedical events. The chapter by Alberto Cambrosio looks at the intriguing possibility that the notion of a uniform, linear time underpinning historical accounts could be replaced by a sociological understanding of time, one postulating the co-existence of multiple time frames. The chapter ends with a tentative list of the consequences for historical practices of adopting such an approach. Finally, Thomas Söderqvist, on the basis of his work on the life of the immunologist and Nobel Prize winner Niels Jerne, explores yet

another possibility for subverting established historical genres, namely the development of a biographical approach grounded in the explicit promotion of ethical themes, not from a deontological or consequentialist point of view (how should I act) but from the point of view of a virtue ethicist (how should I live).

In order to make their point, all chapters in this final section resort, in their different ways, to examples drawn from the history of immunology. Yet, they tackle issues of concern, more broadly, to the history of science, that could also have been examined by resorting to examples from other scientific or biomedical disciplines. In an important sense, section IV thus constitutes a proper ending to this collection, for our intention in editing it has never been to institute the history of immunology into a separate, Self-contained specialty but, rather, to contribute to its development by establishing creative links with all those who, in other fields, share our fascination with the history of that peculiar human activity known as science.

Finally, the co-editors would like to thank Peter Keating for his help in translating and revising several chapters.

Bibliography

1 Cambrosio A, Keating P, Tauber AI, Eds. Immunology as a historical object. J Hist Biol 1994; 27.
2 Keating P, Balaban M, Cambrosio A, Tauber AI, Eds. Historical studies on immunology. J Hist Biol 1997; 30.
3 Mazumdar PMH, Ed. Immunology 1930–1980. Toronto: Wall & Thompson; 1989.
4 Gallagher RB, Gilder J, Nossal GJV, Salvatore G, Eds. Immunology: the making of a modern science. London: Academic Press; 1995.
5 Söderqvist T. How to write the recent history of immunology: is the time really ripe for a narrative synthesis? Immunol Today 1993; 14: 565-8.
6 Tansey EM, Catterall PP. Monoclonal antibodies: a witness seminar in contemporary medical history. Med Hist 1994; 38: 322-7.
7 Bibel D, Ed. Milestones in immunology. San Diego: Science Tech Publishers; 1988.
8 Silverstein A. A history of immunology. San Diego: Academic Press; 1989.
9 Corbellini G. L'evoluzione del pensiero immunologico. Torino: Boringhieri; 1990.
10 Moulin AM. Le dernier langage de la médecine. Histoire de l'immunologie de Pasteur au Sida. Paris: Pressses Universitaires de France; 1991.
11 Cazenave PA, Talwar P, Eds. Immunology: Pasteur's inheritance. New Delhi: Wiley Eastern Publishers; 1991.
12 Tauber AI, Chernyak L. Metchnikoff and the origins of immunology. New York. Oxford University Press; 1991.
13 Tauber AI. The immune self. Theory or metaphor? Cambridge: Cambridge University Press; 1994.
14 Martin E. Flexible bodies. Tracking immunity in American culture from the days of polio to the age of AIDS. Boston: Beacon Press; 1994.
15 Mazumdar PMH. Species and specificity. An interpretation of the history of immunology. Cambridge: Cambridge University Press; 1995.
16 Cambrosio A, Keating P. Exquisite specificity. The monoclonal antibody revolution. New York: Oxford University Press; 1995.
17 Moulin A, Ed. L'aventure de la vaccination. Paris: Fayard; 1996.
18 Löwy I. Between bench and bedside. Science, healing and interleukin-2 in a cancer ward. Cambridge, MA: Harvard University Press; 1996.
19 Daëron M, Fougereau M, Fridman WH, Moulin AM, Revillard JP. Le système immunitaire ou l'immunité cent ans après Pasteur. Paris: Nathan; 1996.
20 Podolsky AH, Tauber AI. The generation of diversity. Clonal selection theory and the rise of molecular immunology Cambridge, MA: Harvard University Press; 1997.

21 Söderqvist T. Hvilken kamp for at undslippe: en biografi om immunologen og nobelpristageren Niels Kaj Jerne [What struggle to escape: a biography of the immunologist and Nobel Prize winner Niels Kaj Jerne]. Copenhagen: Borgens Forlag; 1998.
22 Szentivanyi A, Fridman H, Eds. The immunologic revolution. Boca Raton: CRC Press; 1994.
23 Brent L. A history of transplantation immunology. New York: Academic Press; 1997.
24 Fridman WH. La révolution immunologique. Les défenses naturelles contre le cancer. Paris: Lattès; 2000.
25 Söderqvist T, Stillwell C. The historiography of immunology is still in its infancy. J Hist Biol 1999; 32: 205-15.
26 Moulin AM. La métaphore du soi et le tabou de l'autoimmunité. In : Bessis M, Bernard J, Debru C, Eds. Le soi et le non-soi. Le Seuil: Paris; 1990. p. 55-68.
27 Moulin AM. The immune system; a key concept for the history of immunology. Hist Philos Life Sci 1989; 11: 13-28.
28 Canguilhem G. Études d'histoire et de philosophie des sciences. Paris: Vrin; 1968.
29 Moulin AM, Löwy I. La double nature de l'immunologie. Fundamenta Scientiae 1983; 3: 201-18.
30 Rheinberger HJ. Toward a history of epistemic things. Synthesizing proteins in the test tube. Stanford: Stanford University Press; 1997.
31 Silverstein A, Rose NR. On the mystique of the immunological self. Immunol Rev 1997; 159: 197-206.
32 Fujimura JH. Crafting science: standardized packages, boundary objects, and "translation." In: Pickering A, Ed. Science as practice and culture. Chicago: The University of Chicago Press; 1992. p. 169-21.
33 Fleck L. Entstehung und Entwicklung einer Wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv. Basel: Benno Schawbe and Co.; 1935 [Genesis and development of a scientific fact. Chicago: The University of Chicago Press; 1979].

SCIENTISTS LOOK AT IMMUNOLOGY, PAST AND PRESENT

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Myeloma proteins and antibodies

Michael Potter

Over the last 60–65 years immunologists have been ever intrigued with the biology of myeloma proteins and using them as tools in a variety of different experiments. The questions raised to explain myeloma proteins – what they are and what their function might be – has led to major insights not only in immunology but also in understanding the pathogenesis of plasma cell tumors. Our understanding of these proteins, however, is far from complete. While current dogma accepts the concept that myeloma proteins are structured normally and therefore should bind antigens, only relatively few antigen binding activities have been described and little is known about the immunogenic stimuli that lurk in their past history. This article will review how myeloma proteins were discovered and how they and the cells that produce them have played a key role in some of the major discoveries in immunology over the last 60–65 years. It will then conclude with a discussion of current unresolved problems relating to the biological activities of myeloma proteins.

> *I have been asked to comment on the reasons why certain experimental approaches were taken in the development of the mouse plasmacytoma model system. It is always difficult to try and recreate one's past mind set many years later, the possibility always being that one might take advantage of present knowledge because errors in memory take credit for ideas that did not prevail at the time. I have tried to describe the questions that motivated this research. They are inserted in italics to set them apart from the historical discussion of myeloma proteins.*

Multiple myeloma and myeloma proteins

Myeloma proteins are immunoglobulin molecules made and secreted by plasma cell tumor cells. The term 'myeloma' comes from the major plasma cell tumor process in humans, multiple myeloma. It literally means a tumor arising in the cells that occupy bone marrow cavities.

Multiple myeloma as a specific disease process has been familiar to physicians for over 150 years. Despite its rarity, it provoked many reports because of its unusual clinical features and severe related morbidity. The most famous historical case was that of Mr. Thomas Alexander McBean, a London grocer who had apparently been in good health until 1844 when, during a vacation in the countryside, he jumped out of a cave and upon falling on his chest experienced excruciating pain [1]. This temporarily subsided, but on his return to London he consulted Drs. Thomas Watson and William MacIntyre, who diagnosed his condition as "Mollities et Fragilitas Ossium" (softening and fragility of the

bones). During the progressive course of his disease, they noted that Mr. McBean began excreting large amounts of protein in his urine. Their analyses of this protein showed that it did not behave like albumin, the protein most usually found in association with kidney diseases. They took a sample to Henry Bence Jones, the most renowned 'clinical chemist' in London. Henry Bence Jones confirmed their findings and gave a very detailed description of the protein which was subsequently named for him as Bence Jones Protein (a name known to every medical student today) [2]. Mr. McBean died on January 1, 1846, and his postmortem revealed that his bone marrow cavities were filled with a reddish tumor like process and some unusual cells. Microscopic drawings were made of these curious cells and, crude as they were, these strikingly resemble plasma cells [3].

From 1845 to the early part of the twentieth century several new extensions about "Mollities et Fragilitas Ossium" became clearer. First, the myeloma tumors could involve different bones, hence the new name multiple myeloma [4], and, second, the cells in these myeloma tumors were identified as a newly recognized cell type – the plasma cell – in 1901 [5]. The histologists Paul Unna and Ramon Cajal had defined plasma cells as a distinct morphological cell type in the 1890s [6, 7]. Third, not only did many of the patients have Bence Jones proteinuria, but some also developed amyloidosis, a pathological accumulation of insoluble proteins in tissues. It was not until 1928 when the first reports of changes in the serum proteins were associated with multiple myeloma [8]. Multiple myeloma as a disease then became stigmatized with pathological protein formation but the reason for this was not known. Nonetheless, these characteristics haunted the early concepts of myeloma proteins.

Revolution started by physical chemists led to new insights on the nature of myeloma proteins

The understanding of the abnormalities in protein formation in multiple myeloma began with a technological revolution in the physical-chemical analysis of proteins. Two monumental achievements in protein chemistry in the twentieth century were the development of the analytical ultracentrifuge by Svedberg in 1925 [9] and the moving boundary electrophoresis by Arne Tiselius, who began his career with Svedberg [10] around 1925 and developed moving boundary electrophoresis between 1930 and 1937 [11]. All this occurred at the Institute of Physical Chemistry in Uppsala, Sweden which was created for Svedberg and became a Mecca for workers all over the world. The ultracentrifuge evolved from the pursuits of colloid chemists to separate small particles of various sizes. This led to attempts to sediment protein molecules according to size and shape by centrifugal force and this required the construction of powerful centrifuges that could spin a solution of protein molecules at speeds of greater than 100,000 revolutions per minute and record the movement of the protein molecules as well. Svedberg, who was ever interested in building bigger and more powerful centrifuges for the separation of macromolecules, did not loose interest in his pursuit of other ways to separate and purify specific proteins. Svedberg had begun to work on electrophoretic methods but had put them aside. When the promising and talented student Arne Tiselius came along in 1925, Svedberg urged him to develop moving boundary electrophoresis to separate proteins by using their net surface charges in an electrical field. Tiselius tackled this problem with tenacity and determination, finally perfecting an elegant but very large and complex instrument that he described in 1937

[11]. An early application of these instruments to physiological and clinical problems was the separation of the proteins in plasma, serum and other body fluids. The group in Uppsala included Kai O. Pedersen and Jan G. Waldenstrom who had clinical interests. Both instruments resolved serum proteins into characteristic fractions. In moving boundary electrophoresis albumin and α, β and γ globulin fractions were identified by net surface electric charge distribution. Biologists from all over the world beat a path to Uppsala to study their cherished proteins with Svedberg, Tiselius and Pedersen. Very early Michael Heidelberger in 1935 [12] and Elvin Kabat in 1937 [13] from New York arrived to analyze the highly purified preparations of antibodies they had prepared to study pneumococcal polysaccharides. As many (but not all) of the pure antibodies migrated electrophoretically with the globulins of gamma mobility, the term 'gamma globulin' came into popular use and became a 'household' word synonymous with antibodies, the proteins of immunity.

Following Tiselius' description of moving boundary electrophoresis, new instruments were built and refined in England and the USA., and the field developed explosively. In 1928 the first reports of changes in the ratio of serum albumin to serum globulins were associated with multiple myeloma [8], and these findings alerted the clinically minded physical chemists to obtain serum from patients with multiple myeloma for analysis. This led to the fascinating discovery that the sera of most patients with myeloma contained massive accumulations of specific kinds of globulins. These appeared as peaks of protein concentration arising within the β and γ globulin fractions. Longsworth et al. using a modified Tiselius apparatus at the Rockefeller Institute in New York first described these abnormal peaks [14] in 1939. It was soon followed by more extensive studies [15, 16] by Keckwick in London and Gutman in New York. World War II now began to delay the rapid progress that was being made. These peaks were the true myeloma proteins (as they are so named today) and were distinguished from Bence Jones proteins by their molecular weights.

M-components

Jan Waldenstrom, working away in Uppsala in the 1940's, discovered that there were two other disease processes that produced these unusual electrophoretic bands or spikes in serum protein analyses besides multiple myeloma. The first of these did not to have excessive numbers of plasma cells in their bone marrow cavities (as in multiple myeloma) but lymphocytes instead. These were associated with protein spikes that were called 'millionaires' because the proteins were found to have sedimentation coefficients 19–20S and estimated molecular weights near a million. In contrast, the usual myeloma proteins had sedimentation coefficients of 7S and had molecular weights of 150.000. Today, the millionaires are known as Waldenstrom Macroglobulins. The second process also discovered by Waldenstrom was that there were some individuals whose serum contained a characteristic spike but apparently had no evidence of multiple myeloma or other lymphoproliferative disease. These individuals appeared to have benign proliferations of immunoglobulin secreting cells. The M, (*M*yeloma or *M*acroglobulin) spikes collectively were called "M-components" [17] by some or "paraproteins" by others.

> *In 1954 I came to the National Cancer Institute to work in Lloyd Law's laboratory which was focused on mouse leukemia. Lloyd Law was interested in*

chemotherapy and had made important discoveries on the development of resistance to chemotherapeutic agents such as methotraxate, 6-mercaptopurine and 8-azaguanine. His work laid the foundation for combination chemotherapy of acute leukemia of childhood. Lloyd Law was also fascinated with the problem of trying to explain the pathogenesis of mouse leukemia and this field was in a turmoil because of the experiments of Ludvik Gross, who produced evidence that a vertically transmissible virus was the responsible cause. This had greatly challenged the prevailing concepts of E.C. MacDowall and Jacob Furth that leukemia in mice had a genetic basis. Mouse leukemia in the late 1950's was predominantly a study of lymphocytic neoplasms, with particular emphasis on those tumors of thymic origin. In 1954 Thelma Dunn published her classic monograft on reticular neoplasms in mice (i.e., tumors of the hematopoietic system), and this opened the eyes of many to the variety of different tumor types within this system [18].

The great successes of the transplantable lymphocytic neoplasms in mice as model systems for detecting new chemotherapeutic agents (particularly Lloyd Laws L1210 model) prompted Lloyd Law to suggest to me that I try and develop some of the other hematopoietic tumor types as model systems. This was most appealing to me, and I set about trying to establish in transplant examples of the tumors in Dunn's review. She was most helpful in this respect, and we worked together very closely.

The principal source of these tumors were the DBA/2 lymphomas induced by aromatic polycyclic hydrocarbons. Extant today are P815 (a mastocytoma) and P388 (a pre-B cell lymphoma) which came from this study. The mast cell tumor was associated with characteristic differentiation products such as heparin, histamine and serotonin. When I came to the plasma cell tumors, there was only one possible transplantable plasma cell tumor, a C3H tumor called 70429. Thelma Dunn thought this might have originated as an ileocecal plasmacytoma. I began working with this tumor which was quite fascinating by itself, principally to search for a myeloma protein production. Paper electrophoresis was now available, and John Fahey at the NCI was actively studying the serum proteins in patients with lymphomas. I asked him to run serum from mice with 70429, but he found no evidence of a myeloma band. Since I had many tumors in transplant from the Dunn study, I obtained my own electrophoresis apparatus to systematically study these tumors. By a great stroke of good fortune the first week we began running these samples, the two plasmacytomas transplanted by Ira Pilgrim (which he had sent to. Thelma Dunn for histological diagnosis) arrived from California, and Alvado Campbell, my assistant, and I ran their sera immediately. I could not believe the X5563 paper electrophoresis pattern on the first run because of the huge band that appeared at the origin. Subsequent runs, though, repeated the same result, and it became very obvious this was a myeloma protein.

About this time my fascination with plasma cell tumors was greatly enhanced by a lecture given by Frank Putnam on Bence Jones proteins in 1956. This was presented in Wilson Hall, a room lined by beautiful portraits of past

directors of the NIH and then the major auditorium of the old NIH. Frank Putnam began studying the structures of Bence Jones proteins in the early 1950's and discovered that each one was unique [19]. He had determined the amino acids at the amino terminal ends of the proteins and found these varied between proteins. To my memory what seemed most exciting was the question about the reason for all these variations. Were they due to genetic polymorphisms in the human population or did these different proteins reflect a mutagenic process that resulted from exposure to chemical carcinogens or other mutagenic agents during the development of multiple myeloma? One approach to answer these questions was to find myeloma protein producing plasma cell tumors in inbred mice where the host genotype was all the same. The presumption here was that if the myeloma proteins were different, then this might suggest a mutational process was affecting the proteins made by the tumors. Needless to say, I had no idea how complicated this question would become.

Clonal selection theory

In 1948 immunologists generally accepted evidence that plasma cells secreted antibodies [20], but the intriguing biological significance of the nature of plasma cells was not appreciated until 1957 when the cellular basis for the origin of plasma cells was first understood and explained through the instrument of a brilliant hypothesis, Frank MacFarlane Burnet's clonal selection theory of antibody formation [21]. The implication of this theory was that antibody producing cells were of two related cell types: first, a lymphocyte precursor cell that displayed the antibody molecules on the plasma membrane where it acted as a receptor for antigen, and second, the cellular maturation product of these lymphocytes the plasma cells which developed the ability to secrete these molecules. Both cells were different developmental stages of a single founder B lymphocyte (*figure 1*). Today we refer to this as the B-lymphocytic lineage. A second revolutionary implication of Burnet's theory was that each immunoglobulin producing cell had undergone an unexplained developmental process that limited the cell to producing only one species of antibody molecule. This permitted antigen to select appropriate cells (lymphocytes which contained antibody molecules on their plasma membrane) for clonal expansion and eventual differentiation into antibody secreting plasma cells. The theory also provided a glimpse of the great diversity of the antibody producing lymphocyte population that is required for the workings of this system of cells. Gustav Nossal and Joshua Lederberg established experimental proof of this remarkable division of labor in antibody producing cells by immunizing rats simultaneously to two different antigens and demonstrating that single antibody-producing cells could only make one of the antibodies [22].

Transplantation studies in 1958–1959 on eight available plasma cell tumors showed very clearly that the electrophoretic migration pattern of the myeloma protein was unique for each tumor and was clearly a stable consistent marker for each tumor [23]. In 1955–1956 I met Theodore S. Hauschka for the first time and this began a series of exciting conversations with him about the genetics of somatic cells usually at scientific meetings or visits to Buffalo, NY. In discussing the plasma cell tumors with him, the subject came

up about how to explain electrophoretic variations of myeloma proteins from plasmacytomas originating in the same inbred strain. Were they somatic mutations? Could they be due to some kind of differentiation process in plasma cells? Ted Hauschka told me about the recently published Nossal/ Lederberg experiments [22] that showed single plasma cells were specialized to produce only one kind of antibody molecule. This was a turning point for me, and I began thinking about the individuality of myeloma proteins as differentiation products rather than mutagenic ones. It is embarrassing in retrospect that in our paper of 1960 [23] we did not site Burnet's clonal selection hypothesis [21], which was to play such an important conceptual part of this field. Possibly, it had less of an impact at the time than the Nossal/ Lederberg experiment.

The physico-chemical heterogeneity of pure antibody fractions could now be explained by Burnet who postulated that the immunizing antigen could select out different lymphocyte clones, each of which produced a different (mono-) clonal antibody molecule that reacted with that antigen. The B lymphocyte interacting with antigen was triggered to divide, thus forming a clone of cells that produced the same kind of antibody molecule. Some of the cells were destined to mature to become antibody-secreting plasma cells. Antibody, then, was apparently a collection of structurally- and genetically-related (immunoglobulin) molecules that bound the same antigen. Antibody molecules produced by different clones could have different chemical properties. A tumor (neoplastic proliferation) of plasma cells such as multiple myeloma in humans or the plasma cell tumors in mice represented a megaclonal expansion of one plasma cell. The cells in one tumor originated by a neoplastic developmental process or transformation of a single antibody producing cell and was limited by differentiation to synthesizing a single species of immunoglobulin molecule. Myeloma protein represented a massive accumulation of identical

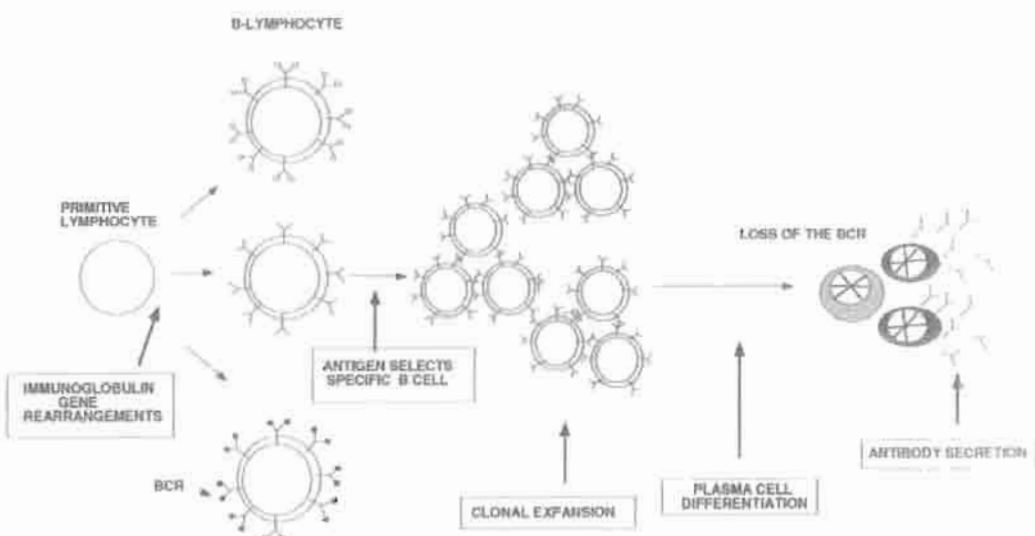


Figure 1. Scheme of B lymphocyte plasma cell formation.

antibody (immunoglobulin) molecules. The next few years after Burnet's theory was published, cellular, genetic and biochemical data supported these basic postulates.

An experimental source of myeloma protein producing plasma cell tumors in mice

By the mid 1950's no means for obtaining plasma cell tumors in laboratory animals was available. The mouse was a particularly attractive experimental species as tumors could be propagated in inbred strains in this species by transplantation. In 1954 Thelma Dunn showed for the first time that mice could develop plasma cell tumors spontaneously but unfortunately these tumors were very rare and most were diagnosed retrospectively from tissue sections [24]. The ileocecal plasma cell tumors appeared to develop in a chronic inflammatory tissue underlying ulcers in the cecal mucosa. A way to induce these lesions and plasma cell tumors was not available. By very good fortune in 1957 H. Ira Pilgrim, a graduate student at Berkeley working on mammary tumors in mice, found two unusual abdominal tumors of probable ileocecal origin that developed in old C3H mice, and he sent these tumors to Thelma Dunn in Bethesda for diagnosis. Both were found to be plasma cell tumors that secreted myeloma proteins [25, 26].

Thelma Dunn again played a key role in plasmacytoma research in1958. Ruth Merwin and Glenn Algire were studying the survival of foreign tissue cells in potentially hostile tissue environments. To carry out these experiments they used a device called a Millipore Diffusion Chamber (MDC). Richmond Prehn along with Algire had developed this device [27] which consisted of a plastic ring onto which two millipore membranes were glued. Cells could be introduced into the space between the membranes. The MDC which was around the size of an American nickel was implanted into the peritoneal space of a mouse. The tissue fluids from the host mouse could now diffuse through the membrane and support the viability and proliferation of the cells. The question Merwin and Algire sought to answer in 1958 was how long potentially incompatible cells could survive in these chambers. The design of this experiment was logical but would produce an unexpected result. They used mammary tumor cells of C3H origin inside the chambers. There was an additional reason for this choice as C3H Mammary Tumors carried the Mammary Tumor Virus (MTV). The millipore membranes employed in these chambers excluded the passage of cells through the membranes but would allow the MTV virus to pass through and infect the host. This led to the obvious choice of the recipient strain BALB/cAn for two reasons. First, the cells of BALB/c mice were incompatible with those of C3H, and if the immune cells of BALB/c came in contact with the C3H cells they would be destroyed. Second, BALB/cAn did not carry the MTV but was quite susceptible to this virus and the females would develop mammary tumors if exposed as adults. Thus, Merwin and Algire had a potential assay system for MTV. Ruth Merwin began building the chambers and setting up the experiment. Glenn Algire was fatally ill at this time and did not live to see the results. An extraordinary set of circumstances and coincidences came in place to produce this far reaching result.

The unexpected result of this project was the development of abdominal tumors in the BALB/c mice. As was the common practice in the NCI in those days, the tumorus mice were taken to Thelma Dunn for pathological diagnosis. Thelma Dunn was the reputed world authority on the histopathology of tumors in mice and particularly tumors of the hematopoietic system. She found a most unusual set of findings in Ruth Merwin's mice.

There were two kind of tumors in the peritoneal cavities of these mice. The first were fibrosarcomas that developed around MDC, but the second kind of tumor was unprecedented in her experience. These were plasma cell tumors that were growing on distant peritoneal surfaces and not in any obvious association with the MDC. The Merwin/Algire experiments were most important, though, because they provided a means for inducing plasma cell tumors in inbred mice and, further, they revealed a genetically susceptible target strain BALB/c [28].

> *As Ruth Merwin was exploring which components of the Millipore Diffusion Chambers (MDC) were responsible for plasma cell tumor formation, I sought other ways to induce plasma cell tumors in mice. This work was based on the hypothesis that prolonged exposure to foreign antigens would stimulate plasma cell formation and hyperplasia which might be a necessary preliminary step in plasmacytoma formation. One obvious possibility suggested by the MDC experiments was that the genetically incompatible C3H cells inside the chambers were releasing potential antigens into the BALB/c mouse. Such an idea inevitably led to using other antigens and to chronically immunize the mice. I consulted several immunologists in a nearby laboratory, and they told me horse serum was a powerful antigen. Multiple injections of horse serum met with rapid failure, as the mice all died of anaphylaxis in a few weeks, long before plasma cell tumors develop. The way around this was to use of immunological adjuvants which were known to slowly release antigens from water-in-oil emulsions [29] without provoking anaphylaxis. In the late 1950's Rose Lieberman at the NIH developed an adjuvant mixture that produced ascites in mice; this consisted of incomplete Freund's Adjuvants (a mineral oil containing adjuvant) and heat killed staphylococci [30]. BALB/c mice were injected intraperitoneally with this adjuvant mixture. We added a few 'condiments' such as the inclusion of a protein antigen (horse serum), and painting of the skin with the carcinogen methylcholanthrene (a technique known to induce hematopoietic tumors). After 6 months some of these mice developed myeloma protein producing peritoneal plasmacytomas [31]. The induction system contained numerous variables, and new experiments were immediately begun with Charlotte Robertson Boyce to simplify the procedure and to determine which components were essential. It soon became apparent that the single critical component required for plasmacytoma induction was the paraffin (mineral) oil that made up 85% of Freund's adjuvant [32]. In the next few years we tested many different brands of mineral oil including USP grade oils, and all were active. Later, Paul Anderson found that a chemically pure component of mineral oil was available. This was pristane, which proved to be the most effective of all the oils tested. Although we lost faith in the hyperimmunization hypothesis, we were in possession of a relatively easy method for inducing an unlimited number of myeloma protein producing plasma cell tumors all derived from genetically similar (close to identical) BALB/c mice. This opened to door in 1959–1960 to begin to systematically characterize and classify myeloma proteins from a biochemical point of view.*

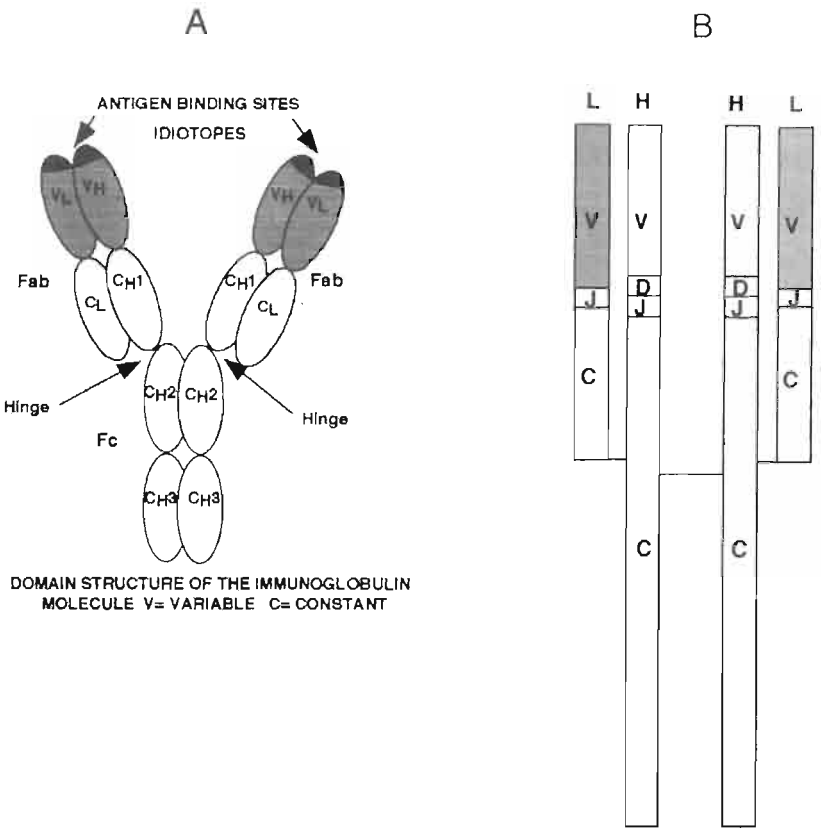
Architecture of the immunoglobulin molecule solved

In 1962 the general structure of the immunoglobulin (Ig) molecule was finally established [33]. Two laboratories taking differing approaches led the way. First, Rodney Porter's group using highly purified rabbit antibody molecules demonstrated that a typical Ig molecule contained three globular structures that could be separated by proteolytic enzyme digestion [33, 34]. Two fragments called Fabs contained antigen binding sites, and a third fragment, Fc, roughly equal in size to an Fab did not bind antigen (see *figure 2*). Gerald Edelman and his associates then unfolded the Ig molecule in powerful denaturing conditions that disrupted all of the covalent and hydrogen bonds and showed that the Ig molecule was composed of two different polypeptide chains, the light (L) and heavy (H) chains [35]. Edelman was then able to solve the mystery of the Bence Jones protein which turned out to be Ig L chain [36]. Some plasma cell tumors and multiple myelomas produced excessive amounts of L chains or only L chains.

Myeloma proteins: key to the antibody diversity mechanism

While the basic structural features (shape, large fragments and molecular weights of the various components) of the Ig molecule were known in the early 1960s the immunoglobulin genes that coded for this molecule were not established (*figure 2*). Theorists, always ahead of the actual data, had become intrigued with the problem whose essence was how to explain how an individual could generate in a genetically economical way an untold number of antibody specificities. It was estimated that this number was very high, greater than 10,000, possibly even 100,000 or more, which would mean that a large part of the genome would consist of Ig genes. What kind of genetic system could code for such a capability? Intriguing hypotheses emerged such as somatic mutations of a few inherited Ig genes [37] and chromosomal recombinations between Ig genes [38]. Myeloma proteins from humans and mice now provided the structures that could be analyzed or deciphered, but to do this the amino acid sequences would have to be determined. Protein chemists had sequenced insulin (51 amino acids) and ribonuclease (149 amino acids), but a molecule the size of an Ig molecule with L chains of 250 amino acids and H chains with 450–500 amino acids presented a formidable problem at the very limits of the available technology. Probably, the protein chemists were unaware of the zeal of the immunologists who now sailed into this problem with the fury of an advancing army. For once the technology rose to the occasion, and streamlined protein sequencing became available [39]. The first analyses were made on Ig L chains (Bence Jones proteins), which were readily available in humans and more recently in mice [40]. Frank Putnam had pioneered many of the studies of the structure of Bence Jones proteins in humans and demonstrated that each Bence Jones protein was unique, that is, part of its structure differed from all others but other parts were identical. The variable and constant parts were recognized [19, 41].

> *In the spring of 1962 the Federation of Experimental Biologists met in Atlantic City, NJ, for the annual meeting. This was usually just after the meetings for the American Association for Cancer Research. As many as 20,000 scientists were gathered. It was a great opportunity to meet new people and discuss ideas. At this meeting Frank Putnam described his continuing studies on the structural individuality of Bence Jones proteins, this*



Legend A An immunoglobulin molecule is composed of two light (L) chains and two heavy (H)chains. These form a monomeric unit. Some immunoglobulins such as IgM have 5 of these monomers linked together to form a pentameric structure known as a macroglobulin. The IgA molecules are usually dimers. IgG molecules are monomers. In physiological conditions the chains are folded into domains, 2 for a light chain the VL and CL, 4 or 5 for a heavy chain VH and CH1-3 or 4. Domains interact with each other to form regions such as the variable region or the different sectors of the C regions. The two antigen binding sites on a typical monomer are parts of the variable regions and each is formed by a combination of VL and VH segments. As a general rule which has exceptions most idiotopes contain parts of antigen binding sites.

B. The L chain is controlled by three structural genes V, J and C while heavy chains are controlled by 4 structural genes V,D,J and C. In the mouse there are approximately 100 VH genes, 17 to 19 DH genes , 5 JH genes and 8 CH genes . For light chains there are again around 100 VL genes , 5 JL genes and 1 CL gene.

Figure 2. Structure of the immunoglobulin.

time using tryptic peptide maps, also called fingerprints. In this procedure a protein molecule was digested with trypsin which split the protein molecule into a series of peptides at all the positions along the chain where a lysine or arginine recurred. The resulting peptides were chromatographed in one dimension and electrophoresed in another. Struck by the beauty of Frank Putnam's work and its potential for expanding the criteria for variability, I returned to Bethesda determined to apply this technology to our growing

library of Bence Jones proteins, but I was very naive in protein chemistry. While I was looking around NIH for help, Bill Dreyer struck up a conversation with me about mouse plasmacytomas which interested him as a potential system. It did not take long to learn he worked with Chris Anfinsen and had developed a high voltage paper electrophoretic method for separating peptides and making beautiful fingerprints. At this point I became engrossed in the diversity problem. I immediately became an apprentice to Dreyer. We made fingerprints for 20 different mouse Ig L chains (Bence Jones proteins) or L chains isolated from myeloma proteins and found that each of them was different. Each protein, however, contained a set of common peptides (the same in all) and a set of variable peptides. Though we had extensive discussions, we could not come up with a genetic mechanism to explain this phenomenon. Bill Dreyer left NIH for Cal Tech, and I worked for another year with Claude Bennett, his post-doctoral student. In 1965 Bennett and Dreyer (then at Cal Tech) published their paper on the two gene-one polypeptide chain model. I thought that this solved the problem of variability – there was one C variability region gene and multiple V (variable) region genes in the mouse. But this turned out to be only a part of the story as hypermutation and recombination were shown to contribute substantially to variability.

By 1965 the necessity of having complete amino acid sequences of Bence Jones proteins of either the kappa or lambda type became the imperative for further progress. In 1965 Melvin Cohn at the Salk Institute organized an Antibody Workshop that was held February 8–11, 1965 at Warner Springs, CA. This was probably the most famous meeting in the field of Immunology since the Geneva Congress in 1882 when Pasteur and Koch exchanged unpleasantries. Melvin Cohn had decided that the time had come for molecular biologists to become interested in the antibody problem, and he invited molecular biologists: Francis Crick, James Watson, Seymour Benzer, Max Delbruch; protein chemists: Chris Anfinsen, Frank Putnam, Gerald Edelman, Edgar Haber, Bill Dreyer, Jon Singer, Russ Doolittle; and a large impressive cast of immunologists to come to this Workshop. Symbolic of its Olympian aura, the meeting was held on a mountain top. There was a natural spring there that fed hot water into a very large pool. In the daytime the pool was warm, but as evening approached the water cooled and those who remained in it gathered as darkness descended at the warm water source, and there many interesting conversations took place chiefly among the immunologists.

In an exchange of letters (personal communication from Melvin Cohn) with his co-planner Rodney Porter (who was unable to attend), Mel Cohn heard about an investigator at the Rockefeller Institute who had quietly been sequencing Bence Jones proteins, although no one knew what he had found. As the meeting progressed the molecular biologists were exhausted by 4 days of immunological phenomenology. Charles Todd presented his serological study of the rabbit immunoglobulin heavy chain and evidence that it was controlled by more than one gene. He was chastised for his heresy. But there was one last session that peaked the molecular biologists' interest, and this was the presentation by the mystery guest, Norbert Hilschmann, the person who had sequenced one complete and one nearly complete Bence Jones protein. When his turn came to present, Hilschmann described the sequences and then showed the sequence slide for one minute

for all to see. Of course, no one was fast enough to copy it. There was no doubt about its authentic alignment of the peptides and its profound significance. His manner of presentation, however, enraged many present who wanted to study, savor and digest this famous 'first'. Hilschmann made it quite clear there were two covalently continuous peptide segments to the immunoglobulin light chain (BJ protein), a constant one and a variable one. The two proteins differed from each other in the variable segment but were alike in the constant region. Hilschmann felt insecure about revealing his hard wrought secrets prior to publication as he was a young post-doctoral student. In the audience were several who had been working feverishly to accomplish this task but were not as far advanced; thus, his alignment would permit them to rapidly catch up. His actions in my opinion were defensible. He told everyone what he knew. Where would he ever find such an appropriate audience? The molecular biologists took the other position – that the sequence should not be shown at all unless it was openly presented. The theoretical repercussions of these electrifying few minutes were profound. The molecular biologists thought that there was no mechanism for joining two independently synthesized chains together in tandem. Thus, Hilschmann's data shattered the existing molecular dogma of the "one gene-one polypeptide" chain [42]. Priority in science has always been a 'touchy subject' that does not bring out altruism, humility, generosity, graciousness and all of the fine characteristics of the human spirit. I do not remember hearing the important question put to Hilschmann about the genetic explanation of this iconoclastic phenomenon. But others recall he speculated prophetically and instinctively it would have to be due to some kind of recombination.

Claude Bennett and William Dreyer proposed in 1965 their now famous and beautiful hypothesis on how the two regions the V and C, each being controlled by a separate gene, could be joined to make a single polypeptide chain (the "two gene-one polypeptide chain hypothesis)" [43]. They proposed that the V and C genes were first joined together at the DNA level. At that time this was quite revolutionary, even though today it is accepted in principle without reservation.

Essentially, they had solved one part of the antibody diversity problem. Simply stated there was one gene for the C region and multiple genes for variable regions. The evolutionary process led to the amplification of V gene libraries. Each V and C gene was free to diverge through the accumulation of random germ line mutations. This type of V to C joining was also proposed and found for heavy chains. Both V_L and V_H were required for forming the antigen-binding site. Structural diversity of antigen-binding sites was generated by utilizing different combinations of V_L and V_H genes. Not everyone thought this was the solution to the diversity problem, because it still required a large number of V genes – 100 V_L and 100 V_H genes to generate 10,000 different binding sites. Many theorists in the mid 1960's were concerned with genetic economy and devised other ideas to explain how a few genes could generate mechanisms or somatic hypermutational processes. Great discussions were held at meetings about which mechanism was correct. By 1967 when the grand Cold Spring Harbor Symposium on Immunology was held the diversity question was in full flower, though still unsettled. The interested reader should peruse this volume and the recorded discussions between molecular biologists and immunologists [44]. In the end when Ig V-genes and their products were sequenced, both at the nucleotide levels and protein levels, multiple germ line V_L- and V_H -genes, recombinational and somatic hypermutational mechanisms were all

found to be factors in generating antibody diversity. Mother nature pulled out all the stops to generate structural variability of immunoglobulin molecules.

The ideas of the 1960's were superseded by molecular DNA data, beginning in 1977 with Susumu Tonegawa and his collaborators [45, 46] when they isolated the immunoglobulin genes from mouse plasma cell tumors and showed that the L chains were coded not by two genes but three genes (V, J and C) and the heavy chains by four genes (V, D, J and C) (*figure 2*). The new elements were J and D genes. These were small mini-genes that coded for joining sequences between the V and C genes. It has been estimated that there are 100–150 V_H, 19 D_H, five J_H, eight C_H, 100 or more V_L, 5–10 J_L, and two or more C_L genes. This generated a vastly larger number of combinations (in the millions) that could be made by rearranging the genes on the chromosome. Plasma cell tumor cells and the myeloma proteins that they produced, made this work possible.

The quest to find antigens for myeloma proteins in mice. Melvin Cohn discovers the first match: a myeloma protein that bound to the pneumococcal C polysaccharide

Melvin Cohn at the Salk Institute and our laboratory in Bethesda began assembling libraries of myeloma sera between 1965 and 1967 for the purpose of finding myeloma proteins that bound antigens. A direct and simplistic approach on how to show that myeloma proteins could bind antigens was clear to non-immunologists such as Seymour Benzer, who suggested to Melvin Cohn in 1964 [47] that it was essentially a 'numbers game'. With the availability of a large library of myelomas (and their proteins) it seemed one need only to screen these with an equally vast number of available antigens to find a match. Jan Waldenstrom had done something like this in 1964 with human myeloma proteins and found some antigen-binding proteins (see below). Melvin Cohn and his colleagues now proceeded to build such a system in the mouse and established a large library of several hundred BALB/c myeloma proteins at the Salk Institute. By 1967 they found the first example of a match involving the IgA S63 myeloma protein that precipitated the pneumococcal C polysaccharide [47], and the finding was presented for the first time at the Cold Spring Harbor Symposium on Immunology in May 1967. The primary evidence was a precipitin band that was very convincing, but he warned the audience to be cautious of the C polysaccharide as an antigen, as it was reactive with other serum proteins, i.e., the C-reactive protein. This skeptical remark of his did not inspire an immediate rush to search our collection of myeloma protein for reactivity to C polysaccharide. We were to subsequently rediscover by another line of investigation that Melvin Cohn was on the right track.

In the fall of 1967 Myron Leon, whose home base was in Cleveland, OH, had arranged a trip to Prague to carry out some experiments on complement. He flew first to Washington to spend a day before catching his flight to Europe and to use the occasion to carry out a 'quickie' experiment. He had made arrangements to screen a bank macroglobulinemia sera against a battery of pneumococcal polysaccharide antigens (Myron Leon had worked in Michael Heidelberger's laboratory when he was a high school student). He brought large bags containing a hundred or so Ouchterlony plates. The day turned out to be frustrating, and after digging around in the freezer he identified only three candidate sera, and none of them reacted. Even though the effort over in Virginia had proved unproductive, Myron still thought he might salvage something of the day by calling a few people across the Potomac river at NIH in Bethesda. By chance he called the lab, and it did not take long

to establish a number of mutual interests. He arrived at about 4:30 P.M., and we sat down for a chat. Describing his frustrating search of ice chests, I said it was too bad he hadn't brought his antigens with him because I had a hundred myeloma samples in search of an antigen, and I opened the refrigerator door, revealing the test tubes. Myron said with a flourish "Oh, but I did," and opening his coat he produced three test tubes from his inside pocket one with a mixture of six of Michael Heidelberger's purest pneumococcal polysaccharides (types II, III, IV, VI, X and XIV), a second with a cocktail of levans from Allene Jeanes, and a third with a series of dextrans. We immediately had some micro-Ouchterlony plates cut, and after the myeloma proteins were loaded in the appropriate wells and the antigens in place we retired to a local restaurant for supper. When we returned to the laboratory there were four beautiful precipitin bands on the plates, three with the pneumococcal antigen mixture and MOPC299, MOPC167, and McPC603, and one with the dextrans and MOPC104E. Visions of glory danced in our heads... Had we discovered myeloma proteins that could bind to the type specific polysaccharides of *Pneumococcus*? Antibodies to these polysaccharides prepared in rabbits were used therapeutically to cure deadly pneumococcal infections in humans.

We, of course, thought that we had found three different specificities, but proof would have to wait until the testing of the separated polysaccharide antigens. Myron departed behind the Iron Curtain, and the next morning the samples were sent to Myron's associate Martin Young in Cleveland. Weeks passed by; then a letter came from Prague – MOPC299, MOPC167 and McPC603 each reacted with all six of the individual Heidelberger (pure) preparations.

Disappointed, I told Myron it was probably due to a contamination of the capsular polysaccharide preparations with C polysaccharide and that we had discovered the same reaction Mel Cohn had described at the Cold Spring Harbor Meeting in the spring. "Unthinkable!" he said. Michael Heidelberger's capsular polyssaccharide preparations were as pure as virgin snow. I acquired some pure C polysaccharide from Emil Gotschlich which had been isolated from the cell walls of R36A non-encapsulated pneumococci, and this reacted vigorously with the mycloma sera. Alex Tomasz at the Rockefeller Institute in 1967 showed that the pneumococcal C polysaccharide contained choline that was linked to ribitol [48]. Myron Leon then became intrigued with this antigen whose structure was just being worked out, and he and Martin Young made an important contribution by identifying phosphorylcholine (PC) as the chemical structure on the complex pneumococcal C polysaccharide molecule to which the specific myeloma proteins bound [49]. By this time eight different PC binding myeloma proteins had been found (e.g., McPC603, MOPC167, S63, S107, TEPC15, HOPC8). Since this original discovery (PC) has been found in a great variety of antigens that span a broad evolutionary scale from fungi to 'Ascaris'.

An alternative approach to screening myeloma proteins was to use chemically-conjugated protein antigens [50-52]. This was particularly attractive because the binding specificity could be ascribed to the chemical structure that had been attached to the protein carrier molecule (i.e., antibodies recognize only the attached chemical group or hapten). Karl Landsteiner had extensively developed this line of experimentation by covalently attaching a variety of different small molecules such as the nitrophenol compounds to the amino acids on the solvent surface of protein carriers. When rabbits were immunized with these conjugated proteins, antibodies were induced that reacted primarily with the hapten,

and some of these anti hapten antibodies had exquisite specificity by distinguishing subtle changes in the chemical groupings on the hapten [53]. Using 5-acetyluracil conjugated to BSA (bovine serum albumine), they found myeloma proteins that bound to DNA [50, 52]. In May 1967 at the Cold Spring Harbor Meeting Herman Eisen, who had been studying the immune responses to 2,4 dinitrophenol (DNP) haptens, decided to screen mouse myeloma libraries to find a monoclonal antigen-binding antibody and discovered to everyone's amazement the MOPC315 IgA myeloma protein in the mouse [54]. The binding affinity of the M315 myeloma protein was in the range of normal induced antibodies to DNP-conjugated proteins. The M315 became a valuable tool for immunochemists studying antigen-antibody interactions, but still there is no clear biological explanation as to what kind of immune response might expand these clones in an unimmunized host.

> *The collaboration with Myron Leon suggested that polysaccharides and other natural antigens with repetitive motifs would be a rich source of antigens that myeloma proteins might bind. This sent us off on a new and fertile direction. Betty Mushinski and I began testing all kinds of available polysaccharides. Otto Luderitz kindly sent us samples of 40* Salmonella *lipopolysaccharides. Using precipitation in agar we found precipitin bands with exotic species of* Salmonella, *such as* Salmonella Tel Aviv *and* Salmonella tranoroa. *But what relevance did these reactions have in the mouse? Betty Mushinski and I then began to isolate antigens from the environment of the mouse which included: non-living antigens that could be extracted from the mouse food or even the woodshavings from the cage bedding and antigens produced by micro-organisms that we isolated from the gastrointestinal tracts of mice in our colony [55, 56, 57] (table I). We were able to find environmental antigens for many of the myeloma proteins. When we prepared highly specific idiotypic sera in appropriate strains of mice for the antigen binding myeloma proteins, we were usually able to show that normal serum contained antibody molecules with these specificities, indicating that BALB/c mice were making natural antibodies with the same specificities which we found for the corresponding myeloma protein. It has been well established that the predominant heavy chain class expressed in BALB/c mouse myeloma proteins was IgA. Collectively, these findings suggested that one source of antigens was through the gastrointestinal tract.*

In general, the myeloma proteins, though having great specificity, had low binding affinities for their respective antigens. In most of these screening studies the active myeloma proteins were culled out as new tumor sera became available; however, when the percentage of positive myeloma proteins compared to the total number of samples screened was compared, only 5–10% of all the myeloma sera were active [51, 52, 58].

1975 Hybridoma technology, revolution in antibody research

Biologists had been experimenting with cell fusion for some time, but the applications to immunological research had proven elusive. In this method two different cell types could be joined together to make one single cell by fusing their plasma membranes. The new cell had the combined genetic information of the two parental cells. Cesar Milstein's laboratory had done the most advanced work and succeeded in fusing two plasma cell tumor cells

Table I. Antigen-binding activities of mouse myelomas proteins [69].

Foreign antigens		*Antigens of autogenous origin*
Relevant antigens (food, bedding, microbes) [56, 55]		Nuclear antigens J509, S176, S23 ss and ds DNA (many new myeloma proteins) 5-acetyluracil
Lactobaccillus sp.	PC choline: TEPC15, S107	
Trichoderma sp.	MOPC167	
Proteus morganii	LPS PC choline: McPC603	
Proteus mirabilis sp2	α-methyl D galactosides: MOPC384, McPC870 [56, 55]	
Arabino galactan [57]	β-1-6 D-galactans: J539, X24 and six others	
Probably in mouse environment		
N-acetylglucosamine	S117	
Dextrans (1,3; 1,6)	MOPC104E	
Plant levans (β-2,1 fructofuran)	UPC61, J606 and many others	
Bacterial levans (β-6 fructofuran)	UPC10	
Synthetic antigens		
Nitrophenol (DNP, TNP)	MOPC315 [54] MOPC460 XRPC25	

each from a different tumor [59]. The hybrid cell produced both myeloma proteins and the mixture of hybrid molecules containing L and H chains from the different parental cell types. These experiments were discontinued. Georges Kohler joined Milstein's laboratory a year or so later, and Cesar Milstein, who wanted to have an antigen-binding myeloma protein producing cells line in tissue culture, asked him to screen the MOPC21 myeloma protein with as many antigens as necessary to find a match [60]. No antigen-binding activity for MOPC 21 had or has been discovered. Georges Kohler did not want to do this experiment and came up with an alternative approach for finding a cell line that produced an antigen-binding protein. He asked permission to try out his alternative experiment which was to fuse a plasmacytoma cell with normal immunoglobulin producing cells that had been induced with antigen. He chose the classical antigen sheep red blood cells (SRBC). Mice were immunized and the antibody producing cells were isolated from the spleen and fused with plasmacytoma cells [61]. The P3 (MOPC21) plasmacytoma cells had been adapted to grow in tissue culture, that is, they were immortalized while the normal Ig producing (plasma cells) could survive for only a limited time in culture. The fusion product, half tumor/half normal cell, was called a hybridoma; hopefully, it would be an immortalized tumor cell that now secreted its own myeloma protein but would also secrete the antibody to SRBC. In one of those rare 'eureka' moments in science [60], the first experiment worked, and Georges Kohler and Cesar Milstein now had in their posses-

sion a cell line that produced a monoclonal antibody to SRBC. Later, Kohler made a mutant MOPC21 cell line that had lost its ability to make the myeloma protein but still retained the potential to form hybridomas. This cell line would now produce only the antibody of the normal cell. This technology changed the world of antibodies by providing immunologists an almost unlimited source of homogeneous monclonal antibody molecules with high binding specificity and affinity and shifted interest away from seeking new antigen binding activities in myeloma proteins.

The unresolved question: what antigens do myeloma proteins bind? New insights from studies in humans

Most IgG and IgA, immunoglobulins (myeloma proteins) secreted by plasma cell tumors cells, are structurally normal and should be capable of binding to antigens. What kinds of antigens do these proteins bind? The first reports of antigen binding activity in human Ig-secreting tumors were made in 1957 by Christenson and Dacie [62] who identified IgM M-components that were autoreactive against antigens on human red blood cells. The patients with these monoclonal immunoglobulins had lymphoproliferative disorders and suffered from hemolytic anemias. In the ensuing years many antigen binding activities have been associated with monoclonal IgM lymphoproliferative states. However, our focus here is the antigen-binding activities in human multiple myeloma proteins which produce IgG and IgA M components. The various antigens that have been identified are listed in *table II* and are grouped according to their biological origin. Two predominating categories are foreign antigens and antigens of autogenous origin. In many of the studies hundreds of human myeloma proteins have been screened. This is a remarkable list, chiefly because of the preponderance of myeloma proteins that bind antigens of autogenous origin.

There are a fewer clear examples of foreign antigens among the human myeloma proteins. Nonetheless, these are intriguing because in several of these cases the individuals who developed multiple myeloma were repeatedly exposed to the same antigen and became immunized to them years before the onset of multiople myeloma. In several there was clinical evidence of repeated infections with streptococci. This suggests that the repeated exposure to the microbial antigen resulted in the development of a deviant clone of cells that subsequently progressed to become a malignant tumor of plasma cells (multiple myeloma). The best documented example are myeloma proteins that bind streptolysin O (a hemolysin that is liberated by β-hemolytic streptococci during an infection). Approximately 20 examples of myeloma proteins with this activity have been collected [63].

The autoantigens, however, are of special interest and include DNA, histones, ribonucleoproteins (ribonucleic acids complexed with proteins), chromatin, cytoskeletal proteins such as actin, myosin and tubulin, and components of plasma membranes. Complex antigens such as these might be liberated from dying cells and greatly increased during inflammation and tissue destruction. It is possible that these complexes (containing potentially different antigenic structures) are autoimmunogenic; this point has not been proven.

One of the surprises evolving from the hybridoma technology was the proof that some monoclonal Igs could indeed be polyreactive (multireactive or highly crossreactive), that is, bind to multiple unrelated antigens [64, 65]. This had been suspected by serologists for many years who regarded these antibodies at first as crossreactive but then suspected they

Table II. Antigens to which human myeloma proteins bind.

Foreign antigens	*Antigens of autogenous origin*
Streptolysin O [43, 70, 63]	Human red blood cell antigens: Pr
Phosphoryl choline	IgG (Rheumatoid Factors) [71]
α-2 manoglobulin (horse) [70]	α-ID [72]
Mycobacterial glycolipids [73]	Thyroglobulin auto antigen [74]
	Antigens associated with nucleoproteins
	ANA
	Sn RNP
	Sm [75]
	Ro/SSA [76, 77]
	La/SSB [77]
	Histones [78]
	ds and ss DNA
	Membrane associated antigens
	Cardiolipin (phospholipid)
	Lipoprotein associated antigens
	α-,β-lipoproteins [79]
	LDH, VLDH, HDL [see [80] for refs]
	Blood coagulation associated antigens
	Glycoprotein IIIa [81]
	Fibrin monomer [82]
	Cytoskeletal antigens
	Actin [67]
	Myosin [67]
	Low molecular weight ligands
	Riboflavin [83]
	Heparan [84]

were more degenerated [66]. Polyreactivity of human myeloma proteins has been shown by Guillaume Dighiero [67] and others. Myeloma proteins that bind DNA, cardiolipin, actin and myosin are frequently polyreactive (see *table II*). It appears from studies by Abner Notkins and his collaborators that a substantial number of normal B cells in the recirculating pool in humans and mice are producing polyreactive antibodies [68]. Unfortunately, the property of polyreactivity in myeloma proteins makes it more difficult to define a predominating antigen. However, polyreactivity of the BCR potentially permit these B lymphocytes to be continuously in contact with antigen; this chronically stimulates B lymphocytes and may prolong their clonal life. The result may be an increase in the opportunities to sustain and accumulate oncogenic mutations. The available

knowledge on the autoantigen-binding activities of myeloma proteins in mice is still sketchy, but our laboratory is now pursuing this problem. The antigen binding activities of myeloma proteins may provide us with new insights about the cells that mutate to become multiple myeloma or plasmacytoma cells. Thus, the odyssey of the myeloma protein is far from complete..

Bibliography

1 MacIntyre W. Case of mollities and fragilitas ossium accompanied with urine strongly charged with animal matter. Med Chir Soc Trans. 1850; 32: 211-32.
2 Jones HBI. On a new substance occurring in the urine of a patient with Mollities Ossium. Phil Transact Royal Soc London 1847; 138: 55-62.
3 Dalrymple J. On the microscopical character of Mollities Ossium. Dublin Quarterly J Mod Sci 1846; 2: 85-95.
4 Von Rustizky J. Multiple myeloma. Dtsch Z Chir 1873; 3: 162
5 Wright JH. A case of multiple myeloma. Johns Hopkins Hospital Rpts. 1906; 9: 359-66.
6 Cajal SR. Quelques antécédents historiques ignorés sur les Plasmazellen. Anat Anz 1906; 29: 666-73.
7 Unna PG. Über Plasmazellen, insbesondere beim Lupus. Monatshf Prakt Dermat 1891; 12: 296-317.
8 Perlzweig WA, Delrue G, Geschicter C. Hyperproteinemia associated with multiple myeloma. J Am Med Asso. 1928; 90: 755-7.
9 Tiselius A, Claesson S. The Svedberg and fifty years of physical chemistry in Sweden. Biochem Physical Chem 1967; 18: 1-8.
10 Tiselius A. Reflections from both sides of the counter. Annu Rev Biochem. 1968; 37: 1-24.
11 Hjerten S. The history of the development of electrophoresis in Uppsala. Electrophoresis 1988; 9: 3-15.
12 Heidelberger M, Pedersen KO, Tiselius A. Ultracentrifugal and electrophoretic studies on antibodies. Nature 1936; 138: 165.
13 Tiselius A, Kabat EA. An electrophoretic study of immune sera and purified antibody preparations. J Exp Med. 1939; 69: 119-31.
14 Longsworth LG, Shedlovsky T, MacInnes DA. Electrophoretic patterns of normal and pathological human blood serum and plasma. J Exp Med 1939; 70: 399-413.
15 Kekwick RA. The serum proteins in multiple myelomatosis. Biochem J 1940; 34: 1248-57.
16 Gutman AB, Moore DH, Gutman EB, McClellan V, Kabat EA. Fractionation of serum proteins in hyperproteinemia, with special reference to multiple myeloma. J Clin Invest 1941; 20: 765-83.
17 Waldenstrom J, Winblad S, Hallen H, Liungman S. The occurrence of serological 'antibody' reagins or similar gammaglobulins in conditions with monoclonal hypergammaglobulinemia, such as myeloma, macroglobulinemia etc. Acta Med Scand 1964; 176: 619-31.
18 Dunn TB. Normal and pathological anatomy of the reticular tissue in laboratory mice with a classification and discussion of neoplasms. J Natl Cancer Inst. 1954; 14: 1281-433.
19 Putnam FW. N-terminal groups of normal gamma globulin and of myeloma proteins. J Am Chem Soc 1953; 72: 2785
20 Fagraeus A. Antibody production in relation to the development of plasma cells. Stockholm: State Bacteriol.Lab. 1948; p.1-127.
21 Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. Australian J Sci 1957; 20: 67-9.
22 Nossal GJV, Lederberg J. Antibody production by single cells. Nature 1958; 181: 1419-20.
23 Potter M, Fahey JL. Studies on eight transplantable plasma-cell neoplasms of mice. J Natl Cancer Inst 1960; 24: 1153-65.
24 Dunn TB. Plasma cell neoplasms beginning in the ileocecal area in strain C3H mice. J Natl Cancer Inst 1957; 19: 371-91.
25 Potter M, Fahey JL, Pilgrim HI. Abnormal serum protein and bone destruction in a transmissable mouse plasma cell neoplasm. Proc Soc Exp Biol Med 1957; 94: 327-33.
26 Fahey JL, Potter M, Gutter FJ, Dunn TB. Distinctive myeloma globulins associated with a new plasma cell strain of C3H mice. Blood 1960; 15: 103-13.
27 Prehn RT, Weaver JM, Algire GH. The diffusion-chamber technique applied to a study of the nature of homograft resistance. J Natl Cancer Inst 1954; 15: 509-17.
28 Merwin RM, Algire GH. Induction of plasma cell neoplasms and fibrosarcomas in BALB/c mice carrying diffusion chambers. Proc Soc Exp Biol Med 1959; 101: 437-9.
29 Freund J. The mode of action of immunologic adjuvants. Adv Tuberc Res 1956; 7: 130-48.

30 Lieberman R, Douglas JOA, Humphrey W, Jr. Ascities induced in mice by Staphylococcus. Science 1959; 129: 775.
31 Potter M, Robertson CL. Development of plasma-cell neoplasms in BALB/c mice after intraperitoneal injection of paraffin-oil adjuvant, heat-killed staphylococcus mixtures. J Natl Cancer Inst 1960; 25: 847-61.
32 Potter M, Boyce C. Induction of plasma cell neoplasms in strain BALB/c mice with mineral oil and mineral oil adjuvants. Nature 1962; 193: 1086-7.
33 Fleischman JB, Pain RH, Porter RR. Reduction of gamma-globulins. Arch Biochem Biophys 1962; (Supp 1): 174-80.
34 Porter RR. The hydrolysis of rabbit gamma-globulin and antibodies with crystalline papain. Biochem J 1959; 73: 119-27.
35 Edelman GM, Poulik MD. Studies on the structural units of the gamma globulins. J Exp Med. 1961; 113: 861
36 Edelman GM, Gally JA. The nature of Bence-Jones proteins. J Exp Med. 1962; 116: 207-27.
37 Lederberg J. Genes and antibodies. Science 1959; 129: 1649-53.
38 Smithies O. Gamma-globulin variability: a genetic hypothesis. Nature 1963; 199: 1231-6.
39 Edman P, Begg G. A protein sequenator. Eur J Biochem 1967; 1: 80-91.
40 McIntire KR, Potter M. Studies of thirty different Bence Jones protein-producing plasma cell neoplasms in an inbred strain of mouse. J Natl Cancer Inst 1964; 33: 633-48.
41 Putnam FW, Easley CW. Structural studies of the immunoglobulins. I. The tryptic peptides of Bence Jones proteins. J Biol Chem 1965; 240: 1626
42 Hilschmann N, Craig LC. Amino acid sequence studies with Bence-Jones proteins. Proc Natl Acad Sci USA 1965; 53: 1403-9.
43 Dreyer WJ, Bennett JC. The molecular basis of antibody formation: a paradox. Proc Natl Acad Sci USA 1965; 54: 864-9.
44 Crick FHC. General discussion on theories of antibody variability. Cold Spring Harbor Symp Quant Biol 1967; 32: 169-72.
45 Tonegawa S, Brack C, Hozumi N, Schuller R. Cloning of an immunoglobulin variable region gene from mouse embryo. Proc Natl Acad Sci USA 1977; 74: 3518-22.
46 Bernard O, Hozumi N, Tonegawa S. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell 1978; 15: 1133-44.
47 Cohn M. Natural history of the myeloma. Quant Biol 1967; 32: 211-21.
48 Tomasz A. Choline in the cell wall of a bacterium: novel type of polymer-linked choline in Pneumococcus. Science 1967; 157: 694-7.
49 Leon MA, Young NM. Specificity for phosphorylcholine of six murine myeloma proteins reactive with Pneumococcus C polysaccharide and beta-lipoprotein. Biochemistry 1971; 10: 1424-9.
50 Schubert D, Roman A, Cohn M. Anti-nucleic acid specificities of mouse myeloma immunoglobulins. Nature 1970; 225: 154-8.
51 Vicari G, Sher A, Cohn M, Kabat EA. Immunochemical studies on a mouse myeloma protein with specificity for certain beta-linked terminal residues of N-acetyl-D-glucosamine. Immunochemistry. 1970; 7: 829-38.
52 Schubert D, Jobe A, Cohn M. Mouse myelomas producing precipitating antibody to nucleic acid bases and-or nitrophenyl derivatives. Nature 1968; 220: 882-5.
53 Landsteiner K. The specificity of serological reactions. New York: Dover Publications, Inc., 1962;
54 Eisen HN, Simms ES, Potter M. Mouse myeloma proteins with antihapten antibody acitivity. The protein produced by plasma cell tumor MOPC-315. Biochemistry 1968; 7: 4126-34.
55 Potter M. Mouse IgA myeloma proteins that bind polysaccharide antigens of enterobacterial origin. Fed Proc 1970; 29: 85-91.
56 Potter M. Antigen-binding myeloma proteins in mice. Ann NY Acad Sci 1971; 190: 306-21.
57 Potter M, Mushinski EB, Glaudemans CPJ. Antigen-binding IgA myeloma proteins in mice: specificities to antigens containing 3-D 1-6 linked galactose side chains and a protein antigen in wheat. J Immunol 1972; 108: 295-300.
58 Potter M. Immunoglobulin-producing tumors and myeloma proteins in mice. Physiol Rev 1972; 52: 631-719.
59 Cotton RG, Milstein C. Letter: Fusion of two immunoglobulin-producing myeloma cells. Nature 1973; 244: 42-3.
60 Wade N. Hybridomas: the making of a revolution [news]. Science 1982; 215: 1073-5.
61 Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 1975; 256: 495-7.
62 Christenson WN, Dacie JV. Serum proteins in acquired haemolytic anaemia (auto-antibody type). Brit J Haemat 1957; 3: 153-64.

63 Riesen WF, Braun DG, Skvaril F, Mansa B. Idiotypic and structural analysis of monoclonal human immunoglobulins with anti-streptolysin O activity. Int Arch Allergy Appl Immunol 1982; 67: 86-92.
64 Casali P, Notkins AL. CD5+ B lymphocytes, polyreactive antibodies and the human B-cell repertoire [see comments]. Immunol Today 1989; 10: 364-8.
65 Ternynck T, Avrameas S. Murine natural monoclonal autoantibodies: a study of their polyspecificities and their affinities. Immunol Rev 1986; 94: 99-112.
66 Morris RJ. Antigen-antibody interactions: how affinity and kinetics affect assay design and selection procedures. In: Ritter MA, Ladyman HM, Eds. Monoclonal antibodies. Cambridge: University Press; 1995. p. 34-59.
67 Dighiero G, Guilbert B, Fermand JP, Lymberi P, Danon F, Avrameas S. Thirty-six human monoclonal immunoglobulins with antibody activity against cytoskeleton proteins, thyroglobulin, and native DNA: immunologic studies and clinical correlations. Blood 1983; 62: 264-70.
68 Chen ZJ, Wheeler J, Notkins AL. Antigen-binding B cells and polyreactive antibodies. Eur J Immunol 1995; 25: 579-86.
69 Potter M. Antigen-binding myeloma proteins of mice. In: Kunkel HG, Dixon FJ, Eds.. Advances in immunology. New York: Academic Press; 1977. p. 141-211.
70 Seligmann M, Sassy C, Chevalier A. A human IgG myeloma protein with anti- 2 macroglobulin antibody activity. J Immunol 1973; 110: 85-90.
71 Hardiman KL, Horn S, Manoharan, et al. Rheumatic autoantibodies in the sera of patients with paraproteins. Clin Exp Rheumatol 1994; 12: 363-8.
72 Grey HM, Kohler PF, Terry WD, Franklin EC. Human monoclonal gamma G-cryoglobulins with anti-gamma-globulin activity. J Clin Invest 1968; 47: 1875-84.
73 Buskila D, Abu-Shakra M, Amital-Teplizki H, et al. Serum monoclonal antibodies derived from patients with multiple myeloma react with mycobacterial phosphoinositides and nuclear antigens. Clin Exp Immunol 1989; 76: 378-83.
74 Zouali M, Fine JM, Eyquem A. A human monoclonal IgG1 with anti-idiotypic activity against anti-human thyroglobulin autoantibody. J Immunol 1984; 133: 190-4.
75 Abu-Shakrah M, Krupp M, Argov S, Buskila D, Slor H, Shoenfeld Y. The detection of anti-Sm-RNP activity in sera of patients with monoclonal gammopathies. Clin Exp Immunol 1989; 75: 349-353.
76 Sestak AL, Harley JB, Yoshida S, Reichlin M. Lupus/Sjögren's autoantibody specificities in sera with paraproteins. J Clin Invest 1987; 80: 138-44.
77 Pereira LF, Marco FM, Boimorto R, et al. Histones interact with anionic phospholipids with high avidity; its relevance for the binding of histone-antihistone immune complexes. Clin Exp Immunol 1994; 97: 175-80.
78 Shoenfeld Y, el-Roeiy A, Ben-Yehuda O, Pick AI. Detection of anti-histone activity in sera of patients with monoclonal gammopathies. Clin Immunol Immunopathol 1987; 42: 250-8.
79 Riesen W, Noseda G, Butler R. Anti-lipoprotein activity of human monoclonal immunoglobulins. Vox Sang 1972; 22: 420-31.
80 Merlini G, Farhangi M, Osserman EF. Monoclonal immunoglobulins with antibody activity in myeloma, macroglobulinemia and related plasma cell dyscrasias. Semin Oncol 1986; 13: 350-65.
81 DiMinno G, Coraggio F, Cerbone AM, et al. A myeloma paraprotein with specificity for platelet glycoprotein IIIa in a patient with a fatal bleeding disorder. J Clin Invest 1986; 77: 157-64.
82 Coleman M, Vigliano EM, Weksler ME, Nachman RL. Inhibition of fibrin monomer polymerization by lambda myeloma globulins. Blood 1972; 39: 210-23.
83 Farhangi M, Osserman EF. Myeloma with xanthoderma due to an IgG lambda monoclonal anti-flavin antibody. N Engl J Med 1976; 294: 177-83.
84 Freedman M, Merrett R, Pruzanski W. Human monoclonal immunoglobulins with antibody-like activity. Immunochemistry. 1976; 13: 193-202.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Tolerance revisited

Leslie B. Brent

Tolerance was shown to be an important biological phenomenon in 1953, when Rupert Billingham, Peter Medawar and I published our first data in *Nature* [1]. The concept of tolerance had already been foreshadowed by Ray Owen in 1945 [2] when he showed that cattle dizygotic twins were red blood cell chimeras (or 'mosaics', to use his expression), and by Frank Burnet and Frank Fenner, who postulated the existence of tolerance in their brilliant speculative 1949 monograph on 'The Production of Antibodies' [3]. In it they rescued Owen's data from oblivion, drew attention to their great potential importance, and speculated on the existence of 'self-markers' on the cells of all individuals that would enable the immune system to distinguish between self and non-self. Later, in 1959, after the experimental verification of tolerance, Burnet made tolerance one of the cornerstones of his clonal selection hypothesis [4] and the clonal deletion of T lymphocytes seemed to offer one important possible mechanism to account for tolerance induction. It is of some interest to me that Mel Cohn, one of the more rigorous theoretical immunologists since Burnet, felt able to discuss the concept of self and non-self at considerable length in his review 'With The Wisdom of Hindsight' [5] without once mentioning the experimental tolerance studies of Medawar's group, let alone any others (with the exception of Owen's). Tolerance continues to be a very live topic and I have devoted two chapters to it in my 'History of Transplantation Immunology' [6]. It is on the threshold of clinical application half a century after its discovery. Here I will take another look at the phenomenon in the light of recent data concerning tolerance mechanisms and discuss it in relation to the concept of self and non-self. The vital question is whether the notion of self-tolerance, built as it is on the body's ability to distinguish between self and non-self, should continue to occupy a central position in immunological theory. To this end I will briefly summarize a few of the salient historical facts.

Some salient facts

It is well known that normal adult mice will reject allogeneic skin grafts in 9–11 days, and this is the usual fate of allografts bearing foreign major histocompatibility molecules (MHC) when transplanted to adult, non-immunosuppressed animals or humans. Thus, adult mammals clearly "can" distinguish between their own histocompatibility molecules (self) and those of other individuals (non-self), and such a subtle act of discrimination is likewise applied to microbial or other extraneous antigens, whether these be pathogenic or non-pathogenic.

By contrast, a mouse inoculated in fetal life or at birth with spleen cells from an allogeneic donor – at a time when its immune system is still relatively immature – will not acquire the capacity to recognize the foreign MHC molecules to which it has been exposed as immunogenic and it will therefore not only fail to reject the cells of the inoculum but display a dramatic inability to reject tissues (or organs) from the same donor or inbred donor strain [1, 7]. There are two important qualifications. First, the inoculated tolerance-inducing cells must be viable, so that they can survive and perpetuate themselves in the tissues of their host, which usually becomes a cellular chimera. Second, they must not be immunocompetent, for if they are they would be able to mount a response against the MHC molecules of the host (a graft-versus-host reaction) [8-10] that can be lethal in both its acute and chronic forms. Experimentally one can circumvent this problem by using cells from F1 hybrid donors carrying both donor and host MHC molecules, or by using cells from a fetus or from adult tissues that are relatively free of immunocompetent T lymphocytes – for example, bone marrow in the case of the mouse. In clinical bone marrow transplantation this continues to be a great hazard, for human bone contains significant numbers of T lymphocytes.

It is significant that tolerance can be induced by a great variety of antigens, from soluble proteins to bacteria and viruses, and even by xenogeneic antigens. However, in the case of soluble proteins it is necessary to administer the tolerogen repeatedly.

Tolerance is individual- or strain-specific, a property that immediately tells us that we are dealing with an immunological phenomenon. Thus, tissue grafts from a third-party strain, unrelated to either the donor or the host strain, will be rejected normally. This was clearly established in the early studies and it has been amply confirmed by numerous other workers. For example, more recently V. Holáň [11] made a special study of it, using a large number of mouse strain combinations, and he was able to confirm that neonatal tolerance is 'absolutely specific'. There have, however, been a few reports suggesting that in addition there may be a nonspecific element in some cases, and I shall return to this presently.

It is relevant that tolerance induction is not confined to mammalian species. Thus, Hašek [12] found that the parabiotic union of two chick embryos prevented the partners from producing a full-blown antibody response against each other's red blood cell antigens, and Billingham et al. [7] extended these observations to the acceptance of skin allografts transplanted between parabiotic partners later in life and to the establishment of erythrocyte as well as white cell chimerism. This represented a faithful reenactment of the situation prevailing in dizygotic cattle twins, which had likewise been shown to accept each other's skin grafts.

On self and non-self

The power of these animals to distinguish between 'self' and 'non-self' could thus be fatally undermined by early exposure to histocompatibility antigens, and I have to ask A.M. Silverstein and N.R. Rose [13]: where is the 'mystique'? Perhaps mystique, like beauty, 'lies in the eyes of the beholder'! Nor can I help wondering whether highly philosophical definitions of self and non-self (e.g., [14]) are necessary or, for that matter, helpful.

One other set of observations embodied in our 1956 [7] paper should be mentioned here. We wanted to ascertain whether tolerance was brought about by a central adaptation within the immune system or by some form of regulatory mechanism. Suppressor T cells had not yet been described and the most likely contender at that time would have been the 'enhancing antibodies', so named because these antisera directed against allogeneic MHC molecules can specifically enhance the growth of allogeneic tumors (see [15]). We therefore took fully tolerant mice and inoculated into them either normal or specifically presensitized lymphoid cells that were syngeneic with the recipient strain cells that, by definition, would survive and reestablish themselves in their adoptive hosts. Our reasoning was that if immunoregulatory mechanisms were involved, the freshly injected cells would come under the influence of such mechanisms and the tolerance would continue unabated. Conversely, if tolerance had been brought about by a central adaptation of the immune system, possibly by the adaptation of cells of the lymphocyte pool, one could reasonably expect the grafts to be destroyed by the freshly introduced non-tolerant cells. The grafts were in fact destroyed within 10–12 days when the cells had been presensitized and within about 3–4 weeks when normal cells were employed. The answer could not have been clearer, then: tolerance was caused by some form of central and highly specific inhibition or inactivation and not by peripheral suppressive mechanisms [7].

This information was avidly seized upon by Burnet when he later enunciated his Clonal Selection Hypothesis [4]. (The way had already been pointed by others such as N. Jerne, D.W. Talmage and J. Lederberg (see [6], p. 27-30). The hypothesis provided an insight into the possible mechanism underlying the development of tolerance, i.e., the removal from the host's pool of lymphocytes of the specific clone carrying receptors for the donor's MHC molecules. It should be stressed that Medawar's group (and Burnet!) had been lucky in that we had fortuitously chosen for our experiments two strains of mice that were many years later shown to differ only for MHC class I molecules. Had we chosen class II-disparate strains we might well have discovered suppressor T lymphocytes some 20 years before Gershon and Kondo, for in such a situation suppressor T cells rather than clonal deletion provide the key to tolerance induction and maintenance; yet crucial evidence for the clonal selection hypothesis would not have become evident.

In their 1953 paper [1] Billingham et al. stated that Burnet and Fenner's prediction had been proved right. We drew attention to the fact that we had 'no reason to doubt that at least some of the cells of the foetal inoculum survived as long as the tolerant state itself,' and that 'any complete theory of antibody formation must be competent to explain' the salient features of tolerance. At the same time we drew attention to the earlier and seminal studies of R.D. Owen and alluded to several other phenomena involving specific unresponsiveness in adult animals. Acquired immunological tolerance, as we called it, was 'due to a specific failure of the host's immunological response'.

A comprehensive account was presented in our 1956 'Philosphical Transaction' monograph [7]. Tolerance was now defined as 'the outcome of an induced specific central failure of the mechanism of immunological response brought about by the exposure of animals to antigenic stimuli before the maturation of the faculty of immunological response', a definition that had to be adapted some years later when it became apparent that even adult animals could be subverted into tolerance, though here the use of irradiation, drugs or alloantibodies is usually mandatory.

Burnet and Fenner's predictions had been shown to be correct, Billingham et al. went on to say:

> [For] their argument applies with particular force to the specialized and complex substances that are formed in the later stages of cellular differentiation. Many such substances could be antigenic were it not for the fact that the future antibody-forming system is exposed to their influence at a sufficiently early stage.

Exceptions cited included lens proteins (which are sequestered from the developing immune system), and spermatozoa and milk casein (which are formed too late for tolerance induction to have occurred whilst at the same time not being readily accessible to the immune system). Billingham et al. concluded that 'there may well exist a wide variety of bodily constituents which are potentially autoantigenic, and therefore isoantigenic, simply because the antibody-forming system has no normal opportunity to become tolerant of their action'. However, because isoantigens are present in all nucleated cells, including those of the immune system, they do not fall into this category: 'a future antibody-forming cell needs to learn not to react against substances which are part of its own fabric'. (It is of some interest that Billingham et al. at this stage, in 1956, still thought of immune responses in terms of antibody formation rather than cellular reactivity, even though they had already provided evidence for the participation of lymphocytes. The field of cellular immunology was only just about to develop.)

Likewise, "self" and "non-self" was not at that time part of the immunological vocabulary. Yet in the 1956 paper we suggested the hazards of autoimmune responses as the *raison d'être* for tolerance, a theme that Medawar and I developed at some length 2 years later [16].

Once the clonal selection hypothesis had been developed it was widely assumed that tolerance induction involved the deletion of specific clones of lymphocytes (T or B). It proved to be true for some murine strain combinations but not for others (see above). Thus it became apparent that in strain combinations differing for the whole of the MHC spectrum, or only for class II molecules, suppressor T lymphocytes could be identified in the peripheral tissues, and this also turned out to be true for many models of tolerance induction in adults. Peripherally-induced tolerance can clearly play an important part in regulating immune responses.

Anyone familiar with the tolerance literature over the last two decades will know that there is a strong case to be made for the participation of several regulatory mechanisms and that these need not be mutually exclusive. The prevention of autoimmune disease is patently of such vital importance to the survival of any one species that a variety of mechanisms evolved: clonal deletion, suppressor T lymphocytes, anergy, veto cells and others. Even P. Matzinger, before her 'danger' period, presented evidence with A. Bonomo for negative selection of T lymphocytes by the dendritic cells of the thymic epithelium, the selection process appearing to depend largely on the affinity that T cell receptors have for self-proteins – cells with low affinity escaping into the periphery [17].

There are other possible explanations for negative selection and I will confine myself to one other put forward by S. Schneider and N.A. Mitchison [18]. They described self-reactive T cell hybridomas that recognized a self-protein present in the liver but only at a high concentration, higher than that needed for allo-recognition. They suggested that, in addi-

tion to –ve selection in the thymus and other forms of peripheral regulation, the frequency of receptor-bearing cells may be sufficient to determine whether a population of T cells with the potential for self-reactivity are switched on. They did not, however, envisage such a mechanism as doing away with the need for –ve selection based on receptor affinities; rather, it was seen as an auxiliary mechanism enabling the immune system to function with a form of –ve selection that is not wholly rigorous.

It is not my brief to argue the case here for peripheral regulatory mechanisms. I would, however, like to make two points. Jerne's anti-idiotypic network theory [19] created much excitement when it was first enunciated some two decades ago but interest in it has, I think, waned. Suppressor T lymphocytes undoubtedly exist, at least at cell population level, as numerous studies have shown. My group was among the first to extend R.K. Gershon and K. Kondo's groundbreaking studies to the transplantation arena: the specific tolerance induced in adult mice by pre-treatment with antigenic tissue extracts and transient immunosuppression could be reliably transferred to naïve mice with purified T cells [20].

Many similar demonstrations of the power of T cells to transfer tolerance can be found in the literature. Whether they play a role in self-tolerance is still a moot point; they have certainly been identified as a regulatory mechanism after experimental and clinical infections and apparently they help to curtail the immune response when it is no longer required. Suppressor T lymphocytes have been derided by molecular immunologists because gene rearrangements and a specific phenotypic profile have not as yet been convincingly demonstrated. Instead, the anergic cell has been favoured by them – a cell that is physically present but unable to respond to specific antigens. Nonetheless, now that H. Waldmann's [21] and R. Lechler's [22, 23] groups have shown that anergic T cells can transfer unresponsiveness and that this is brought about by their ability to suppress the responses of naïve T cells, the distinction between anergy and suppression clearly has become semantic; anergic cells that transfer unresponsiveness can, in my view, be reasonably described as suppressor cells, though some may prefer the term 'immunoregulatory'.

That thymic deletion and peripheral tolerance mechanisms are not mutually exclusive has been shown by A. Lanoue et al. [24]. They revealed that in transgenic mice bearing a class II-restricted T cell receptor for the hemagglutinin of the influenza virus some mature T cells that had escaped thymic deletion in antigen-containing recipients were first weakly activated and then became anergic.

Another interesting mechanism of self-tolerance, in this case of B lymphocytes, is suggested by the study of Rathmell et al. [25]. Mice deficient in CD95 (Fas/APO-1) and Fas-ligand developed an autoimmunity resembling human systemic lupus erythematosus (SLE) and these workers studied the response to a transgene encoded autoantigen, hen egg lysozyme (HEL), using cells from mice transgenic for immunoglobulin and T cell receptor genes. Naïve B cells were triggered by HEL after interaction with antigen and HEL-specific T cells. B cells chronically exposed to HEL during their development were anergic and did not produce antibody; indeed, they were eliminated in the presence of HEL-specific CD4+ve T cells. CD95-deficient anergic B cells were, on the other hand, not eliminated and triggered to proliferate. Thus the elimination of auto-reactive B cells depended on the presence of CD95 – a novel immunoregulatory mechanism.

Negative selection in the thymus and peripheral mechanisms therefore need not be in conflict with each other. Deletional thymic mechanisms are clearly of the greatest importance – how else would one make sense of the demonstration, originally made by B. Waksman and his colleagues in the 1960s [26] and more recently confirmed and extended [27], that the inoculation of antigens directly into the thymus of adult rodents is an excellent way of inducing systemic tolerance? Nevertheless, there have to be peripheral "fail-safe" mechanisms if autoimmune disease is to be a rare rather than a common event. I am attracted to the notion developed by B. Arnold and his colleagues [28] that peripheral tolerance is a stepwise process driving T cells progressively into deeper levels of unresponsiveness.

I would like to make two other points relating to transplantation tolerance. First, there is the by now well-known observation by T.E. Starzl's group that many if not all human long-term liver and kidney transplant recipients have a low level of donor cell chimerism in their tissues [29]. According to them it is the establishment of this micro-chimerism that drives the immune system into tolerance and they believe that tolerance is the outcome of an interaction between an anti-graft response by the host and a graft-versus-host reaction by the donor cells – the "two-way" paradigm [30]. This proposition is still disputed by some who believe that chimerism is the consequence rather than the cause of tolerance, and there are certainly experimental models, such as my own [31], in which a long-lasting tolerance was induced by methods that would make micro-chimerism highly unlikely.

The second point concerns the specificity of tolerance, which has generally been found to be absolute (see above) – so much so that it has been axiomatic that "all" MHC antigens present in a test graft have to be represented in the cells of the tolerance-inducing inoculum. (This is quite different from the process of sensitization, where a single epitope difference can be enough to trigger a response against a transplant.) The data from K. Wood's laboratory [32] are therefore of special interest, for they seem to be turning the axiom on its head. They found that, with the help of a short course of anti-CD4 antibody, recipient-type bone marrow cells transfected with a gene coding for a single donor class I MHC molecule could induce tolerance in adult mice to fully allogeneic cardiac allografts. Surprisingly, therefore, a single epitope induced tolerance across the full MHC spectrum.

The mechanism of this 'linked epitope suppression' is still far from clear though the Oxford group has suggested that the single epitope acts as a 'key' that will override responsiveness to other antigens present in the graft, possibly thanks to some form of dominance [33]. It is believed to involve either direct contact between the epitope-tolerant cells and naïve cells or else the action of localized cytokines. The phenomenon was first described in 1983 by V. Holáň and N. A. Mitchison, who believe that haplotype-specific suppressor T cells mediate linked suppression of responses elicited by third-party alloantigens [34]. H. Waldmann's group has reported similar findings [35]. From the point of view of tolerance induction to human cadaveric organ transplants the concept is very beguiling, but it remains to be seen whether the murine model can be translated to large and outbred animals such as non-human primates.

Back to self/non-self

This topic was given a good airing in the 1997 volume of 'Immunological Reviews' [36], following the provocative contributions of P. Matzinger and her colleagues (see below).

A debate was sparked by a paper by A.M. Silverstein and N.R. Rose who, whilst disagreeing with Matzinger's danger hypothesis, felt that the old notion of self/non-self as originally conceived by Burnet was outdated and in need of considerable revision [13]. They expressed the view that the concept of the immunological self was wrapped in an undesirable and obscurantist aura of "mystique." In their comments on this paper, published in reference [36], several immunologists (D.W. Talmage, L. Brent and R.E. Langman and M. Cohn) have rebutted some of the arguments used by Silverstein and Rose and little point would be served by going over the same ground here. I will, however, repeat part of my conclusion: '...they [Burnet's 'self-markers'] have in fact become more rather than less plausible in the light of what we now know about histocompatibility molecules, peptide presentation and selection processes – both negative and positive – within the thymus [37].'

Finally, I must comment on the work and theoretical contributions of P. Matzinger. She developed a theoretical framework according to which lymphocytes require not only two signals, as suggested long ago by P. Bretscher and M. Cohn [38], but a third – the danger signal – produced by potentially pathogenic organisms or molecular structures that could be harmful to the organism [39]. Self/non-self discrimination and the built-in notion of self-tolerance would thus become superfluous. In a more recent experimental paper her group believes to have found supporting evidence [40]. This is based on their finding that tolerance induction to the male-specific minor antigen encoded by the Y chromosome could be prevented in the majority of 7-day-old female mice by providing the recipients soon after birth with syngeneic male dendritic cells. To conclude from this that 'neonatal and adult T cells have similar options', and that 'tolerance occurred not because the neonate is inherently tolerizable...' is entirely unsafe. What Ridge et al. are saying, in essence, is that there is nothing very special about the neonatal or fetal period in the mouse, as claimed by Billingham, Brent and Medawar [7]. They ignore the fact that: 1) their manipulation of the neonatal immune system was artefactual; 2) the antigen utilized by them was relatively weak compared with MHC antigens; 3) M. Malkowsky and P.B. Medawar observed almost two decades before them that the neonatal immune system can be manipulated, i.e., matured, by the inoculation of IL-2 into newborn mice, thus subverting the induction of tolerance [41] – a finding that these workers did not find at all at odds with the special nature of the undeveloped immune system so far as tolerance induction is concerned; 4) using the full spectrum of MHC differences (H-2), L. Brent and G. Gowland demonstrated this special nature in 1961 with their finding that the post-natal tolerance window closed within a few days of birth even when the cellular antigen dose was adjusted to take into account the rapid increase in body weight [42]; 5) tolerance induction can indeed occur in young "mature" mice (2 weeks) by the inoculation of repeated intravenous weight-adjusted doses of allogeneic F1 hybrid spleen cells [43]; and, finally, 6) tolerance to non-MHC, minor histocompatibility antigens can be induced in fully adult mice either by a fairly high intravenous dose of spleen cells or even by the application of a single large skin graft [44, 45]. There is, therefore, nothing new in claims that the neonatal immune system can be manipulated towards sensitivity or tolerance and the known findings can be readily accommodated by the self/ non-self theory. One of the questions Matzinger et al. will have to address is how the immune system can so readily respond to a plethora of

antigens that cannot conceivably pose a threat to the organism, not least among them the MHC and minor histocompatibility antigens of allogeneic organs such as kidneys or livers.

Whilst it could be argued that I have a vested interest in the continuation of tolerance and its *raison d'être* in preventing autoimmune disorders, I do not believe that either Matzinger's group or, in their different way, Silverstein and Rose have made a convincing case for a major re-assessment. The fact that negative selection allows some cells to slip through the net does not do away with the necessity for the elimination of potentially autoreactive lymphocytes. That elimination must clearly depend on the recognition, by the developing immune system, of what is self. Those who argue that there is nothing special about the embryonic or neonatal environment of the mouse are blind to the fact that there is an absence (or, in the case of the neonatal mouse, a relative absence) of mature lymphocytes, and that this provides a special milieu for the tolerization of T cells as they mature. That is why, in adult animals, we have to employ highly complex and potentially harmful strategies to reduce the adult immune system to a quasi-embryonic or neonatal level. Silverstein and Rose [13] regard the route, quantity and physico-chemical nature of the antigen, as well as the context in which it is presented (e.g., adjuvanticity), as paramount in determining the nature of the immune response; had they added 'and the presence or absence of mature T lymphocytes and antigen-presenting cells' I could have agreed with them more readily.

Conclusion

The concepts of tolerance and self/non-self remain essentially unscathed by recent attempts to replace them with new and superficially attractive notions. That is not to say that they may not need be amended in the light of compelling new evidence.

Bibliography

1 Billingham RE, Brent L, Medawar PB. 'Actively acquired tolerance' of foreign cells. Nature 1953; 172: 603-6.
2 Owen RD. Immunogenetic consequences of vascular anastomoses between bovine twins. Science 1945; 102: 400-1.
3 Burnet FM, Fenner F. The production of antibodies. Melbourne: Macmillan; 1949.
4 Burnet FM. The clonal selection theory of acquired immunity. Cambridge: Cambridge Univ. Press; 1959.
5 Cohn M. The wisdom of hindsight. Annu Rev Immunol 1994; 12: 1-62.
6 Brent L. A history of transplantation immunology. London: Academic Press, 1997.
7 Billingham RE, Brent L, Medawar PB. Quantitative studies on tissue transplantation immunity. III. Actively acquired immunity. Phil Trans Roy Soc Lond 1956; 239: 357-414.
8 Billingham RE, Brent L. A simple method for inducing tolerance of skin homografts in mice. Transpl Bull 1957; 4: 67-71.
9 Simonsen M. The impact on the developing embryo and newborn animal of adult homologous cells. Acta Path Microbiol Scand 1957; 40: 480-500.
10 Billingham RE, Brent L. Quantitative studies on transplantation immunity. IV. Induction of tolerance in newborn mice and studies on the phenomenon of runt disease. Phil Trans Roy Soc Lond 1959; 242: 439-77.
11 Holáň V. Absolute specificity of neonatally induced transplantation tolerance. Transplantation 1990; 50: 1072-4.
12 Hašek M. Vegetative hybridisation of animals by joining their blood circulation during embryonic life. CS Biol 1953; 2: 265-77.
13 Silverstein AM, Rose NR. On the mystique of the immunological self. Immunol Rev 1997; 159: 197-206.
14 Tauber AI. The immune self: theory or metaphor? Cambridge: Cambridge Univ. Press; 1949.
15 Batchelor JR. The use of enhancement in studying tumor antigens. Cancer Res 1968; 28: 1410-4.

16 Brent L, Medawar PB. Tolerance and autoimmune phenomena. In: Tunevall G, Ed. Recent progress in microbiology. Stockholm: Almqvist and Wiksell; 1959. p. 181-90.
17 Bonomo A, Matzinger P. Thymus epithelium induces tissue-specific tolerance. J Exp Med 1993; 177: 1153-64.
18 Schneider SC, Mitchison NA. Self-reactive T cell hybridomas and tolerance. J Immunol 1995; 154: 3796-805.
19 Jerne N. Towards a network theory of the immune system. Ann Immunol 1974; 125C: 373-89.
20 Kilshaw PJ, Brent L, Pinto M. Suppressor T cells in mice made unresponsive to skin allografts. Nature 1975; 255: 489-91.
21 Quin S, Cobbold SP, Waldmann H. 'Infectious' transplantation tolerance. Science 1993; 259: 974-7
22 Lombardi G, Sidhu S, Batchelor JR, Lechler R. Anergic T cells as suppressor cells in vitro. Science 1994; 264: 1587-9.
23 Chai JG, Bartok I, Chandler P, Vendetti S, Antoniu A, Dyson J, et al. Anergic T cells act as suppressor cells in vitro and in vivo. Eur J Immunol 1999; 26: 686-92.
24 Lanoue A, Bona C, von Boehmer H, Sarukhan A. Conditions that induce tolerance in mature CD4+ T cells. J Exp Med 1997; 185: 405-14.
25 Rathmell JC, Cooke MP, Ho WY, Grein J, Townsend SE, Davis MM, et al. CD95 (Fas)-dependent elimination of self-reactive B cells upon interaction with CD4+ T cells. Nature 1995; 376: 181-4.
26 Horiuchi A, Waksman BH. Role of the thymus in tolerance. VI. Tolerance to bovine gamma-globulin in rats given a low dose of irradiation and injection of non-aggregated or aggregated antigen into the shielded thymus. J Immunol 1968; 100: 974-8.
27 Posselt AM, Barker CF, Tomaszewski JE, Markmann JF, Choti MA, Naji A. Induction of donor-specific unresponsiveness by intrathymic islet transplantation. Science 1990; 249: 1293-5.
28 Arnold B, Schönrich G, Hämmerling GJ. Multiple levels of peripheral tolerance. Immunol Today 1993; 14: 12-4.
29 Starzl TE, Demetris AJ, Murase N, Ildstad S, Ricordi C, Trucco M. Cell migration, chimerism, and graft acceptance. Lancet 1992; 339: 1579-82.
30 Rao AS, Starzl TE, Demetris AJ, Trucco M, Thomson A, Quian S, et al. The two-way paradigm of transplantation immunology. Clin Immunol Immunopath 1996; 80: S46-51.
31 Kilshaw PJ, Brent L. Prolongation of skin allograft survival with spleen extracts and anti-lymphocytic serum. Nature 1970; 227: 898-900.
32 Madsen JC, Superina RA, Wood KJ, Morris PJ. Immunological unresponsiveness induced by recipient cells transfected with donor MHC genes. Nature 1988; 332: 161-4.
33 Wong W, Morris PJ, Wood KJ. Pretransplant administration of a single class I MHC molecule is sufficient for the indefinite survival of fully allogeneic cardiac allografts: evidence for linked epitope suppression. Transplantation 1997; 63: 1490-4.
34 Holáň V, Mitchison NA. Haplotype-specific suppressor T cells mediating linked suppression of immune responses elicited by third-party alloantigens. Eur J Immunol 1983; 13: 652-7.
35 Davies JD, Leong LYW, Mellor A, Cobbold SP, Waldmann H. T cell suppression in transplantation tolerance through linked recognition. J Immunol 1996; 156: 3602-7.
36 A postscript on the immunological self. Immunol Rev 1997; 159: 197-218.
37 Brent L. Commentary on Silverstein and Rose "On the mystique of the immunological self." Immunol Rev 1997; 159: 211-3
38 Bretscher P, Cohn M. A theory of self-nonself discrimination. Science 1970; 169: 1042-9.
39 Matzinger P. Tolerance, danger and the extended family. Annu Rev Immunol 1994; 12: 991-1045.
40 Ridge JP, Fuchs EJ, Matzinger P. Neonatal tolerance revisited: turning on newborn T cells with dendritic cells. Science 1996; 271: 1723-6.
41 Malkowsky M, Medawar PB, Thatcher DR, Toy J, Hunt R, Rayfield LS, et al. Acquired immunological tolerance of foreign cells is impaired by recombinant interleukin 2 or vitamin A acetate. Proc Natl Acad Sci U S A 1985; 82: 536-8.
42 Brent L, Gowland G. Cellular dose and age of host in the induction of tolerance. Nature 191; 192: 1265-7.
43 Brent L, Gowland G. Induction of tolerance of skin homografts in immunologically competent mice. Nature 1962; 196: 1298-301.
44 McKhann CF. Transplantation studies of strong and weak histocompatibility barriers in mice. II. Tolerance. Transplantation 1964; 2: 620-6.
45 Berrian JH, McKhann CF. Transplantation immunity involving the H-3 locus: graft survival times. J Natl Cancer Inst 1960; 25: 111-23.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Logic of the Self-NonSelf discrimination: principles and history

Melvin Cohn

When I read the history of science, I often think how much more satisfying it would be to learn from it the optimal pathway to choosing and solving problems. This does not mean that I am uninterested in the gossipy side of history, the priority squabbles, the romance of arriving at eternal truths by taking a hot bath or a stroll in the night air of Copenhagen (frankly, I have tried both, unsuccessfully). The dark side of me loves all that, but as a scientist I know that we should take the personality out of science when we seek understanding. Historians, as well as scientists-turned-historians, love to put personality back into science and more than often trace a sentimentalized pathway of discovery of little value as a map of the roadway to understanding.

It is in the nature of scientific endeavor that each individual's contribution is assimilated into a body of knowledge with which the next generation works. Eventually this body of knowledge composed of accepted postulates stands as unique in that no heuristic competing set of postulates exists. Unlike composers, novelists, poets, and artists, the scientists who contributed to this armamentarium of postulates are dèstined to anonymity. Present day scientists work from a body of information held together and ordered by conceptualization. It is not ordered by author and opus number as would be Mozart's symphonies or a Bartok's quartets. Contemporary scientists seem to function adequately with no long range knowledge of who did what in the history of their subject. Today's scientists start with a body of coherent inherited knowledge that serves as a foundation upon which they build. This body of knowledge has assimilated that which is sense in the products of past workers and has winnowed out the nonsense. In science this is called "progress" and the price paid for progress is the rapid eclipse of its participants into anonymity.

The more senior contemporary immunologists write history inevitably influenced by trying to keep their own contributions from dissolving in this sea of anonymity. Historians of immunology write history valiantly trying to extract from the sea of anonymity its component contributors. As this is very difficult for them to do, they often rely on the value judgments of others, for example, prize-awarding committees, the Nobel being the most influential, or the Citation Index in one form or another.

The "scientist-turned-historian" decries the present day bench scientist's ignorance of the historical background of their subject for such reasons as:

1) a knowledge of history will decrease the number of cases of rediscovery of the wheel (decrease the waste of repetition); or

2) history is of interest in its own right, 'a source of wonder and of insight'.

However, for me the so-called "rediscovering the wheel" is most often salutary because the repeat of the past increases confidence in the observation and, in any case, often looks at it from a higher level of understanding. We don't reinvent in a circle, we rediscover in a helical path, a spiral staircase that permits us to look at the past down below from a higher plane. For example, Silverstein criticizes the authors of the three papers in *Science* (1997) heralded as putting "immunology on the brink of danger," because they simply reinvented the wheel. This may be true (albeit discussible, because they really reinvented the helix) but for what is that an argument? Why should I care? I want to know if these authors were justified in their interpretation of the totality of the data, and for what reasons. This is what should be revealed by the history of the concept that the evolving data purported to support? As a general rule "the reinventing of the wheel" argument means that there is agreement on the interpretation and disagreement on the priority.

'Interest' and 'wonder' are sources of pleasure to the select few who, like myself, enjoy history, as though it were a novel, but the key word 'insight' ends up as a sound bite never analyzed or illustrated.

I would like history to clarify 'insight' by uncovering the process by which creative individuals from the past arrived at their understanding. How are concepts generated and established? This has to include the why of the failure as well as the why of the success in arriving at a general law. If I could extract from the past the laws of creativity, I would be a better scientist. Terribly naive you say! True! But, then, why should I pay attention to history? The annual reviews should be a sufficient source of wonderment.

Klausner has remarked that, "[t]here are many good historians but precious few good prophets." This is why I called my only incursion into history "The Wisdom of Hindsight." [1]

When I look back into the immunology that I have lived, I see many central facts that should have been so obvious in hindsight that I cannot see why they were not predicted. All of the arguments one makes today for their a priori necessity existed long before the experimentalism stumbled on to them, and, even then, often from an unrelated study.

For example, why did nobody predict:

– the joining of V to C. It awaited Hilschmann's 1963 demonstration by sequencing;

– the role of developmental time in the Self-NonSelf (S-NS) discrimination. It awaited Owen's 1945 chimera observation;

– the requirement for the T cell to recognize a peptide-restricting element (MHC [Major Histocompatibility Complex]) complex. It awaited a family of experiments from biochemists, X-ray crystallographers and physical chemists trying to determine what was the ligand for the MHC restricted T cell antigen-receptor.

Then there are the correct predictions made for the wrong reasons, like Talmage-Burnet's haplotype exclusion (one cell one antibody), or Jerne's germ line encoding of the recognition by the T cell receptor (TCR) of allele-specific determinants on restricting elements. These end up as being derouting fortuities because they make more difficult the acceptance of the correct conclusion for the right reasons.

There are two to three articles in immunology published every hour, roughly one every 20 minutes. If there is such a thing as "information overload," this is it. Each publication contains an observation that the author feels fits into the jigsaw puzzle. However, with no idea of an emerging picture these pieces are lost or ignored or forced into place in the

wrong pattern. Information becomes useful knowledge only when it is coherently cemented together by a valid conceptual framework.

It is usually argued in answering "Why history?," that one must know the past to understand the present. I have found in examining the past that the converse is also true. One must understand the present in order to interpret the past. If we want to plot the pathway from the past to the present we must know what the present is. Or even better, we must have a vision of the future in order to extrapolate back to what transpired in the past. This thought guides the structure of this essay.

Is there something we can learn from history about the process of understanding and can we optimize this process?

Historians of immunology are quick to tell us 'what to think', but their products rarely tell us 'how to think'. Maybe they feel that this is not the province of history or that there is no answer or too many answers to the 'how to think' question. Whatever their answer, by not even posing the question, the historians of science fail us and leave the problem to ill-equipped dabblers like myself who search for the 'insight' heralded to be extractable from past experience.

Evolution and understanding are both historical processes. This defines two aspects of immunology that I will develop. The first will put 'immunology' in its evolutionary context by asking what was the pathway of sequential stepwise events that led to the appearance of such a unique function? The second will put 'understanding' in its historical context by asking what was the pathway of sequential selection on the theories that led to our present day understanding of immunology.

This will encourage us to seek the most general laws that encompass the largest segment of information. Those who revealed important truths did just that; those who used inductive extrapolation from the body of empirical observation had a checkered success and those who extrapolated by analogy with other systems (e.g., nervous) or from metaphysical, sociological or intuitive postulates, or those who were guided by elegance and parsimony, were almost always misled.

The criteria for analyzing past accomplishments should be the same as the criteria for analyzing present accomplishments. Rewriting these criteria because the creative event, experimental or conceptual, occurred many years ago is both unproductive and misleading.

I will illustrate this using the S-NS discrimination as my subject. It is a fundamental problem that cannot be dismissed as metaphor or fustian as is so popular today. An understanding (conceptualization) of the S-NS discrimination is the underpinning for immunobiology. How you think about the S-NS discrimination determines how you view all other aspects of immunology.

In order to develop the problem I decided to return to basics. My plan is simple; it has two parts:

In part I I will place the immune system in a broad biological context. Any attempt at constructing a history of a subject presupposes a knowledge of the subject. Summarizing this is the goal of Part I.

In part II I will then trace the history of my present understanding of the mechanism of the S-NS discrimination in the light of this knowledge.

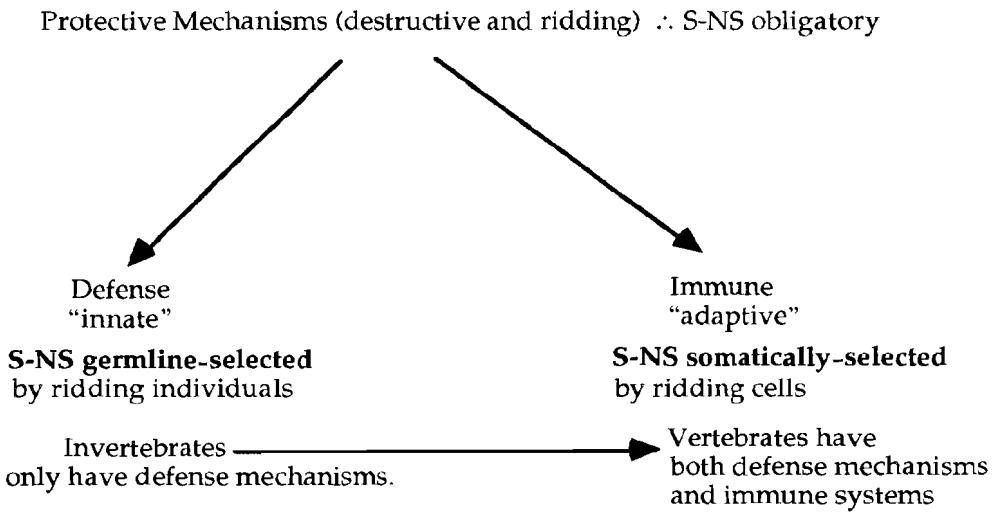


Figure 1. The comparison of defense mechanisms with immune systems.

The framework of the analysis

What is an immune system?

All living creatures have mechanisms to protect themselves against parasitism. For example, bacteria have restriction enzymes that cleave the DNA of invading viruses but do not cleave the host DNA. Insects have an armamentarium of lytic peptides that kill invading bacteria and fungi but do not kill host cells. Plants have a chemical warfare potential that enables the ridding of pathogens without attacking host cells.

I refer to these mechanisms of protection used by invertebrates as "defense mechanisms." Immunologists prefer to call them "innate" immune systems, a term to which I would have no objection provided the meaning of "innate" were clear. The term "innate" means that the recognitive repertoire and, consequently, the S-NS discrimination is 'germ line-selected'. The germ line selection is on the recognitive specificity, which must be sufficient to interact effectively with the selecting target (e.g., a pathogen) to the exclusion of the host species (the S-of-the-species).

Vertebrates also have defense mechanisms but, in addition, they have immune systems. Here immunologists like to refer to "adaptive" immune systems to contrast them to "innate" immune systems. If "adaptive" is taken to mean that the recognitive repertoire is somatically generated and, as a consequence, the S-NS discrimination must be 'somatically selected', then the terms "innate" and "adaptive" are usable by me. Since this is not the case, I will use the terms "defense mechanisms" and "immune systems" based solely on this difference, a germ line-selected versus a somatically-selected S-NS discrimination (*figure 1*).

Whether "defense" or "immune," a S-NS discrimination is obligatory!

These devices, protective against parasitism, have a biodestructive and ridding effector output. If this output operated indiscriminately on host and pathogen, the host would be destroyed. This is what makes a S-NS discrimination obligatory.

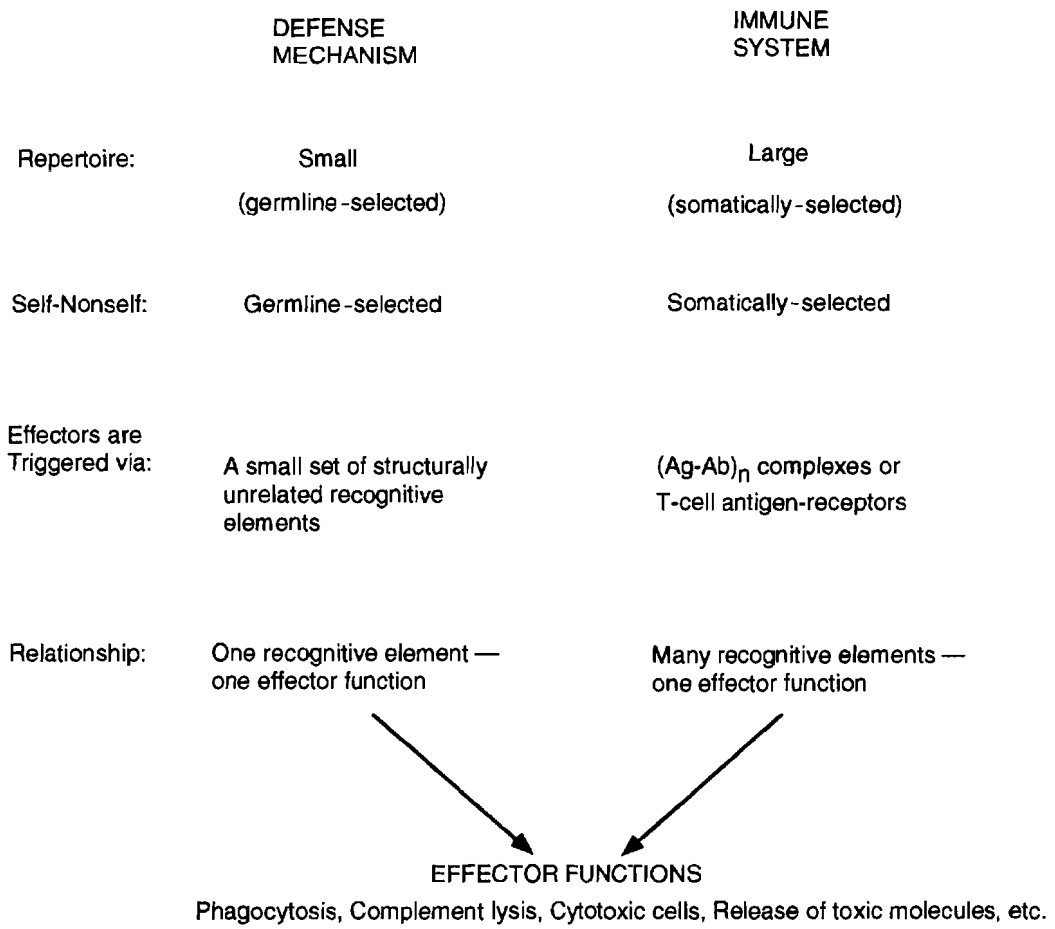


Figure 2. The relationship of the recognitive elements of defense mechanisms and immune systems to effector function.

'The S-NS discrimination is the mechanism by which a biodestructive and ridding effector function is brought to bear on a "pathogen" without debilitating the host'.

From where did immune systems arise?

The invertebrates early in evolution put into place most of the effector mechanisms used by today's immune systems (*figure 2*). The vertebrate immune system, in large measure, hijacked these mechanisms by coupling them to a somatically-selected larger recognitive repertoire. A general statement might be that there is 'one recognitive element–one effector function' in the case of defense mechanisms, whereas there are 'many recognitive elements–one effector function' in the case of immune systems. When these effector mechanisms are triggered by a defense mechanism pathway that does not involve the recognitive repertoire of the immune system, immunologists refer to this as the "alternate" pathway. In other words, the effector output can be triggered either by interactions of

pathogen with the antigen-receptors of defense mechanisms (i.e., alternate pathway) or of immune systems [2-4].

What was the selection pressure for adding immune systems to the existing armamentarium of defense mechanisms?

As invertebrates live in relatively defined ecological niches, are cold-blooded and short-lived, a stable steady state is set up between the host and its small pathogenic load. Under these conditions defense mechanisms are adequate. On the one hand, it is of no evolutionary advantage for the pathogen to kill all hosts as that would lead to self-destruction. On the other hand, defense mechanisms dependent on genomic evolution cannot outrun the potential of the pathogenic load to escape by mutation. A steady state is established between the pathogenic load and the small recognitive repertoire of defense mechanisms. There is no selective advantage for the host to increase the size of the repertoire when it is no longer limiting to the survival of most of the population through the age of procreation. Other factors become limiting like the probability of being eaten by a predator. Further, there is no selective advantage for the pathogen to debilitate or kill more hosts as that limits its own survival.

However, as vertebrates evolved to be land animals that were long-lived, warm blooded and wandered over vast areas, they encountered a much larger variety of "pathogens." This made the recognitive repertoire of defense mechanisms inadequate and set up the selection pressure for the evolution of the qualitatively different recognitive repertoires of immune systems. This immune repertoire coupled its output to the effector functions of defense mechanisms providing them with the ability to be triggered by a much larger variety of different pathogens and at a much lower concentration of antigen, a precious fringe benefit as pathogens grow faster at the body temperature of warm blooded animals.

The recognitive repertoire of defense mechanisms is selected in the germ line by ridding individuals. The S-NS discrimination is dependent on this repertoire recognizing determinants on "pathogens" that are absent in the host species. For example, a restriction enzyme recognizes a DNA sequence present in an invading virus but absent in the host species. The selection is such that any individual expressing a target recognized also to be on the pathogen would be eliminated.

In order for a germ line-selected recognitive repertoire to track a pathogen, the determinant being recognized by the repertoire must vary slowly enough. There is no way, for example, that genomic evolution could track a virus escaping recognition by mutation of a coat protein. A typical target determinant of a germ line-selected recognitive element would be a carbohydrate. In order to change the structure of a carbohydrate so that it escapes recognition, the specificity of a synthesizing enzyme must be changed and this is sufficiently rare to permit tracking of the carbohydrate during genomic evolution. The repertoire of defense mechanisms is small and unique for each target driving evolutionary selection. The repertoire of an individual (equivalent to that of the species) might be made up of a lytic enzyme, a lectin, an antibiotic peptide, and a necrosis factor.

The only way to establish the large recognitive repertoires of immune systems was to develop a receptor structure that recognized shape, not chemistry. The repertoire of recognitive combining sites or paratopes divided the antigenic universe into 'epitopes' (units of shape recognized by paratopes) that are distributed combinatorially on antigens. An

antigen defined as a 'unit of elimination' by the effector output, is made up of a collection of linked 'epitopes'. The paratopes were coupled to an adaptor structure that interacted with the pre-existing effector mechanisms of defense systems converting them into all inclusive immune systems. This, in turn, required that the S-NS discrimination be somatically learned by ridding cells. As a result, the S-NS discrimination became a property of the individual, not of the species. Tissue grafts between individuals of the species are rejected by immune systems, unlike the case for defense mechanisms where such grafts are accepted. For defense mechanisms the S-NS discrimination is germ line-selected by ridding individuals and is, as a consequence, a property of the species.

A somatically selected S-NS discrimination is required by a large, somatically generated, recognitive paratopic repertoire that looks at shape determinants on the "pathogen." The paratopic repertoire (equal in size to the repertoire of epitopes) divides the antigenic universe into *e* epitopes that are combinatorially distributed on antigens *n* at a time. For example, a paratopic repertoire of 10^5 that views the antigenic universe as collections of linked epitopes taken 10 at a time, would distinguish ${}_eC_n = {}_{10^5}C_{10} = 3 \times 10^{43}$ antigens [5-7]. This paratopic repertoire looking at combinatorials (antigens or units of elimination) must see most antigens (e.g., miss only of the order of 1 in 1000).

An aside on somatic evolution

We are all familiar with the principles of germ line evolution but, when considering somatic evolution, ground rules seem to disappear. Somatic evolution is the only solution to generating a large enough functional recognitive repertoire to cope with the protective needs of the individual.

Although there are several quite different mechanisms for generating the somatic repertoire, the principles do not change [8]. Somatic evolution has two components, a set of variants with respect to recognitive specificity and a selection pressure to sort out the variants that are functional. The variation in recognitive specificity is generated at two levels, germ line and somatic. I will illustrate my points using mouse and man as prototypes. The germ line encodes $N\mathrm{V_L}$ and $N\mathrm{V_H}$ gene segments, the products of which complement to generate $\leq N^2\mathrm{V_L}\mathrm{V_H}$ germ line encoded specificities. $N\mathrm{V_L}\mathrm{V_H}$ pairs are directly selected for the protective specificities they encode. The rules for this germ line selection are the same as those for any defense mechanism, a specificity recognized as present on the selecting agent but absent in the host species. This fixed the complementing subunit pool resulting in $N\mathrm{V_L}\mathrm{V_H}$ germ line selected specificities and $\leq (N^2 - N)\mathrm{V_L}\mathrm{V_H}$ germ line encoded but unselected specificities. This $\leq N^2\ \mathrm{V_L}\mathrm{V_H}$ family of germ line encoded specificities is expressed in high copy number (10^2–10^3 copies per 10^7 B cells) and is the substrate for a mechanism of somatic hypermutation that generates an additional roughly 10^5 specificities (single amino acid replacement mutants) each in single copy per 10^7 B cells. This is the population upon which somatic selection due to antigenic encounter operates.

The somatic selection pressure is the necessity to produce an effective effector response to the "pathogen" in a short enough time, without debilitating the host. This selection has several targets on which it operates and that become limiting at different points during the course of evolution:

1) the specificity of the antigen-receptor itself (see discussion of K in ref. [3, 6, 7]);

2) the number of antigen-receptors expressed per cell (haplotype [allelic] exclusion or one cell-one antigen-receptor);

3) the number of paratopes per antigen-receptor (one antigen-receptor-one paratope);

4) the range over which inductive signaling occurs (short enough to permit associative recognition of antigen [see later]);

5) the range over which effector function operates (short enough to minimize effects on innocent bystanders).

Somatic selection could not operate if one cell produced all antibodies or even many antibodies. A single evolutionary advance solved this and several other problems. Unlike defense mechanisms, where the germ line selection operates on a single gene encoding both recognition and effector function (an enzyme, a lectin, an antibiotic, a particle sensor), the immune system separated these two properties into V-gene segments specifying recognition and C-gene segments specifying the link to effector function. Thus, each effector function could be initiated by a large number of different paratopes but, equally important, a key step in developing a mechanism for expressing one antigen-receptor per cell was realized. The fusion efficiency with which the V-gene segment is joined in frame to the C-gene segment accounts for roughly 80% of Ig^+ cells expressing one antigen-receptor. Additional regulatory controls increase this to 99%; about 1% of cells are double producers, the limit to one cell one antigen-receptor and a potential source of autoimmunity [6, 7, 9, 10].

The point to stress is that the somatic selection by immune systems requires a clonal distribution of antigen-receptors (referred to as allelic or haplotype exclusion), whereas the germ line selection operating on defense mechanisms does not require allelic exclusion.

In summary, germ line selection of the recognitive repertoire proceeds antigen-by-antigen. This is limited by two factors:

1) when the rate of escape of the pathogen from recognition exceeds the generation time of the host, germ line selection cannot keep up and the high metabolic cost of somatic evolution becomes worthwhile;

2) as there are more pathogens in the environment of vertebrate land animals than there are genes that can be committed to defense, it becomes impossible for the genome to decide which pathogen to track.

The only solution for the vertebrates was to develop a mechanism for somatic evolution of the recognitive repertoire that matched the ability of the pathogen to escape recognition. Totally new in its principle was to divide the antigenic universe up into shapes called epitopes that were combinatorially distributed. An antigen, as defined by the immune system, is a linked group of N epitopes. A somatically evolving recognitive repertoire makes it virtually impossible for the pathogen to escape recognition. To escape it must foil the effector mechanisms. Lastly, a somatically evolving recognitive repertoire requires a somatically-selected S-NS discrimination that acts at the level of the epitope-specific combining site (paratope). Anti-S must be deleted and anti-NS saved.

An experimental illustration of a S-NS discrimination resulting from germ line or somatic selection

How does the behavior of immune systems and defense mechanisms differ when challenged with antigen?

THE F_1 TEST

Host	Graft	Outcome	Interpretation
AA, BB	AB	Rejected	AA expresses anti-B, not anti-A. BB expresses anti A, not anti B. The S-NS discrimination is somatically selected and antigen-specific. (Immune system)
AA, BB	AB	Accepted	AA recognizes A as a "self-marker." BB recognizes B as a "self-marker." The S-NS discrimination is germline-selected; recognition is self-marker-, not antigen-specific. (Defense mechanism; see text for qualifications).

Figure 3. The F_1 test defines an immune system.

The F_1 test (*figure 3*) is to determine whether or not AA or BB (parental) will accept $(AB)F_1$ grafts.

Under Case I (immune system), the $(AB)F_1$ graft is rejected. In this case, I recall that an AA animal will accept AA grafts but reject BB grafts, as well as CC, DD, EE, etc. Similarly, a BB animal will accept BB grafts but reject AA grafts, as well as CC, DD, EE, etc. The $(AB)F_1$ does not self-destruct; further it accepts AA and BB grafts, while still rejecting CC, DD, EE, etc.

The F1 test shows that somatic selection depends on the specificity of recognition of S, and by difference, NS. S-recognition is deleted, leaving NS-recognition. The immune system responds to the difference between the S of the individual and NS. If an antigen-related factor (or NS-marker like "pathogenicity" or "danger") were the determining element, we would not expect the graft of AB onto AA to be rejected when the graft of AA onto AA is accepted. Clearly, specificity of recognition of NS-B by AA is a key factor. The ability to distinguish a S- from a NS-epitope is equivalent to distinguishing two NS-epitopes. This is what gives us a uniquely meaningful definition of specificity [3, 6, 7] and explains why tissue grafts between individuals in a species with immune systems are rejected.

Under Case II (defense mechanism), the $(AB)F_1$ graft is accepted. I raise here the question of S-markers only because the first attempt to explain the S-NS discrimination was a

S-marker model. S-markers are a special case. In essence, protective mechanisms cannot use S-markers. A S-marker system can only be used in order to exclude inappropriate interactions between colonies or to housekeep or to forage for food. As a defense against intracellular pathogens a S-marker system would be useless as intracellular pathogens hide in the S and would be protected against destruction by the recognition of the S-marker. As a defense against extracellular pathogens, S-markers are of limited value as no single effector mechanism can protect against all pathogens and multiple effector mechanisms that require independent and unrelated recognitive elements obviate the need for a S-marker, making it unselectable. The recognitive element, in this case, would be germ line-selected to recognize a given NS to the exclusion of species-S. Lastly and this is key, the reliance on a specific S-marker system obviates the need for another independent antigen-specific recognitive system.

In sum, a S-NS discrimination that is germ line-selected looks at the difference between the target antigen and the 'S-of-the-species'. A S-NS discrimination that is somatically-selected looks at the difference between the target antigen and the 'S-of-the-individual'. The F_1 test illustrates this as immune systems reject allogeneic grafts, whereas defense mechanisms accept them.

The F_1 test tells us that for immune systems

1) 'S cannot be distinguished from NS by any physical or chemical property of antigens as classes'. What is S for AA is NS for BB. AA rejects AB by recognition of B and BB rejects AB by recognition of A. AB expresses neither and accepts AA and BB grafts; further, AB does not self-destruct;

2)'the S-NS discrimination by immune systems must be learned', a true historical process. AA expresses anti-B, not anti-A; BB expresses anti-A, not anti-B; AB expresses neither anti-A nor anti-B; and CC, DD, etc. express both anti-A and anti-B. A learning process requires that anti-S be inactivated and anti-NS be activated; this in turn requires that S be separated from NS by some mechanism; one, possibly the only reasonable suggestion depends on developmental time. Key here is that there must be an antigen-specific step where a discrimination is made, no matter what other antigen-unspecific factors are introduced that regulate expression of effector function (i.e., responsiveness).

Any attempt to explain the S-NS discrimination by invoking "S-markers" implies by symmetry, the possibility of explaining the S-NS discrimination by invoking "NS-markers." In fact, in recent years "NS-markers" have achieved a unique popularity. Over the years, many investigators have invoked NS-markers under such rubrics as markers of adjuvanticity, danger, pathogenicity, harm, disintegrity, localization, etc. These "NS-markers" like the "S-markers" imply a germ line-selected S-NS discrimination and, as such, do not require antigen-specific recognition; recognition of the "NS-marker" should be sufficient. NS-marker models have been analyzed elsewhere[11-13].

Broadening the field; the wide angle view

It is important to appreciate that evolutionary selection is operating on the factors contributing to the specificity of the effector output, which, in the end, must distinguish host from pathogen. There are many steps between encounter with antigen and the destructive and ridding effector output, any one of which can be and has been viewed as contributing

to the S-NS discrimination. This, in fact, is the origin of all of the present day ambiguities and polemics. 'However, one step is indisputable and that is the initial antigen-specific decision step'.

Were it not for the necessity to make a S-NS discrimination, the antibody molecule (or antigen-receptor) could be a universal glue. Antibody specificity is driven solely by the necessity to make a S-NS discrimination. It is the antigen-specific step that permits anti-S to be separated from anti-NS and this, in turn, requires that S be distinguished from NS, a problem we will face later.

However, it is not unreasonable to consider a more general picture, which I will refer to as the 'Unresponsiveness-Responsiveness' decision. Many immunologists consider S and NS as they are defined by all of the steps resulting in a triggered effector function. After all, this is what they measure when they determine whether an animal is responsive or unresponsive. Further, as I just pointed out, evolution is assaying the specificity of the effector output and selecting for a response able to rid NS without ridding S and debilitating the host. Consequently, there is a school of immunologists who introduce into the discussion of the S-NS discrimination, antigen-*un*specific or antigen-related factors (e.g., adjuvanticity, danger, harm, pathogenicity, etc.) required for the expression of effector function. These antigen-related factors are important when put in the proper context, but they are independent of the antigen-specific decision step (S-NS discrimination), which, even in this more general framework, is an essential part of the Unresponsiveness-Responsiveness mechanism because the destructive and ridding effector output must target the "pathogen" to the exclusion of the host. The Unresponsiveness-Responsiveness decision is not without its ambiguities, because an assay interpreted as "tolerance" (unresponsiveness) could well be due to responsiveness in an ineffective effector class or to an insufficiency in the level of an effective effector class or to one not assayed.

A novel cell must be added to make a somatically selected S-NS discrimination

The driving force for a somatically selected S-NS discrimination was the need to increase the size (and concomitantly the specificity for the antigen) of the recognitive repertoire over that expressed by defense mechanisms. As the somatically selected S-NS discrimination operates by deleting cells, not individuals, evolution had to introduce an 'intermediate cell' into the differentiation pathway, which is the cell upon which somatic selection operates (*figure 4*).

In the case of defense mechanisms, the cells are born as effectors (e-cells). As the S-NS discrimination is germ line-selected, all effectors are anti-NS. 'However, there would be no way to make a somatically selected S-NS discrimination if cells were born as effectors.' An intermediate stage or 'i-cell' stage had to be introduced between the stem cell and effectors. The i-cell stage is where the S-NS discrimination based on antigen-specific events is made. As we will see, the i-cell has two pathways open to it upon appropriate interaction with antigen, inactivation and activation.

The i-cell differs from the e-cell in that it has no effector function, it receives signals, it does not send them; it listens, it does not talk. The e-cell sends signals; it does not receive them. The e-cell talks; it does not listen. Immunologists refer to what I am calling the i-cell as the antigen-sensitive or antigen-responsive cell without putting it in this precise context as a decision step.

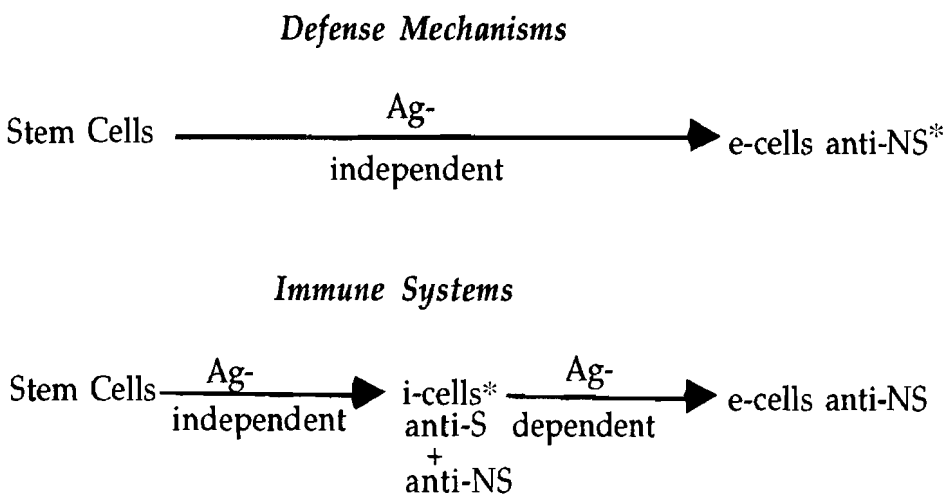


Figure 4. The introduction of the i-cell into the pathway of differentiation of immune systems.

* For definitions of i- and e-cells, see text.

Putting the S-NS discrimination into the big picture

Evolution is selecting on the effector output, which must rid and destroy NS without being debilitating for S. In order to do this, the immune system ("adaptive") makes two successive decisions (*figure 5*) when it encounters an antigen.

The antigen-specific step of the S-NS discrimination is made at Decision 1, which operates at the level of the i-cell. Once this decision has been made and the further steps to effectors put under the control of Decision 2, no antigen-specific S-NS discrimination is possible. An anti-S cell that sneaks through Decision 1 or arises by mutation during Decision 2 would go on to become an effector. The deleterious consequences of such an event is limited both by the probability of sneak-through or of mutating to anti-S, and the fact that effectors are either dead-end or cycle back to the i-state where a S-NS discrimination step is re-initiated.

What is key to realize is that the activated cell is not yet an effector. It takes a coordinated set of regulatory events for it to proliferate and differentiate to effectors. These regulatory events are what is argued by many to be part of the S-NS discrimination, and well they might be, but they do not involve antigen-specific decisions with respect to a S-NS discrimination based on somatic selection; they involve decisions about germ line-selected, antigen-related properties like localization, "pathogenicity," mode of entry, dose, "danger," etc. As a consequence they are independent of and have no way to affect the antigen-specific Decision 1. Antigen-related properties are, in this framework, used by Decision 2. The pathway followed by Decision 2 has antigen-specific steps required to relate effective effector function to "pathogen," but these follow the same rules for S and NS, that is, they are designed to optimize the destruction and ridding of the target independent of whether it is S or NS.

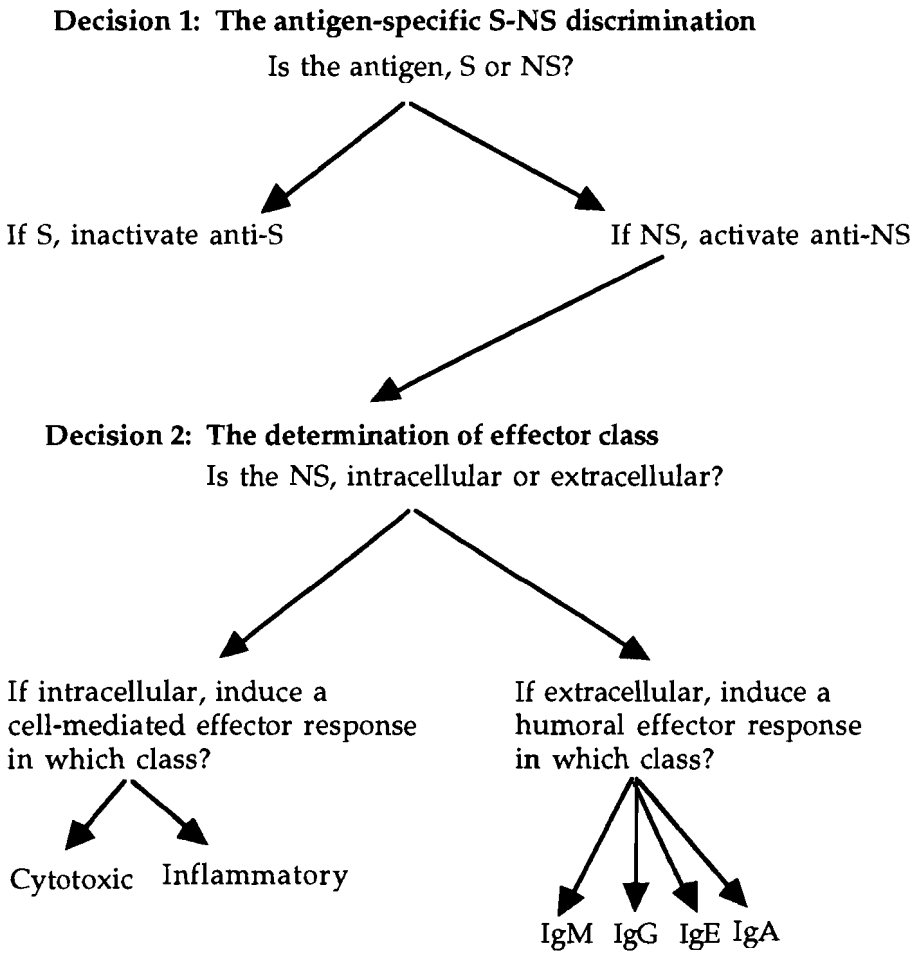


Figure 5. The Decision steps of the immune system.

When I refer to the 'S-NS discrimination', an antigen-specific Decision 1 step is implied; when I refer to the 'Unresponsiveness-Responsiveness' discrimination, in addition to the antigen-specific Decision 1 step, a set of antigen-related steps (Decision 2) are implied.

So much for the conceptual framework. Now let us consider its translation into mechanism and how did we get there?

History of ideas as to mechanism of the S-NS discrimination

The first stab at an explanation

Burnet and Fenner [14] in 1949 should be credited as being the first to insist that the S-NS discrimination had to have an explanation.

Burnet in 1984 reminiscing on this period felt that the book "Production of Antibodies (1949)" is a collector's item because of a prediction made on p. 103. "If, in embryonic life,

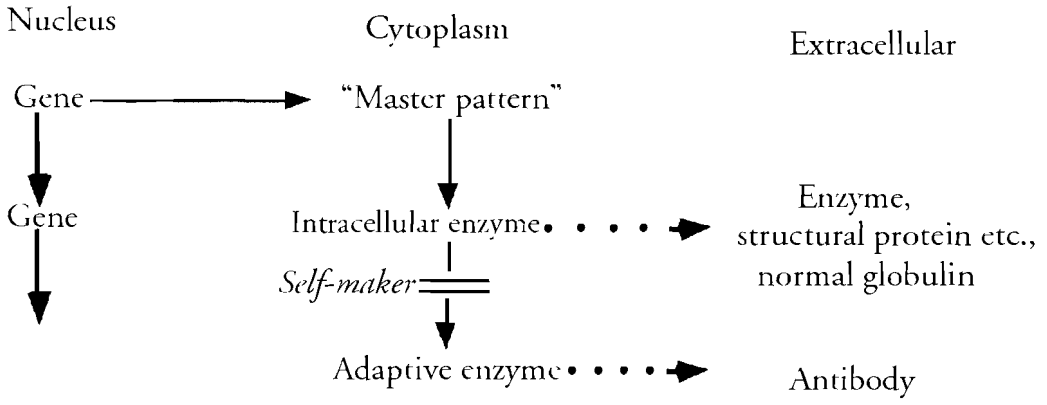


Figure 6. The Self-marker template theory of Burnet and Fenner (1949).
Heavy arrow: replication of pattern.
Light arrow: partial replication of pattern.
Dotted arrow: liberation of protein unit bearing specific pattern but lacking any further capacity for replication.
All processes will necessarily be dependent on intracellular mechanisms for supply of energy and "bausteine".

expendable cells from a genetically distinct race are implanted and established, no antibody response should develop against the foreign cell antigen when the animal takes on an independent existence" [15].

Burnet and Fenner [14] based this statement on Owen's observation [16] that dizygotic calves sharing placental circulation, expressed throughout life hematopoietic cells of the fraternal twin without having the genes that encode the given cell markers. These same dizygotic calves (fraternal twins) would have rejected each others' hematopoietic cells (and skin grafts) if they had developed in separate mothers or in the same mother at different times. Owen's observation was extended by Medawar's group in 1951 when they showed that, in calves, skin grafts could be exchanged between fraternal twins as well as between identical twins. These observations implied a learning period during embryonic life when 'what-is-S' is defined. They show that the immune system has no way of knowing what is encoded in the germ line. Today these observations have been extended by constructing allophenic mice and even xenophenic animals, like goat-sheep, by parabiosis (reviewed in [17]) and by remarkably skilled transplantation experimentation at very early stages of embryogenesis[18].

To explain the Owen phenomenon, Burnet and Fenner (Ref. [14]; p. 94) proposed the following (*figure 6*).

A constitutive "intracellular enzyme" synthesized on a master template, was envisaged to rid senescent and effete constituents; an "adaptive enzyme" derived by antigenic modification of the master template was postulated to rid invaders (pathogens). The function of the intracellular enzyme is "housekeeping." The recognition of its target is 'germ line-specified' and, therefore, the intracellular enzyme is part of a "defense mechanism" that protects the individual from becoming a toxic dump site. Consequently, it is irrelevant to the somatically-selected S-NS discrimination. The remainder of the antigenic universe

induces nonspecifically a destructive and ridding "adaptive" response. This antigenic universe is composed of both S and NS, each of which would activate the same destructive and ridding response. Consequently, Burnet and Fenner, in order to correct the failure of the Pauling template model to deal with the S-NS discrimination, proposed that all S is associated with a S-marker that blocks the "adaptive" template, thereby protecting S from destruction. Anything not associated with the S-marker (i.e., NS) is destroyed and ridded. Under this model, specificity of recognition of antigen is not required. Consequently, not even a template is required because the role of the template, is to provide a mechanism that results in antibody specificity. The S-marker obviates this requirement. It only needs "turn-off" a non-specific activity that destroys and rids anything not associated with the S-marker. Of course, the S-marker must associate with 'all' S, and 'no' NS. It is the specificity of this interaction that determines the S-NS discrimination. A S-marker theory does not imply an antigen-specific somatically-selected process; it is a theory accounting for a germ line-selected process, as I pointed out earlier. A S-marker theory would not pass the F_1 test.

Burnet and Fenner [14] made three conceptual errors:

1) protective mechanisms cannot use S-markers to make a somatically selected S-NS discrimination and the postulate of a S-marker does not imply an antigen-specific recognitive system;

2) the S-marker assumption does not predict (in fact, it is in contradiction with) the "Owen finding" that they were trying to explain.

Any antigen not associated with the S-marker, no matter when it was introduced into the animal, would be treated as NS and if associated with the S-marker as S. Therefore, the S-marker model did not explain the Owen finding simply because it is a germ line-selected mechanism and the Owen finding requires a somatically-selected mechanism. The "F_1 test" (*figure 3*) and the Owen finding are telling us the same thing. 'The S-NS discrimination made by immune systems is an antigen-specific somatically-selected process.

3) in no way could this formulation be translated into a reasonable mechanism.

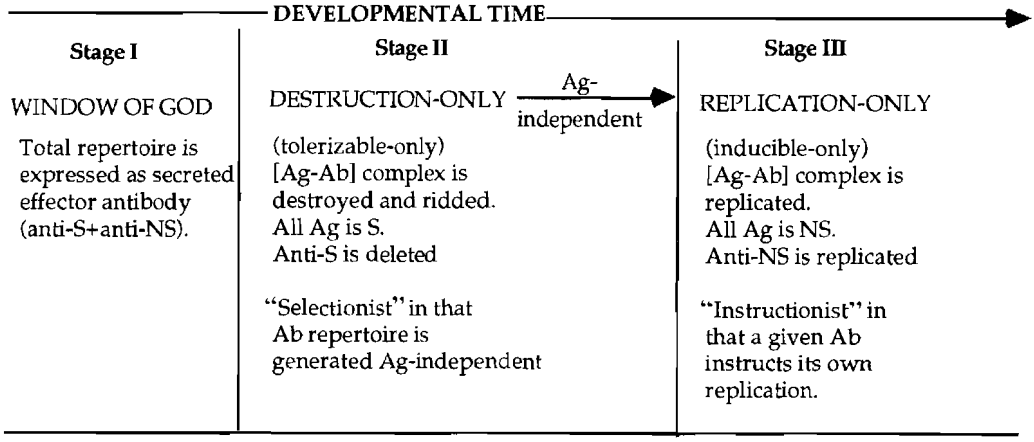
A meaningful model should have the potential to be envisioned in terms of molecular mechanism at the less complex level and in terms of evolution (and physiology) at the more complex level.

Lederberg [19] in 1959 tried to save the S-marker theory [14] for an antigen-specific, somatically-selected S-NS discrimination from being viewed as simply irrational by transforming the meaning of "S-marker." He wrote, "[t]here are no more plausible candidates for S-markers than the S-antigens themselves."

The S-NS discrimination as viewed by Jerne, Burnet and Lederberg

Jerne (1955) is usually credited with the first selectionist theory [20]. I have always viewed it as a hybrid instructionist-selectionist theory. As Jerne was close to Delbruck's phage group, he was aware of the work showing that heritable variation precedes selection. Evolution selects upon pre-existing heritable variants; the selection pressure does not induce the variation. By 1955, the 1943 Luria-Delbruck fluctuation analysis, the 1949 Newcombe spreading experiment, and the 1952 Lederberg replica plating technique were a standard part of every microbial genetics course. No biologist would have considered a template mechanism in which antigenic selection induced the variation. Jerne's selective theory

A. JERNE (1955) - NATURAL ANTIBODY MODEL



B. BURNET (1959) - CELLULAR SELECTION MODEL

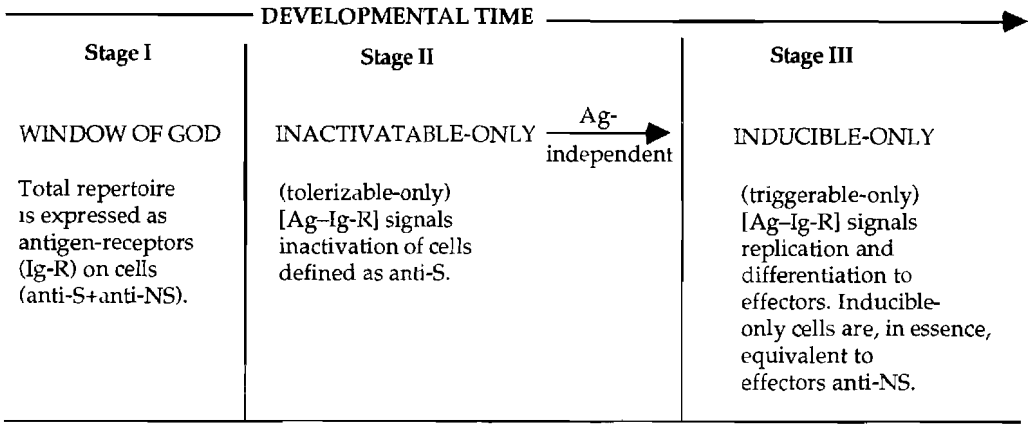


Figure 7. A. The hybrid instructionist-selectionist theory of Jerne (1955). **B.**The developmental stage-dependent, cellular selection model of Burnet (1959).

reflects the direct influence of the phage group. However, Jerne could not make a clean break with instructionist thinking. As a card-carrying immunologist he felt that he had the right to treat the subject as an eccentricity in biology.

Jerne [20] envisaged a sequence of three stages in developmental time, a decisive advance in thinking, albeit for the wrong reason (*figure 7A*).

Jerne (1955) proposed a defined period during embryogenesis when the total repertoire of antibody was generated "big bang" and secreted as "natural antibody" (Stage I). He then suggested that the interaction of antibody with S during stage II of ontogeny resulted in

anti-S being subtracted. This was the selectionist step. Then, in order to account for responsiveness, Jerne proposed that, at stage III of ontogeny, the remaining anti-NS functioned as a self-replicating entity when it interacts with NS. This was the instructionist step.

Jerne, therefore, substituted a developmental stage-dependent deletion of anti-S for Burnet and Fenner's S-marker while retaining the template mechanism. It is important to realize that the stage II development-dependent deletion of anti-S was not driven by a learning or historical process, but was a corollary of the postulate that the repertoire was generated "big bang" at Stage I. Jerne erroneously mapped the necessity for a developmental stage dependent S-NS discrimination onto the mechanism for generating diversity, an error to be perpetuated by Burnet.

Jerne made three conceptual errors. The first error, seen by everyone, was the assumption of self-replication of a protein that had bound antigen. By way of review, the structure of DNA was known to the phage group in 1952; it was published in 1953. By 1955, every molecular biologist's blackboard had written on it, "DNA → RNA → Protein." This says something about the state of Delbruck's thinking about the subject of immunology in 1955, as it was he who submitted the paper to PNAS without requiring Jerne to comment on it. Had Delbruck done that, Jerne would certainly have preempted Talmage [21] and Burnet [22]. One should always ask one's enemies, not one's friends, to submit papers to PNAS.

The second error was that the "big bang" generation of the total repertoire as secreted antibody would result in no given antibody being at a sufficient concentration to function (react with antigen and be processed). Jerne aggressively ignored effector function in all of his thinking, throughout his career. It took Protecton theory to deal with this problem [6, 7].

The third error was more subtle and was to be repeated by Jerne during the "Idiotype Network era" (1975-1985). There is no way to make a somatically selected S-NS discrimination at the level of effector antibody or cells. If the destruction and ridding of anti-S were dependent on its being at a concentration that interacted effectively with S (stage II in *figure 6A*), then S itself would be under attack by the immune system. The destructive and ridding effector output of the immune system cannot be used to regulate the S-NS discrimination (Decision 1). This is the selection pressure that forced evolution to introduce i-cells into the pathway to effectors (*figure 4*).

The next step is owed to Talmage [21] and Burnet [22] who in 1957 introduced the theory of cellular selection by antigen. Talmage [21] and Lederberg [19] correctly refer to a theory of cellular selection. Burnet [22] uses the term "clonal selection" leaving it ambiguous as to "how clonal" and why? For example, in 1961 he [23] wrote, "[t]he theory is called clonal-selection theory because the action of antigen is simply to select for proliferation of that particular clone of cells which can react with it." Thus he meant "cellular selection" by the term "clonal selection." Burnet [22] appreciated that template or instructionist theories require that each cell be toti-specific; selectionist theories as they were viewed in the late 1950s only required that cells be oligospecific. However, both Talmage [21] and Burnet [22] seemed to prefer a unispecific cell (clonal distribution of antigen-receptors). Talmage gives us no reason for this preference. Burnet argues that, "[t]he clonal selection theory requires at some stage in early embryonic development...a 'randomiza-

tion' of the coding responsible for part of the specification of gamma globulin molecules, so that...there are specifications in the genomes for virtually every variant that can exist..." Thus, Burnet tries to answer the right question, but inverts the logic in his answer. The statement, "If it rains, the fire hydrant will be wet" does not imply that "if the fire hydrant is wet, it has rained."

If the generation of diversity were due to "randomization" of genes, then unispecific clonality might result depending on a great many additional (and unlikely) assumptions Burnet would need to have added. However, if unispecific clonality obtains, it does not mean that the generation of diversity is due to somatic "randomization"; that is only one of several pathways to unispecific clonality as the subsequent discussions on "germ line" versus "somatic" theories illustrated. In the end, Burnet's argument was wrong; unispecific clonality is driven by the necessity to make a S-NS discrimination, not by any requirement or fortuity of the generator of diversity, as the Protecton theory [6, 7] has now clarified.

The 1957 papers of Talmage [21] and Burnet [22] resolved both Jerne's first error, (a self-replicating protein) and his third error (a S-NS discrimination at the level of effector antibody) by putting antigen-receptors on cells and by having the cells replicate and differentiate to effectors or be deleted. The second error, (i.e., the "big bang" expression of a random vast repertoire would leave no specificity at a sufficient level to function) was not appreciated by Burnet or Talmage in 1957 (and most immunologists in 1999). As a consequence, Burnet accepted Jerne's proposal for a "window of Generation Of Diversity (GOD)." Talmage [21] is silent on this question of repertoire. However, in 1959, when Burnet unveiled his model for the S-NS discrimination, both Talmage [5] and Lederberg [19] had faced Jerne's second error in two remarkable back-to-back papers in *Science*. In Lederberg's words, "It would embarrass a theory of cellular selection only if the repertoire were large compared to the number of potential antibody-forming cells in the organism." (As an aside, Lederberg did not quite get it right. By using the organism as his unit immune system, he implied that a vast repertoire (e.g., 10^{14}) would allow cellular selection to operate in elephants but not in mice, an obvious absurdity. This paradox inspired Protecton theory [6, 7] as a unique solution.)

In 1959, Burnet [24] proposed a theory for the S-NS discrimination (*figure 7B*). He began his analysis, as did Jerne, by delineating three stages in the development of immune responsiveness. At stage I he postulated a Jernerian "big bang" Window of GOD after which the repertoire expressed by the cell population is fixed. At stage II any cell interacting with antigen is deleted. This cell population is characterized by Burnet as "hyperreactive" to antigen, or, in my translation, inactivatable-only or "tolerizable-only." Stage II removes from the population all anti-S cells. The remaining population, anti-NS, differentiates antigen (Ag)-independently to "inducible-only." The anti-NS cells present at stage III respond to antigen by proliferation and differentiation to effectors. Formally speaking "inducible-only" is a state not functionally distinguishable from effectors.

The key here is that the cell population goes from "tolerizable"-only to inducible-only, antigen-independently, as a function of the developmental stage of the organism. This shift in responsive behavior is a once in a lifetime developmental event making S a static entity defined at stage II (*figure 7B*), a correct conclusion for the wrong reason. Further, such a process cannot cope with somatic mutation to anti-S in the stage III population.

Silverstein and Rose [25] tell us that, "[w]ith this theory, Burnet [24], accomplished two things: he made the acquisition of tolerance the sole province of the immature mammalian fetus and neonate and he implied that an initial S-NS discrimination is the most important event in the development of the immune system."

The assumption of a defined period during embryogenesis when diversity is generated and anti-S is deleted might be viewed as a clear statement that Burnet [24] considered tolerance to be a property of the animal, not the cell. However, I have my doubts because it is the "Window of GOD" that is the property of the animal, not the S-NS discrimination, the mechanism of which is dictated by that assumption. Burnet, like Jerne, did not understand that the assumption of a role for the developmental stage of the animal in making the S-NS discrimination, had to be part of a historical somatic learning process, not an indirect consequence of the postulated stage I "big bang" origin of the repertoire.

Silverstein and Rose [25] feel that the assumption of a role of the developmental stage in establishing the S-NS discrimination is wrong, based on experimental disproof. They [25] totally agree with Miller and Basten [26], who wrote:

> Although Burnet suggested that antigen encountered in early life selectively deletes specific clones, whereas it activates them in later life, it has been realized for close to four decades that there is nothing intrinsically unique about the prenatal or neonatal period, as far as tolerance is concerned. One key factor in determining the nature of the response, whether tolerance or immunity, is, therefore, not the developmental stage of the individual but rather the state of differentiation of the lymphocyte at the time it encounters antigen.

Thus Miller and Basten (1996) took a purely Lederbergian (1959) position as I will come to next.

However, both [25, 26] were wrong and Burnet [24] was right, but for the wrong reason. Sometimes the wrong assumption (Window of GOD) yields the right conclusion. However, I might stress, the reason for a conclusion is as important as the conclusion itself. The requirement for a somatically learned, developmentally established S-NS discrimination is 'conceptual' and, to date, remains unchallenged by any extant experiment or competing proposal (see discussion in Ref. [11]).

Jerne's (1955) proposal for a Window of GOD expressed during fetal life was rational, given that he had placed the S-NS decision at the level of the secreted effector (the so-called "natural") antibody. He had to end up with a self-replicating anti-NS population, devoid of anti-S, that functioned to protect throughout the life of the individual. However, Burnet (1959) by assuming cellular selection, no longer required this assumption; in fact, in the framework of the selectionist theory, a Window of GOD is ad hoc. His failure to see this led him to treat clonality of antigen-receptors on cells as an evolutionary by-product of the mechanism for generating diversity, 'not' as a requirement of the S-NS discrimination, a major error in thinking corrected only by the Protecton theory [6, 7].

In 1959, Lederberg felt that this window for generating diversity (*figure 7*) was unreasonable as lost clones would be irreplaceable. So he modified the Jerne [20]-Burnet [24] proposal by having the generator of diversity operate throughout life. Incidentally, Burnet changed his 1959 position [24] in 1964 (in effect, *un pli cacheté*) [27] by accepting Lederberg's assumption without realizing that by doing this he no longer had a theory of the

S-NS discrimination; and, in the end, Burnet left us with no theory of the S-NS discrimination.

Now, let us consider Lederberg's model for the S-NS discrimination (*figure 8*). He postulated the following:

1) cells are born tolerizable-only, symbolized t-cells;

2) the repertoire is derived by mutation that occurs only during the generation of t-cells, not in e-cells;

3) after a period of time when no antigen is encountered, t-cells differentiate to e-cells. This is a default antigen-independent pathway;

4) as stated by Lederberg, this pathway applies to all cells, T and B.

Several comments are relevant:

1) the S-NS discrimination is a property of the state of differentiation of the cell, not of the developmental stage of the animal. Thus he differed from Burnet [25] who had proposed that the shift from tolerizable-only to inducible-only was a property of the developmental maturity of the animal.

Once Lederberg [19] had abandoned the Window of GOD, he saw no reason to have an animal level control of the S-NS discrimination, and he put the discrimination at the level of cell differentiation, a steady state lifelong process. At first glance, one might say that a S-NS discrimination at either level could have provided an adequate mechanism. However, this turns out to be incorrect as the cell differentiation Lederberg model did not correct the inability of the Burnet Model to deal with unavoidable spontaneous mutation to anti-S in the dividing e-cell population. Further, the Lederberg model could not deal with persistent (non-riddable) antigens like carbohydrate (see later).

The Coutinho et al. [28] argument is simply unjustified that, because development of cells and organisms had been confused, the Lederberg model can be refuted by simple logic. I do not believe that Lederberg confused the two levels of development; he chose between them. Further, the model has not been refuted by simple logic. A t-cell, true enough, cannot distinguish S from NS. It would be deleted by interaction with either. However, the system responds to NS because it has accumulated e-cells anti-NS previously generated in the absence of NS, which is transient. The Lederberg model remains rational, albeit wrong. There is a difference between irrational and erroneous. Burnet's (1949) S-marker theory was irrational; Lederberg's (1959) cell differentiation model was erroneous. In the end scientific progress reflects the sad fact that most rational models are erroneous;

2) the proportion of anti-S cells that sneak through is a function of the ratio of the time to encounter antigen while in the t-state to the time it takes to differentiate Ag-independently to the e-state.

As the time for t-cells to differentiate to the e-state becomes shorter, more and more anti-S cells sneak through. As this time becomes longer, the sneak through decreases but the repertoire of e-cells that protect the animal becomes smaller. There is a trade-off that Lederberg did not face, namely, at 'what level is anti-S sneak through acceptable and how can one determine it'? This was a problem first faced in 1989 in the development of Protecton theory [6, 7];

3) mutants to anti-S in the e-cell population would be uncontrollable. Further assumptions are required;

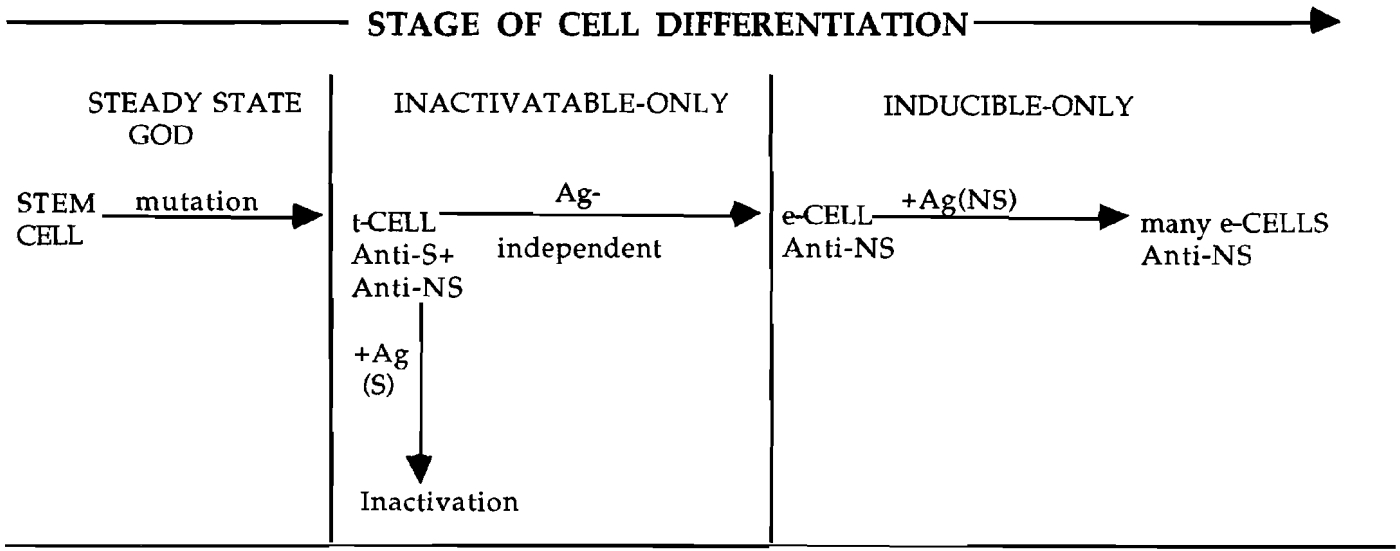


Figure 8. The developmental stage-independent, cellular differentiation-dependent model of Lederberg (1959).

4) there is no predicted role for effector T-helpers (eTh) in the induction pathway and consequently no associative (linked) recognition of antigen essential to a coherent response. All antigen-specific events are on an epitope-by-epitope basis; no 'supraclonal' regulation is implied, to use Coutinho's language. The Lederberg system cannot distinguish antigens from epitopes, and, therefore, has no way to provide a coherent effector response to antigen. The same, of course, would have also been true of Jerne's and Burnet's models. These three models [19, 20, 24] had no way to translate epitopes into antigens, which, after all, are the targets (units of elimination) of immune effector function. The response to all epitopes on an antigen must be the same (i.e., coherent), if the immune system is to function optimally;

5) the signal to the cell upon interaction with antigen is interpreted as inactivation when in the t-state and as triggering when in the e-state. The same signal is interpreted in two different ways depending on the state of differentiation of the cell. This would also be true of Burnet's (1959) model [24].

Lederberg's (1959) paper was totally ignored by the immunological community. Even Burnet, in whose laboratory Lederberg had worked, cites the paper for everything but a model of the S-NS discrimination. In 1968, Bretscher and I [29] disinterred the paper as a starting point for our model.

Bretscher and I [29-31] were aware of the difficulties inherent in both the Lederberg [19] and Burnet [24] models, which nonetheless, had set the stage for clear thinking about the problem. For the first time, these models posed and dealt with the three questions that must be answered: how are cells anti-S inactivated? How are cells anti-NS activated? How is the distinction between S and NS established (learned) and maintained (remembered)?

Associative recognition of antigen (ARA): the "two signal" or ARA model

The ARA Model made four clean breaks with all previous formulations:

1) it introduced a decision step between anti-S and anti-NS at the level of the i-cell (Decision 1, *figure 5*). Cells were not born tolerizable-only but had two pathways open to them, inactivation and activation;

2) it made the origin of the distinction between S and NS, developmental stage-dependent, in that it defined S as 'prior and persistent' and NS as 'posterior and transient.' Unlike Lederberg, then, we postulated that the appearance of responsiveness was 'not' a property of the state of differentiation of the cell, but of the developmental stage of the animal. We differed from Burnet [24] in that the i-cell population did not differentiate from inactivable-only to activable-only as a function of developmental maturity. The behavior of the i-cell depended on the presence or absence of effector T-helpers (eTh) anti-NS. Burnet's assumption was no longer required because the generator of diversity was postulated to operate throughout life;

3) it emphasized the central role of Associative Recognition of Antigen by a regulatory effector T-cell ("helper") and an i-cell in order to make the response to a given antigen, coherent;

4) it raised [29-32] but did not answer until 1983 [33] the critical question of the origin of the regulatory effector T-helper cell (eTh).

The Associative Recognition of Antigen model (ARA model) has been updated continuously over the years [4, 34-37] and is outlined in *figure 9* as we view it today. A major

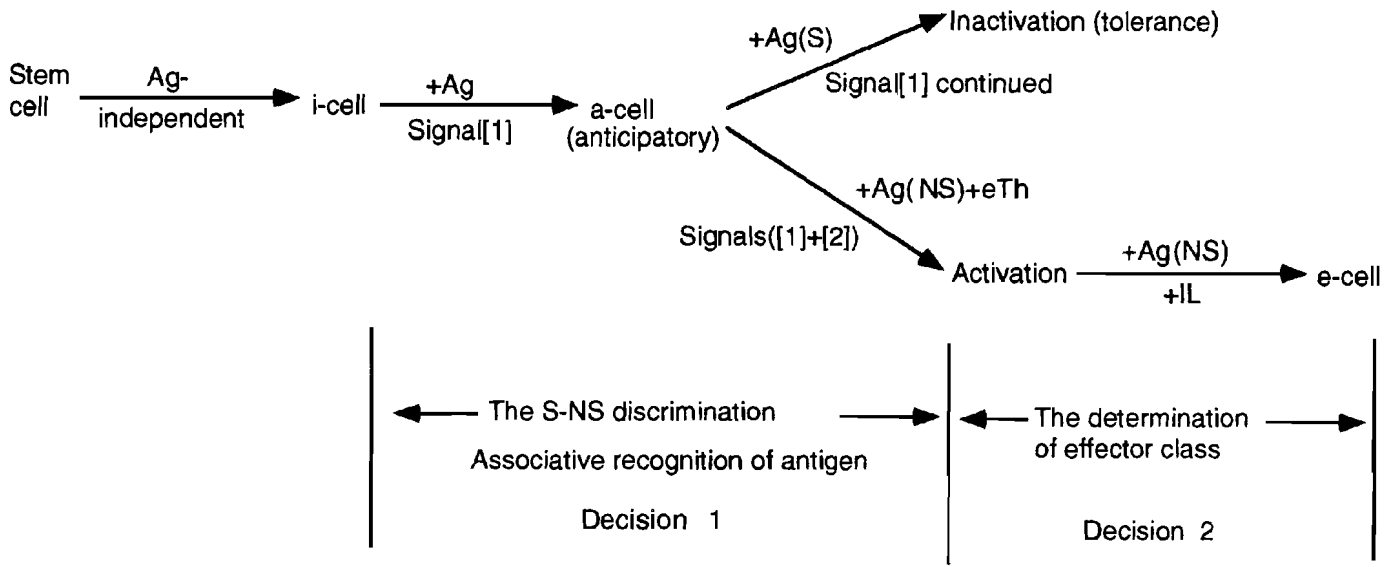


Figure 9. The Associative Recognition of Antigen (ARA) model for the Self-NonSelf discrimination.

contribution to the clarification and justification of this model in its present form, is owed to Langman.

We appreciated that interaction with Ag had to be common to both pathways. The distinction between inactivation and activation requires an additional event that itself has made a S-NS discrimination. It is the nature of this second event (Signal[2]) that was and still is under polemic; the division is sharp between whether Signal[2] is antigen-specific or antigen-nonspecific and if antigen-specific how strict must be the associative (linked) recognition be? We have always argued that Signal[2] is antigen-specific and strict. In fact, a large part of our 1970 paper [31] was devoted to the evidence that iTh are inducible and tolerizable, both Ag-specific properties!

Why does Signal[2] have to be Ag-specific?

As Signal[1] cannot distinguish S from NS, the discrimination had to be a property of Signal[2]. As classes, S cannot be distinguished from NS by any physical or chemical property (the F1 test, *figure 3*). Therefore, Signal[2], the arbiter of the S-NS discrimination, had to be Ag-specific. An 'a-cell' receiving Signal[1] has to be told by the receipt of Signal[2] that the antigen is NS.

In developing the ARA Model, we analyzed three possible signaling pathways (discussed in [38]) and chose the one that I will present.

There are several explanatory comments:

1) the a-cell, unlike the i-cell, is responsive to Signal[2]. This assures that no i-cell can be activated that, in principle, could not have been inactivated (Signal[1]). Since interaction with Ag is common to both pathways, a second signal, which is antigen-specific and delivered by an effector cell that itself has undergone a S-NS discrimination, is required;

2) Signal[2] is delivered by an eTh to an a-cell by associative (linked) recognition of antigen resulting in its being activated. Activation means that it becomes responsive to soluble components, interleukins, which regulate proliferation and differentiation to effectors of a given class (Decision 2).

This transfer of regulation from a cell-cell interaction to an interleukin (IL) or soluble factor-cell interaction is required to permit a rapid enough response. If the immune system imposed a cell-cell interaction at every division in order to monitor the S-NS discrimination it could never respond fast enough to a growing pathogen. So once activated, the sneak through of or mutation to anti-S, is dealt with by requiring e-cells to recycle through the i-stage when antigen is ridded or be short-lived and dead-end;

3) antigen-receptors recognize antigens, epitope-by-epitope. All previous models (those of Jerne, Burnet, Lederberg) proposed that both tolerance and induction were mediated epitope-by-epitope. The ARA Model proposed a regulation that was epitope-by-epitope for tolerance, but antigen-by-antigen for induction. Associative recognition of antigen is required to coordinate the effector response to the linked epitopes defining an antigen. The response to an antigen is optimally effective when the response to each of its epitopes is in the same effector class. Associative recognition of antigen is the mechanism by which the immune system translates epitopes into antigens. This is essential as the immune system is selected upon to rid antigens not epitopes. The regulatory effector cell interacting with one epitope on the antigen tells every other i-cell interacting with that antigen what its response should be. The interaction of an eTh with one epitope on an antigen tells the a-cell interacting with any other epitope to go to activation. This is the "supraclonal"

property correctly insisted upon by Coutinho, and, in our formulation, is a role for the effector T-helper in activation, not the effector suppresser in inactivation, as in his formulation.

4) while associative recognition of antigen permits coordination of the response to antigen, it has a limitation set by the existence of NS-antigens sharing epitopes with S-antigens ([S+NS]–antigens) that present the problem of breaking tolerance resulting in autoimmunity [31,32]. Thus, one key property of the model needs stressing. There is 'competition' between inactivation and activation at the level of the a-cell. This is essential to keep the frequency of autoimmunity low in the face of about 10% of NS-antigens that share epitopes with S-antigens ([S+NS]-Antigens) [6, 7]. The immune system responds to the difference between S and [S+NS], and explains why tissue grafts between allogeneic individuals are rejected;

5) once activated the pathway to expression of effector function (Decision 2) requires a set of germ line-selected, antigen-related factors that a large number of immunologists argue should be considered as part of the S-NS discrimination (e.g., adjuvanticity, danger, pathogenicity, harm, localization, integrity, etc.). For purposes of discussion, I will accept their argument, but insist that these factors play no role in the antigen-specific Decision 1. If the existence of an antigen-specific Decision 1 step is denied, then I would have to shrug it off, as it is not possible to refute an absurdity. If it is not denied, then these antigen-related factors play a role in regulation of class, both type and magnitude.

The origin of the first eTh: the primer question

It is the sufficiency or insufficiency of eTh that determines whether an a-cell will be activated or inactivated. Since all cells are born as i-cells, where does the first eTh, necessary to get the response started, come from? This question posed by the ARA model when first proposed in 1969 [30, 32] took us until 1983 [33] to come up with a reasonable answer. Over the years up to the present, the Two Signal or ARA model has been attacked as ruled out because we were unable to come up with a credible solution to what we called the "primer problem" but everyone else called the "chicken and egg" problem. In fact every immunologist who voiced this criticism acted as though he/she had discovered the problem. Therefore, it is worth stressing that it is not a weakness, but rather a strength of the ARA model that it highlighted the question of the origin of the first or "primer" eTh. All theories must answer this question. Janeway [39] and Matzinger [40] have appreciated this and, unaware of our solution, proposed a mechanism competing with it.

Our answer [33-35, 37] was that there must be an antigen-independent pathway unique to helpers that provides the primer eTh (*figure 10*). These, then, autocatalytically drive the response by associative recognition of antigen via the APC with iTh (i.e., iTh-APC-eTh).

This pathway, although fundamentally different from the Burnet-Lederberg pathway, shares with the latter the assumption of an antigen-independent differentiation affecting uniquely the pathway from iTh to eTh. Further, in our case, the pathway is used 'solely' as a source of "primer" eTh (*figure 11*). The steady state level of eTh must be kept low if the frequency of autoimmunity is to be minimized. For a detailed discussion of this pathway, see reference [37].

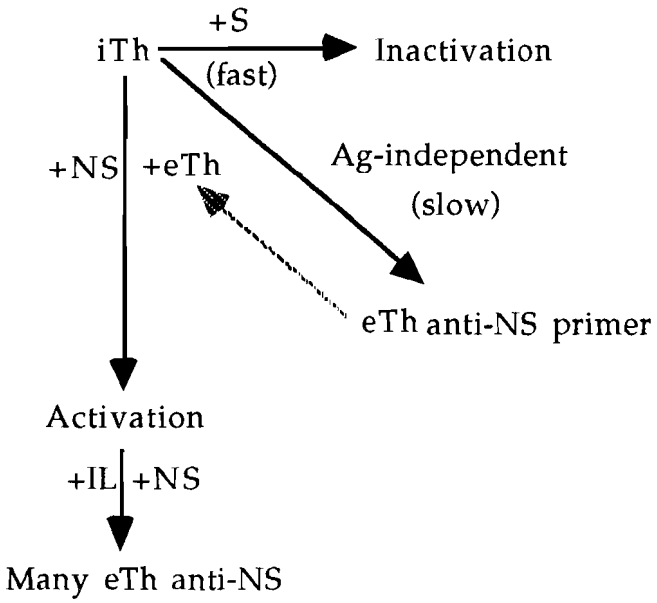


Figure 10. The Primer pathway for effector T-helpers (eTh).

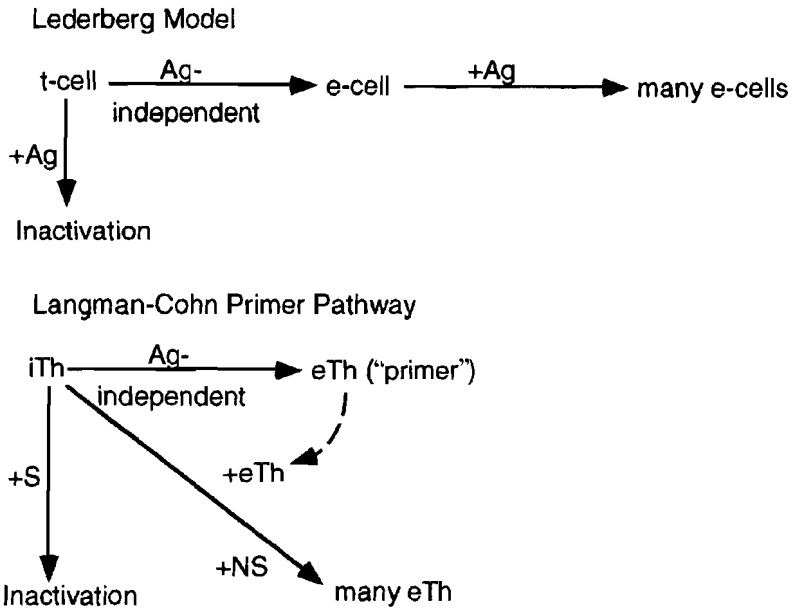
The fundamental concepts governing the S-NS discrimination as a learning process

There are many factors that determine whether the immune system will respond to an antigen (meaning, generate effectors anti-Ag). One of these factors is whether or not the specificity elements necessary to recognize the antigen in question are present. This depends on a somatically-selected S-NS discrimination that is antigen-specific. It is this property that is under analysis here. Whatever other factors that one wishes to add under the umbrella of another definition of the S-NS discrimination, a consideration of the antigen-specific element remains unavoidable. The steps between i-cells and e-cells are driven by the antigen-specific events with respect to S versus NS comprising Decision 1 plus a set of antigen-related events with respect to determination of effector class comprising Decision 2.

Given the steady state generation of a repertoire of anti-S plus anti-NS combining sites (paratopes) and a collection of S-antigens and NS-antigens, how might the somatic selection process sort them out? A given cell expressing an antigen-specific receptor must be told whether it is anti-S or anti-NS. This means that the universe of antigens must be separated in some way into S and NS. One property that is available to make this separation is developmental time. S is 'prior and persistent'. NS is 'posterior and transient'.

Prior or posterior to what?

There is a period during ontogeny when antigens defined as S are present, antigens defined as NS are absent, and antigen-specific cells (i-cells) are absent. Under the ARA model, as i-cells arise, if they interact with antigen (S), they would be deleted because there is an insufficiency of eTh to activate. Those i-cells that do not interact with antigen accumulate



	Lederberg	Langman-Cohn
Born as:	t-cells	i-cells
Tolerance is:	property of cell (t)	property of animal (±eTh)
Model for:	all cells	iTh→eTh primer only
Self is:	persistent	prior and persistent
Nonself is:	transient	posterior and transient
Signals:	Same signal interpreted as [1] by *t* and as [2] by *e*	Signal[1] via antigen-receptor Signal[2] via eTh→a-cell
Tolerance:	epitope by epitope	epitope by epitope
t--------->e:	Ag-independent	not relevant
iTh----->eTh:	not relevant	Ag-independent
Induction of effectors:	epitope by epitope	antigen by antigen (associative recognition)
Mutation:	in e-cell to anti-S is "lethal"	in e-cell to anti-S is controlled

Figure 11. A comparison between the Lederberg (1959) and the Langman-Cohn (1983) models.

Authors	During developmental time Self must be: Prior	Persistent
*Germline-Selected**		
Burnet and Fenner (1948-1957) (Self-marker)	–	–
TNTC (e.g., 38, 39) (1948-present) (Nonself-Marker)	–	–
Somatically selected		
Jerne (1955) - Burnet (1958-1964) (Window of GOD)	+	–
Lederberg (1959) (Tolerizable-only cell)	–	+
Bretscher-Langman-Cohn (1968-present)	+	+
	– not required	+ required

Figure 12. The role of 'prior and persistent' under the various models of the S-NS discrimination.
* Self must be prior and persistent with respect to germline.

as anti-NS. When the system becomes responsive, which means that a sufficiency of eTh anti-NS arises to prime the response, these i-cells anti-NS can be activated to respond to NS. Thus, S is defined as an antigen that the immune system encounters prior to its being activable (during the period when there is an insufficiency of eTh). As long as S persists the state of unresponsiveness to it is maintained because iTh anti-S are deleted as they arise.

By definition, S is that which is present 'prior' to the appearance of activatability when there is an insufficiency of eTh, whereas NS appears after the presence of a sufficiency of eTh and is 'transient' due to an effector response that rids it.

This is a true learning or historical process in that the response of the immune system to an antigen depends on the previous experience of the system with respect to that antigen.

How did the various models deal with the problem of 'prior and persistent'? (figure 12)

The Pauling Template model could not deal with the S-NS discrimination. NS or S-marker theories are based on a germ line-specified S-NS discrimination and, therefore, do not depend on either the 'prior' or 'persistence' of antigen. A somatically-selected, antigen-

specific S-NS discrimination does depend on 'prior' and/or 'persistence' of antigen. The Jerne [20]-Burnet [24] model only requires that S be present prior to the inducible-only Stage III. S need not be persistent because no anti-S specificities are assumed (can be permitted) to appear after closing the Window of GOD. The Lederberg [19] model only requires that S be persistent because the t-cell cannot distinguish S from NS. Persistence of S prevents accumulation of e-cells anti-S. The Associative Recognition of Antigen Model, which starts with i-cells and is driven by eTh, requires that S be 'prior and persistent'. The normal establishment of tolerance is a once in a lifetime event determined during a developmental stage when all S is present and no eTh are present. After that tolerance is maintained by the persistence of S. This makes the S-NS discrimination a historical process. The response of the immune system depends on its prior experience with respect to that antigen. Under the Jerne/Burnet/ARA models, S is defined during a developmental window and, once defined, is fixed ("the static S").

The competing contemporary landscape: the bottom line

My analysis has not dealt with the present day competing frameworks. I would like to close with a brief commentary on them. They range from a minor modification of the ARA model to a complete rejection of its principles. The competing models are three types, Suppresser, NS-Marker and Contextualist.

Suppresser models

These models (e.g., [41]) are based on the same principles as the ARA model but differ in how the principles are translated into mechanism. Suppresser models are the mirror image of the ARA model requiring suppression to inactivate the i-cell and only a ligand driven signal via the antigen-receptor to activate it [38]. For the ARA model, inactivation ("tolerance") is mediated epitope-by-epitope, while activation is mediated antigen-by-antigen (ARA). For suppresser models, inactivation ("tolerance") is mediated antigen-by-antigen (ARA), while activation is mediated epitope-by-epitope, unless a role for eTh is appended ad hoc (i.e., unsupported by an argument of principle). This class of models has been analyzed elsewhere [42] and while not ruled out are certainly less explicative than the ARA model.

NS-marker models

There has been a recent flurry of such models based on one principle. There is a germ line-selected property (i.e., a NS-marker) common to all foreign antigens that determines whether the immune system will respond. These NS-markers are variously described as pathogenicity ('stranger signal'), danger, localization, modes of entry, dose, etc. These NS-marker models challenge the generalization that there is no chemical or physical property of antigens that can be used by the immune system to separate S from NS as classes. This is why the ARA model is based on a separation in developmental time. While these factors are important in another context, they play no role in the antigen-specific somatically-learned S-NS discrimination (Decision 1) of immune systems. Rather, they should be mapped onto Decision 2, the choice and magnitude of effector class responses.

The argument of principle can be stated as follows:
– the germ line-selected recognitive elements of defense mechanisms have a specificity that distinguishes the selecting agent from the 'S-of-the-species'. The somatically-selected recognitive element of immune systems have a specificity that distinguishes the 'S-of-the-individual' from all else. As the repertoire of recognitive elements of immune systems is much larger than that of defense mechanisms (i.e., largely non-overlapping), there is no way that a S-NS discrimination operating on the large repertoire of somatically-generated recognitive elements could be solved by the use of germ line-selected recognitive elements that overlap with only a small proportion of the immune system's paratopic repertoire. Clearly, if the defense mechanism repertoire recognized everything that the immune system recognized, there would be no selection pressure for the latter. Therefore, we reject all S [14]-and NS [39,40]-marker models that are claimed to operate at the level of the S-NS discrimination (Decision 1). For detailed analysis of these models, see References [11-13].

Contextualist models

These are the least well formulated at the level of mechanism but present an enormous appeal to historians, philosophers and mathematicians. The Contextualist models are driven by the assumption that S is ever changing (an assumption shared with NS-marker models) and that the immune system is essentially S-oriented. This class of models is an outpouring from Jerne's preoccupation with S-recognition and its translation into idiotype networks. This class of model can be rejected on the grounds that the immune system has no way to track an ever changing S. It requires S to be 'prior and persistent' (i.e., static). Further, there is no way to select for recognition of S in the germ line, if a somatic S-NS discrimination deletes its expression. Lastly, recognition without consequence (the so-called "formal" idiotype network) is also unselectable. A representative proposal [43] by the leading contextualists, Atlan and Cohen, has been analyzed elsewhere [44].

Conclusion

The pathway of development of the models

The models for the S-NS discrimination unfolded one step at a time. Interestingly, between Pauling (1940) and Bretscher-Cohn (1970), these stepwise changes resulted in a total paradigm shift (*figure 13*) due to the introduction of a Decision 1 step between anti-S and anti-NS at the level of the i-cell ("antigen-responsive cell") (*figure 9*) instead of the linear ontogenic pathway to filter out anti-S activity (*figures 7, 8*).

Propositions that must be faced

The following are the conclusions to be challenged in considering competing models for the S-NS discrimination.

1) a S-NS discrimination is obligatory whenever the effector output of the protective mechanism is biodestructive and ridding for both host and "pathogen;"

2) the S-NS discrimination is germ line-selected in the case of defense mechanisms ("innate"). The S-NS discrimination is somatically-selected in the case of immune systems ("adaptive");

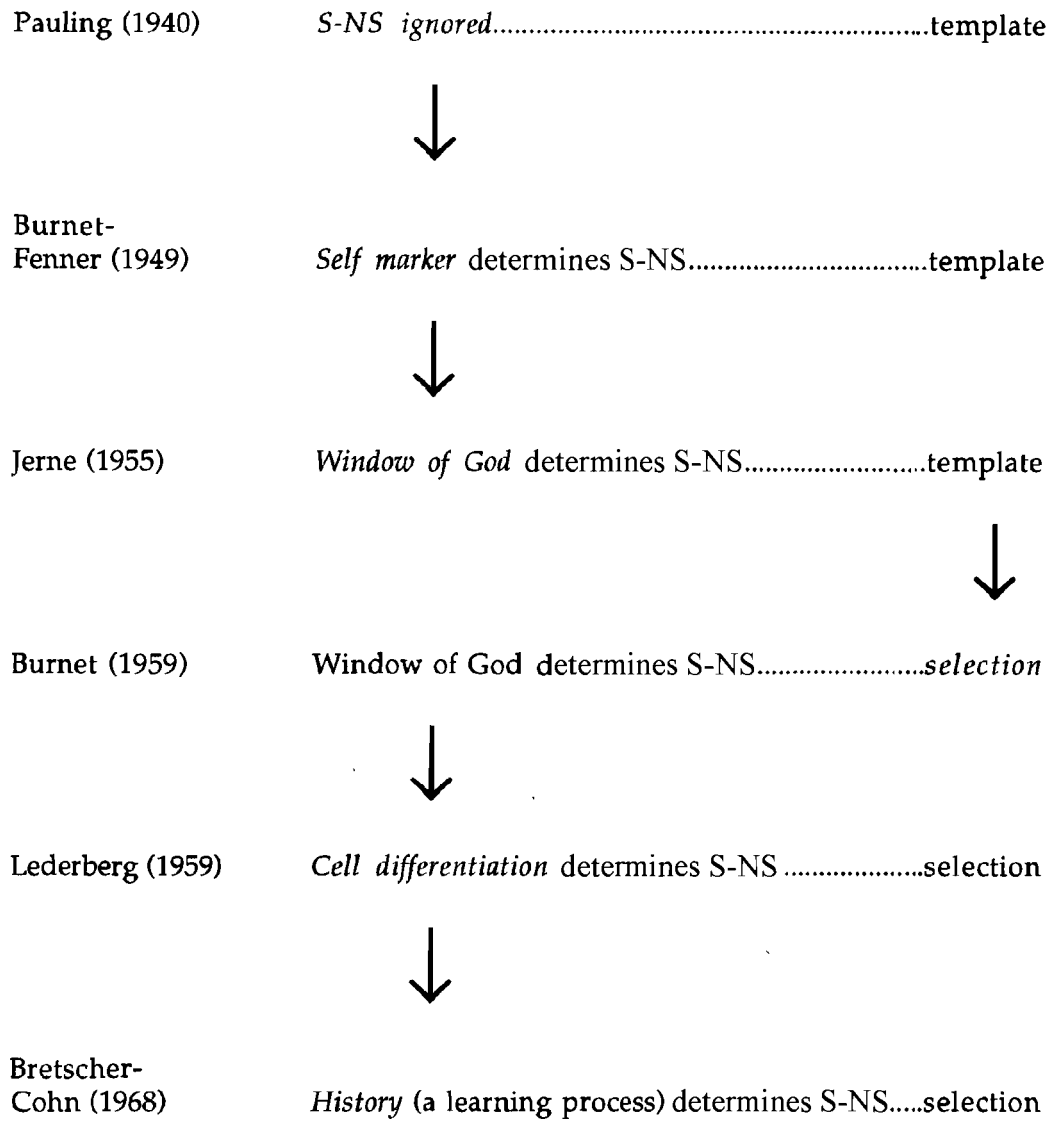


Figure 13. The evolution of the models of the S-NS discrimination.

3) the S-NS discrimination by either defense mechanisms or immune systems is the sole evolutionary selective pressure driving the degree of specificity of the recognitive elements for antigen;

4) there is no physical or chemical property of antigens that can be used by the immune system ("adaptive") to distinguish S and NS as classes;

5) only distinction that is based on developmental time has led to meaningful models of how the somatic selection process distinguishes S ('prior and persistent') from NS ('posterior and transient');

6) the antigen-specific decision step in the steady state process of somatic selection results in anti-S being inactivated ("tolerance") and anti-NS accumulating, poised for activation upon appropriate encounter with NS;

7) responsiveness to NS depends on the appearance of "primer" eTh anti-NS that are derived via an antigen-independent pathway;

8) inactivation ("tolerance") is mediated epitope-by-epitope (Signal[1]), whereas activation is mediated antigen-by-antigen (Associative Recognition of Antigen);

9) the steps between activation of cells anti-NS and the appearance of effector function are influenced by a set of germ line-selected factors that are dependent on the physical and chemical nature of the antigen (antigen-related properties that are not antigen-specific) and its interactions with host cells;

10) any errors in the antigen-specific Decision 1 step of somatic selection that allows anti-S to sneak through and be activated cannot be specifically corrected by antigen-unspecific or antigen-related germ line-selected factors (danger, integrity, localization, route, quantity, pathogenicity, etc.) acting at the level of Decision 2 (see *figure 5*).

Acknowledgements

This paper was written in 1998 while Melvin Cohn was a Scholar-in-Residence at the Fogarty International Center for Advanced Study in the Health Sciences, National Institutes of Health, Bethesda, Maryland, USA.

Bibliography

1 Cohn M. 1994.The wisdom of hindsight. Ann Rev Immunol 1994; 12:1-62.
2 Cohn M. Some thoughts on the response to antigens that are effector T-helper independent ("thymus-independence"). Scan J Immunol 1997; 46: 565-71.
3 Cohn M. A new concept of immune specificity emerges from a consideration of the Self-Nonself discrimination. Cell Immunol 1997; 181: 103-8.
4 Cohn M. At the feet of the Master: The search for universalities. Divining the evolutionary selection pressures that resulted in an immune system. Cytogenet Cell Genet 1998; 80: 54-60.
5 Talmage DW. Immunological specificity: An alternative to the classical concept. Science 1959; 129: 1643-8.
6 Langman RE, Cohn M. The E-T (elephant-tadpole) paradox necessitates the concept of a unit of B-cell function: The Protecton. Mol Immunol 1987; 24: 675-97.
7 Cohn M, Langman RE. The Protecton: the evolutionarily selected unit of humoral immunity. Immunol Rev 1990; 115: 1-131.
8 Langman RE, Cohn M. A theory of the ontogeny of the chicken humoral immune system: The consequences of diversification by gene hyperconversion and its extension to rabbit. Res Immunol 1993; 144: 421-46.
9 Langman RE, Cohn M. What determines k/L ratio? Res Immunol 1992; 143: 803-11.
10 Langman RE, Cohn M. Reply to Takeda et al. Immunol Today 1996; 17: 200-1.
11 Langman RE, Cohn M. A short history of time and space in immune discrimination: reply to the commentaries. Scand J Immunol 1997; 46: 113-6.
12 Langman RE, Cohn M. Terra Firma: A step back from 'danger'. J Immunol 1996; 157: 4273-6.
13 Langman RE, Cohn M. The essential Self: a commentary on Silverstein and Rose "On the mystique of the immunological Self." Immunol Rev 1997; 159: 214-7.
14 Burnet FM, Fenner F. The Production of Antibodies. Melbourne: Macmillan and Company Limited; 1949.
15 Sexton C. The Seeds of Time. Melbourne: Oxford University Press; 1991.
16 Owen RD. Immunogenetic consequences of vascular anastomoses between bovine twins. Science 1945; 102: 400-1.
17 Brent L. A history of transplantation immunology. San Diego: Academic Press; 1997.
18 Le Douarin N, Corbel C, Bandeira A, et al. Evidence for a thymus-dependent form of tolerance that is not based on elimination or anergy of reactive T cells. Immunol Rev 1996; 149: 35-53.
19 Lederberg J. Genes and antibodies. Science 1959; 129: 1649-53.

20 Jerne NK. The natural selection theory of antibody formation. Proc Natl Acad Sci USA 1955; 41: 849-56.
21 Talmage DW. Allergy and immunology. Ann Rev Med 1957; 8: 239-56.
22 Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. Aust J Sci 1957; 20: 67-9.
23 Burnet FM. The mechanism of immunity. Sci Am 1961; 204: 158-64.
24 Burnet FM. The Clonal Selection Theory of Acquired Immunity. Nashville: Vanderbilt University Press; 1959.
25 Silverstein A, Rose N. On the mystique of the immunological Self. Immunol Rev 1997; 159: 197-206.
26 Miller JFAP, Basten A. Mechanism of tolerance to Self. Curr Op Immunol 1996; 8: 815-21.
27 Burnet FM. The clonal selection theory of immunity-A Darwinian modification. Aust J Science 1964; 27: 6-7.
28 Coutinho A, Coutinho G, Grandien A, Marcos MAR, Bandiera A. Some reasons why deletion and anergy do not satisfactorily account for natural tolerance. Res Immunol 1992; 143: 345-54.
29 Bretscher PA, Cohn M. Minimal model for the mechanism of antibody induction and paralysis by antigen. Nature 1968; 220: 444-8.
30 Cohn M. Speculation on cellular and molecular mechanisms involved in immunological responsiveness and tolerance. In: Landy M, Braun W, Eds. Immunological Tolerance. New York: Academic Press; 1969. p. 281-338.
31 Bretscher P, Cohn M. A theory of Self-Nonself discrimination. Science 1970; 169: 1042-9.
32 Bretscher P. The control of humoral and associative antibody synthesis. Transplant Rev 1972; 11: 217-67.
33 Cohn M. Antibody diversity 1983: Some elementary considerations. In: Yamamura Y, Tada T, Eds. Progress in Immunology. V. Orlando, Florida: Academic Press; 1983. p. 839-851.
34 Langman RE. Was everyone a "little-bit-right" after all? Res Immunol 1992; 143: 316-22.
35 Langman RE. The Immune System. San Diego: Academic Press; 1989.
36 Bretscher P. The two-signal model of lymphocyte activation twenty-one years later. Immunol Today 1992; 13: 74-6.
37 Cohn M. The Self-Nonself discrimination: Reconstructing a cabbage from sauerkraut. Res Immunol 1992; 143: 323-334.
38. Cohn M. The ground rules determining any solution to the problem of the Self/Nonself discrimination. In: Matzinger P, Flajnik M, Rammensee HG, Stockinger G, Rolink T, Nicklin L, Eds. The Tolerance Workshop Vol. 1,. Basle: Editiones Roche, 1987; p. 3-35.
39 Janeway J, C.A. The immune system evolved to discriminate infectious Nonself from noninfectious Self. Immunol Today 1992; 13: 11-6.
40 Matzinger P. Tolerance, danger and the extended family. Ann Rev Immunol 1994; 12: 991-1045.
41 Modigliani Y, Bandeira A, Coutinho A. A model for developmentally acquired thymus-dependent tolerance to central and peripheral antigens. Immunol Rev 1996; 149:155-74.
42 Langman RE. The Self-Nonself discrimination is not regulated by suppression. Cellular Immunol 1987; 108: 214-19.
43 Atlan H, Cohen IR. Immune information, self-organization and meaning. Int Immunol 1998; 10:711-17.
44 Langman RE, Cohn M. Away with words: Commentary on the Atlan-Cohen essay "Immune Information, Self-Organization and Meaning." Int Immunol. 1999; 11: 865-70.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Assisting immunologists to examine the philosophical foundations and implications of the new theories of tolerance

Kenneth F. Schaffner

The 1990s witnessed a vigorous and still emerging debate about the fundamental nature of immunology that over the past 2 years or so has seized even the popular imagination. Much of this debate is framed around the self/nonself distinction, with a variety of challenges to that notion that has appeared to be part of the foundations of immunology for the past 50 years. The implications of a possible overturning of a central concept of immunology are potentially vast: from new ways to research vaccines and control infectious diseases (including HIV/AIDS), to novel understandings of cancer and autoimmune diseases and new therapies, to better ways to keep transplanted organs from being rejected by the recipient. However, some have expressed serious doubts about this potential revolution, among them Arthur Silverstein, Silverstein writing with Noel Rose, Jacques Miller, and Melvin Cohn. [References are keyed to the debates discussed in the paper and are available on-line and linked to the debates; see reference URLs at the end of this paper.]

More focally, over the past 5 years, an increasingly contentious debate has developed related to the danger theory, championed by Fuchs and Matzinger, but also involving other major alternatives. These contrasting approaches include extensions of Jernean idiotype network theories by Coutinho, Stewart, and Bandeira, as well as "stranger," morphostasis, "integrity," cytokine cascade, and antigen localization accounts by Janeway, Cunliffe, Dembic, Weigle, and Zinkernagel, respectively. The associative recognition model of Bretscher, Cohn and Langman has also evolved in this context to respond to this debate in immunology.

This backdrop set the stage for the on-line "Cutting-Edge" debate that was held on the HMS Beagle web site in May 1997. Having a debate in which many of the major proponents of competing theories could actively present their views and criticisms of alternative approaches was an idea that emerged from the HMS Beagle staff. Beth Schachter, then the editor, was able to recruit a stellar group of participants, who each suggested several possible questions and themes for debate, and who then slugged it out over the course of a week responding to questions and to each other. As a historian and philosopher of immunology, but not a proponent of any specific theory, I was asked to serve as the reasonably objective moderator, and distilled and prioritized the submitted questions. In addition, at the end of each day, I attempted to summarize where the discussion had gone that day, and

posed what seemed to be the next logical question to the participants. This paper is a summary, and a reflective interpretation, of that debate.

The participants included one of the originators of the "danger" theory, Ephraim Fuchs from Johns Hopkins. (Polly Matzinger could not participate in the May 1997 debate because of previous travel conflict, but did so in the more recent Salk Institute debate cited in the reference section.) The May 1997 debate also included a proponent of the "integrity" theory, Zlatko Dembic of the University of Oslo, and a spokesperson for the network theory, Antonio Bandiera, an associate of Antonio Coutinho, both at the Institut Pasteur. Also involved in the debate were Bill Weigle of the Scripps Institute, an experimental immunologist and proponent of an experimentally-based cell-cytokine approach, Doug Green from the La Jolla Institute for Allergy and Immunology, who initially tried to take a nontheoretical stance, but ended up arguing against a self/nonself distinction and in favor of a "local damage" theory that "walked" like a danger theory, and Rod Langman from the Salk Institute, a vigorous defender and developer of the Bretscher-Cohn-Langman associative recognition (and self/nonself) model. Melvin Cohn of the Salk Institute did not participate, but did observe the debate, and provided some valuable background to the moderator prior to the debate. Arthur Silverstein at Hopkins also had a travel conflict and could not participate, but circulated a draft essay by himself and his colleague Noel Rose on the subject area of the debate to all of the participants, and that article was subsequently published in the October 1997 issue of *Immunological Reviews*.

The May 1997 on-line debate began with a general question that asked what were the main functions of the immune system, or to put it another way, why was the immune system developed by evolution? This question was intentionally designed to look for whatever consensus we felt might emerge, before polar positions were taken on the viability of the self/nonself distinction, and on the competing models. In addition to a question each day, a rationale, and for later questions, a summary as part of that rationale, was posted.

Day 1 (first) question

The first day's question was the following: what has the immune system evolved to do? In other words, what is (are) the main function(s) of the immune response? Can we reach any consensus on this? At this initial stage, please keep the answers fairly general – we will get into details and mechanisms (both cellular and molecular) later.

This question was immediately followed by a brief background and rationale paragraph.

Background and rationale for the forum and the first question

The past few years have seen exciting challenges to widely accepted immunological knowledge. Perhaps the most recent challenge is the "danger" theory championed by Fuchs and Matzinger, but other major alternatives have been advanced. These include extensions of Jernean idiotype network theories by Coutinho and Bandeira, as well as "stranger," morphostasis, "integrity," and antigen localization accounts by Janeway, Cunliffe, Dembic, and Zinkernagel, respectively. The associative recognition model of Cohn and Langman has also evolved in this context to respond to this debate in immunology. These contrasting views will probably see the main function(s) of the immune system differently, but perhaps we can come to some agreement about the main function(s), and

then move on to look at the specifics of the alternative approaches, and their strengths and weaknesses. The import of these various approaches for understanding and controlling responses to pathogens, to organ transplantation, and perhaps to cancer, may be significant.

Day 2 question

For day 2, the more provocative question and rationale, which also summarized what consensus we seemed to have reached, read as follows.

Is the self/nonself distinction (still) important in immunology, and if so, what do these terms (self/nonself) mean? If not, what are better terms (e.g., integrity, danger, etc.), and why are they better?

Rationale

All postings thus far accept a major role for the immune system in detecting and eliminating "pathogens," while not attacking the body or the immune system (self-destruction of the immune system). In recognizing some things as to-be-eliminated, and others as not, is this tantamount to an 'implicit' definition of the self/nonself distinction? If not, is there a better way to conceptualize this difference at a general level, or must the discussion proceed to mechanisms (cellular and molecular)? If some alternative to the self/nonself distinction is favored, why has self/nonself been so important for the past 50 years?

Responses to this question quickly revealed the anticipated fractures among the group. First, Dembic suggested that self/nonself was a good approximation to the truth, but that the "danger" and "integrity" approaches were better. Langman, however, argued that if the immune system can kill pathogens and not the host, it 'must' [my emphasis] make a self/nonself distinction, and maybe the language of self/nonself should be replaced with something less freighted, though the issue would be the same. Fuchs suggested self/nonself is useful in immunology as long as it generates novel and testable theories of tolerance, that though it had done so in the hands of Burnet, Lederberg, and Bretscher and Cohn, the self/nonself distinction should now be ditched. Weigle said self/nonself was useful as long as we recognized its limitations, but we should not be looking for just 'one' general mechanism. Green weighed in with what he said would be an "inflammatory" view: that self/nonself is no longer useful to the study of immunology, and we would do better to look at a more specific level. Bandiera's view was that if we understand it right – that is, from the network point of view – the self/nonself terms are acceptable.

This increasingly vigorous exchange foreshadowed day 3 on which we had the most extensive series of postings (49 in all!), which did deal with the 'specifics' of the competing models. In the space available for this paper, it is impossible to summarize the day 3 discussion, though you can obtain some of the flavor of this dispute from the authors in this volume, and the details of the debate exchange are available on-line.

Suffice it to say by the end of day 3, with the theoretical arguments shooting and ricocheting every which way, that it seemed like some experimental testing and grounding needed to be brought into the picture. Thus, day 4's question asked the participants to look more closely at novel experimental evidence, at important classical experiments from their perspective, and also at potential clinical implications of the new and developing views. Day 5 asked how we might choose between the theories, and if yet-to-be-done exper-

imental results could make that decision. The results from these final 2 days were most interesting, and though the more vociferous theoretical contributors were less vocal on these two topics, they were clearly not silent.

I was particularly struck by Doug Green's response to day 5's question, which was my most explicit attempt to introduce a philosophical dimension into the debate.

Day 5 question

The question that Green replied to, and its rationale were as follows.

Can you think of any experiment that would distinguish between alternative tolerance models – that might clearly support one and not the others?

Rationale

The debate has identified some importantly different accounts of how tolerance (and the immune response) is induced, and has clarified these differences. The different accounts say that there are different things at work. These different things include a danger signal, a primer (effector) T-cell, a loss of integrity signal, an antigen-presenting cell provoked cytokine cascade, a temporally evolving network, among others. Without necessarily asking you to commit to any one 'unitary' model, can each of you propose one or more "crucial experiments" that would have different results depending on which model holds, and that thus might confirm or disconfirm these different models? Relatedly, though it is difficult to prove a negative, how long is it reasonable to search for experimental support for a hypothetical signal or a process, given the likelihood that not 'all' these things are likely to exist (or work the way the model proposes)? Are there other non-experimental (maybe theoretical?) factors that need to be considered in connection with the model that one accepts, And if so, what might they be?

Green replied as follows, and it is worth quoting him in extenso on this:

> Our final question in this debate is probably the toughest, and yet what could be fairer than to ask for the experiments that could possibly distinguish between conflicting theories. Isn't this what we do in science?
>
> Well, yes and no. Certainly we test hypotheses (e.g., IL[interleukin] 2 is required for T-cell proliferation), but we do not generally do this with theories – not in the same sense. In immunology, for example, no experiments have rigorously tested the clonal selection theory or the network theory. Before this statement raises 'alarms,' I don't mean that the theories have not generated lots of testable ideas that have been explored to great effect. Its just that they haven't been tested in the same sense that we test a hypothesis.
>
> This is partly because theories are so 'slippery.' As soon as we propose a test, and that test fails, the proponents of the theory will usually retort that it wasn't a definitive test. Not because it wasn't properly designed, but because experiments rely on concrete entities and theories do rather well with abstract things. If we propose to test the theory that co-stimulation is an integral part of tolerance, we could examine a mouse in which a co-stimulatory molecule is absent. If this does not have the predicted effect, we do not scrap the theory, we simply look for another co-stimulatory molecule. And we'd be right to do so.

> So what good are these theories if they're not testable? Several of us have stated here, and I think it's worth repeating that theories have value if they generate interesting and testable hypotheses. When the well of new hypotheses dries up, or when they are no longer interesting, the theory may either be replaced or become dogma. For some of us, the theory itself holds a certain beauty in its own right, for others of a more practical bent, the theory is useful because it leads us to new findings.
>
> Which is why I've worked to try to convert a theory, rife with constructs and models, into a story with real-life players. It's kind of a middle ground. It's a story of how the immune system might work, and since it's built of real players, all of the elements are testable. But just as with any theory, if any testable idea fails, we don't scrap the story, we just swap players. But the value lies in the questions that are asked. So what are the questions? Here are a few.
>
> I've stressed the role of cellular trauma in initiating immune responses. What are the components released from a necrotic cell that actually recruit an inflammatory response? (There have been a few suggestions, but surprisingly little work has been done to identify this well-known response to cellular injury.) How does cellular stress (e.g., of epithelial cells) trigger cytokine release? How important are mast cells in initiating the vascular changes that occur early in inflammation, and how critical are these for initiating an immune response? (We can envision knocking out genes that are critical for mast cell differentiation or function.) And of course I've noted that trauma isn't always completely necessary. The system is clearly setup to respond to molecules that are commonly associated with parasites – one example is the activation of complement via the "alternate" pathway. How many such signals are "hard-wired" into the system? How critical are such signals (under what circumstances will tissue damage be sufficient to give us responses in the absence of exogenous adjuvants)?

Doug Green's comment here seems right to me, and also resonates with how I interpret the earlier debate between the clonal selection theory and its competitor, the instructive template theory. (Although on that earlier topic, both Art Silverstein and Fred Tauber may still disagree with my interpretation, though I think for different reasons.)

But let me make two general philosophical points that I think are implicit in Green's comment about theories, models, and experiments, and also illustrate it by referring to an exchange between Fuchs and Langman on day 5 of the debate.

My first philosophical point is that whereas Green uses the term "slippery" to describe theories, philosophers would use the terms "flexible" and "experimentally underdetermined." These are insights based on the work of Pierre Duhem, the French philosopher-physicist of the early twentieth century, and go against Karl Popper's later views about strong and simple falsificationism. The Duhemian themes were readdressed and generalized by the American philosopher Quine, and reappear in Tom Kuhn's notions of the resistance of paradigms to experimental falsification. Thus, the flexible and indeterminate character of theories is well recognized and accepted in recent and contemporary philosophy of science.

What is more contentious, and needs further work, is the fine-structure and interrelations among theories, more specific models, auxiliary assumptions, and experiments. There have been a number of attempts to look at these issues, and they must be done so in a diachronic or temporally developing way if they are to make any sense of the science.

Philosophers from Kuhn, through Lakatos, Laudan, Shapere, and Kitcher, and a number of others, have proposed schemas for describing these disputes over time. My own preferred approach, found primarily in chapter 5 of my 1993 book on *Discovery and Explanation in Biology and Medicine,* develops the notion of a 'temporally extended theory', in which there are levels of generality, or, what Lindley Darden urges, of abstractness (for references to the philosophy of science literature see the bibliography in my 1993 book, p. 569-609). The most general or abstract level typically involves those concepts and principles that are central to the theory – that distinguish it from its competitors. And this highest level tends to 'persist' throughout a debate. An example would be an associative recognition 2-signal model or a commitment to a "danger" signal. At more specific levels, details are introduced – and there can be a number of ways, mutually inconsistent successive ways, of specifying these elements. Here is where we get into model specifics and then into specific mechanisms. These can change, or evolve, or be given up and replaced, but without giving up the highest or most persistent assumptions. It is this highest or most abstract level that confers unity on the approach, and also individuates that approach from its competitors. Thus to my mind, it is not only heuristic, but of critical importance as well.

What forces change, at both specific and general or more abstract levels, and ultimately brings closure, is a topic for another paper, but suffice it to say it a complex set of arguments that I like to think of as a logic of comparative theory evaluation, involving the fit among other theories and mechanisms, the match with experimental results, and appeals to vaguer notions such as simplicity and the lack of ad hoc hypotheses. Such a "logic" is not a prospective decision instrument, or a crystal ball, in part because biological theories involve additions at the more specific levels indicated, and thinking up those additions is an exercise in an ampliative discovery or problem-solving process. But such a logic can be used retrospectively to rationalize the process of comparative theory debate, and it may have some utility in identifying contemporaneous areas that need more attention and discussion.

Perhaps Fuchs's replies to the day 5 questions illustrate some of Green's points as well as my more philosophical gloss. Fuchs on day 5 spoke from a kind of Popperian point of view, though he himself did not cite Karl Popper. Fuchs wrote:

> It has been said that, although it is possible to disprove a theory, it is never possible to prove it. As the official advocate of the danger model in this discussion, I will therefore propose experiments that would disprove the danger model. The two key features of the danger model are 1) only "professional" antigen presenting cells, such as dendritic cells, can initiate immune responses among naive T cells; and 2) the 'professional' antigen presenting cell must be activated by an exogenous or endogenous 'danger' signal to become immunogenic for naive T cells. Thus, one experiment that would prove the danger model wrong is if it could be shown that either a B cell or a parenchymal (non-hemopoietic) cell could activate a naive T cell.

This looks very straightforward – and a bold claim begging for falsification, but Fuchs immediately began to add some qualifiers:

> Many may say that this has been done already, but I must stress the word 'naive' in the previous sentence. There are indeed experiments that show that a purified pop-

> ulation of CD4+ T cells with putative markers of the naive phenotype (such as high expression of the CD45R molecule, low expression of the CD44 molecule, etc.) do respond to B cell-presented antigens, but we would claim that these markers do not reliably distinguish naive from previously activated T cells. There are other studies showing that purified T cells from unprimed T cell receptor transgenic mice are activated by B cells, but such T cells may actually express two distinct types of receptors on their surface, one pairing the transgenic alpha and beta chains, and one in which transgenic beta chains are paired with endogenous alpha chains. The TcR combining the endogenous alpha with the transgenic beta could be specific for environmental antigens, providing the stimulus for this bispecific T cell to become a memory cell. The bispecific cell could then respond in vitro to the antigen for which the transgenic pair is specific, even when presented by a B cell.

Other participants, however, did not find this *prima facie* falsificationist view quite so clear, perhaps for reasons that Green articulated in the long quote above. For example, Langman replied to Fuchs that:

> It's tough to devise a test that depends on 'naive' cell populations without some objective assay of this property. It is not too helpful to argue, for example, that the behavior of cells as predicted by the model is our only assay of naiveté. Is there a clear definition we all can use experimentally before going further to test the model?

And Langman further added that:

> As an aside, I would like to think that there is more to immune regulation than the two signals that allow i-state cells to be directed to tolerance or immunity. Host and parasite derived factors must play a huge role in modulating the progress of cells after the self/nonself decision has been taken and they head towards becoming effectors. And, there are more modulating effects that influence the non-immune effector mechanism that actually perform the ridding function. The self/nonself discrimination is just the first small step in the life of a T or B cell.

If I and other philosophers are right, there is, as Lakatos once wrote, "no instant rationality." The on-line debate caught a narrow time slice of a still-evolving larger debate. So it is not surprising that overall, none of the debaters convinced any of the others to give up their favored theory. However, the differences between the theories, as well as some common families of theory, were identified. Day 5's discussion indicated in fairly explicit ways why changing minds is, and should be, difficult. There was a consensus that emerged over the course of days 1 and 2 that the debate had to move to the level of specific models (as it did in day 3), but some disagreement continued as to whether higher level principles, such as self/nonself or integrity, added to immunological knowledge at this point. A particularly valuable aspect of the interchange was the clear and concise formulation of the various competing models in day 3, and their follow-up in days 4 and 5. Several new experiments or experimental directions were proposed for further study during days 2 to 5, and day 4 gave these prominent theorists in immunology a chance to articulate some potentially momentous clinical implications of their various models.

To be involved as a moderator in this debate was a most interesting, and privileged, experience. In this paper I have tried to share some of the debate with you, but I am still – after 3 years – digesting it and attempting to draw conclusions. I would encourage readers to access it and its follow-up Salk Institute debate on the web, and hopefully, continue this exciting exchange in still further forums.

Bibliography

The full 1997 debate is available on-line at: <http://biomednet.com/hmsbeagle/12/cutedge/synopsis.htm>.
An extensive set of references, including background readings and all works cited in the debate, is available at: <http://news.bumn.com/hmsbeagle/12/cutedge/iirefs.htm>.
The more recent 1999 debate mentioned in the text above as involving a number of the participants from the 1997 debate as well as some important additional contributors is available on-line at: <http://www.cig.salk.edu/seminars.htm>.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

The immunological approach to classification of leukemias

Patrick Triadou

Hematology is a fascinating branch of medicine. The last 30 years or so have witnessed an astonishing increase in knowledge, so much so that even the vocabulary of hematology has been transformed. With each new technique of investigation that has been introduced, the inevitable consequence has been a further splitting-up of diseases previously considered to be well circumscribed, even if ill-understood. Leukemias provide an example of this process. Immunological techniques have changed ideas on leukemias profoundly. On the one hand, the question of the nature of the leukemia cell has undergone analysis at the molecular level with monoclonal antibodies (mAb), and on the other hand, the understanding of the genomic abnormalities of Burkitt's lymphoma has provided a model for classifying blood cancers.

During more than a century after the first description of leukemia major advances have depended upon the use of the microscope and various staining procedures. Classification for this period mainly consisted of grouping together clinical and hematological features, not according to any selected properties but according to their most important resemblances. The two main criteria used in the classification were the clinical course of the disease and the type and maturity of the predominant leukemia cell. This has resulted in some confusion in the terminology.

To understand to what extent the immunological approach to leukemias changed hematology in its modes of concept formation and ontological implications, it is essential to pay attention to its temporal dimension by tracing the sequences by which concepts are gradually modified and the manner in which a model can contribute to help guide further research. This paper examines the role of immunology in modifying the interpretation of leukemias by introducing a molecular paradigm that implies a shift in the meaning of the observation statement.

Leukemias before immunology

If we rapidly glance at the beginning of the history of leukemia, we notice that all attention was focused on the cell. With the application of Virchow's cell theory to human pathology during the second half of the nineteenth century, the point of immediate importance was the manner by which cells entered into the definition of leukemia. The discovery of leukemia depended upon the presence in the observer's mind of a clear conception by which observed clinical and morphological facts may be analyzed and connected.

Leukemia was described as a clinical entity almost simultaneously by Craigie, Bennett and Virchow after the forerunner work of Donné. Bennett's paper, published in October 1845, described the autopsy findings on a patient presenting two major features: enlargement of the spleen and changes in the color and consistency of the blood. Definitive identification of leukemia required the use of a microscope, both as an instrument of research and as a diagnostic tool. Without attention turned to the visible globule it would have been difficult to distinguish an alteration affecting the blood system itself from the already known pus and inflammatory pathology. Virchow also knew that blood contains colorless corpuscles and noticed that the ratio of pigmented to colorless corpuscles seemed to be reversed. He used the term "leukemia" or "white blood" to describe the disease. Bennett chose the more exact term of "leucocythemia" or "white cell blood." During the second half of the nineteenth century, Paul Ehrlich developed a stain procedure that produced the morphological hematology. This led to the sub-classification of leukemias into myeloid and lymphoid forms and later, when taken together with their associated clinical pictures, into acute and chronic leukemias. Although Neumann was the first to identify leukemia as a disease involving the marrow, acute leukemia was defined as a separate clinical entity in 1889 by Ebstein. Further morphological characterizations were made when polychrome stains became available for the study of blood cells. From that time on ideas about leukemia and hematopoiesis could not be dissociated.

The history of hematopoietic stem cells is in fact almost as old as cellular hematology itself. In the early decades of this century, battle raged between the great hematological schools regarding the morphological identity of these cells. This explained, on the one hand, the interest of Ferrata in defining the stem cell which he termed "hemocytoblast" and, on the other hand, the debates between hematologists proposing various schemes of phylogeny. In 1911 Pappenheim, who based his idea on observations of the nucleus, regarded genesis of leukemia cells as a pathology of normal maturation with abnormal development leading to the appearance of a new kind of cell [1-6].

Briefly, during the 80 years preceding the 1920 advances in the understanding of blood cell differentiation, the location of hematopoietic organ precursors as well as the classification of leukemias were dependent on progress in histology and staining procedures. This was the result of the creation of a group of skilled microscopists corresponding to Virchow's and Bennet's initial conception of the bases of rational clinical practice. Microscope and Virchow's cell theory are the two handles of the paradigm by which leukemia was apprehended, although some practitioners remained obsessed by the disease's clinical aspects. This not only reflected a traditional controversy of the second part of the nineteenth century between two antagonistic camps, but also underlined the need for coherence between the pragmatic and rational aspects of medical classification.

Late period of morphology, ontological conception of leukemia etiology

Crucial for the evaluation of treatment later in the twentieth century was a uniform system of classification once chemotherapy had become increasingly effective. In the mid 1970s various classifications were introduced ; however, the same terms were used to name different entities, and different terms to define the same findings. For these reasons, in 1976 a group including seven French, American and British hematologists (FAB) proposed in 1976 a uniform system of classification and nomenclature that would permit more ac-

curate recording of cases entered into clinical trials. These proposals were based on conventional morphological and cytochemical methods. It became possible on a large scale to define subgroups and to see whether there were any correlations between the subgroups and clinical and laboratory findings, and in the response to treatment [7].

Until this period immunological tools for classifying leukemias were lacking, but lymphoproliferative diseases were identified as specific entities. In the 1960s the diagnosis of lymphoproliferative diseases was essentially based on the lymphocyte count, the aspect of blood and bone marrow, lymph node histology, analysis of serum immunoglobulins, and clinical manifestations. The formation and function of lymphocytes were known only from the studies performed in the chicken and in certain mouse strains. Methods distinguishing T and B cells in humans were regarded as under development [8]. The etiology of leukemia was unknown, but a number of factors, both environmental and inherited, were under consideration.

The bacteriological notion of cause gave rise to the theory that infectious diseases, particularly tuberculosis, caused leukemia. When this hypothesis turned out to be incorrect, unidentified microorganisms were suspected. The description of transmission of chicken sarcoma with cell-free extracts in 1911 by Peyton Rous remained ignored for a long time. In 1953 Gross demonstrated cell-free transmission of murine leukemia. While during this period the infectious theory had a major boost, except for some tumors in birds and some mammals there was no evidence of leukemogenic viruses in human [9-12].

The other important theory was heredity, as a few authors reported families with either a high incidence of leukemia or hereditary diseases carrying a high risk for leukemia. The occurrence of leukemia is increased in Down's syndrome, which is associated with abnormalities of chromosome 21. Muller's demonstration of mutagenic properties of X-rays also suggested that radiation might alter genetic information. Nevertheless, until the 1960s thought about the etiology of blood malignancy remained, as with other forms of cancer, very diffuse.

mAb and the molecular biology of Ig (Immunoglobulins) and TCR (T cell receptor) genes

Until 1970, leukemia cells were generally more or less regarded as an envelope containing a nucleus. However, subsequent studies on the immunophenotype and structure of genes resulted in a complete revision of ideas about the relationships between leukemic cells and normal cells, and about the etiology of leukemia. The cell came to be considered as a complex of multiple molecular markers. While it is possible to see elements of continuity running through medical ideas, it is only with this modern vision that the leukemia cell was explicitly related to molecular defects. During the course of this inquiry, the definition of leukemia cells changed, but the relation between the new and traditional conception remained ambiguous. Some revisions of the existing interpretation and classification were made necessary simply by the emergence of a much wider body of primary information.

The hybridoma technique described in 1975 by Kohler and Milstein for production of mAb revolutionized the development of human T-lymphocyte antibodies and offered the possibility of defining precisely the stages of lymphocyte and granulocyte differentiation. In addition to the traditional cell surface markers defining B and T lymphocytes and cytochemical staining of myeloid cells, highly specific mAb that distinguish cell surface

membrane antigens were used to classify leukemia cells [13]. These advances have led to important insights into leukocyte differentiation and the cellular origin of leukemia.

As reviewed in different papers in the early 1980s, several investigators have described mAb that react not only with B and T lymphocytes but also with granulocytes, monocytes, platelets and leukemia cells. This allowed ALL (acute lymphoblastic leukemia) to be divided into five major subgroups: unclassified ALL, common ALL, pre-B-ALL, B-ALL and T-ALL. The immune classification has therefore to be compared with the previous classification of ALL based on morphological features of leukemia cells under light microscopy. The FAB cooperative group has proposed, as mentioned, a cytomorphological classification of ALL that suddivides patients into L1, L2 and L3 subgroups. The L1 subgroup, for example, is characterized by a homogeneous population of predominantly small cells with a high nuclear-cytoplasmic ratio and few nucleoli. Except for the L3 subgroup, attempts to correlate morphologic subgroups and immune classification have been largely unsuccessful [14, 15].

Another important matter of debate was the nature of the precursors of blood cells in hematopoiesis and of their relation to leukemic cells. mAb provided a challenge to the traditional view and were used extensively in order to propose schemes of myeloid and lymphoid differentiation according to the cell phenotype. These schemes were based on the concept that the phenotype of normal lymphoid cells at each level of differentiation was comparable to the phenotype of its malignant counterpart. The reason for such a hypothesis was founded on substantial evidence that the phenotypes of most malignant leukemia cells were not unique but instead reflected the phenotypes of normal cells. None of the surface markers described were leukemia specific [16]. The common acute lymphoblastic leukemia antigen – cALL or CALLA, for example – was initially defined by antisera produced in rabbits by immunization with E-rosette negative, surface-membrane Ig-negative ALL cells. A J-5 monoclonal antibody that reacts with the cALL antigen present on normal bone marrow cells has been shown to also react with renal tubular cells or small intestine epithelial cells.

With DNA analysis it became possible to follow a whole series of modifications taking place during the differentiation process. Chromosomal location and genomic organization of the genes encoding the Ig and TCR polypeptides were described in the early 1980s for the Ig and later in this decade for T-cell receptors. These genes are arranged in a germ line configuration as discontinuous segments of DNA. During B and T cell lineage commitment, the Ig and TCR genes undergo an orderly sequence of somatic rearrangements, which ultimately leads to a functional recombined gene with transcription. With the availability of molecular probes for these genes, two important questions could be investigated by gene rearrangement studies: the clonality and lineage of certain putative B and T cell malignancies [17, 18]. A last threat to the traditional classification of leukemia emerged with the combined use of mAb and molecular probes that identify the rearrangement of Ig and T-cell receptor genes. Non-T-ALL could be subdivided into six distinct subgroups and T-ALL into three primary subgroups.

Although mAb were fruitful in distinguishing acute myeloid from acute lymphoid leukemias, they have been less so in the classification of acute myelogenous leukemia. Seven subtypes of acute myelogenous leukemia (AML) were identified with the widely-used FAB group classification relating the morphological appearance of leukemia cells to

presumed normal hematopoietic counterparts. Because there was a controversy over prognostic and therapeutic significance provided by the FAB system, alternative classifications based on antigenic marker expression have been proposed. Thus four types of AML were distinguished according to differentiation-associated phenotypes as identified by mAb reactive with the differentiation antigens of myeloid cells. But within the four groups that have been isolated, each phenotype contained more than one morphological type of AML. Thus there was considerable morphological heterogeneity, showing that attempts to correlate surface marker phenotype and FAB classification were unsuccessful [19].

Controversy over classification highlighted the limits of molecular markers to offer a comprehensive alternative definition of leukemia cell.

Burkitt's lymphoma model and the mechanism of leukemogenesis

With the development of methods for analyzing chromosomes and the discovery of oncogenes the understanding of the biology of leukemia has changed. The new conception of leukemogenesis is not content just to mirror a biochemical science that has developed independently. It plays an essential role in building up the new representation of the process leading to leukemia that involves genes and proteins and replaces the earlier notion of cause. Expressed in this way one does not expect to answer the question of the initial cause of leukemia, but to explain the biochemical mechanism responsible for producing leukemia cells. In this view, Burkitt's lymphoma translocations could be regarded as a model that has to be taken into consideration for an explanation of other blood cancers.

The stem-line conception of tumor evolution as well as the proposition that mutations were responsible for neoplasia opened up possibilities that simply could not have occurred to scientists before the beginning of the twentieth century [20]. This was not generally accepted for some time. A combination of molecular biology and cytogenetics has now settled the issue. However, it was not until 1960 that biologists could describe the first specific abnormal chromosome in leukemia cells from patients with chronic myelocytic leukemia (CML). The Philadelphia chromosome laid the foundation for the view of the clonal origin of cancers. With more sophisticated banding techniques, it became apparent that chromosome changes were widespread among leukemia cells. Furthermore, the different forms of leukemia seem to have particular cytogenetic abnormalities. It was demonstrated that the designated t (9;22) aberration observed in CML resulted from a reciprocal translocation between chromosomes 9 and 22 [21-24]. Burkitt' s lymphoma, a malignancy of B cell origin, first described in 1958, was initially associated with Epstein Barr virus. Some 30 years later it has become one of the models explaining the mechanism of leukemogenesis and is used as a reference for the molecular classification of hematological cancers. Burkitt's lymphoma is known to be characterized by one of three reciprocal translocations, t (8;14), t (2;8) and t (8;22), involving respectively the Ig heavy and light chain genes [25, 26].

Translocations have two main consequences. Either the gene for TCR or an Ig protein comes to lie near a proto-oncogene, thereby activating it, or the breaks occur within a gene on each chromosome involved, creating a fusion gene encoding a chimeric protein. Research on the biology of retroviruses led in 1976 to one of the most important findings in cancer, which is that sequences homologous to viral oncogenes are found in human DNA [27-29]. It was discovered that cellular oncogenes are involved in different aspects

of cellular differentiation and regulation, and that mutation of these genes plays a major part in tumorigenesis. In recent years their chromosomal location have been determined as well as the way their structures could be modified [30].

In Burkitt's lymphoma the c-MYC gene translocation typifies the situation in which a cellular oncogene is juxtaposed to an Ig or TCR gene by chromosomal fusion, thereby activating the oncogene [31-32]. On the other hand, it turns out that the translocation between chromosomes 9 and 22 observed in patients with CML involves the movement of c-ABL, which is normally located on chromosome 9 to chromosome 22. The fusion of BCR and c-ABL genes on the Philadelphia chromosome symbolizes the situation in which breakage occurs within the intron, producing fusion genes [33].

Thus, the two first chromosomal translocations studied at the molecular level became the references for a molecular classification of hematopoietic cancers, which can be divided into non-fusion and gene-fusion hematopoietic tumors [34]. Continuing with the model of Burkitt's lymphoma various translocation genes have been found near chromosomal breakpoints in chronic forms of leukemia and lymphoma, as well as in acute T leukemias. In this context, BCL2 protein and HOX 11 protein have received much attention because the first one blocks apoptosis and the second, not expressed in normal T cells, may well activate target genes in leukemic T cells with translocation t (10;14) [35, 36].

If we ask today what these advances tell us about the pathogenesis of leukemia, the answer clearly is that there is no single cause. It seems likely that most cases of leukemia result from endogenous damage to DNA with the appearance of mutations or chromosomal translocations that initiate the pathway of leukemic transformation. Although a great deal has been learnt about the c-Myc protein, we still do not understand why c-MYC, rather than other oncogenes, is important for the appearance of Burkitt's lymphoma. While it is impossible to give an answer to this kind of question and to explain the genesis of specific translocations, interest has been focused on the alterations of proteins encoded by genes located at the translocation breakpoints, on their functions and on the mechanisms of their interactions.

c-Myc protein, for example, has several functional domains including a basic region for DNA binding and helix-hop-helix and leucine zipper protein dimerization motifs. This means motifs found in DNA-binding proteins and in protein-protein interactions involved in regulation of transcription. The alteration or activation of transcription seems to be the key element in most of the chromosomal abnormalities described in leukemia. Activation of c-MYC by translocation compromises a transcriptional network involving at least three other factors, all of which also have the same domains. All these proteins are presumably in monomer-dimer equilibrium in the normal cell and it is supposed that excess c-MYC expression stemming from translocation would cause a shift in equilibrium that results in transcription of downstream target genes leading to oncogenesis. Several DNA-binding proteins activated by chromosomal translocation have been reported in T-ALL.

Although as exemplified by Burkitt's model chromosomal aberrations might be thought to activate most proto-oncogenes by juxtaposing them to other genes, the study of the increasing numbers of abnormalities indicates that, as originally observed with the Philadelphia chromosome, gene fusion resulting in the synthesis of chimeric proteins, is the main repercussion. The study of BCR-ABL fusion has paved the way for many more descriptions of fusion proteins often involving transcription factors. An important

example of transcription factor fusion occurs after translocation t (15;17). It can be concluded that proteins in both categories are often transcription factors. Thus, disruption of transcriptional control plays a major role in the etiology of blood cancers [37-38].

Conclusion

The basic object on which hematology was founded at the end of the nineteenth century is the cell, the lowest biological unit. Henceforth, hematologists attempted to classify the different types of leukemia according to their morphological characteristics. Although traditional categories retained their value for the comparison of data and as a guide to appropriate treatment, the main breakthroughs that led to our present molecular understanding of leukemia relies on immunological and biochemical approaches. Immunological and genetic tools, as well as the conceptual apparatus developed by molecular biologists, provided useful insights into the problems raised by leukemogenesis. That point marked the start of a discussion about which aspects of hematology – the classical or the morphological, the immunological or the molecular – could serve as the most adequate means for the construction of a theory of leukemogenesis. Among these two aspects, the second one has been recognized as the most suitable, as it indicates the mechanism accounting for leukemia cell formation. The interest in the cause disappeared and attention has been focused on gene defects and on protein-protein interactions that result in dysregulation of cellular functions. Immunologists and molecular biologists enter into play, characterizing cell markers, decoding the mechanisms by which leukemia is established and sequencing the corresponding genes. With the spread of molecular techniques, the seductive power of the new paradigm is increasing, but we still do not know why the initial genetic defect appears. Apart from this epistemological problem, there are some important benefits from a clinical point of view. Immunophenotypic analysis with the availability of large panels of mAb improved diagnosis and classification of leukemia according to the stage of cell differentiation, especially in children [39]. In addition, chromosomal abnormalities are effectively found in 80 to 90% of children with acute leukemia and the identification of specific genetic lesions has proven to be important in the management of this disease.

Leukemia is not a marginal field dominated by a few medical experts. It provides an overarching rationale to different concerns about medical practice. Very similar ideas circulate about other types of cancer. Many aspects of the understanding of disease are dependent upon the available techniques and the prevailing paradigm. The comparative history of biomedicine highlights the extent to which medical knowledge is transformed by a distinct set of ideas. Definitions of diseases, hypotheses concerning their etiology and classifications for therapeutic purposes are closely connected with the techniques used for analyzing the biological characteristics of the same diseases. Morphology, immunology and molecular genetics represent three different approaches to cell pathology that illustrate the historical process of the development of medical knowledge.

Bibliography

1 Piller GJ. John Hughes Bennet, his life and the identification of leukemia. Proc R Cool Physicians Edinb 1997; 27 (Suppl 3): 1-11.

2 Debru C. Identification et définition des leucémies: Donné, Bennett, Virchow. Classer les leucémies et l'inclassable. In: Debru C, Eds. Philosophie de l'inconnu: le vivant et la recherche. Paris: Presses universitaires de France; 1998. p.196-215.

3 Donné A. Cours de microscopie complémentaire des études médicales. Anatomie microscopique et physiologique des fluides de l'économie, Paris: Baillère; 1844. p. 135.
4 Wintrobe MM. Leukemia. In: Clinical Hematology. Philadelphia: Lea & Febriger; 1967. p. 983-1098.
5 Bennet JH. Case of hypertrophy of the spleen and liver, in which death took place from suppuration of the blood. Edin Med Surg J 1845; 64: 413-23.
6 Ebstein W. Ueber die acute Leukämie und pseudoleukämie, Dtsch Arch Klin Med 1889; 44: 343-8.
7 Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DAG, Gralnick HR, Sultan C. Proposals for classification of the acute leukemias. Br J Haematol 1976; 33: 451-9.
8 Wieser RS, Myric QN, Pearsall NN. Fundamentals of immunology for students of medicine and related sciences. London: Lea & Febiger; 1969.
9 Vecchio G. Oncogenes of DNA and RNA tumor viruses and the origin of cellular oncogenes. Hist Phil Life Sci 1993; 15: 59-74.
10 Gaudillière JP. Le cancer entre infection et hérédité: gènes, virus et souris au National Cancer Institute (1937-1977). Rev Hist Sci 1994; XL VII/1: 57-89.
11 Rous P. A sarcoma of the fowl transmissible by an agent separable from tumor cells. J Exp Med 1911; 13: 397-411.
12 Gross L. Spontaneus leukemia developing in C3H mice following inoculation in infancy with A-K leukemic extracts, or AK embryos. Proc Soc Exp Biol Med 1951; 76: 27-32.
13 Kohler G, Milstein S. Continuous cultures of fused cells secreting antibody of undefined specificity. Nature 1975; 256: 495- 8.
14 Schroff RW, Foon KA, Billing RJ, Fahey JL. Immunological classification of lymphocytic leukemias based on monoclonal antibody-defined cell surface antigens. Blood 1982; 59: 207-15.
15 Foon KA, Schroff RW, Gale RP. Surface markers in leukemia and lymphoma cells: Recent advances. Blood 1982; 60: 1-19.
16 Foon KA, Todd R.F. Immunological classification of leukemia and lymphoma. Blood 1986; 68: 1-31.
17 Tonegawa S. Somatic generation of antibody diversity. Nature 1983; 302: 575-81.
18 Griesser H, Tkachuk D, Reis MD, Mak WT. Gene rearrangements and translocations in lymphoproliferative diseases. Blood 1989; 73: 1402-15.
19 Griffin JD, Mayer RJ, Weinstein HJ, Rosenthal DS, Coral FS, Beveridge RP, Shlossman SF. Surface marker analysis of acute myeloblastic leukemia: Identification of differentiation-associated phenotypes. Blood 1983; 62: 557-63.
20 Chaganti RSK. Significance of chromosome change to hematopoietic neoplasms. Blood 1983; 62: 515-24.
21 Nowell PC, Hungerford DA. A minute chromosome in human chronic granulocytic leukemia. Science 1960; 132: 1497-9.
22 Rowley JD. A new consistent chromosomal abnormality in chronic myelogenous leukemia identified by quinacrine fluorescence and Giemsa staining. Nature 1973; 243: 290-3.
23 Groffen J, Stephenson JR, Heisterkamp N, de Klein A, Bartam CR, Grosveld G. Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell 1984; 36: 93-9.
24 Melo J.V. The molecular biology of chronic myeloid leukaemia. Leukemia 1996; 10: 751-6.
25 Burkitt D. A sarcoma involving the jaw in African children. Br J Surg 1958; 46: 218-23.
26 Lenoir GM., Preud'homme JL, Berheim A, Berger R., Correlation between immunoglobulin ligth chain expression and variant translocation in Burkitts lymphoma. Nature 1982; 298 : 474-6.
27 Stehelin D, Varmus HE, Bishop JM, Vogt PK. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 1976; 260: 170-3.
28 Spector DH, Varmus HE, Bishop JM. Nucleotide sequences related to the transforming gene of avian sarcoma virus are present in DNA of uninfected vertebrates. Proc Natl Acad Sci (USA) 1978; 75: 4102-6.
29 Morange M. The discovery of cellular oncogenes. Hist Phil Life Sci 1993; 15: 45-58.
30 Rabitts TH. Chromosomal translocations in human cancer. Nature 1994; 372: 143-9.
31 Erikson J., Ar Rushdi A., Drwinga HL, Nowell PC, Croce CM. Transcriptional activation of translocated c-Myc oncogene in Burkitt lymphoma. Proc Natl Acad Sci (USA) 1983; 80: 820-4.
32 Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, Aaronson S, Leder P. Activation of the c-Myc gene into the immunoglobulin heavy chain locus in human Burkitt's lymphoma and murine plasmocytoma cells. Proc Natl Acad Sci (USA) 1982; 79: 7837-41.
33 Shtivelman E, Lifshitz B, Gale RP, Canaani E. Fused transcripts of abl and bcr genes in chronic myelogenous leukemia. Nature 1985; 315: 550-4.
34 Cline MJ. The molecular basis of leukemia. New Engl J Med 1994; 330: 328-36.
35 Drexler HG, LacLeod RAF, Borkhardt A, Jansen JWG. Recurrent chromosomal translocations and fusion genes in leukemia-lymphoma cell lines. Leukemia 1995; 9: 480-500.
36 Solomon E, Borrow J, Goddard AD. Chromosome aberrations and cancer. Science 1991; 254: 1153-60.

37 Grignani F, Fagioli M, Alcalay M, Longo L, Pandolfi PP, Donti E, Biondi A, Lo Coco F, Grignani F, Pelicci PG. Acute promyelocytic leukemia: from genetics to treatment. Blood 1994; 83: 10-25.
38 Tenen DG, Hromas R, Licht JD, Zhang DE. Transcritption factors, normal myeloid development and leukemia. Blood 1997; 90: 489-519.
39 Pui CH. Acute leukemia in children. Curr Opin Hematol 1996; 3: 249-58.

HISTORIANS LOOK AT IMMUNOLOGY

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Receptor immunology: the origins of Paul Ehrlich's guiding principle

Arthur M. Silverstein

The study of specific receptors is, in many respects, the central issue in modern immunology and its allied disciplines. One thinks not only of the important role of B- and T-cell receptors and of Fc receptors, but of receptors which transduce the signals of the many cytokines, of receptors for certain pharmacological agents against which we may develop autoimmune disease, and even of the receptors that mediate the penetration of viruses into cells. Indeed, it has become clear in recent decades that many important physiological processes are mediated by the binding of stereochemical groupings to their specific receptors, just as substrates bind to their respective enzymes. It is uncommon that the general concept underlying a major breakthrough in the biological sciences will appear suddenly and fully formed. It results more often from the slow accretion of many varied facts that leads to a formulation that is tested and re-tested, each time slightly modified, until a mature system of thought is recognized and made explicit. This was the way with Darwin's theory of evolution and with Virchow's cellular pathology; so also was it the way that Paul Ehrlich developed his concept of the role of receptors. He then applied this concept broadly, most notably in his side-chain theory of antibody formation.

In his chapter in the 1914 *Festschrift* celebrating Ehrlich's 60th birthday [1], Leonor Michaelis wrote in 1914, "...the side-chain theory was established finished and ready in [the 1885 monograph on] 'The oxygen requirements of the organism'"[2] at a time when there was not yet an immunology. This was Michaelis' attempt to show that Ehrlich's theory had grown on an earlier rootstock. But only 5 years later, Michaelis would unearth Ehrlich's long-lost thesis for his M.D. degree [3], entitled "Contributions to the theory and practice of histological staining" [4], written in 1878. The discussion of the mechanism of staining by the 24-year-old Ehrlich put back to an even earlier date the germination of the theory that would guide him in all of his future scientific endeavors.

In this paper, rather than giving a detailed review of Paul Ehrlich's immunology, I would like to outline the early history of the idea that determined all of his scientific research. During this period, Ehrlich made notable contributions not only to histological staining, but also to cell respiration, to the founding of hematology, and to a variety of clinical areas. It goes without saying that the background to Paul Ehrlich's theoretical approach to immunology applies equally well to his noteworthy research in the field of chemotherapy [5].

The early years

Even while Ehrlich was at the gymnasium, his future love of chemistry was presaged by his answer when asked to explain the meaning of his final examination essay in composition, on the obligatory subject of "Life – a Dream". The young student stammered in distress, and finally came out with, "...you know...life is...a chemical incident...a normal oxidation...and the dream...the dream is...a fluorescence of the brain" [6]. Ehrlich would later write, recalling his youth, "I really believe myself that my talents lie in the field of chemistry; I can picture the chemical formulae in my mental vision..." [7].

While still a student, Ehrlich had the opportunity to visit his cousin Carl Weigert in the pathology department in Breslau. It was Weigert, one of the first to introduce the use of the newly discovered aniline dyes into histology, who showed stained tissue preparations to Ehrlich, and who pointed out that some cells stain well with certain dyes, whereas others stain only poorly or not at all – a demonstration that would be remembered later, with important consequences.

Ehrlich entered university in Breslau in 1872, to study medicine. Here he came under the influence of anatomist Wilhelm von Waldeyer, who exposed Ehrlich further to histological methods for the differentiation of cell types. Throughout his medical student days, Ehrlich experimented with the wealth of new dyes emerging from the growing German chemical industry. He would test each dye on a variety of tissues, and under a variety of conditions, so that his bench top presented a spectrum of colorful solutions, and his fingers and occasionally his face were often highly stained.

In 1877, Ehrlich published his first scientific paper on "Contributions to the knowledge of aniline staining and its use in microscopic technics" [8]. In this maiden effort, Ehrlich described the technical aspects of tissue staining, and the variable staining qualities of a variety of tissues and cells. Interestingly, he devoted much of the paper to the study of the distribution of plasma cells in different tissues, and especially in the components of the lymphoid system, tonsil, Peyer patches, peripheral lymph nodes, and spleen. It would be more than 70 years before the importance of these cells for the discipline of immunology was discovered [9].

In 1878, the 24-year-old medical student published his histochemical magnum opus, "Contributions to the theory and practice of histological staining," as a dissertation for the M.D. degree. This was a truly remarkable body of work, especially for a 24-year-old undergraduate. The entire thesis testifies not only to the contemporary absence of chemical science in histologic technology, but also to the self-confidence that would characterize Ehrlich throughout his career. He, an outsider, would dare to introduce rigorous scientific method into this hitherto purely empirical field. He would later be as daring and innovative in immunology and in chemotherapy.

Three important points emerge from this early example of Ehrlich's approach to biological research.

The first point is chemical; staining reactions are chemical in nature rather than physical. (Twenty-odd years later, Ehrlich would argue for the chemical interaction between antibody and antigen against Jules Bordet's and Karl Landsteiner's physical adsorption ideas.)

The second point of the dissertation is also chemical; there is discernible in staining reactions a certain degree of specificity, in that certain dyes react preferentially with

certain cells or structures. (Ehrlich felt that the dye must attach to some sort of receptor, based upon charge or other characteristics.)

Finally, the third point is chemical as well; to a great extent, structure appears to define function. (The nature of the groups attached to the aniline core of the dye define not only molecular charge, but also solubility and strength of attachment.) Here was the seed of a receptor theory that would take fuller form, first in Ehrlich's 1885 cell respiration studies, then in his 1897 side-chain theory of antibodies, and finally in his ultimate triumph, the design of such chemotherapeutic agents as salvarsan, the future magic bullets of chemotherapy. As Michaelis pointed out, one learned from this dissertation how "...the idea of the chemical binding of foreign substances to the protoplasm developed on reflection about the nature of staining, and how from this idea later developed the side-chain theory" [10].

This point is discussed even more fully in 1959 by Maria-Louise Eckmann in her doctoral dissertation on the historical significance of Ehrlich's staining studies. She joins Michaelis in noting that, "This idea [the chemical binding of substances to cells] dominated Ehrlich's life. A straight path led from the doctoral work past the significant publication *Die Sauerstoffbedürfnis* to vital staining, and further to the side-chain theory, whose experimental basis shaped the work on toxins, antitoxins, and immunity" [11].

The Charité decade

Having completed his studies and already well known for his histologic staining, Ehrlich received an invitation to become an assistant in Professor Friedrich Frerichs's Second Medical Clinic at the prestigious Charité Hospital in Berlin. The Charité was a teaching hospital, where the developing relationship between chemistry and medicine was well recognized, and where basic and clinical research was encouraged. This was especially true of Frerichs's department. Once the young Ehrlich had demonstrated his research talent, Frerichs, whose favorite maxim was "Caged birds do not sing," allowed him even more time for research than was permitted to other assistants. Ehrlich never seemed to lack for interesting and important research projects, nor for the time to pursue them.

All of Ehrlich's basic research on staining, hematology [12], and the physiology of respiration was performed in the context of continuing clinical activities; during the same period, he published reports on syphilitic heart infarcts, on the occurrence and metabolism of glycogen in diabetics, on acute splenic tumor, and on phosphorus and iodine poisoning. Meanwhile, he found time to introduce, for the first time, the use of fluorescein to study aqueous humor dynamics in the eye [13]. Of further clinical importance, Ehrlich's diazo reaction for the detection of various substances in the urine [14] found broad acceptance in the diagnosis of a variety of febrile diseases, and his demonstration of supravital staining of peripheral nerve endings with methylene blue [15] was widely employed by neuroanatomists.

Of all of Ehrlich's basic investigations during the 1880s, his monograph on "The oxygen requirement of the organism" [16] must be considered his most important. He prepared it as his *Habilitationsschrift*, or inaugural dissertation required for appointment as a university lecturer. Once again, he introduced an innovative technological contribution to medical research; this was the use of redox dyes to study intracellular physiology. Ehrlich had demonstrated earlier the specificity of dye interactions with cells, as we saw above.

Now he used the color changes that accompany the oxidation-reduction reactions of dyes to assess the oxygen-fixing capacities of various tissue cells in the body. Those cells that possess a high affinity for oxygen will provide a reducing milieu within the cytoplasm, and decolorize the highly colored oxidized form of dyes. On the other hand, those cells that bind oxygen only poorly will provide an oxidizing environment and thus bring color to the colorless reduced form of the dye. Ehrlich ascribed the specific physiological functions of the cell to a chemically-conceived *Leistungskern* (activity- or power-nucleus). The term was not meant to describe the anatomic nucleus of the cell, but rather something akin to the aniline nucleus of a complicated dye, where side-chains account for modifications of specific function. Once again, we see Ehrlich dealing in terms of structurally-based specificity, of affinity [17], and of side-chains (such as amino, nitro, and halogen groups) that determine solubility, color, and specificity. Here in this 1885 paper was, almost fully formed, a close approximation of Ehrlich's famous 1897 side-chain theory of antibody formation. In Ehrlich's own words, "...in living protoplasm a [chemical] nucleus of special structure is responsible for the specific function peculiar to the cell, and... to this nucleus are attached, as side-chains, atoms and atom-complexes which are [important]... for its vital activity in general" [18]. At the end of this monograph was a discussion that presaged another of Ehrlich's later interests: immunity to infectious diseases. The influence of Louis Pasteur's and Robert Koch's contributions to bacteriology was already apparent in Ehrlich's speculations; he discussed the implications of his findings for cellular immunity to pathogenic organisms. Most bacteria require ample oxygen for life, so that those cells which bind oxygen strongly should provide a hostile (i.e., immune) environment for such organisms. Ehrlich pointed out that this thesis is not unlike Ilya Metchnikoff's theory to explain cellular immunity against infection [19], except that Metchnikoff's proposal involved only the mobile phagocytes, whereas Ehrlich's referred to parenchymal cells in general.

In his work on blood, Ehrlich tested numerous aniline dyes; he identified and named basophiles and mast cells whose cytoplasmic granules take up basic dyes, eosinophiles which stain with acidic dyes, and neutrophiles that can only be stained with neutral dyes. Here was further evidence that the intracellular elements of different cells might differ chemically, resulting in a degree of specificity in their staining reactions. Once again, the leitmotiv of Ehrlich's work, that chemical specificity depends upon molecular structure, had been verified with outstanding results.

With the death of Frerichs and the accession of Carl Gerhardt at the Charité, Ehrlich was no longer free to pursue his own self-determined directions. Gerhardt insisted upon greater attention to clinical duties, and Ehrlich became frustrated. He later recalled that period, saying, "When in those days I felt so miserable with Gerhardt, I always went to my dye cabinet and said, 'These are my friends, who will never forsake me'" [20]. He would finally take a position in Robert Koch's Institute for Infectious Diseases, where he would be exposed fully to the immunology of Koch, Gaffky, Behring, Wassermann, and others. In the meantime, he studied the derivatives of cocaine, as he said, "...insofar as is possible, to determine the ultimate relationship between chemical constitution, local damage, and anesthetic activity." He was able to comment once again about these molecules that, "As is evident, the side-chains embody the carrier of specific activity..." [21]. Ehrlich's initial assignment from Koch was to work on tuberculin which, for a time, was thought to repre-

sent the best hope for a therapy of tuberculosis. He presented a paper at the International Congress of Hygiene in London in 1891, in which he summarized his entire therapeutic philosophy [22]. Ehrlich outlined the guidelines that would govern all his future work in immunology, oncology, and pharmacology. He declared that there exists a direct relationship between chemical structure and function; the binding of molecules by receptors mediates most of the functions of physiology. Ehrlich pointed out that knowledge of chemistry must necessarily lead to the desired goal of the synthesis of new drugs that will attack the disease and specifically destroy its agent. Here is the clearest statement thus far of the dream that Ehrlich would realize only some two decades later, in his landmark excursion into scientific pharmacology, the creation of a magic bullet.

Ehrlich's immunology

It is not the purpose of this paper to discuss the fine details of Ehrlich's significant contributions to the nascent field of immunology during the decade of the 1890s. It will suffice for our present purposes to review briefly the results that he obtained in applying his ideas of quantitative chemistry and receptor specificity to the problems of the origins and functions of antibodies.

Ricin and abrin

Ehrlich's first immunological studies employed the plant toxins ricin [23] and abrin [24] as antigens [25]. Behring and Kitasato had demonstrated the immune response to diphtheria and tetanus toxins [26], but these were impure and unstable substances, whereas ricin and abrin could be in pure form, permitting quantitative studies. These studies of the antibody response to ricin and abrin produced a wealth of important data. Ehrlich showed:

1) that antibodies can be formed against other than bacterial toxins;

2) how high-titer antisera may be produced using small, increasing doses of antigen;

3) a new method for the quantitative measure of *Immunitätsgraden* (degrees of immunity), the level of protection in the immunized animal;

4) the difference in the duration of immunity conferred by active immunization and by passive transfer;

5) the discrepancy between the small amount of toxin needed to induce immunity and the large amount of antitoxin formed [27];

6) the nature of immunological specificity by clearly demonstrating that whereas ricin and abrin cannot easily be differentiated by their toxicity, they can readily be distinguished using their respective antisera, showing no cross-reaction; and

7) the precise correlation between ricin's in vitro ability to agglutinate erythrocytes and its in vivo toxicity.

Here was an impressive set of observations that would affect all future studies of the nature and consequences of the immune response.

Immunity in the fetus and neonate [28]

The experiments on the immunological relationship of mother and fetus/newborn that Ehrlich published between 1892 and 1894 are some of the most elegant of the late nineteenth century. As he says in the first paper of this series, "...I have been able to succeed in finding a simple research plan which made it possible to establish in each instance the

mechanism of inherited immunity" [29]. Using mice, Ehrlich mated immune fathers with non-immune mothers and vice versa. He then gave the offspring of immune mothers to be suckled by non-immune foster mothers, and vice versa. With these imaginative experiments, he demonstrated:

1) that immunity is conferred by the mother and not the father;

2) that a degree of immunity is transferred to the fetus during pregnancy;

3) that a high level of immunity is transferred to the suckling neonate in the milk of the actively immunized mother;

4) that passive transfer of antibody to the lactating mother proves that the antibody comes from the blood and is not formed in the mammary tissue [30]; and

5) that the passive immunity derived from the mother is short-lived, disappearing within weeks after weaning. Ever alert to the practical, Ehrlich concludes, "Thus, mothers milk is the most ideal food for the newborn." It would be more than 50 years before this work would be repeated [31].

Kinetics of the immune response

In a paper on milk antibodies published by Ludwig Brieger and Ehrlich [32], they immunized a lactating goat and measured the antibody content of the milk periodically. This permitted them to plot a curve of the kinetics of the primary and booster antibody response. They described the following characteristics: 1) a slow rise in antibody titer after the initial injection of antigen; 2) an initial sharp decline in titer for 2 days following each subsequent booster injection; 3) then a rapid rise in titer over the next several weeks to ever higher values; and 4) a subsequent slow decline in titer after the last booster injection. It is apparent that these investigators described as early as 1893 every principal feature of the immune response curve of antibody formation, the weak primary response, the immune elimination phase, the sharp and enhanced booster response, and the slow waning of the antibody titer. Neither the results nor their interpretation would be improved upon for the next 40 to 60 years.

Diphtheria toxin and antitoxin

The demonstration in 1890 by Emil Behring and Shibasaburo Kitasato [33] that diphtheria antitoxic serum could be administered passively to cure diphtheria infection in children opened up a radical new therapeutic approach. However, initial clinical trials of the new approach met with variable success, because high titer antisera could not be routinely prepared and dosages could not be quantified. In his first paper on the subject, published in 1894 [34], Ehrlich and his colleagues showed how high titer anti-diphtheria sera can be produced, following the ricin protocol. In addition, the authors stressed the critical importance of employing the highest titer antisera available in order to attain success in treating the disease. Then, in an extensive clinical trial of diphtheria serotherapy, they confirmed the importance of using high titer antisera, and showed further the critical importance of starting therapy as early as possible after the onset of disease.

Finally, in 1897, Ehrlich published a paper that would finally solve the vexing problem of how to measure accurately the potency of diphtheria antitoxic sera [35]. This was no simple problem, for both diphtheria toxin and its antitoxin are labile, so that a dependable standard had thus far been unavailable. Ehrlich therefore prepared a desiccated antitoxin to serve as the international standard, against which any solution of toxin (and therefore

any antitoxin) can be titered. Using this standard, Ehrlich showed how antisera could be quantitatively and reproducibly titered, thus finally solving the practical problem of diphtheria serotherapy. In his 1914 review of Ehrlich's quantitative approach to immunology Madsen, by then one of the world's leading diphtheria therapy experts, pointed out that, "Ehrlich's method of measurement [of toxin and antitoxin] is the common property of all civilized nations," and "Ehrlich's immunity unit plays the same role for antitoxin measurement as does the Standard Meter for the measurement of length" [36].

The side-chain theory of antibody formation

This concept constitutes the full and final expression of Paul Ehrlich's life-long preoccupation with the idea of receptors, of stereochemical interactions, and of specificity. In its final form, it presented a picture [37] of the cellular origin of protective antibodies as well as showing the domains on the molecule that mediate its functions. Toxins possess a haptophore group that mediates attachment and a toxophore group that does the damage. Antibodies possess an attachment site for antigen, and certain ones have also a site to which complement attaches. In line with his ideas, Ehrlich would coin a set of names (*Amboceptor*, *Zwischenkörper*, and *Komplement*) that carried with them the full semantic message of his theory [38]. He appended this theory to his publication describing the measurement of diphtheria toxins and antitoxins [39]. Ehrlich postulated that all physiologically active substances, including toxins, function by first attaching to preformed receptors on the surface of cells. He claimed that, "The reactions of immunity, after all, represent only a repetition of the processes of normal metabolism, and their apparently wonderful adjustment to new conditions is only another phase of the *uralte protoplasma Weisheit* (the ancient wisdom of the protoplasm)." When an antigen attaches to its specific receptor and exert its effect, the receptor is usually lost and is regenerated by the cell. When large amounts of antigen or repeated doses are administered, the cell overcompensates for the loss of receptors, producing such an excess that they are cast off into the blood – thus circulating antibodies. This was a selection theory of antibody formation that would anticipate by 80 years Niels Jerne's natural selection theory [40] and Macfarlane Burnet's clonal selection theory [41].

Immune hemolysis

After Ehrlich's success in measuring diphtheria toxin and antitoxin, and especially after the widespread success of his side-chain theory, he apparently felt that there remained few unanswered questions in immunity research. It was time to leave immunology and to move on to important new research challenges – to experimental oncology and to chemotherapy. This is not an unusual phenomenon in science [42]; the temptation of great scientists, after having contributed significantly to a discipline, is to conclude that they have solved all of its problems. Lord Kelvin is supposed to have declared in the 1890s that it was all over in physics – this just before Roentgen, Planck, and Einstein. In immunology, both Macfarlane Burnet and Niels Jerne made similar declarations [43] during the 1960s – just before the immunobiological revolution of B and T cells, of lymphokines, and of immuno-genetics.

We return now to Paul Ehrlich. While his theories were widely admired and accepted in Germany, they came under attack elsewhere. His toxin–antitoxin interpretations were challenged by Madsen and Arrhenius from Copenhagen and by Gruber and Landsteiner

from Vienna, and his humoral theory of immunity was contested by Metchnikoff from Paris [44]. But it was principally Jules Bordet, the discoverer of immune hemolysis [45], who questioned Ehrlich's concept of the mode of action of antibodies, and forced the proud Ehrlich to plan new experiments to counter these criticisms. While Ehrlich complained at one point that these later immunological studies were keeping him from other interests [46], they did in fact impel him to undertake some of his most productive work in immunology, the six papers "On hemolysis" done with Julius Morgenroth from 1899 to 1901 [47]. These publications added importantly to our understanding of the mechanism of immune hemolysis. They increased our knowledge of the nature of immunological specificity, by demonstrating for the first time the existence of cross-reactions among the erythrocytes of related species. Finally, they showed how the partial absorption of an antiserum could be effected by a cross-reacting antigen. Others have recognized the heuristic value of the challenges to Ehrlich's data and theories. In his introductory overview of Ehrlich's immunology in the 1914 *Festschrift*, Georg Gaffky wrote:

> It must appear very fortunate, in looking back on progress in immunity that the side-chain theory was not immediately fully accepted without opposition. The ensuing objections and debates caused Ehrlich and his students to perform a long series of magnificent experiments... [48].

Eventually Ehrlich did move on to other research areas, leaving immunology to such assistants as Morgenroth and Hans Sachs. Then, during the first decade of the twentieth century, he would apply his now-fully-mature receptor theory to the chemotherapy of infectious diseases [49], in an attempt to produce the magic bullet that would attack the parasite while sparing the host. He would succeed in this magnificently with the 606th preparation tested for the treatment of syphilis, salvarsan.

Acknowledgements

I thank Darwin Stapleton and the staff of the Rockefeller Archives Center, North Tarrytown, New York, for their generous aid and hospitality.

Bibliography

1 Michaelis L. Das Sauerstoffbedürfnis des Organismus. In: Paul Ehrlich: Eine Darstellung seines wissenschaftlichen Wirkens. Jena: Gustav Fischer; 1914. p. 25. This volume will hereafter be referred to as the Ehrlich Festschrift.
2 Ehrlich P. Das Sauerstoffbedürfnis des Organismus. Eine Farbenanalytische Studie. Berlin: Hirschwald; 1885. Reprinted in Collected papers of Paul Ehrlich. Oxford: Pergamon Press; 1956. vol I. p. 364-432; English translation, p. 433-96. These volumes will henceforth be referred to as Collected papers.
3 Michaelis L. Zur Errinerung an Paul Ehrlich: seine wiedergefundene Doktor-Dissertation. Naturwissenschaften 1919; 7: 165-8.
4 Ehrlich P. Beiträge zur Theorie und Praxis der histologische Färbung [thesis]. Leipzig: University of Leipzig; 17 June 1878; Collected papers, vol. I, p. 29-64; English translation, p. 65-98.
5 Parascandola J., Jasensky R. pointed out that Ehrlich initially thought that receptors do not play an important role in chemotherapy, but eventually would accept this as a guiding principle in his work in this area. Bull Hist Med 1974;48: 199.
6 Quoted in Marquardt M. Paul Ehrlich. London: William Heinemann; 1949. p. 11-2.
7 From Ehrlich's autobiographical notes included in a letter to his friend Christian Herter, quoted in Marquardt, note 6, p. 15.
8 Ehrlich P. Beiträge zur Kenntnis der Anilinfärbung und ihrer Verwendung in der mikroskopischen Technik. Arch Mikroscop Anat 1877; 13: 263; Collected papers, vol. I, p. 19-28.

9 It was finally Astrid Fagraeus who demonstrated that the mysterious plasma cell was in fact the source of circulating antibodies (Acta Med Scand 1948; Suppl 204), a fact elegantly confirmed with fluorescent antibodies by Coons AH, Leduc EH, Connally JM. J Exp Med 1955;102: 49.
10 Michaelis, note 3, p. 168.
11 Eckmann ML. Die Doktorarbeit Paul Ehrlichs und ihre Bedeutung für die Geschichte der histologischen Färbung [dissertation]. Hamburg: University of Hamburg; 1959. p. 26.
12 In his book Hematology: the blossoming of a science (Philadelphia: Lea & Febiger; 1985), Maxwell Wintrobe named Ehrlich as the "father" of modern hematology. It is significant that in Wintrobe's Blood pure and eloquent (New York: McGraw-Hill; 1980), Ehrlich is more frequently cited in the name index than any other individual.
13 Ehrlich P. Ueber Fluoresceinerscheinungen am Auge. Dtsch. Med Wochenschr 1882; 8: 21, 35, 54; Collected papers, vol. I, p. 344-53.
14 Ehrlich P. Ueber eine neue Harnprobe. Z Klin Med 1882; 5: 285; Collected papers, vol. I, p. 619-29.
15 Ehrlich P. Ueber die Methylenblaureaction der lebenden Nervensubstanz. Dtsch. Med Wochenschr 1886; 12: 49-52; Collected papers, vol. I, p. 500-8.
16 Ehrlich, note 2.
17 Ehrlich took his concept of the affinity of oxygen binding to the cell from Pfluger's treatise, Ueber die physiologische Verbrennung in den lebendigen Organismen. Pflugers Arch Ges Physiol Mensch Tier 1875; 10: 251-367.
18 Ehrlich, note 2; Collected papers, vol. I, p. 436.
19 Metchnikoff I. Ueber eine Sprosspilzkrankheit der Daphnien: Beitrag zur Lehre über dem Kampf der Phagocyten gegen Krankheitserreger. Virchows Arch 1884; 96: 177-94.
20 This comment is quoted by Marquardt, note 6, p. 28.
21 Ehrlich P. Studien über die Cocainreihe. Dtsch. Med Wochenschr 1890; 16: 717-9; Collected papers, vol. I, p. 559-66.
22 Ehrlich P. Ueber die neuere Erfahrungen in der Behandlung der Tuberkulose nach Koch, insbesondere der Lungenschwindsucht. Proc VII Int Congr Hygiene Demography (London) 1891; 2: 211; Collected papers, vol. I, p. 13-20.
23 Ehrlich P. Experimentelle Untersuchungen über Immunität. I. Ueber Ricin. Dtsch. Med Wochenschr 1891; 17: 976; Collected papers, vol. II, p. 21-6.
24 Ehrlich P. Experimentelle Untersuchungen über Immunität II. Ueber Abrin. Dtsch. Med Wochenschr 1891; 17: 1218; Collected papers, vol. II, p. 27-30.
25 Ehrlich's studies on the immunology of plant toxins is discussed at length in several chapters in the Ehrlich Festschrift: Madsen Th. Method and quantitative principles in dealing with problems of immunity. p. 151-8; Aronson H. The constitution of toxins. p. 166-90; and Ritz H. Plant toxins. p. 200-8.
26 Behring E, Kitasato S. Ueber das Zustandekommen der Diphtherie-Immunität und der Tetanus-Immunität bei Tieren. Dtsch. Med Wochenschr 1890; 16: 1113. See also Behring, ibid., p. 1145-8.
27 In spite of these results, Hans Buchner would later suggest (Munch Med Wochenschr 1893; 40: 449, 480, 482) that antitoxin is formed directly from the toxin molecule itself by some fairly simple transformation.
28 Ehrlich's work in this area is presented in greater detail by Silverstein AM. Paul Ehrlich: the founding of pediatric immunology. Cell Immunol 1996; 174: 1-6.
29 Ehrlich P. Ueber Immunität durch Vererbung und Säugung. Z Hygiene 1892; 12: 183; Collected papers, vol. II, p. 31-44.
30 Only 60 years later would experiments by Ita Askonas and coworkers (Biochem J 1954; 56: 597-601) confirm Brieger and Ehrlich's findings, and show that most antibody in rabbit's milk and in goat's colostrum and milk is a transudate from the blood, without degradation and reformation.
31 See, for example, Brambell FWR. The transmission of passive immunity from mother to young. Amsterdam: North Holland; 1970 and Hemmings WA, Ed. Maternofoetal transmission of immunoglobulins. Cambridge: Cambridge University Press; 1976.
32 Brieger L, Ehrlich P. Beiträge zur Kenntnis der Milch immunisierte Tiere. Z Hygiene 1893; 13: 336; Collected papers, vol. II, p. 48-55.
33 Behring and Kitasato, note 26.
34 Ehrlich P, Kossel H, Wassermann A. Ueber Gewinnung und Verwendung des Diphtherieheilserums. Dtsch. Med Wochenschr 1894; 20: 353-5; Collected papers, vol. II, p. 56-60.
35 Ehrlich P. Die Wertbemessung des Diphtherieheilserums. Klin Jahrbuch 1897; 6: 299; English translation. Collected papers, vol. II, p. 107-25.
36 Madsen, note 25, p. 155.

37 Ehrlich literally pictured these molecules in cartoon form, to illustrate his ideas. See, for example, Cambrosio A, Jacobs D, Keating P. Image et controverse scientifique dans les premières théories immunologiques. In: Colloque Usages de l'image au XIX^{e} siècle. Paris: Créaphis; 1991.
38 Jules Bordet would contest both the theory and the names that Ehrlich employed (see Silverstein AM. History of immunology. New York: Academic Press; 1989. p. 193-7). Bordet believed in a physical adsorption theory of antigen–antibody-complement interactions; he called Ehrlich's Ambozeptor a substance sensibilisatrice, and Ehrlich's Komplement by the more neutral term: alexine.
39 Ehrlich, note 35.
40 Jerne NK. A natural selection theory of antibody formation. Proc Natl Acad Sci USA 1955; 41: 849.
41 Burnet FM. A modification of Jerne's theory of antibody formation using the concept of clonal selection. Austr J Sci 1957; 20: 67; see also Burnet's The clonal selection theory of antibody formation. London: Cambridge University Press; 1959.
42 For a review of the phenomenology of the declaration of "the end of science," see Silverstein AM. The end is near! The phenomenon of the declaration of the end of science. Hist Sci 1999; 37: 407-25.
43 At an international symposium, Molecular and Cellular Aspects of Antibody Formation, held in Prague, Czechoslovakia in 1964, attended by this author, Burnet implied that his clonal selection theory had carried the day and substantially solved the outstanding problems in immunology. Jerne's Summary: waiting for the end. (In: Antibodies: Cold Spring Harbor symposium on quantitative biology 1967; 32: 601) made this conclusion somewhat more explicitly. Finally, Jerne made it official in his tribute to Burnet, The complete solution of immunology. Austr Ann Med 1969; 4: 347.
44 See Ehrlich P. Die Seitenkettentheorie und ihre Gegner. Münch Med Wochenschr 1901; 48: 2123; Berlin Klin Wochenschr 1902; 28: 18.
45 Bordet J. Sur l'agglutination et la dissolution des globules rouges par le sérum d'animaux injectés de sang défibriné. Ann Inst Pasteur 1898; 12: 688-95; Belfanti S, Carbone T had earlier reported the phenomenon, but in less detail (G R Acad Torino 1898; 46: 321).
46 In a letter to a Dr. Clemens in 1899, Ehrlich complained that his preoccupation with immunity was keeping him from other interests (Rockefeller University Archives Center, Tarrytown, New York, Ehrlich Collection, Ref. No. 650 Eh89, Box 6, Copirbuch #4, p. 59-60).
47 Ehrlich P, Morgenroth J. Zur Theorie der Lysinwirkung (Ueber Hämolysine). Berlin Klin Wochenschr 1899; 36: 6, 481; 1900; 37: 453, 681; 1901; 38: 251, 569.
48 Gaffky G. Einleitender Ueberblick (Immunitätsforschung). In the Ehrlich Festschrift, p. 131.
49 See Parascandola and Jasensky, note 5.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

The phagocyte, the antibody, and agency in immunity: contending turn-of-the-century approaches

Eileen Crist, Alfred I. Tauber

It is conventional to regard the science of immunology as emerging through a battle between cellular and humoral approaches around the turn of this century [1]. The immunochemical perspective of the humoral school dominated immunology for nearly half a century, while the organismal orientation of cellular immunology became marginalized, finally making a come-back in immunology in the 1950s and 60s. In examining the clash between cellularists and humoralists, the question that arises is "Why was the humoral tradition apparently victorious over the cellular"? This question cannot be satisfactorily answered on the basis of uneven fruitfulness of their respective research, for both were theoretically and experimentally productive, promising to resolve a host of immunological problems. Indeed, promising research directions in cellular immunology were ignored until the middle of this century, in favor of studies championed by the early humoralists [1, 2]. And yet the ideas advanced by the ostensibly losing camp – the cellular approach – were quietly incorporated into the thought of the scientific discipline. As the resurgence of cellular immunology in the 1950s demonstrates, ideas themselves do not 'lose', for they persist and resurface in both foreseeable and unpredictable ways. So rather than frame the debate in terms of victory and defeat, we regard the divergence between cellularists and humoralists as more akin to a dialectic, leading to their mutual formation and conceptual entanglement.

At the heart of their divergence, the constitution of immunity was at stake: its medium, mechanisms, and central components. To show how they rivalled over which perspective could best represent immunity, we focus our analysis on a pivotal opposition that emerged between 'phagocyte' and 'antibody'. These entities, championed by cellularists and humoralists respectively, embodied different visions of biology. By examining the distinct portrayals of these entities in cellular and humoral immunology, we glimpse the conceptual architecture of competing designs of life's reality.

The novel view of inflammation

Edward Jenner (1749–1823) is acclaimed as the first immunologist to advocate the protective function of vaccination – against smallpox, the deadly disease of his era. Despite the momentous implications of inoculation, nearly 100 years followed Jenner's experimenta-

tion with little advance in the understanding, or therapeutic potential, of immune action. According to Paul Ehrlich, a central representative of the humoral school, "Jenner's discovery remained so isolated...due essentially to the fact that the theoretical conceptions of the cause and nature of infectious diseases made no advance during the subsequent decades" [3, p.178]. Only after Pasteur's and Koch's understanding of virulent, hitherto invisible, micro-organisms as causes of disease would the modern conception of immunity be articulated.

Once the microbes were exposed, "those invisible enemies... who hide in the air we breathe and the water we drink" [4, p.27], techniques could be designed to investigate how they maim the host organism and how the organism responds. Pasteur not only validated the efficacy of immunization with his instantly seized upon vaccines against cholera and rabies, he went well beyond Jenner in providing explanations of what causes cholera, rabies, and other diseases. Ehrlich's assessment of the absence of a theoretical context for Jenner's precocious discovery seems on the mark: the birth of immunology awaited the pivotal idea of the 'microbe' as the harbinger of disease. At the heels of Pasteur's and Koch's discoveries, Elie Metchnikoff first articulated the concept of immunology in its modern expression [5].

If immunity is defined in terms of applications of its therapeutic potential, then Edward Jenner is the first immunologist; if immunity is defined in terms of how its nature and actions are conceived, then Elie Metchnikoff is the first proper immunologist. As Arthur Silverstein has noted, the watershed between previous and modern understandings of immunity – between pre- and post-Pasteurian science – was the contrast between a passive and an active conception of the host's response to the threat of disease. Earlier views regarded the organism as a passive vehicle within which disease simply ran its course. According to the modern conception of immunity, the body mounts an active response. This dynamic view is especially stark in the pervasive language of warfare, which constitutes immunity as 'defense against', and 'resistance to', invading microbes.[1]

Metchnikoff's pioneering studies inaugurated this view of immunity – as the body's protective defense and resistance. In his magnum opus "Immunity in Infective Diseases," Metchnikoff defined "immunity against infective diseases... as the group of phenomena in virtue of which an organism is able to resist the attack of the micro-organisms that produce these diseases" [6, p. 10]. First articulated by Metchnikoff, this conception of immune phenomena as matters of 'attack' and 'resistance' subsequently became deeply rooted in the science of immunology. In the debate between Metchnikoff and the humoralists, the view of immunity as 'active defense' was not contested, but accepted by all as a shared tenet. Instead, the debate centered on the relative significance of cellular and humoral mechanisms involved in host defense, as well as on the relative conceptual virtues of biological and chemical formulations.

Metchnikoff formulated the modern idea of immunity, as defensive activity, in its evolutionary context in his "Lectures on the Comparative Pathology of Inflammation" [7], where he introduced the "biological theory of inflammation." Before Metchnikoff, the

1. Contemporary textbooks canonically define immunology as "the study of the body's 'defense' against infection," according to "The Penguin Dictionary of Biology," "immunity is the ability of an animal or plant to 'resist' infection by parasitic microorganisms" [8, 9].

phenomenon of inflammation had been regarded pathological, but he re-registered it as a defensive and restorative process, basing this view on a new understanding of the role of cells he called "phagocytes." In the inflammatory response, he observed that phagocytes – or "amoeboid cells" – accumulate at the site of injury, engulfing and destroying microbes. After this interpretation, inflammation ceased to be regarded as an expression of disease, but on the contrary, was understood as the body's curative response against external and endogenous threats. His "Lectures" may be characterized as "one long argument," in which he furnished abundant evidence for the proposition that "the essential and primary element in typical inflammation consists in a reaction of the phagocytes against a harmful agent" [7, p.187]. With a systematic examination of immune processes in a diversity of species, he showed that inflammation always involves phagocytic action. Phagocytes not only protect the organism by ingesting bacteria and other potentially harmful entities, but also have a central role in morphological metamorphoses during development and in scavenging damaged or malignant tissues.[2]

Metchnikoff's empirical demonstration of the universality of phagocytic action in inflammation was couched within an evolutionary argument. He relied on several aspects of the Darwinian evolution to comprehend infectious disease and immunity, the most encompassing dimension being genealogical common descent. Observations and discoveries about inflammation in one class of organisms acquired more general significance as they became corroborated in others. According to Metchnikoff, only after understanding the inflammatory process in all major phyla could the biologist apprehend general properties and arrive at a global explanation of inflammation. Through a comprehensive canvassing of life forms and stages of development in his "Lectures," he traced and described phagocytic action from unicellular protozoa to metazoa, from invertebrates to vertebrates, and from embryos to mature organisms.

On the basis of this comparative perspective, he revealed "the starting point, the *primum movens*" of inflammation [7, p. 8] The *primum movens*, both metaphorically and literally, was the motile, protective response of phagocytic cells, for "all other phenomena are merely accessory to this process, and may be regarded as means to facilitate the access of phagocytes to the injured part" [7, p.109-110]. The comparative approach was Metchnikoff's chief method toward a global understanding of infection and immunity. He frequently deplored that pathologists were not yet agreed on a comprehensive view of inflammation. His ambition was to fill that void by uncovering "a genealogical tree of inflammation" [7, p. 103] – a project both inspired by Darwinism and a contribution to it.

Evolutionary thought provided Metchnikoff's science with its core rationale. Against the background of the genealogical unity of organisms, his study of different phyla acquired scientific force. With a comparative methodology and the underlying support of the evolutionary standpoint, Metchnikoff exposed the conceptual shortcomings of his contemporaries' understanding of inflammation. Whether targeting mere descriptivism or failed explanatory attempts [7, p.180], his critiques always aimed to expose "the lack of a compre-

2. With the implicit themes of host identity and integrity, Metchnikoff adumbrated the problematic of "selfhood" in immunology that would be explicitly re-addressed by Sir Frank Macfarlane Burnet 50 years later [2, 10].

hensive understanding of inflammation."[3] He regarded his own analyses as achieving this objective: "As the evolution of inflammation shows, it is... both the most general and the most active means of defense among the members of the animal kingdom" [7, p.184].

Metchnikoff's elucidation of inflammation had the virtue of offering a general, single explanatory principle to encompass the diverse expressions of the phenomenon. His view of inflammation resonated with the evolutionary perspective, for he found it recurring in the major groups of animals. Metchnikoff's understanding of phagocytosis as defensive action evinced the doubly compelling features of singularity and comprehensiveness, and his argument was thus both eloquent and biologically cogent. He was not only substantively influenced by Darwin, but also in style of scientific argumentation – which, like Darwin's, was at once global and detailed in scope. His force as a biological thinker resided in his panoramic command of a wealth and diversity of facts, which he marshalled toward the articulation of one powerful principle of immunity.

Metchnikoff invoked the evolutionary perspective not only as the "systematic theory"[4] – a comparative method unifying the entire biological terrain – but in its more specific tenets as well. The "struggle between organisms" was a central motif in his thought, for he understood parasitic infection and host defense as manifestations of predatory aggression and inter-specific competition.[5] Beginning with the simple yet startling insight that "from active aggression to infection, there is but a small step" [7, p. 2], after extensive empirical expositions, Metchnikoff arrived at the conclusion that "the essential phenomena of inflammation represent an actual struggle between the phagocytes and the irritant agent" [7, p. 189]. The view of disease and health as consequences of "the struggle for survival between living beings" (within internal environments) was revolutionary in his day.

The identification of infection as a form of aggression, and parasitism as predation, were not metaphorical for Metchnikoff, but biological assessments. Upon completing his survey of the animal kingdom, he concluded that "in all these cases it is the phagocytes which war against the aggressor by devouring, englobing and digesting it" [7, p. 109]. Aggression and infection, predation and parasitism, are evolutionarily identical processes, their only divergence being that they transpire in different environmental contexts. Since he viewed infection as microbial aggression, Metchnikoff suggested that from a broad biological standpoint pathology is a branch of the natural history of the behavior of micro-organisms [7, p. 3]. Just as predatory attack against animals meets with some form of defensive response, so the predatory action of micro-organisms meets with the defensive reaction of phagocytes inside organisms' bodies.[6] Metchnikoff's training as a zoologist

3. Metchnikoff criticized the views on inflammation of his contemporaries, especially those of the influential pathologist Julius Cohnheim [1, 11]. Cohnheim believed that the inflammatory response was a result of the passive movement of "corpuscular elements" of the blood – that is, Metchnikoff's phagocytes – through vascular lesions. In contrast, Metchnikoff reinterpreted inflammation as a physiological remedial active response found throughout the animal kingdom–including organisms lacking a vascular system.

4. See Hacking 1992:45 [21].

5. Metchnikoff writes that "these phenomena [of infection and defense]... are more or less directly connected with the struggle for survival that is always going on between the representatives of the different orders of living beings" [7].

6. He captured the afflictions of protozoa – 'first animals', viz., unicellular eukaryotic organisms – in the same terms. For example, he described the amoeba's susceptibility to the micro-organism *Microsphaera* in terms of the amoeba becoming "less and less active, showing that it is not in a healthy condition" [7]. Regarding analogously infected *Infusoria* (or ciliates), he wrote that "the affected individual presents unmistakable signs of disease".

was indispensable in seeing parasitism as predatory action and infection as microbial aggression: he deployed a zoological idiom to characterize bacterial parasites of protozoa as "carnivorous, voracious organisms, and aggressors." His emphasis on a struggle between pathogens and phagocytes constituted an organismal understanding of immunity, rooted in a zoological perspective which he brought to bear on the behaviors of single-celled organisms. The transposition of an organismal viewpoint from single-celled organisms to phagocytic cells, paved by the tenet of common descent, was a cogent step. Indeed, Metchnikoff provided detailed arguments in support of the biological unity of phagocytes and protozoa. He conceptualized the phagocyte as a quasi-autonomous biological entity – almost an organism onto itself.

Metchnikoff's phagocyte

The term "phagocyte" was coined by Metchnikoff and Karl Claus from the Greek roots *phagos* (to eat) and *cyte* (cell). The conception of phagocytes as "eating cells" was pervasively stressed by Metchnikoff: he linked phagocytosis with a nutritive function, called phagocytes "digestive cells," and often described phagocytes as "devouring" bacteria and other entities. The term phagocyte was immediately compelling, and became established in the conceptual repertoire of immunology. It is an etymological expression of Metchnikoff's zoological approach to immunity, for it profiles what phagocytes do in phenomenological terms of animal activity. The phagocyte encompassed: an evolutionary connection between nutritive and protective functions; a picture of dynamic, spontaneous cellular action; and the life and death struggle between protective cells and pathogens. The characteristics of phagocytes that he focused on were those of motility, sensibility, engulfing and ingesting actions, and digesting and excreting functions – "capacities" and manifest "behaviors" of unicellular organisms. The affinity between phagocytes and independently subsisting organisms, especially amoebae, was created simultaneously on the levels of theoretical argumentation and observation. Metchnikoff's protagonist of immunity was a rich nexus of theoretical and phenomenological perspectives, crystallizing a new way of understanding and seeing immune action.

The phagocyte referred to amoeboid cells that literally eat potentially dangerous cells, including foreign micro-organisms and damaged cells of the host's own body. The theoretical dimension of the concept was connected to the protective and restorative functions Metchnikoff understood phagocytes to serve. At the same time, with the use of a microscope, phagocytic cells could be seen engulfing, for example, bacteria. The observational availability of phagocytosis was never itself in question, but with Metchnikoff its interpretation radically shifted as it became the core purveyor of immunity. With this new interpretation of phagocytic function, the perceptual configuration of phagocytes' actions was rearranged (cf. [13, 14]). No longer were phagocytic cells to be seen as passive carriers of bacteria, nor was phagocytosis to be understood as action deleterious to the organism. After Metchnikoff, immunologists and microbiologists looking down the microscope would witness the hygienic role of phagocytosis as a counter-aggressive response.

Metchnikoff argued that those organisms "which were possessed of mobile cells to englobe and destroy the enemy, survived, whereas others whose phagocytes did not exercise their function were necessarily destined to perish" [7, p.193]. Beyond applying the framework of natural selection, he also evinced the evolution of phagocytic function in the

universality of phagocytes and their actions; in the close connection between the nutritive and protective functions of cellular digestion; and in the manifestation of the evolutionary principle of competition. As a Darwinian, Metchnikoff regarded biological explanations as most compelling, the broader their scope of applicability. A theoretical explanation was trustworthy, for Metchnikoff, on Darwin's criterion of whether it is "applicable in allied cases; and especially, (when) the same general principles can be applied with satisfactory results, both to man and the lower animals" [15, p. 18].

In his "Lectures," Metchnikoff began his investigations with protozoa, proceeding to simple and complex metazoa. He found phagocytes not only throughout the gamut of animal life, but in the different developmental stages of embryogenesis and ontogenesis as well. From protozoa to metazoa, from vertebrates to invertebrates, from embryos through mature organisms, the form and role of phagocytes – or "amoeboid cells" – were closely allied.[7] The fungibility of the terms "phagocyte" and "amoeboid cell" in Metchnikoff's work underscored the evolutionary link – for more than a superficial resemblance, the interchangeability of these concepts signified common descent. The amoeba represented the organism-prototype of the phagocyte – the simplest and most primitive phagocytic cell. Like amoebae, phagocytes move in a gliding fashion along surfaces, extending protoplasmic processes called "pseudopods" to advance. Using their pseudopods, amoebae and phagocytes engulfed bacteria and other animate or inanimate entities, digesting them in their protoplasm. For amoebae this had nutritional function: it was how they capture their food.

In the case of phagocytes, the digestive function was preserved, but in the service of immunity – protecting the multicellular organism of which phagocytes were component cells. Metchnikoff maintained that in protozoa and the simple metazoa the phagocytic functions of nutrition and protection were fused. Intracellular digestion of microbes served to feed the organism as well as protect it against their potentially harmful parasitism. The function of protection via digestive assimilation was complimented by that of protection via excretion. When microbes were seen to penetrate protozoa, Metchnikoff observed that

> [I]f the intruders are killed and partially digested..., or are expelled as excrementitious matter, the protozoon remains uninjured and continues its normal and routine existence. Here we have an example of natural immunity, due to intracellular digestion [6, p.17].

The nutritive and expulsive functions of cells, starkly observable in the unicellular protozoa, enabled the evolution of immunity in multicellular organisms. In contemporary terminology, Metchnikoff found that the engulfing and digesting actions of amoebae, serving nutritive requirements, were "preadaptations" [16] for the evolution of immunity.

In Metchnikoff's analysis, the boundary between unicellular organisms and phagocytes lacks clear-cut definition. He frequently pointed out the organism-like behaviors of phagocytes: their independent motility, responsiveness to environment, and engulfing and digesting functions [7, p. 47, 118]. With his comparative approach, Metchnikoff discerned a

7. He notes the "general significance of the presence of these amoeboid cells which are able to englobe solid bodies. We have met with them in the various classes of Protozoa, and we find them again even in the most primitive forms of Metazoa" [7].

"general organismal" phenomenon characterizing the action of phagocytes in multicellular bodies, on the one hand, and the action of single-celled animals, on the other. What transpired in organisms through "component phagocytic cells," the protozoa accomplished with their "entire" bodies. Digestive and excretory functions of single-celled organisms were preserved and put to the service of multicellular animals of which phagocytes were a part. Indeed, the organismal nature of phagocytes might countermand its function as a docile component, when occasionally, at the peril of the whole organism, phagocytes were observed to move away from virulent bacteria rather than ingesting them (discussed below).

Emphasis on the phagocyte's "autonomy of movement" underscored its nature as almost an organism-in-itself. In response to a threatening or parasitic presence, phagocytes moved toward it as of their own accord, "accumulat[ing] around the injurious body and either surround[ing] it entirely or englob[ing] it" [7, p. 74]. Metchnikoff was bent on disproving the idea that phagocytic cells – especially, leukocytes (white cells) – required a vascular system to move them. (This view was widely held among pathologists in his day.) He remarked that phagocytes' "reaction is effected through the 'sensibility' of the phagocytic cells themselves, and is in no way influenced by the nervous or vascular system" [ibid., emphasis added; cf. 7, p.144]. To a passive image of phagocytes as "carried" to their needed location, Metchnikoff counterposed that they "actively migrate" to infected sites.

His training as a zoologist allowed him to see inflammation in a way that pathologists of his day were not trained to detect or apprehend. Metchnikoff advanced an image of phagocytic movement as animal-like movement of a predator toward prey. He attributed the unmediated, active locomotion to "the sensibility of the leucocytes (i.e., phagocytes) themselves, which is not always positive in the presence of invading microbes, but may also be negative" [7, p. 155]. He portrayed this sensibility as primitive perception – "cellular perception." Positive and negative chemotaxis (movement toward and away from pathogens), were the evidential index of phagocytic sensibility: differential responsiveness suggested that phagocytes discriminated between objects in their environment.

A compelling sign of autonomous sensibility was phagocytes' occasional "avoidance" of virulent bacteria which, if engulfed, might kill them. In one case, Metchnikoff noted that despite the danger that certain bacteria posed to the host, and despite favorable circumstances for the diapedesis across the vascular walls to the infected site, the leucocytes did not "emigrat[e] toward the invaded spot"; he explained this as "negative chemiotaxis manifested by the intravascular leucocytes" [7, p. 146]. Their organismal nature and autonomous perceptual capacities were expressed in behavior that contravened their protective role. Phagocytes thus occasionally exhibited atavistic behavior – as though they were in solitary struggle against the virulent microbes, avoiding them even at the peril of the host.

In sum, Metchnikoff's organismal understanding of phagocytes emerged on three levels: that of systematic theory, with an evolutionary understanding of the affinity between phagocytic cells and unicellular organisms; that of equivalence (with unicellular organisms) of the biological functions of locomotion, ingestion, digestion and expulsion; and that of phenomenology with the behavioral and morphological similarity between autonomous single-celled organisms and phagocytes. As we discuss below, humoralists

took exception to Metchnikoff's representation of phagocytes, for they interpreted it as signifying teleological and vitalistic thinking.

The humoral factor

The history of the discovery of antibody was quite different from the elucidation of phagocytic cells. The existence of antibodies was postulated as a consequence of experimental findings about immunity which could not be explained by the theory of phagocytosis. George Nuttall's observations in 1888, and Emil von Behring and Shibasaburo Kitasato's 1890 crucial experiments, suggested the existence of a factor in the "cell-free" humors of animals that was protective against both microbes and their toxins. These findings convinced the scientific community that immunity was not due solely to phagocytosis. The humoral factor – soon named "antibody" – seemed to be effective independently of cellular presence. The immunity-conferring powers of the antibody were thus conceived partially through juxtaposition to the phagocyte. The divergence between the cellular and humoral schools emerged in the different nature of the entities they advocated as core constituents of immunity. Dissensus was reinforced by their divergent methods of studying these entities – with differential emphasis on in vivo versus in vitro observations and experiments, and their distinct tools of comparative, observational biology versus chemical test-tube experimentation.

Antibodies could not be seen. Thus in contrast to phagocytic cells, the detection of antibodies did not involve seeing a known entity under the novel auspices of immunological thinking. Their existence was inferred through the immunizing or bactericidal powers of sera, in the absence of cellular intervention. A corollary of the invisibility of antibody was uncertainty about its nature, especially discernible in late nineteenth century humoralist writings. The term antibody itself did not become immediately established. Competing and overlapping terminology proliferated in the late nineteenth and early twentieth century literature, all attempting to capture elusive antibody-related phenomena: "alexine," "stimulins," "cytase," "complement," "antitoxin," and "cytotoxin" [19, 20]. The profusion of terms was a compounded effect of the inherent complexity of humoral phenomena and the antagonism of immunologists affiliated with different research communities and agendas. The term antibody – as well as "antitoxin" that briefly preceded it, and was often used interchangeably with it – belonged to the humoral tradition, encompassing in Paul Ehrlich's words "an altogether new factor" of immunity.

The experimental work of Behring and Kitasato in the early 1890s was a turning point. It provided clear evidence for the existence and significance of a non-cellular, humoral factor of immunity.[8] In their 1890 paper, "On the mechanism of immunity to diphtheria and tetanus in animals," the authors reported experiments demonstrating that the blood of animals immunized against tetanus acquired a property which was subsequently protective against tetanus bacilli or tetanus toxin. After immunizing a rabbit against tetanus – by inoculation with a non-virulent strain or with dead microbes – they found that the rabbit could withstand a dose of virulent tetanus bacilli twenty times what would kill a non-

8. This has already been suggested in 1888 by Nuttall, who explicitly challenged the adequacy of Metchnikoff's theory of phagocytosis: "That phagocytic activity is the most important protective measure must be weighed against the fact that anthrax bacilli under the skin of frogs are destroyed in large amount outside phagocytes as well. It is clear that Metchnikoff's experiments suffer considerably by this finding" [17].

immune rabbit. Furthermore, both the blood and serum of the immune rabbit, when injected into the abdominal cavities of other animals, such as mice, could protect them from tetanus bacilli or toxin. Analogous results were announced for diphtheria [18, 1890, p. 139].

The authors summarized their findings with four points, which became dominant themes in immunological research for the next 50 years: 1) the blood of rabbits immune to tetanus has the ability to neutralize or destroy the tetanus toxin; 2) this property exists also in cell-free serum of those rabbits; 3) this property remains effective even in the body of other animals, so that it is possible, through blood or serum transfusions, to achieve an outstanding therapeutic effect; and 4) the property that destroys tetanus toxin does not exist in the blood of animals that are not immune to tetanus, and when one incorporates tetanus toxin into non-immune animals, the toxin can still be demonstrated in the blood and other body fluids of the animal, even after its death [ibid.]. With points 1 and 4, they enunciated the discovery of the protective property of the blood of animals that have either already been immunized or are naturally immune;[9] with point 2, they implicitly underscored the inadequacy of purely cellular conceptions of immunity; and with point 3, they called attention to the therapeutic implications of their findings.

Behring and Kitasato's work was revolutionary on four levels: theoretical, conceptual, technical, and therapeutic. Their report arrived as the highly charged rivalry between cellular and humoral perspectives was emerging. Prior to Behring and Kitasato's decisive experiments, the existence of a humoral factor in immunity was entertained, yet unconfirmed. Their results established the presence of a protective property in the fluids of immune animals as incontrovertible fact. Their experiments may therefore be regarded as "crucial," in the sense of implicitly having "two theories in question" [22] – namely, cellular and humoral – and providing solid evidence for one. By submitting proof of the existence of a protective factor in the blood and serum of immune animals, Behring and Kitasato gave weighty support to the humoral framework, thereby heralding a theoretical revolution in immunology.

Their 1890 work was also conceptually revolutionary in paving the way toward a chemical register of immunological phenomena. The interactions between toxin and antitoxin could be conceived as chemical, for both substances were extracted and isolated from their respective cellular contexts.[10] Tetanus toxin, for example, was acquired by filtering a tetanus culture so as "to render it free of bacteria." And the antitoxin effects of the cell-free serum of the host body were shown to be protective, if the serum was injected before exposure to infection, and therapeutic, if the serum was injected after exposure. By experimentally working with substances "entirely separable from cells," Behring and Kitasato inaugurated the shift from a biological to a chemical approach, with immune phenomena represented in the framework of chemistry. As will be discussed shortly, this shift was clinched by immunochemist Paul Ehrlich: Behring and Kitasato's findings were

9. The existence of natural immunity applies to diphtheria; for example, rats are naturally immune against both the bacilli and the toxin of diphtheria.

10. The humoralists did not establish where these protective humoral products were produced. Metchnikoff, seeking a unification of the cellular and humoral viewpoints, addressed this etiological question, arguing that such humoral factors were phagocytes products which he called "ferments."

in fact the link that brought Ehrlich, and his attendant commitment to chemical explanation and nomenclature, into the humoral fold.

Behring and Kitasato's 1890 work was also revolutionary on the level of technique, decisively shifting immunology from the cellular approach of a natural-observational, in vivo science, toward experimental, in vitro methods. It was pivotal in transforming immunology into what Ian Hacking has called "laboratory science," in the sense of "study[ing] phenomena that seldom or never occur in a pure state before people have brought them under surveillance"[21, p. 33]. Hacking maintains that "exaggerating a little, I say that the phenomena under study are 'created' in the laboratory" [ibid.]. Behring and Kitasato's techniques were indeed crucial in creating the phenomena under study. The "toxin" of bacilli must be experimentally extracted and divorced from the microbes; the fibrin-free fluid "serum" is extracted after blood is allowed to coagulate; and last but not least, the serum of inoculated or naturally immune animals is added (usually) to the abdominal cavity of other animals, including other species [18, p.139-140]. For the first time so thoroughly in immunology, Behring and Kitasato exhibited the total interventionist strategy of laboratory science.

The final revolutionary dimension of Behring and Kitasato's experiments was their therapeutic implications. The researchers themselves placed great emphasis on practical applications, introducing their paper with the highly promising claim of being "able to cure infected animals, as well as to pretreat healthy animals so that later they will not succumb to diphtheria or tetanus" [18, p. 138]. In the same issue of the journal that published the famous Behring and Kitasato report, another paper on diphtheria, by Behring alone, ended with the pronouncement that "the possibility for the cure of very acute diseases can... no longer be denied" [p. 144]. This promise was fulfilled for the case of diphtheria, a lethal disease known colloquially as "the strangling angel of children." Serum therapy against diphtheria was brought to hospitals in major cities of Europe by 1894 and onward with dramatic results: mortality was cut in half.[11]

The therapeutic applications of early immunochemistry were touted by humoralists. Ehrlich wrote that Behring's "remarkable discovery seemed at one stroke to open up an entirely new and extremely promising prospect of immunizing mankind against the majority of infectious diseases" [3, p. 179]. The therapeutic promise of "immunizing mankind" may have been the decisive factor in the apparent victory of the humoral over the cellular approach [1, p. 215]. Indeed, Behring and Kitasato's experimental work instigated alliances between science institutes, industry, and government for the production, testing, and dissemination of vaccinations [4, 16].

Behring and Kitasato's solid experimental intimation of a protective humoral factor had a powerful effect on immunology. As Silverstein put it, for the next 50 years or so, "few questions about cells in immunity were asked within a discipline comfortable with the dogma that circulating antibody would provide all essential answers to the problems of

11. Grundbacher notes that "the rapid reduction in child mortality from diphtheria at the turn of the century was one of the largest decreases in mortality ever achieved by any therapeutic intervention. In Germany alone, an estimated 45,000 lives were saved per year with the antitoxin therapy for diphtheria" [24]. Due to eventual immunization procedures, the occurrence of the disease also declined. Behring, who worked on diphtheria in particular (while Kitasato's work was on tetanus), was awarded the first Nobel Prize for Medicine in 1901 for "his work on serum therapy, especially its application against diphtheria" (quoted in Grundbacher [24]).

immunity" [18, p.55 ; 17]. How the discovery of antibody unfolded over approximately a century of scientific investigation is a topic in itself. For the moment it might be noted that an intriguing aspect of the story of the antibody is that when first conceived it was akin to an abstract idea, inciting further research toward the discovery of its concrete embodiment. In the dawn of its discovery, the antibody was as intangible and nebulous as "phlogiston": but obversely to "the increasing vagueness and decreasing utility of the phlogiston theory" that contributed, according to Thomas Kuhn, to its eventual demise [14, p.71], from Behring and Kitasato's 1890 experiments, to Ehrlich's contributions, through to its present-day molecular-biological elucidation, the antibody became "decreasingly vague" and "increasingly useful."

Ehrlich's antibody

Paul Ehrlich had a great impact on the science of the antibody, and on immunology in general. He coined the term antibody in 1891, and proceeded to give it conceptual, theoretical, and imagistic definition. He argued explicitly for a "chemical understanding" of the antibody and its action; he advanced his famous "side-chain theory" to account for its production by cells and its proliferation in the humors; and he modelled, on a key-and-lock analogy, the chemical affinities between antibodies and toxins with controversial, yet rhetorically effective, representational "drawings" [19]. Ehrlich's approach to biological phenomena, and to immunity in particular, was in every respect from a chemical standpoint. If his early research in histological staining, and later contributions to chemotherapy [1, 22], suggested that it was the specific nature of his subject-matter that dictated his chemical mindset, his work in immunology clearly showed that Ehrlich did not regard chemistry as the handmaiden of biology: he viewed biological function as thoroughly chemical in nature. The most encompassing framework in his studies of immunity was chemistry – in his methods, theories, and terminology.

In his important 1897 paper "The assay of the activity of diphtheria-curative serum and its theoretical basis," Ehrlich introduced the antibody by means of a chemical explication:

> [A] toxin and antitoxin influence one another by a direct chemical interaction... I have been able to demonstrate by test-tube experiments... that the interaction of toxin and antibody is much more rapid in concentrated than in dilute solutions, and also that heat accelerates the action and cold retards it. Similar phenomena are frequently observed in pure chemistry, notably in the formation of double salts. Be this as it may, all the observations suggest that the reaction between toxin and antitoxin takes place in accordance with the proportions of simple equivalence... A molecule of toxin combines with a definite and unalterable quantity of antibody [25, p.114, emphasis in original]. The terms antitoxin and antibody are used interchangeably.

The assessment of the toxin-antibody interaction as straightforwardly chemical was typical of Ehrlich's thought. As he put it in another paper, "the action of antitoxins is accentuated or diminished under the influence of the same factors which bring about similar modifications in chemical processes – warmth accelerates, cold retards the reaction, and 'this' proceeds more rapidly in concentrated than in dilute solutions" [3, p.180, emphasis in original]). The parameters of "concentration" and "temperature" were chemical parameters. That these were the identical sort observed "in the formation of

double salts" was offered in the 1897 paper as speculation about the specific chemical compounds involved. However, in following this diagnosis with "be that as it may," Ehrlich implied that the "specific" chemical nature of the exchange was, for the moment, irrelevant: what he considered important was that the quantitative relation of "proportions of simple equivalence" revealed the toxin-antibody interaction as molecular, and thus, a matter of "pure chemistry."

Ehrlich's emphatic claim that "a molecule of toxin combines with a definite and unalterable quantity of antibody" invoked features of physico-chemical systems often juxtaposed to biological processes. The interaction exhibited constancy; it could be expressed in quantifiable ratios; it was a phenomenon that could be repeatedly witnessed, and therefore was predictable and consistently recordable. These features ensured that the toxin-antibody interaction had the law-like inexorability of chemical phenomena. Ehrlich did not conceptualize his chemical perspective as "reductionist" – as decomposing higher levels of organization into more basic elements. Rather he regarded the mapping of chemical and biological phenomena as one of "identity." On the basis of this identity, he advocated the superiority of in vitro methods.

> It was necessary for me to try to eliminate as far as possible, the varying factor of the animal body, and bring the investigations more nearly into line with the conditions necessary for experiments of a chemical nature. In the course of these endeavours it was shown that it was possible to obtain in a comparatively simple manner an insight into the theoretical considerations necessary to a proper understanding of immunity, by means of test-tube experiments with suspended animal tissues [3, p. 179-180].

"Eliminating the varying factor of the body," far from being a potentially adulterating procedure, was understood as the very prerequisite for objectivity. For Ehrlich, in vitro experiments simplified and made visible in vivo phenomena, and so results garnered from test-tube trials could safely be taken to portray what transpired in the body. The contrast with Metchnikoff's understanding of the core phenomena of immunity is already visible. Ehrlich's quantitative methods, stable and predictable ratios, law-like properties, and in vitro experiments, contrasted with Metchnikoff's natural-historical studies, comparative methodology, variable observations, and unpredictable consequences of the struggle between organisms.

The experimental setting and procedures were steadfastly regarded by Ehrlich as the neutral medium that revealed in a way "reproduced at will," what is obscured by the complexity and opacity of the body. "Each test-tube," Ehrlich averred, "represents as it were a research animal" [ibid.]. What was observed in experimental procedure directly reflected what transpired in the body, and thus the results of test-tube experiments were, without demurral, reinstated back into the body. Ehrlich himself described his strategy in what might be characterized as "synecdochal" (rather than reductionist) terms: what was observed on a small, modified, non-living context was taken as a faithful representation of what transpired in the unaltered, living body. The synecdochal form of reasoning – the qualitative equation of the stable part with the varying whole – echoed the methodology of the constant traffic of materials taken out of the body, experimented with, and put back into the same or a different body, out of which materials might again be extracted [3, p. 185].

Test-tube experiments allowed for maximum control of materials and quantities used. Blood corpuscles, fluids, or tissues were extracted from animal bodies, and altered with the use of special techniques, as for example in the production of serum or plasma. Bodily materials and their derivatives were thereby rendered passive and manageable, amenable to exacting quantitative expression. Under the auspices of chemical methodology, they were now legitimately describable in the language of chemistry. Ehrlich regarded his methods as perfectly tailored to the nature of the toxin-antibody interactions. The possibility that his theoretical deductions were "method-laden" – that his methodology aggressively contributed to the constitution of the phenomena as chemical – did not enter Ehrlich's considerations. The phenomena of immunity were comprehended as purely chemical, and in Ehrlich's studies indeed they were: laboratory techniques, measurement methods and nomenclature guaranteed that they were transformed into chemical phenomena, even if they did not quite have this status to begin with [12].

With his "side-chain theory" Ehrlich formulated a hypothesis about how the antibody interacts with toxin. He pictured the antibodies (or antitoxins) as groups of atoms found in the protoplasm of particular cells. "Adopting the nomenclature of organic chemistry," he wrote, "these groups may be designated 'side-chains'" [3, p. 185]. The stereochemical locking of toxin and side-chains resulted in an association considered debilitating to the cell. Ehrlich speculated that by the toxin's occupying the side-chains their normal, possibly nutritive, physiological functions were blocked. As a consequence, the cell produced more side-chains, which might also become stereochemically linked with toxin, if toxin continued to circulate. According to Ehrlich :

> [The cell then overcompensated for the continued binding of its side-chains by toxin, and] [w]ith great increases in the doses of toxin there must finally come a point at which such an excess of side-chains will be produced that, to use a trivial expression, they will become too much for the cell itself and will be discharged into the blood, like an excretion, as unwanted ballast. *According to this view, the antibodies represent side-chains of the cell protoplasm which have been produced in excess and therefore thrust off* [25, p.115, emphasis in original).

The side-chains, in their attached and free modes, were the etiology of toxic and protective effects, respectively. Their affinity with toxin allowed the latter to bind to the cell, thereby subverting its functions. The cell then produced an excess of side-chains, released into the humors. As a consequence, cells were "protected" by circulating side-chains. These side-chains were the antibodies.

The side-chains of cells that bound toxin allowed for its toxic effect on the cell and ultimately the organism. For example, on the basis of experiments showing the binding of tetanus toxin by brain substances, Ehrlich asserted that "the presence of such groups is the necessary preliminary and cause of the poisonous action of the tetanus toxin in the living animal" [ibid. p. 185]. With this view of disease, he proceeded to define natural immunity as "the absence of any chemical affinity" between cell side-chains and toxins. "If the cells... lack side-chains fitted to unite with them, the toxophore group (of the toxin) cannot become fixed to the cell, which therefore suffers no injury, i.e., the organism is naturally immune" [ibid., p. 186]. Natural immunity was thus a "passive" condition for Ehrlich, the consequence of the absence of any chemical interaction between cellular and toxic compo-

nents. In contrast, according to Metchnikoff, the inner environment of the body was subject to the "active surveillance" of phagocytic cells. Whereas Ehrlich viewed natural immunity as the result of chemical inactivity, Metchnikoff saw natural immunity as the consequence of organismal action.

With Ehrlich's theory, the issue of "the production of antibody" was tentatively resolved. It was a matter of pure chemistry, of inexorable interactions.

> It must be assumed that this ability to combine with antitoxin is attributable to the presence in the toxin complex of a specific group of atoms with a maximum specific affinity to another group of atoms in the antitoxin complex, the first fitting the second easily, as a key does a lock, to quote Emil Fisher's well-known simile [25].

Ehrlich's immunochemistry may indeed be regarded as the direct ancestor of the contemporary molecular biological imagery of immune phenomena, for there are deep strands of continuity in thinking and terminology. Ehrlich's "lock and key" side-chain theory and contemporary receptors have a clear conceptual family resemblance. In their 1994 textbook "Immunobiology," Janeway and Travers provide a bibliographic note about Ehrlich that acknowledges this continuity: "Ehrlich (1854–1915) was an early champion of humoral theories of immunity, and proposed a famous side-chain theory of antibody formation that bears a striking resemblance to current thinking about surface receptors" [8]. The authors imply that current discoveries confirm Ehrlich's prescience. What this retrospective reading elides is that Ehrlich's contributions were crucial in formulating the agenda for the experimental search for the antibody. His side-chain theory conceptualized the potential molecular configuration of the antibody and its circulation in the humors. He formulated a research program and a way of thinking about, and manipulating, the antibody. Thus it is not as if the antibody, as currently understood, "surprisingly" turned out to look a lot like Ehrlich's conception. Ehrlich was a key, and historically not-so-distant, link in the chain of the material, theoretical, and literary technologies through which the antibody was discovered.

The 1908 Nobel Award

In 1908, Elie Metchnikoff and Paul Ehrlich shared the Nobel Award for their respective contributions to immunology. The joint award was not only in acknowledgement of their respective contributions, but an institutional attempt to encourage the cellular and humoral scientific communities to concede that their approaches and findings were not necessarily antagonistic. Rather than using the occasion to make overtures to one another's approaches, in their addresses the chief proponents of the phagocyte and the antibody reiterated central aspects of their divergent theoretical and methodological frameworks.

In his Nobel Lecture, Metchnikoff proclaimed that "the white corpuscles" – the phagocytes in the bloodstream of vertebrates –"are microscopic organisms" [26, p.289]. Further on he reiterated: "Now the white corpuscles are living organisms, hypersensitive to external conditions and which admit of very great variation" [26, p.292]. He was intent on making the point that the performance of phagocytes is highly variable, due to the fact that they were, themselves, organisms which could be "strong" or rendered "powerless" [ibid.]. As a biologist he celebrated such individuality. Ehrlich's concern, on the other hand, was to "contain" variability. While he understood that variability was intrinsic to biological phenomena, as a chemist he considered that his task was not so much to account

for this variability as to find methods to eliminate it. The inevitable consequence of this discrepant appreciation of variability is that Metchnikoff placed epistemic emphasis on in vivo observation, while Ehrlich favored knowledge garnered from the controllable environment of experimental set-ups.

Metchnikoff aspired to a comprehensive understanding of health, disease, and immunity. Indirectly criticizing the application-oriented and narrowly-based research of the humoral tradition, he observed that "disease is not the prerogative of man and the domesticated animals." He based this view on the evolutionary unity of all species, maintaining that "lower animals, with very simple organizations, showed pathological phenomena, and if so, infection, cure and immunity could be observed among them" [26, p. 282]. The study of simple organisms was justified by Darwinism, and allowed for the observation of immunity-related phenomena under the microscope. Metchnikoff valued in vivo observation, noting that "certain of the lower animals, 'transparent enough to be observed alive', clearly show in their midst a host of small cells with moving extensions (i.e., phagocytes). In these animals the smallest lesion brings an accumulation of these elements at the point of damage" [26, p. 282, emphasis added]. To test the hypothesis that these cells played a defensive role, he sought "some higher animal... small and transparent enough to be observed living under the microscope and yet subject to microbial disease" [26, p. 283]. Metchnikoff found such a subject in the species *Daphnia*, commonly known as water-fleas. Discussing the "battle" between *Daphnia* phagocytes and certain infective spores, he commended the method of natural observation openly, averring that "this description is from a living animal and can be observed at each stage under the microscope with such precision as could hardly be bettered" [26, p. 284].[12]

Ehrlich's approach could not differ more. In his Nobel Lecture, "Partial cell functions," he advocated "break[ing] down the concept of the cell 'as a unit' into that of a 'great number' of individual specific 'partial functions'." With the move from "unit" to "partial function" – from the visible to the invisible realm – he drew the border between biology and chemistry: "But since what happens in the cell is chiefly of a 'chemical' nature and since the configuration of chemical structures lies beyond the limits of the eye's perception we shall have to find other methods of investigation for this" [26, p. 304, emphasis in original]. Ehrlich implicitly disparaged Metchnikoff's predilection for direct observation. He proclaimed that the limits of knowledge from microscopic observation had been reached, and that for "further penetration into... cell life even the most refined optical aids will be of no use to us" [ibid.]. Metchnikoff of course was aware of limits of the microscope. But he expressed an epistemic and aesthetic predilection for natural observation, writing, for example, that "controlled observations on living organism (sic) can not be wrong" [26, p. 284]. And while admitting the necessity of in vitro experiments, he expressed a deep suspicion toward their results:

> Given the impossibility of submitting a vertebrate, even the smallest such as a newborn mouse, to direct examination by microscope, a more complicated way had to be taken, by combining the results of research on the blood and organs extracted from

12. On the award of the Nobel prize to Elie Metchnikoff, see also Tauber 1994 [10].

> the organism, and thinking out the interconnection. *In such circumstances the door is wide open to mistakes of all sorts* ([26, p. 284, emphasis added]).

While Ehrlich was suspicious of the living body as an object of study, regarding its complexities as a hindrance to productive research, Metchnikoff was skeptical about experimental work using bodily extracts, and always reticent to generalize from in vitro results to in vivo processes [6, p. 284, 233]. Interestingly, Ehrlich and Metchnikoff used the Award occasion to argue over 'methods'. Their divergence on theoretical frameworks – evolutionary biology versus chemistry – was not broached.

Metchnikoff's and Ehrlich's Nobel addresses underscored the deep incompatibilities between cellularists and humoralists. Their central entities—phagocyte and antibody—competed for primary mechanism of immunity and were, moreover, altogether different 'types' associated with disparate approaches. The cellular and immunochemical perspectives differed in their systematic theories, topical theories, methodologies, and, relatedly, in nomenclatures and uses of language. They also differed in more fundamental, philosophical ways which we explore in the next sections.

The phagocyte as agent: an organismal vision of biological realism

Even though largely implicit, the theme of "agency" was central in the dispute between cellularists and humoralists. The question of agency in immunity was pervasive during this period, surfacing in the language of writings, the tenor of observations and theories, and the metascientific commitments of different research schools. The divergence between cellularists and humoralists was inseparable from the question of what sorts of entities have agency, a question openly posed in discussions about teleology and vitalism. Metchnikoff was seen as imbuing phagocytes with immanent purpose, a conception intolerable to immunochemists. According to the latter, not only was Metchnikoff's phagocytosis insignificant as immune mechanism, but his portrayal of phagocytes was vitalist and teleological – and so even more deeply problematic.

Key differences between cellular and humoral conceptions are apparent in juxtaposing the characteristics of Metchnikoff's phagocyte and Ehrlich's antibody. Phagocytes were visible, cellular and organism-like in their amoeboid motility. Metchnikoff attributed their spontaneous movement to a sensibility mediating their positive and negative chemiotaxis. He ascribed inchoate perceptual capacities to the phagocyte, on the basis of its discriminating between various entities and elements in its environment; it exhibited purposiveness in its actions, whether protective of the host or atavistically protective of itself; it occasionally exhibited unpredictable but meaningful – from the viewpoint of the "struggle between organisms" – behaviors in attacking other phagocytes or even healthy cells of the host. The phagocyte thus emerged as self-contained, almost in charge of its own 'destiny', autonomous, alive, and quasi-sentient. The existential contours of the antibody were very different: invisible and soluble; not a living entity, but a chemical component; prophylactic in its function, but non-purposive; exhibiting no chemiotaxis, but only stereochemical affinities for certain substances; having no intrinsic motility, but ejected by cells and drifting in fluid mediums; fluid-contained, not self-propelled. The antibody was understood as the body's "inadvertent" prophylactic tool, rather than "purposive" protective gendarme. In juxtaposition to the phagocyte, the conceptual elements through which the antibody was composed did not delegate agency to it.

Metchnikoff, however, constituted the phagocyte as agent.[13] This was a consequence of regarding phagocytes as autonomous organismal entities that possess what may be characterized as a "protean nature." Their understanding as organisms emerged through observation and phylogenetic comparison to protozoa; their portrayal as autonomous was derivative from the linked features of sensibility and locomotion; and their protean nature emerged in the fusion of unpredictability and meaningfulness occasionally characterizing their behavior. The play of these features assembled an entity irreducibly at the center of its own actions – an agent. Its agency was both expressed through, and reinforced by, Metchnikoff's language of action and interaction. This language was integrated into the science, forming the beginnings of the familiar idioms in immunology of intentionality, on the one hand, and warfare, on the other.

Far from an imposed construct, phagocytic agency was assembled from the ground up, through various mutually elaborating means. All aspects summarized above worked synergestically together, allowing phagocytic agency to supervene as a compelling image.

For Metchnikoff the organismal nature of phagocytes was visible in the near identity of their activities with those of micro-organisms: sensibility, locomotion, engulfment, ingestion, digestion, and excretion. These capacities served to protect the organism against pathogens, and testified to an evolutionary history of life-and-death struggle between organisms. Phagocytic cells (like unicellular organisms) continued to "eat," but their eating now had a new immunity-conferring function. Darwinian themes formed the background of the organismal view of immunity. For Metchnikoff immunity was: i) a protective function that co-evolved in close functional proximity with nutritive organismal processes; ii) a general phenomenon across the entire spectrum of life; and iii) a set of phenomena manifesting the witnessable form, and unpredictability of expression, of the struggle for survival. The capacities and activities ascribed to phagocytes were, for the most part, observationally available – though not in a theoretically 'innocent' sense, but from a zoological and Darwinian vantage.

The agency of phagocytes came into sharper focus with the two features of "sensibility" and "motility," underlying the autonomy of phagocytes – their representation as independent centers of activity. Locomotion and sensibility were interconnected in Metchnikoff's portrayals, for movement toward or away from pathogens originated in phagocytes' discriminatory abilities. Since locomotion was viewed as stemming from a discriminating response to the environment, it was linguistically captured as "action." In a typical example of rendering phagocytic function through action terminology, Metchnikoff wrote: "in too many cases the phagocytes 'flee' before the enemy or 'destroy' the cells of the body to which they belong" [7, p. 194, emphasis added]. Motion delivered as action – fleeing, destroying, seeking, or attacking – expressed a peculiar way of seeing phagocytes move, strikingly different from passive movement in fluid mediums. Moreover, motion delivered as action was a very different conception from the mechanistic inexorability of chemical interactions.

In addition to organismal origins and perceptually-guided motility, the "protean nature" of the phagocyte played into its understanding as agent. Phagocytes' spontaneity of action,

13. Following Donald Davidson [28], we regard an "agent" as an actor that is the cause of its own actions. An agent brings about changes in the world as a matter of endogenously generated and directed behaviors.

coupled with their struggle with pathogens, preempted the full-proof predictability of immune response, either in terms of perfect adaptation or pure chemistry. Metchnikoff described such unpredictable phagocytic behavior as meaningful within the framework of phagocytosis theory. Phagocytic behaviors that were unpredictable, yet once observed understandable, we refer to as "protean." Protean behaviors – both variable and coherent, unexpected yet consistent–contributed significantly to the image of the phagocyte as agent. Two types of observation reported by Metchnikoff exemplified the protean face of phagocytes: negative chemiotaxis and phagocytic attacks on non-pathogenic cells of the organism.

Metchnikoff articulated a novel understanding of immunity with his view of phagocytic response as active and protective against pathogens. On the heels of this new understanding, he came across an odd observation: sometimes phagocytes abrogate their prophylactic duties, moving away from virulent microbes rather than destroying them. Since phagocytes have been selected for their prophylactic functions, biologists would not predict that phagocytes would "move away" from pathogens; yet given their perceptual and locomotor capacities, and their quasi-autonomous engagement in inter-cellular struggle, the avoidance became understandable as phagocytes acting in a "self-serving" fashion. In negative chemotaxis, phagocytes did not exhibit their functional role, but reverted to a primordial, organismal modality. Metchnikoff thus did not interpret negative chemotaxis as aberrant, nor explained it away in an ad hoc fashion, but integrated it into the very fabric of his theory. He saw it as a natural corollary of phagocytes' organismal origins and their autonomous involvement in conflict; and it demonstrated the existence and importance of phagocyte sensibility.[14] His reasoning in integrating phagocytes' occasional avoidance of microbes into his theory – when the latter rested on the significance of their destroying microbes – was circular, though not viciously so.

Indeed, Metchnikoff harmonized negative chemotaxis with his very definition of immunity. He did not regard immunity as definitionally bound to specific function, but identified it in the global and abstract terms of an active, protective response. Hence, while the prototypic expression of immune action in multicellular organisms was phagocytic "attack" against pathogens and their products, this expression need not be incorporated into the definition of immunity: active "retreat" from pathogens was also protective – of the phagocytic cells themselves– and hence an immune act. For Metchnikoff the core of immunity was about resistance, manifest through attack and sometimes through retreat. In unicellular organisms, immunity was effected through the avoidance of threatening circumstances, viz., immunity was precisely achieved through negative chemotaxis [6, p. 31,32]. Analogously, phagocytes' retreat from virulent microbes – for example, their avoidance of cholera vibrios in non-immunized animals – was a cellular act of immunity, even while it resulted in the animal's infection and possible death.

So one expression of the protean nature of phagocytes was their occasional avoidance of pathogens and their products. Another was the near obverse: phagocytic attack on

14. On the basis of this phenomenon, Metchnikoff's critics attempted to undermine the immune function of phagocytosis. If phagocytes are the effectors of immunity, should they not destroy pathogens rather than avoid them? Metchnikoff was undeterred by this critique, and concluded from "negative chemotaxis" that "we must look upon the 'sensibility' of leucocytes as the most important factor in inflammatory diseases" ([7], emphasis added).

normal cells of the organism's body. Metchnikoff described a striking example of phagocytes turning against other phagocytes. In one experiment, the red blood corpuscles of a goose were injected into the body of a snail for the purpose of observing the snail's immune response. The phagocytic cells of the snail responded by engulfing and eventually digesting the blood cells. Metchnikoff observed that "the next day red blood corpuscles are still to be found intact in the blood plasma, but the great majority have been devoured by the leucocytes" [6, p.70]; on his theory, this protective response was not surprising. However what also occurred, unpredictably, was certain phagocytes engulfing "other phagocytes" that have already ingested goose blood cells [ibid.]. While unexpected, this phagocytic action was entirely meaningful on Metchnikoff's understanding of immunity as the active response of cells locked in struggle with harmful micro-organisms.

The same principles that accounted for the role of phagocytes as the body's gendarme, also accounted for their protean behaviors – avoidance of pathogens or attack on normal cells. With the essential characteristics of phagocytes as near-autonomous organisms that redeployed their primitive nutritional functions for protection, Metchnikoff could comprehend the entire gamut of phagocytic action – from their normal prophylactic response to their self-serving behaviors or misplaced attacks. From the cellular perspective, therefore, what would soon be called "autoimmune" phenomena – the immune system turning against elements of its own organism – would be expected to arise. Here Metchnikoff diverged from Ehrlich, who introduced the expression "horror autoxicus" for antibodies turning against normal cells, and emphasized the improbability of this occurring [11]. Yet on Metchnikoff's conceptualization of the phagocyte as agent, the possibility of "horror autoxicus" could be anticipated.

The facets of organismal origins, autonomous motility, and protean nature represented phagocytes as agents. Agency emerged through both expected and surprising observations, and was constituted through a language of action and interaction. This language of (inter)action was rich, for the interconnection between various action terms afforded an expansion of predication. The movement of phagocytes was variously described, for example, as "traveling," "migrating," and "seeking"; leucocytes were said to "escape" from vessels "in order to arrive at" or "endeavoring to arrive at" infected sites. Engulfing pathogenic entities became a "phagocytic duty," and was also described as an act of "seizing" and "eagerly" or "greedily devouring"; phagocytes were said to "gorge themselves" with bacteria and "devour them alive." Upon "killing" invading parasites, the phagocytes were "the masters of the day." This colorful use of language enhanced the picture of agency, yet in itself was insufficient to constitute phagocytes as agents: the credible application of such vocabulary required that it be built upon a more fundamental form of action – discriminating, directed movement – which was observationally and theoretically grounded. Because it was well-grounded, Metchnikoff's evocative use of language was almost unobtrusive. It served to animate his writing, but also strengthened the picture of phagocytic agency advanced.

The language of action is the key that unlocks the door to intentionality. Action is oriented in a discriminating, meaningful fashion. With this orientation of action in the background, attributions of intention work to describe visible scenery, rather than to ascribe mental, internal life. Metchnikoff revealed the phenomenological connection between "observing action" and "witnessing intention" at the level of micro-organismal

life. In the passage below, for example, the "visible connection" between action and intention is evident.

> Anyone who will watch the maneuvers of *Amoebae* or of certain *Infusoria* in the midst of a rich microscopic flora and fauna, will at once be struck by the preferences which these Protozoa exhibit in their choice of food. *Amoebae* are often seen making search for Diatoms only, disdaining all other Algae, or again they may single out one species of *Palmellaceae* from a very varied flora. The *Infusoria* also have likes and dislikes in the matter of food. Many of the ciliated *Infusoria* choose Bacteria to the exclusion of everything else; others, as *Nassula*, have a special partiality for the *Oscillariae*. A most striking example of this is afforded in *Amphileptus claparedie*, a voracious ciliate, which chooses *Vorticellae* to the exclusion of all other animalcules; these it devours, and then becomes transformed into a cyst upon the peduncle of the *Vorticellae* it has devoured... [A]lmost all the ciliated *Infusoria*, on becoming aware of the proximity of dead bodies of kindred organisms, rapidly draw away, thus manifesting a very marked negative chemiotaxis. This property must, it is evident, protect them from any contamination by the parasites contained in the bodies of *Infusoria* that have succumbed to infective diseases [6, p. 18, 19].

For Metchnikoff observations of the behavior of microorganisms – in terms of preferring, disdaining, or searching – were available for corroboration to "anyone who will watch." The attribution of intentionality to microorganisms was thus rooted in phenomenological thinking. Predicates of action and intention were seamlessly connected, their simultaneous articulation founded on observation. "Preferences" among protozoa were rooted in differential behaviors toward objects in their environment. Against the background of "making a search," attitudes of "disdaining" or "singling out" were seen in the behaviors of *Amoebae*. Concepts like "searching" and "choosing" – attributed to the ciliated *Infusoria* – entangled both behavior and intention in single words. And the action of "rapidly drawing away" anchored the almost casual description of *Infusoria* as "becoming aware." Metchnikoff's language of action thus entrained intentionality at an observational level. The intentional idiom in immunology was inaugurated through Metchnikoff's reasoning about phenomena of cellular behavior. The organismal portrayal of phagocytic cells was the basis of his introduction of intentionality into immune function.

The inclusion of the language of intentionality was a bona fide move – rather than metaphorical or evocative – as Metchnikoff deliberately created an interpenetrability between technical terms and their ordinary language explications. He wrote of the "sensitiveness [of phagocytic cells] as a 'chemiotaxis', that is to say a 'perception' of the chemical composition of the surrounding medium" [6, p.167, emphasis added]. "Sensitive chemiotaxis" was thus defined as equivalent with movement based on "perception." Elsewhere, Metchnikoff wrote that "phagocytes possess a 'kind of taste or chemiotaxis' which enables them to distinguish the chemical composition of substances ([6, p. 79, emphasis added]. His semantic maneuvers incorporated intentionality by mixing technical terminology with ordinary language descriptions. "A kind of taste" and "chemotaxis" are presented as equivalent, even though "taste" is a sensuous experience of a sentient agent, and "chemotaxis" is a technical term describing movement guided by chemical stimulation. Metchnikoff's equalization canceled out the potentially reductionist implications of the term chemotaxis.

The confluence between ordinary language depiction and technical terms affirmed both the appropriateness of a vernacular grasp of phagocytic sensibility and the right of a technical-scientific terminology to encompass it. This "mixed semantics" was entirely typical of Metchnikoff's writing, and it allowed him to accomplish two tasks: to clarify the nature of cellular phenomena by elucidating them in the vernacular; and to express his understanding of evolution, by affirming the universality of sentience, its continuity from cellular through organismal levels. This latter point became especially relevant in his response to the charge of 'vitalism' leveled against him by the humoralists.

Philosophical divergence

At the core of immunity, Metchnikoff identified struggle which he often portrayed as a visible affair: for example, he depicted ingested microbes "struggling" inside the protoplasm of phagocytic cells, and he even described microorganisms as "being eaten alive" by phagocytes.[15] Metchnikoff's imagery was charged with vitality, incarnate interaction, purposive aggression, and life and death confrontation. Humoralists charged that Metchnikoff's representation of phagocytes was vitalistic and teleological. The theme of "agency" was thus implicitly invoked, with a philosophical dimension in the exchanges between cellularists and humoralists concerning "vitalism" and "teleology."

Metchnikoff did not espouse a vitalist philosophy,[16] and objected to the charge. For example, he cited one of his critics who maintained that "the phagocyte theory presupposes extraordinary powers on the part of the protoplasm of leucocytes, to which are attributed sensations, thoughts and actions, in fact a kind of psychical activity" [7, p. 192]. Metchnikoff's rejoinder to this criticism was bold. Rather than being conciliatory, and denying that he attributed "sensations, thoughts and actions" to phagocytes, Metchnikoff countered that the rejection of psychical activity at the unicellular level was not a scientific claim, but founded on an a priori presentiment. He opposed his critics with an evolutionary defense of the view they rejected as "too vitalistic":

> [W]hether (phagocytes) possess powers of thought and volition, as this author accuses me of assuming, is quite beside the question, though we are justified in considering that they possess a germ of these qualities and that their sensibility, like that of various vegetable and animal unicellular organisms represents the lowest stage in the long series of phenomena which culminate in the psychical activities of man [ibid].

The identification of "sensibility" as a form of sentient proto-experience, supported through a fungibility of technical and vernacular terms in Metchnikoff's work, was theoretically validated here from an evolutionary perspective.

---

15. For example, in experiments involving the injection of spermatozoa of different mammal species into the peritoneal cavity of guinea-pigs, Metchnikoff remarked that "the macrophage seizes the spermatozoa which sometimes, by the active movements of their flagella, exhibit a great vitality" [6]; elsewhere he stated that "bacteria, living in the blood plasma, become the prey of phagocytes which render them inoffensive and kill them" [6].

16. Yet Metchnikoff was profoundly influenced by Virchow who regarded himself as a 'cellular vitalist'; Metchnikoff took the concept of 'irritability' from Virchow. He declared that Virchow was one of the three most influential thinkers on his work (the others being Darwin and Pasteur). Thus it may be said that through Virchow, Metchnikoff was exposed to and influenced by a vitalist tradition in physiology and cellular biology. It is also important to remember, however, that 'vitalism' did not mean one thing, and certain forms of vitalist thinking were compatible with materialism (for example, Virchow's vitalism). See [11, 29,].

Metchnikoff refused to preclude the existence of "psychical powers" at a cellular level, on no evidence other than an a priori judgement. He countered this standpoint with an encompassing appreciation of evolutionary continuity. In a masterful move, he reversed the charge of vitalism, arguing that "psychical acts" became vitalist when they are assumed to make an ex nihilo appearance: "the accusation of vitalism and animism, which is unjustly cast at the phagocyte theory, might really be more appropriately applied to my opponents, who maintain that the psychical acts of the higher animals are fundamentally different in their nature from the more simple phenomena peculiar to the lower organisms" [7].

Metchnikoff redirected the charge of vitalism back at his opponents by exposing their tacit belief in the sudden emergence of "psychical acts" in the higher animals. He suggested that evolutionary thinking invited a monistic understanding of capacities across the gamut of organisms – including different forms and degrees of psychical qualities.

Metchnikoff also used the evolutionary theory to counter the charge of "teleology." He invoked natural selection to account for the purposive nature of phagocytic action, arguing that organisms that evolved protective "phagocytosis" were better equipped to survive than those which did not. But phagocytosis, he reasoned, was not teleological, for it was not the "final" product of perfect design or adaptation. Evolution was an on-going and imperfect process, signifying that "the inflammatory reaction is not yet perfectly adapted to its object" [7]. To drive this point home he reconsidered the avoidance behavior of phagocytic cells: "it is because the defense by the phagocytes is developed according to the law of natural selection and is not a designed adaptation to a particular end, that cases naturally occur where the phagocytes do not fulfil their functions, a neglect followed by the most serious danger to or death of the organism [7, p. 194].

Metchnikoff stressed organisms' ceaseless struggle for life, a struggle carried out with imperfect means. The counter-adaptiveness of negative chemotaxis thus revealed that immune function was far from guaranteed, and indicated that his theory of phagocytosis was not teleological.

By interpreting phagocytic avoidance behavior as an "imperfection in the curative forces of nature," he distinguished clearly between life processes as products of natural selection versus design. Metchnikoff drew a conceptual distinction that would resurface in evolutionary thinking, in the differentiation between the teleology implicated in evolutionary adaptation and the teleology of natural-theological views as immanent purpose evidencing deliberate design [30]. Humoralists did not draw this distinction. Evolutionary arguments of adaptation were strikingly absent from Ehrlich's immunochemistry, suggesting that in his rejection of teleology he did not distinguish between "teleological" phenomena conceived as consequences of natural selection versus purposeful design.

For humoralists, the interactions between elements in the environment of the body – cellular, molecular, or atomic – were chemical: passive and inexorable and, thus, predictable and quantifiable. Understanding biological phenomena as chemical was premised on a negation of agency. Humoralists were deliberate in their opposition to teleology and vitalism, which they regarded as outmoded ways of thinking. Ehrlich's desire to eschew teleological reasoning partly drove his formulation of the side-chain theory of antibody. He sought to preempt the idea that antibodies were pre-fashioned for protection, or ejected into the body's humors for prophylactic purposes. To abrogate "purpose" from immune

function, he postulated that side-chains must be receptors that, ordinarily, had a physiological role "other" than conferring immunity; he postulated that side-chains fulfil the cell's nutritive needs. He then reasoned that when the receptors, still attached to the cells, become bound by toxin, the nutritive function was blocked and the cell overcompensated by producing a plethora of chains, which eventually were automatically ejected into the fluids as "antibodies." The free-floating antibodies, "as a purely serendipitous consequence of chemical affinity," bound the toxin circulating in the fluids, thereby serving a protective function for the organism.

The design of Ehrlich's theory speaks to his commitment to extirpate all traces of teleology from his thinking. The antibodies that neutralized toxin, thereby averting its catastrophic effects, must be either formed in response to the toxin or already performed. Ehrlich opted for the second alternative, for he rejected the metaphysical implications of the former:

> [T]o attribute what could be called inventive activity to the body or to its cells, enabling them to produce new groups of atoms as required, would involve a return to the concepts current in the days of (an obsolete) natural philosophy. Our knowledge of cell function and especially of synthetic processes would lead us rather to assume that in the formation of antibodies, we are dealing with the enhancement of a normal cell function, and not with the creation at need of new groups of atoms. Physiological analogues of the group of the specifically combining antibodies must exist beforehand in the organism or in its cells [25, p. 114].

Side-chains were both antibodies (when free-floating) and the receptive anchors of toxin as components of the cell. They were thus inadvertently curative in one context and pathogenic in another. In making the relationship between antibody and pathogen thoroughly fortuitous and chemical, Ehrlich eschewed the slightest hint of purposiveness from his side-chain theory.

Conclusion

What entities exhibit "agency" was an animating, contentious theme in the debate between cellularists and humoralists, a theme barely beneath the surface of how the phenomena were witnessed and what conceptual frameworks were chosen to deliver them. Through a language of (inter)action, on the basis of its organismal nature, autonomy, and protean behaviors, Metchnikoff portrayed phagocytes as agents. The picture of agency allowed for the introduction of an intentional idiom which, in turn, amplified this view. Humoralists objected to what they regarded as a vitalist and teleological conception, to which they counterposed chemical depiction. The language of chemistry was well-suited to the "invisibility" of antibody-mediated immune phenomena they focused on; however, this language also served as a weapon against what they saw as a metaphysical conceptualization of cell life. Metchnikoff, while not a vitalist, defended the priority of cellular action in immunity, placing chemical phenomena in a secondary position as products by means of which cells fulfilled defensive functions.

In their discussions, then, the participants of this debate openly addressed what they themselves considered "metascientific themes": specifically, whether there is immanent purpose in life processes and whether cellular life could be sentient life. These themes were extremely significant, for even though they were not resolvable within frameworks

of scientific practice they were integral to, and formative of, scientific conceptions. With the themes of vitalism and teleology – implicitly and overtly invoked – two schools of the immunological community essentially engaged in a philosophical exchange: they argued about what sorts of entities, actions, and images are real and what are not. This debate illustrates that scientists do not "naively experience the natural world."[17] The mixed discourse of these immunologists, the continuities between their philosophical and scientific reasoning, reveal that they were not naïve realists, but willing to recognize and engage the significance of their metascientific thinking.

In closing, we return to our opening claim that it would be not accurate to assess the outcome of the debate between cellularists and humoralists in terms of "winners" and "losers." It is true that Metchnikoff's cellular school was eclipsed for a number of decades as a research tradition, but his "conceptual representations" of immune phenomena were quietly assimilated into the science. In particular, the ideas of defense, resistance, and their affiliates swiftly became "dead metaphors" in immunology [33]. With their origins in a picture of agency forgotten, these pervasively used conceptions go unnoticed in their effects of modelling intra-organismal events as intentional acts of aggression and defense. Despite the scientific community's enthralment with humoral immunity (or serology) until the 1950s, Metchnikoff's reasoning left an indelible conceptual mark on the constitution of immunity. His constitution of immunity in terms of intentionality and warfare became deeply rooted within the science of immunology despite the apparent "defeat" of his cellular school.

Bibliography

1 Silverstein AM. A history of immunology. San Diego: Academic Press, Inc.; 1989.
2 Tauber AI. The immune Self: theory or metaphor? Cambridge: Cambridge University Press, 1994.
3 Ehrlich P. On immunity with special reference to cell life. In: F. Himmelweit, Ed. 1900; Op. cit., p. 178-195.
4 Latour B. The pasteurization of France. Cambridge: Harvard University Press: 1988.
5 Heifets L. Centennial of Metchnikoff's discovery. J Reticuloendothel Soc 1982; 31: 381-91.
6 Metchnikoff E. Immunity in infective diseases. New York: Johnson Reprint Corporation; 1968[1905].
7 Metchnikoff E. Lectures on the comparative pathology of inflammation. New York: Dover Publications Inc.; 1968[1892].
8 Janeway CA, Travers P. Immunobiology. London: Current Biology Ltd./Garland Publishing Inc.; 1994.
9 Abercrombie M, Hickman CJ, Johnson ML. The Penguin dictionary of biology. London: Penguin Books; 1987.
10 Tauber AI, Podolsky SH. Frank Macfarlane Burnet and the immune self. J Hist Biol 1994; 27: 531-73.
11 Tauber AI, Chernyak L. Metchnikoff and the origins of immunology. From metaphor to theory. Oxford: Oxford University Press; 1991.
12 Hacking I. Representing and intervening. Introductory topics in the philosophy of natural science. Cambridge: Cambridge University Press; 1990[1983].
13 Hanson NR. Patterns of discovery. Cambridge: Cambridge University Press; 1958.
14 Kuhn TS. The structure of scientific revolutions. Second edition. Cambridge: Harvard University Press; 1970[1962].
15 Darwin C. The Expression of emotions in man and animals. Chicago: Chicago University Press; 1965[1872].
16 Gould SJ. Ever since Darwin. New York: W.W. Norton and Company; 1977.
17 Nuttall GHF. Experiments on the antibacterial influence of animal substances. In: D.J. Bibel, Ed., Milestones in immunology, p. Madison: Science Tech Publishers; 1988[1888]. p. 161-66.

17. [32]. See the debate between H.M. Collins and Steven Yearley and Michel Callon and Bruno Latour, for a discussion of the natural scientists as "naive realists" [33].

18 Von Behring E, Kitasato S. The mechanism of immunity in animals to diphtheria and tetanus. In: Brock, TD, Ed. Milestones in microbiology. Englewood Cliffs: Prentice-Hall, Inc.; 1961[1890]. p. 138-141.
19 Cambrosio A, Jacobi D, Keating P. Ehrlich's 'beautiful pictures' and the controversial beginnings of immunological imagery. Isis 1993; 4: 662-99.
20 Keating P, Ousman A. The problem of natural antibodies, 1894-1905. J Hist Biol 1991; 24: 245-63.
21 Hacking I. The self-vindication of the laboratory sciences. In: Pickering A, Ed. Science as practice and culture. Chicago: The University of Chicago Press, 1992. Op. cit., p. 29-64.
22 Baumler E. Paul Ehrlich. Scientist for life. New York: Holmes and Meier; 1984.
23 Mazumdar PMH. Species and specificity. An interpretation of the history of immunology. Cambridge: Cambridge University Press; 1995.
24 Grundbacher FJ. Behring's discovery of diphtheria and tetanus antitoxins. Immunol Today 1992; 13: 188-90.
25 Ehrlich P. The assay of the activity of diphtheria-curative serum and its theoretical basis. In: Himmelweit F, Ed. The collected papers of Paul Ehrlich. London: Pergamon Press; 1956[1897]. p. 107-25.
26 Metchnikoff E. On the present state of the question of immunity in infectious diseases. In: Nobel lectures in physiology or medicine (1901–1921). Amsterdam: Elsevier; 1967[1908]; p. 281-300.
27 Ehrlich P. Partial cell functions. In: Nobel lectures in physiology or medicine (1901–1921). Amsterdam: Elsevier; 1967[1908]. p. 304-24.
28 Davidson D. Essays on actions and events. Oxford: Clarendon Press; 1980.
29 Lenoir T. The strategy of life: teleology and mechanics in 19th Century German biology. Chicago: the University of Chicago press; 1989.
30 Mayr E. The growth of biological thought. Cambridge: Boston : The Belknap Press of Harvard University Press; 1982.
31 Collins HM, Yearley S. Epistemological chicken. In: Pickering A, Ed. Science as practice and culture. Chicago: The University of Chicago Press, 1992. Op.cit., p. 301-26.
32 Discussion between HM Collins, Steven Yearley, Michel Callon and Bruno Latour. See the debate in: Pickering A, Ed. Science as practice and culture. Chicago: The University of Chicago Press, 1992.
33 Fraser B. The interpretation of novel metaphors. In: Ortony A, Ed. Metaphor and thought. Cambridge: Cambridge University Press; 1979.
34 De Kruif P. The microbe hunters. San Diego: Harcourt, Brace, Jovanovich; 1954[1926].

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

From blood fractions to antibody structure: gamma globulin research growing out of World War II

Angela N.H. Creager

Arthur Silverstein has offered a sweeping periodization for twentieth-century immunology that commences with the Age of Bacteriology (1900–1920), followed by the Age of Immunochemistry (1920–1950), and culminates in a still-dominant Age of Immunobiology, which emerged in the late 1950s and is epitomized by the Clonal Selection Theory [1, p. 114]. Other prominent scholars of immunology have emphasized different scientific transitions (such as the emergence of notions of immune self or immune system) and controversies (e.g., between proponents of cellular and humoral immunity), but there has been a general historiographical commitment to concepts and metaphors as motivating scientific change in the field [2–9][1]. Since the beginning of the twentieth century, however, immunology has also included the development of therapeutics, serums, vaccines, and chemotherapies. My essay will explore the role of research materials and therapeutic needs in shaping scientific developments. Immunology will be viewed here in terms of its therapeutic interventions as much as its biological theories.

My case focuses on an aspect of Silverstein's Age of Immunochemistry, which was marked not only by instructionist theories of antibody formation, but also by the mass production of antibodies as gamma globulin from human plasma. Spurred by the medical needs of World War II, gamma globulins were developed to prevent or treat a variety of diseases, including measles, hepatitis, and poliomyelitis. By the 1950s, gamma globulin research began to shape fundamental conceptions of antibodies and their diversity in the blood. Thus, one can see through the story of gamma globulins the way that products from the Age of Immunochemistry bled into the Age of Immunobiology, giving rise to new questions and approaches in immunology.

Gamma globulin from blood protein research to military medicine

Gamma globulin as a scientific entity came into being through the physico-chemical characterization of blood proteins in the laboratory of Swedish colloidal chemist Theodor Svedberg. When Arne Tiselius developed his electrophoresis apparatus there in the 1930s

1. Experimentation has also received attention by historians of immunology, as seen in the recent books by Podolsky and Tauber [8], Cambrosio and Keating [59], and several contributions to "Immunology as a historical object" [62]. Nonetheless, the more panoramic accounts of immunology have been organized around conceptual developments.

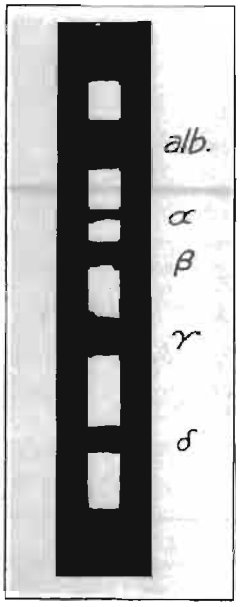


Figure 1. Electrophoresis of horse serum into its protein constituents. (Tiselius A. Electrophoretic analysis and the constitution of native fluids. Harvey Lectures 1939-40; 35: 37-70, 49.) Copyright © 1940 Science Press Priting Company. Reprinted by permission of Wiley-Liss Inc., a subsidiary of John Wiley & Sons, Inc.

[10, 11], his model demonstration was the separation of horse blood serum into several discrete protein components: albumin and α, β, γ, and δ globulins [12, 13] (see *figure 1*). Subsequently, Tiselius and Elvin Kabat demonstrated that the γ (gamma) globulin in rabbit serum contained antibodies (in this case, those raised against egg albumin) [14].

The identification of antibodies with this particular protein fraction in blood – gamma globulin – not only reinforced humoralist explanations of immunity, but opened up antibody research to a range of physico-chemical tools beyond electrophoresis[2]. Studies of gamma globulin in the analytical ultracentrifuge allowed researchers to assign a molecular weight of 157,000 and a Svedberg sedimentation constant of 7S to the entity [14; 15, p. 2]; however, not all gamma globulin preparations appeared to be composed of homogeneous macromolecules of this size and shape. Horses immunized against pneumococci, a common source of serum for laboratory study, produced a much larger antibody species, of molecular weight near 1 million. Moreover, only 90% of the particles in human gamma globulin exhibited the 7S sedimentation constant. Thus, it was unclear how many species of antibodies comprised the electrophoretic fraction called gamma globulin [16].

The diversity of antibodies, not only in terms of size but, more importantly, in terms of binding specificity, seemed to set them apart from other purifiable proteins, such as hemoglobin and ovalbumin. This difference was especially evident by the end of the 1930s, once even large proteins were thought to be simply linear sequences of amino acids joined by

2. On humoral versus cellular theories of immunity, see Silverstein [1], chapter 3 and Podolsky and Tauber [8].

peptide linkages [17; 18, p. 87–179; 19, p. 104–20]. The question of how such large proteins were formed intracellularly remained the subject of much speculation. Because antibodies with very similar amino acid composition could possess very different specificities, instructionist theories of antibody formation – that antigens (not antibodies) carried the information for immunological specificity – were favored. According to Linus Pauling's 1940 theory of antibody formation, all antibodies shared the same composition (being serum globulins) but differed in the structural configuration of the antigen-binding site. The antigen itself, once it came into contact with a nascent antibody polypeptide, served as the molecular template for the folding of the antibody polypeptide into a specific antigen-binding structure. This configuration was subsequently maintained by interatomic (especially hydrogen) bonds [20, 21, 22][3]. Such instructionist explanations for antibody formation continued to be elaborated (e.g., by Macfarlane Burnet [23]), even in the absence of compelling empirical evidence.

World War II provided an unexpected impetus towards putting the field of blood protein research on a firmer biochemical footing. A request from the U.S. Navy in 1940 for assistance from Harvard protein chemist Edwin Cohn in developing a transfusion material from bovine blood motivated Cohn's laboratory to develop a method for separating blood plasma, from humans or cows, into various fractions. By World War II, physiological shock was attributed to a loss in blood volume. Serum albumin was known to be the component of plasma that maintains blood volume, so this protein was targeted for purification as a transfusion material. Cohn's method separated plasma into five fractions on the basis of solubility in ethanol/water mixtures at low temperature. (Cohn's fractions are schematically illustrated in *figure 2*). Fraction V contained highly purified serum albumin. Clinical testing revealed that bovine serum albumin caused adverse reactions, but human serum albumin was found to be an effective transfusion material to prevent shock. From 1942, human serum albumin was mass-produced using Cohn's method by seven pharmaceutical houses. This purified protein, like dried human plasma, was widely transfused into wounded U.S. soldiers during the war [24–27].

Cohn's fractionation method also attracted the attention of physicians interested in the medical use of blood beyond transfusion. Elliott Robinson, head of the Red Cross, as well as Joseph Stokes, a professor of pediatrics at the University of Pennsylvania, suggested to Cohn that the by-product "globulin" fraction might also be mass-produced and used for the "control of measles and other infectious diseases" [24, p. 412]. Late in 1941 and through early 1942, Cohn's laboratory developed a method to further purify the antibodies from the combined Fractions II and III. As with serum albumin, Cohn's preparation of gamma globulin was soon put to the test as a new therapeutic product.

In testing the effectiveness of his gamma globulin preparations in disease protection, Cohn collaborated extensively with two clinicians: Stokes, whose research affiliation was with the Children's Hospital of Philadelphia, and Charles Janeway, a pediatrician at Harvard Medical School and Children's Hospital in Boston. The project drew on the long-standing use of serum in treating infectious disease (see [9]). In 1935, Stokes worked with colleagues at Penn to develop a method for storing human serum in a lyophilized (freeze-

3. For more on the history of theories of antibody formation, see Silverstein [1], chapter 4.

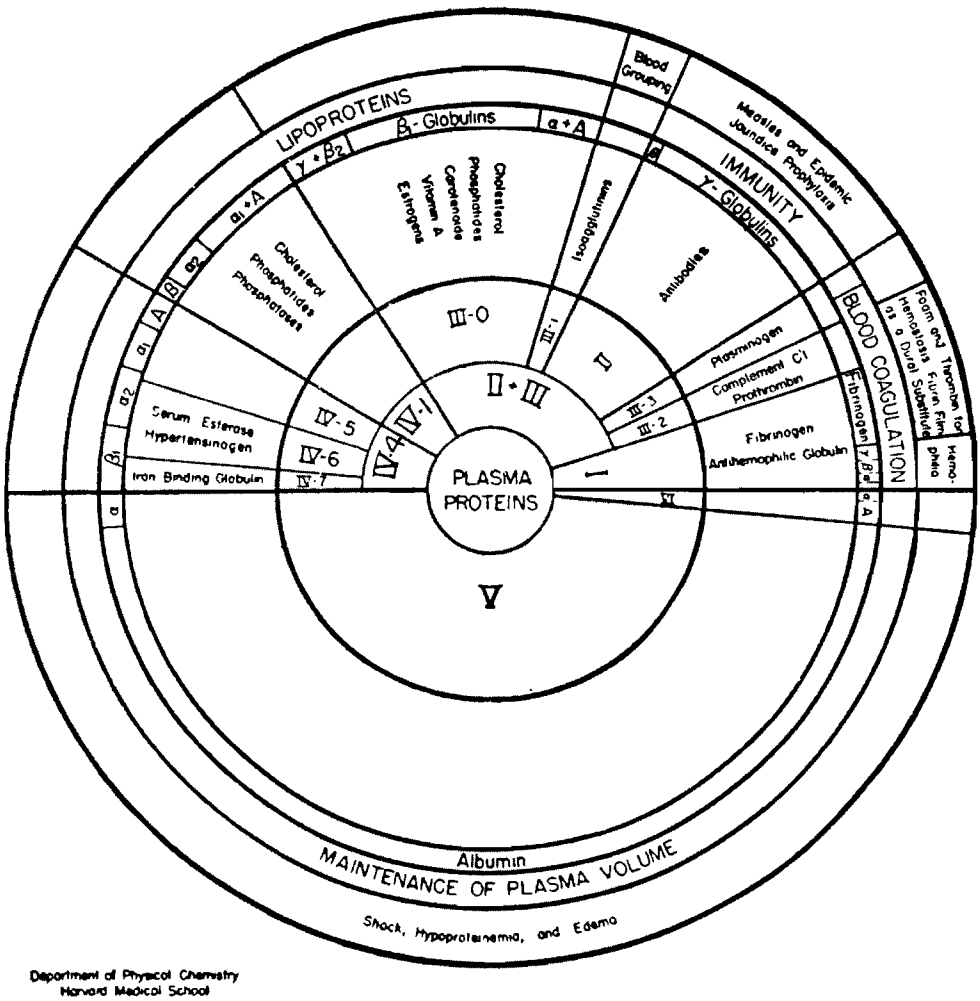


Figure 2. Edwin Cohn's diagram depicting the separation of the various components of plasma into its major fractions, indicating their natural functions and clinical uses. (Cohn EJ. The history of plasma fractionation. In: Andrus EC, Ed. Advances in military medicine, vol. 1. Boston: Little, Brown and Company; 1948. p. 364-443, 391.)

dried) form [28], and he also helped found the Philadelphia Serum Exchange at Children's Hospital, which used this new technical advance in their handling of sera. Serum centers were increasingly common in the United States during the 1930s (they had been established in Paris and Stockholm earlier), with early institutions being located in Chicago, Detroit, Providence, and Milwaukee[4]. The Philadelphia Serum Exchange produced hyperimmune sera against *Pneumococcus bacillus* and *Staphylococcus aureus*, and distributed

4. "The Philadelphia Serum Exchange," Stokes papers, American Philosophical Society Library, Manuscripts Division (hereafter APS), B: St65p, folder Lyophile Serum #2 (1935).

pooled normal human serum as well as convalescent serum for use against measles, mumps, scarlet fever, impetigo, chicken pox, and poliomyelitis[5]. By 1937, the Philadelphia Serum Exchange had entered into an agreement with Sharp & Dohme for the mass production of these lyophilized sera[6]. Stokes' interest in passive immunization extended beyond sera: in the early 1930s he tested whether adult whole blood transfusion (from parents) was an effective means of immunizing children against poliomyelitis [29]. His research projects were characteristic of a general mood of medical optimism about the power of immunotherapy, passive serotherapy, as well as vaccination [9, p. 480].

By the time Stokes became involved with Cohn's plasma fractionation project, he was serving as the Director of the Army Epidemiological Board's Commission on Measles and Mumps. Testing of Cohn's Fraction II + III for immunizing potency soon became part of the commission's war work [30]. The researchers associated with the commission focused on three lines of investigation. On the laboratory side, John Enders assayed the fractions for antibodies to known diseases (particularly viral diseases) to determine the immunologic properties. Second, Stokes and collaborating pediatricians did a field study in Philadelphia of the potency of gamma globulin in immunizing children against measles[7]. Their study included 450 children intimately exposed to measles in families or institutions. Because the measles epidemic of February and March 1943 was so severe, Stokes noted that "families would not permit the division of susceptible exposed children into injected and control groups."[8] Nonetheless, their trials indicated the efficacy of gamma globulin, not only in preventing measles but also in altering the course of the disease to produce a more mild case. Similar field studies on a smaller scale were carried out by Enders and Janeway in Boston[9]. A third line of inquiry concerned the "local and general reactions in human subjects in comparison with the reactions caused by other preparations of human immune bodies."[10] Cohn's gamma globulin was found to cause less severe reactions than immune bodies isolated from placenta and to be more potent than pooled adult of convalescent serum. Gamma globulin was further tested by the commission as a means of passively immunizing against mumps, but the results were not as promising as those for measles[11].

By the fall of 1942, epidemic hepatitis stood out among other infectious diseases as a serious military medical problem. In addition to the increased incidence of epidemics of

5. "Clinical Applications of the Lyophile Process," Stokes papers, APS B: St65p, folder Lyophile Serum #1 (1935).
6. See documents in Stokes papers, APS B: St65p, folder Lyophile Serum: Sharp & Dohme.
7. This study built on Stokes' work in 1941 with the Philadelphia Serum Exchange on use of pooled normal adult serum (lyophilized) in passive immunization against measles; they showed it to be "more uniformly successful in the prevention or attenuation of measles than placental globulin." Stokes to Stuart P. Mudd, Oct 7, 1941, Stokes papers, APS B: St65p, folder Mudd, Stuart P. #1.
8. Joseph Stokes, Jr., "Report of the Commission on Measles and Mumps, Board for the Investigation and Control of Influenza and Other Epidemic Diseases, United States Army, July 1, 1942 to April 20, 1943," Stokes papers, APS B: St65p, folder Commission on Measles and Mumps #18, p. 2.
9. Joseph Stokes, Jr., "Report of the Commission on Measles and Mumps, Epidemiological Board, U.S. Army, April 20, 1944 to April 1, 1945," Stokes papers, APS B: St65p, folder Army Epidemiological Board #13. Janeway also helped coordinate field studies with physicians in other different regions of the United States.
10. Joseph Stokes, Jr., "Report of the Commission on Measles and Mumps, Board for the Investigation and Control of Influenza and Other Epidemic Diseases, United States Army, July 1, 1942 to April 20, 1943," Stokes papers, APS B: St65p, folder Commission on Measles and Mumps #18, p. 1.
11. Ibid., p. 9. The commission also investigated active immunization against measles with a vaccine developed by Squibb, but many of the vaccinated children contracted measles (p. 6).

jaundice associated with the war, hepatitis was increasingly associated with transfusions of whole blood and plasma [31][12]. Consequently, Stokes sought to test the clinical effectiveness of gamma globulin against hepatitis. Over the course of the war, medical researchers differentiated the viral agents of homologous serum hepatitis (now called hepatitis B) from infectious hepatitis (hepatitis A) on the basis that immunity to one did not give immunity to the other. Clinical trials carried out in 1945 by Stokes and John Neefe (on children in summer camps and on soldiers in the Mediterranean Theater of Operations) demonstrated the protective effects of gamma globulin against infectious hepatitis [32; see also 33]. Moreover, their studies showed that under the epidemic conditions prevalent in military service "active immunity was superimposed upon a waning passive immunity following a single injection of gamma globulin," such that a single dose of gamma globulin, presumably followed by mild infection, induced "a solid immunity."[13] Significantly, unlike other blood products, gamma globulin was never found to cause homologous serum hepatitis. In fact, based on some studies suggesting that gamma globulin protected against serum hepatitis as well, many clinicians felt that gamma globulin should be administered to all patients receiving transfusions to prevent the possibility of serum hepatitis[14]. A human blood product itself, it served to protect patients against the dangers of these new therapeutic agents, pooled from increasingly larger groups of (possibly infectious) donors.

Gamma globulin therapy in the postwar years

Gamma globulin continued to be prepared industrially in the postwar period, often from surplus or outdated plasma collected during the war effort, and was usually pooled from between 3,000 to 4,000 donors[15]. As even normal gamma globulin showed a 25-fold enrichment of antibodies over pooled normal human serum, small doses could be effective in conferring immunity for particular diseases. Donors were also solicited for the preparation of special hyperimmune globulin, in which the concentration of antibodies against a particular pathogen was greatly enhanced. Building on earlier use of convalescent serum, Stokes developed a hyperimmune pertussis globulin from selected and paid donors who were bled regularly and whose antibody levels were maintained by repeated injections of antigen[16]. The customized nature of this therapeutic did not prevent its industrial production – Cutter Laboratories handled the packaging and distribution of this biologic product in 1945.

---

12. By 1945 it was recognized that a severe outbreak of jaundice in the Armed Forces in 1942 was actually attributable to contaminated yellow fever vaccine which had been diluted with pooled human serum.
13. Stokes to Harry Weaver, Jan. 13, 1950, Stokes papers, APS B: St65p, folder National Foundation for Infantile Paralysis #5.
14. "Requirements for Gamma Globulin," undated [~1962], Stokes papers, APS B: St65p, folder NRC–Plasma #5. Other studies showed that gamma globulin did not offer protection against homologous serum hepatitis [63].
15. The pooling of blood or plasma could combine from 500 to 20,000 donations. Charles Janeway, "Nature and Properties of Human Serum Gamma Globulin," March 21, 1963, prepared for National Research Council Subcommittee on Plasma's Report on Gamma Globulin, Stokes papers, APS B: St65p, folders National Research Council #6, #7 (continuation of draft report), p. 14-15.
16. "Report to the Commission on Plasma Fractionation and Related Processes from the Advisory Committee on Standards for the Appraisal of Hyperimmune Pertussis Globulin," March 15, 1945, Stokes papers, APS B: St65p, folder Army Epidemiological Board #10 (1945).

Developments in biomechanical blood separation equipment in Cohn's Harvard laboratory further stimulated Stokes' interest in developing individualized plasma-derived products. Seeking ways to preserve the most labile components of blood in the early 1950s, Cohn developed a portable blood fractionation machine, which could be directly connected to a donor's arm in order to separate the plasma components as the blood was drawn [34][17]. He patented his portable fractionation system and negotiated in 1951 with a scientific instruments company to mass-produce his miniature blood fractionator[18]. Cohn intended to oversee the production of a whole fleet of these machines, which were to be used at blood donation centers, but died unexpectedly in 1953, as instrument manufacturer Arthur D. Little prepared to make their first customized models.

Stokes acquired one of the first five of Cohn's commercial miniature plasma fractionator machines. He and coworker Joseph Smolens were the first to use the apparatus for plasmapheresis, the "removal of the donor's plasma while returning his red blood cells" [35] (See *figure 3*). Showing that weekly plasmapheresis was safe for healthy individuals and did not lead to a loss in antibody titer, Stokes and Smolens advocated use of the Cohn portable fractionator for "separation of the gamma globulin alone, which we have termed immunophoresis."[19] In fact, plasmapheresis of the same donor in concert with immunization or hyperimmunization could yield custom hyperimmune sera [36].

In 1953, Stokes and Smolens were particularly interested in obtaining hyperimmune plasma from donors who had been inoculated with Jonas Salk's poliomyelitis vaccine[20]. Cohn's group at Harvard was once again eager to collaborate with Stokes and Smolens to test the effectiveness of preparations from this hyperimmune plasma for passive immunization against polio[21]. In a 1954 grant application to the National Foundation for Infantile Paralysis, Smolens and Stokes justified their proposal on the basis "that immune globulin will be needed to supplement the use of vaccines, [it being] highly probable that the two methods can be developed jointly for the production of more permanent immunity, in much the same manner as we have demonstrated in man in epidemic hepatitis."[22] Stokes, however, was to be disappointed; by 1955, his earlier success with gamma globulin against hepatitis was viewed as a poor model for poliomyelitis, and Jonas Salk

17. Cohn received funding from the National Institutes of Health for this project and in 1951 opened a new laboratory, the Blood Characterization and Preservation Laboratory at the Harvard-associated Bussey Institution of Applied Biology. In this way he addressed the recognized superiority of whole blood over blood derivatives for transfusion. Arguing that the only way to guard against the next national emergency would be to find a way to stockpile blood, his research program was to investigate how to preserve the "formed elements" of blood, red blood cells, white blood cells, and platelets [34].

18. Edwin J. Cohn, "Implications of New Knowledge about Blood for Transfusion Services and for the Public Health," unpublished mss., Cohn papers, Rare Books Room, Francis Countway Library of Harvard Medical School, folder Patent Matters 1952.

19. Stokes to Harry M. Weaver, Jan. 17, 1953, Stokes papers, APS B: St65p, folder National Foundation for Infantile Paralysis #9.

20. Stokes to T. Duckett Jones, July 7, 1953, Stokes papers, APS B: St65p, folder National Foundation for Infantile Paralysis #8.

21. See Stokes to Kumm, July 24, 1953, Cohn to Stokes, 23 July 1953, and Stokes to Kumm, July 31, 1953, Stokes papers, APS B: St65p, Plasmapheresis Grants #9.

22. Joseph Stokes, Jr. and Joseph Smolens, "Adaptation of Biomechanical Equipment for Immunopheresis," draft mss., Stokes papers, APS B: St65p, Plasmapheresis Grants #2, p. 3.

Figure 3. ADL Cohn fractionator in use. The apparatus is returning red blood cells to the donor on the left while blood is withdrawn from the donor on the right. (Stokes J Jr, Smolens J. Repeated plasmapheresis in the same person – a rationale for modern bloodletting. Proc Am Phil Soc 1957; 101: 330-5, 331.) Reproduced with permission of the American Philosophical Society.

was too busy with his vaccine development to devote time to contribute toward their studies. Interest in the use of hyperimmune plasma then shifted to mumps and pertussis[23].

Developments in new immunization measures against poliomyelitis in fact served to dampen medical hopes for gamma globulin generally. In the 1940s, David Kramer, Stokes, and other clinicians had urged the National Foundation for Infantile Paralysis to

23. Hyperimmune gamma globulins continued to be of medical interest for some diseases, such as measles, mumps, rabies, tetanus, pertussis, malaria, vaccinia, and staphylococcal infections. See presentation title by Elvin Kabat, National Research Council Symposium on Gamma Globulin, sponsored by the Subcommittee on Plasma, Oct. 26-27, 1962, preliminary program in Stokes papers, APS B: St65p, folder National Research Council—Plasma #3. Some of these gamma globulins (for mumps, pertusis, and tetanus) were still being made commercially available at that time. Minutes, NRC Subcommittee on Plasma, Feb. 7, 1962, Stokes papers, APS B: St65p, folder NRC–Plasma #4, p. 5.

follow up suggestive animal studies with gamma globulin to test its efficacy in humans to protect against polio (and perhaps to also alter the course of the disease). During the war, however, supplies of gamma globulin were limited, and the Red Cross was apparently unwilling to make sufficient quantities available for a field trial [37, p. 472-5, esp. note 18]. Working on a smaller scale, Stokes was able to administer gamma globulin to hundreds of children in summer camps in 1947 and 1948, and subsequent rates of polio observed were low. A less intentional field study of the efficacy of gamma globulin occurred in Houston in the summer 1948, when medical practitioners were dispensing gamma globulin to children in their care during an epidemic. Analysis of polio rates afterwards pointed to the protective effects of gamma globulin.

On the laboratory side, in 1949, David Bodian found that Red Cross gamma globulin neutralized all three strains of poliomyelitis virus in approximately equal titer [38]. After several round table conferences with virologists about "the wisdom of passive immunization" [37, p. 476], and after much resistance, the National Foundation for Infantile Paralysis decided to support William Hammon's field trials of gamma globulin as a protective measure in polio epidemics. A 1951 pilot study in Provo, Utah, in which Hammons and Stokes collaborated, showed the feasibility of a larger-scale field study [39]. Results from the 5,767 children injected in Utah were suggestive, but not statistically significant. A field trial was planned for 1952, involving 50,000 children in Texas, Iowa, and Nebraska (principally around Houston and Sioux City), and the results again showed a protective effect. Children were injected with either gamma globulin or a placebo (gelatin), and six cases occurred among those who received gamma globulin, whereas 38 cases occurred in recipients of gelatin [40, 41]. A surge in demand for gamma globulin followed these results, and in the summer of 1953 the Red Cross released large amounts of gamma globulin to state public health agencies for distribution[24]. A National Foundation's "Report to Physicians" the following year emphasized that:

> [G]amma globulin is the *only* proved weapon *now* available to physicians for prophylaxis against paralytic poliomyelitis. The basic scientific evidence established in 1951 and 1952 field trials that gamma globulin can prevent paralytic polio has recently been not only reaffirmed but strengthened, so that its usefulness seems even greater than we formerly believed[25].

However, coupled with the clinician's advocacy was the epidemiologist's skepticism. In the summer of 1953, an evaluation by the U.S. Public Health Service of large-scale use of gamma globulin prophylaxis was disappointing. The "Report [by the CDC's] National Advisory Committee for the Evaluation of the Efficacy of Gamma Globulin in the Prophylaxis of Paralytic Poliomyelitis" led the Public Health Service to conclude that "the preventive effect [of gamma globulin] in community prophylaxis ... has not been demonstrated[, and] no modification of the severity of paralysis by gamma globulin was shown" [42, p. 21].

---

24. See correspondence in Stokes papers, APS B.St65p, folder Gamma Globulin #6.
25. Kenneth Landauer, "Passive and Active Immunization Against Poliomyelitis: Current Status," Report to Physicians, Summer 1954 [NFIP publication], Stokes papers, APS B: St65p, folder National Foundation for Infantile Paralysis #1, p. 1.

But more saliently, by 1954, hopes had shifted to the Salk vaccine, which in fact displaced gamma globulin as an immunizing agent, even as the successful demonstration of the vaccine had profited from the experience the National Foundation had gained in field trials with gamma globulin. Nonetheless, some clinicians felt that gamma globulin still had a role in vaccination. Worries about live virus which might be present in formalized polio preparations led Stokes to propose that gamma globulin be used in concert with vaccine administration. In fact, Stokes had collaborated with Hilary Koprowski early in the 1950s on attenuated living poliomyelitis virus, which was administered to infants still protected by their mother's antibodies or to older children injected simultaneously with gamma globulin[26]. Stokes was not the only medical researcher to envision such an immunization process; in his Herter lectures at Johns Hopkins Medical School in 1950, Macfarlane Burnet stated, "I feel confident that sooner or later it will become necessary to use living virus vaccine given by mouth in infancy, perhaps under cover of gamma globulin passive protection" [43, as quoted in 44, p. 441]. However, the efficacy of active immunization vaccines alone brought the use of gamma globulin for poliomyelitis to a close. Instead, the main immunization targets of gamma globulin became measles, hepatitis, and tetanus[27]. In addition, gamma globulin became the key therapy for another childhood disease: agammaglobulinemia – a pathological condition in which the patient's blood lacks gamma globulin. The discovery of this condition in 1952, in children who experienced a succession of serious bacterial infections [45], relied on specific medical developments from World War II: newly available antibiotics could be used to keep these children alive and their gamma globulin levels were analyzed by the new biochemical methods[28]. Regular treatment with gamma globulin largely protected these children against otherwise recurrent infections [46]. In 1961, the National Research Council's Subcommittee on Plasma gathered data on current gamma globulin usage and found that about "one-third was being used for management of hypogammaglobulinemia [antibody deficiency including agammaglobulinemia] and for research purposes and the remainder for prophylaxis against measles and hepatitis."[29] The development of an effective measles vaccine by John Enders promised to decrease even this level of demand for gamma globulin. (It might be noted that the Enders measles vaccine was at first used in concert with a gamma globulin injection to decrease reactions, but a further attenuated vaccine made simultaneous administration of gamma globulin unnecessary.) At the same time, new applications for gamma globulin, particularly through the development of preparations suitable for intravenous administration in the 1970s and 1980s, increased demand [47]. Still used to treat some infectious diseases (such as hepatitis A), gamma globulin is increasingly administered to immunocompromised individuals (including those with

26. See Joseph Stokes, Jr., comments for Discussion of Poliomyelitis Panel, June 7, 1955, A.M.A., Atlantic City, Stokes papers, APS B: St65p, folder Poliomyelitis 1955.

27. See package-circular draft for Lederle's Immune Serum Globulin (Human), Stokes papers, APS B: St65p, folder Lederle Laboratories #5.

28. I am indebted to Craig Stillwell for information on this point. Patients with agammaglobulinemia are particularly susceptible to high-grade pyogenic pathogens, such as pneumococci, streptococci, and *Hemophilus* [46].

29. Minutes, National Academy of Sciences–National Research Council, Division of Medical Science, Subcommittee on Plasma of the Committee on Blood and Related Problems, 23rd meeting, Nov. 21-22, 1961, Stokes papers, APS B: St65p, folder National Research Council—Plasma #6, p. 23.

AIDS), and is also now used to treat some autoimmune conditions (on account of its function as a modulator of the immune system). Expanded medical usage has resulted in shortages of gamma globulin in the United States since the mid-1990s [48, 49].

From gamma globulin to immunoglobulin

How did these postwar experiences in the clinic and the field relate to contemporary developments in immunological theory? At one level, the isolation of gamma globulin fractions made some long-standing immunological questions accessible to biochemical investigation. As Charles Janeway asserted:

> The development of human gamma globulin has had important consequences for public health, for medicine, and for research... The availability of large amounts of pooled human antibody in a safe form has provided material for many types of clinical and biologic investigation, and it has served medicine well as a means of preventing or modifying measles and infectious hepatitis in large numbers of susceptible people. [50, p. 6]

For instance, as John Enders noted, immunologists had long been interested in knowing "whether or not the antibodies found in so-called normal individuals and in low titer represent the same factors [that] appear in greatly increased concentration following an overt attack of a disease or after artificial active immunization" [51, p. 175]. Enders cited evidence pointing to the chemical identity of natural and immune antibodies, although consensus was not assured on this question.

Cohn's fractionation methods also made antibodies more accessible for physicochemical study in the postwar period. However, a physical chemist researching gamma globulin from natural rather than hyperimmune plasma found a diversity of proteins, and, as John Oncley noted in 1953, the traditional division by solubility into "pseudoglobulins" (more soluble) versus "euglobulins" (less soluble) shed little light on biological variation among antibodies [16, p. 181-2]. Components of gamma globulin exhibited a Gaussian distribution of isoelectric points from pH 6.3 to 8.4. Some antibodies were concentrated in beta globulin (defined by its more rapid electrophoretic migration as compared to gamma globulin), and this source displayed a similar molecular heterogeneity. Physical biochemists sought to further fractionate Cohn's fractions into specific antibody pools, but had only limited success. As Burnet observed about this kind of investigation:

> The methods chiefly used, electrophoretic analysis, determination of sedimentation constants in the ultracentrifuge, and fractional precipitation by various methods, will each allow the differentiation of certain fractions. It is, however, not usually possible to equate completely the fractions obtained by one method with those obtained by another [52, p. 3].

Thus, biochemical studies did not uncover a natural molecular order within the entity known as gamma globulin. Changing terminology reflected the growing awareness of the molecular heterogeneity of antibodies. The term "immunoglobulin" began to supplant that of "gamma globulin," since globulins with "real or potential antibody activity... display[ed] a wider range of electrophoretic mobility" than conventional gamma

globulins[30]. (This shift became official at the World Health Organization Meeting on Nomenclature of Human Immunoglobulins in 1964 [8, p. 387-8].) Serological precipitation tests began to be used in concert with electrophoretic separation to diagnose antibodies via immunoelectrophoresis, and these assays confirmed the great diversity in structure and specificity among antibodies. Strikingly, differences in biological function between different antibody species (such as fixation of complement or ability to pass across the placenta) were found to correlate with structural differences [1, p. 131].

Using a combination of polypeptide-splitting enzymes and specific antigens to assay antibody binding, researchers also undertook a molecular dissection of antibody molecules. The proteolysis of antibodies by enzymes was first observed in the 1940s [50, p. 7-8]; more precise studies along these lines by Rodney Porter during the 1950s culminated in the classification of the Fc (crystallizable) and Fab (antigen-binding) fragments [53, 54]. Gerald Edelman provided clear evidence that antibodies contain a multi-polypeptide structure and subsequently (with Miroslav Poulik) differentiated heavy from light chains [55, 56]. By 1970, the word "immunoglobulin" was being replaced by designations for more chemically discrete species of antibodies: IgG, IgA, IgM, IgE, and IgD. The individual immunoglobulin polypeptide chains became designated by the symbols κ and λ for light chains and γ, α, μ, etc., for heavy chains. At this point, "Cohn's Fraction II + III, ... the material used clinically for prophylaxis and therapy, and called gamma globulin," was reconceptualized as a mixture of 95% IgG, with some IgA, IgM, and albumin [57].

Both the shifting nomenclature and the new experimental methods played into the larger debates about the value of chemical versus biological approaches to the antibody diversity problem. At the very least, the molecular complexities of gamma globulin as an entity could be seen as indicating the limitations of chemical approaches to understanding antibody structure and formation. Burnet wrote that his own approach to the problem of antibody production was that it "is a biological phenomenon to be interpreted on biological rather than chemical or pseudo-chemical lines"[52, p. v]. He criticized the biochemical purification and characterization of antibodies as "highly fruitful in theoretical and practical results but it has still left immunology as a science almost wholly unrelated to the general pattern of biological knowledge" [52, p. 1]. However, the biochemistry of antibody structure growing out of the technologies for producing and characterizing gamma globulin contributed key structural support for the Clonal Selection Theory [8, p. 58 ff.]. In this respect, the representation of the solution to the antibody diversity problem as a triumph of biological over chemical approaches is historically oversimplified. Chemical methods of analysis, in conjunction with the abundance of gamma globulin as a therapeutic material, contributed powerfully toward the new articulations of antibody structure and genetics.

Conclusion

Robert Kohler asserted some years ago that "immunochemistry was intimately connected historically and socially with the rise of biochemistry" [58, p. 195]. Zymase, for instance,

30. Charles Janeway, "Nature and Properties of Human Serum Gamma Globulin," March 21, 1963, drafted for NRC Report of Gamma Globulins for Subcommittee on Plasma, Stokes papers, APS B: St65p, folder National Research Council—Plasma #6.

was uncovered in the course of experiments on bacterial immunoproteins, and biochemical notions of specificity have long been associated with the antibody–antigen interaction. The synergy between immunology and biochemistry, if a century old, is not obsolete and it continues to work both ways. Antibodies, for instance, have long been laboratory tools as well as objects of investigation, and Cambrosio and Keating have focused attention on the significance of monoclonal antibodies for biochemistry, indeed biomedicine, during the last two decades [59].

Here I have focused attention on the interplay between biochemical experimentation and clinical innovation from 1940 to 1960 by following the fate of blood-based therapies and entities. Rather than focusing on the diverse uses of antibodies as laboratory tools, I have examined research aimed at using antibodies therapeutically and the implications for understanding antibody structure that resulted. By emphasizing the interchanges between clinical and research settings, I have taken inspiration from Ilana Löwy's analysis of the role of the notion of an immune self in mediating between the social words of researchers and physicians [60]. Both of our accounts focus on the transition, in the 1950s, between the "chemical" and "biological" eras of immunology. Yet where Löwy stresses the centrality of a flexible concept, I have sought to emphasize the significance of a material, gamma globulin, which was administered as a therapeutic product and used as the basis for antibody research. Löwy points out that the marked growth of immunology after World War II cannot be attributed to the development of the Clonal Selection Theory, which was accepted after the disciplinary expansion of the 1950s [60, p. 385]. She proposes instead that the notion of an immune self, particularly as articulated by Burnet in 1941 and 1949, served as a flexible and highly useful concept bringing together the concerns of clinicians and researchers. Yet it seems to me that this argument, while highly persuasive, privileges unduly the role of concepts in the history of immunology (even as Löwy pays close attention to experimental developments). Particular objects and materials also helped to consolidate immunology as both a clinical and scientific field of knowledge. In 1950, as older debates between humoralists and cellularists were recast into debates between immunological chemists and biologists (or, in Niels Jerne's clever terms, "between 'cis' and 'trans' immunologists [61])," antibodies provided one such common (and concrete) point of reference for theoretical debates and therapeutic innovation.

Acknowledgements

I thank Gail Schmitt and Suman Seth for invaluable research assistance, Douglas Surgenor for the loan of primary sources in his personal possession, and Martin Weigert for encouraging comments on an earlier draft of this essay. A Princeton University Research Grant supported archival research at the American Philosophical Society, where Scott de Haven offered helpful guidance to the manuscript collection. I am indebted to Anne-Marie Moulin and Alberto Cambrosio for their comments and editorial work on this volume, and to the Musée Claude Bernard-Foundation Marcel Mérieux for their hospitality and support, in conjunction with the Wellcome Trust.

Bibliography

1 Silverstein AM. A history of immunology. New York: Academic Press; 1989.
2 Moulin AM. The immune system: a key concept for the history of immunology. Hist Phil Life Sci 1989; 11: 221-36.
3 Moulin AM. Le dernier langage de la médecine : histoire de l'immunologie de Pasteur au Sida. Paris: Presses Universitaires de France; 1991.

4 Tauber AI, Chernyak L. Metchnikoff and the origins of immunology: from metaphor to theory. Oxford: Oxford University Press; 1991.
5 Tauber AI. The immune self: theory or metaphor? Cambridge: Cambridge University Press; 1994.
6 Martin E. Flexible bodies: tracking immunity in American culture from the days of polio to the age of AIDS. Boston: Beacon Press; 1994.
7 Mazumdar PMH. Species and specificity: an interpretation of the history of immunology. Cambridge: Cambridge University Press; 1995.
8 Podolsky SH, Tauber AI. The generation of diversity: Clonal selection theory and the rise of molecular immunology. Cambridge, MA: Harvard University Press; 1997.
9 Moulin AM. A science 'dans le siècle': Immunology or the science of boundaries. In: Krige J, Pestre D, Eds. Science in the twentieth century. Amsterdam: Overseas Publishers Association for Harwood Academic Publishers; 1997. p. 479-93.
10 Tiselius A. A new apparatus for electrophoretic analysis of colloidal mixtures. Trans Faraday Soc 1937; 33: 524-31.
11 Kay LE. Laboratory technology and biological knowledge: the Tiselius electrophoresis apparatus, 1930–1945. Hist Phil Life Sci 1988; 10: 51-72.
12 Tiselius A. XLV. Electrophoresis of serum globulin. I. Biochem J 1937; 31: 313-7.
13 Tiselius A. CLXXXII. Electrophoresis of serum globulin II. Electrophoretic analysis of normal and immune serum. Biochem J 1937; 31: 1464-77.
14 Tiselius A, Kabat EA. Electrophoresis of immune serum. Science 1938; 87: 416-7.
15 Janeway Ca, Rosen FS, Merler E, Alper CA. The gamma globulins. Boston: Little, Brown and Company; 1967.
16 Oncley JL. Physical characteristics of the gamma globulins. In: Tullis JL, Ed. Blood cells and plasma proteins: their state in nature. New York: Academic Press; 1953. p. 180-6.
17 Edsall JT. Proteins as macromolecules: an essay on the development of the macromolecule concept and some of its vicissitudes. Arch Biochem Biophys 1962; Suppl 1: 12-20.
18 Fruton JS. Molecules and life: historical essays on the interplay of chemistry and biology. New York: Wiley-Interscience; 1972.
19 Kay LE. The molecular vision of life: Caltech, the Rockefeller Foundation, and the rise of the new biology. Oxford: Oxford University Press; 1993.
20 Pauling L. A theory of the structure and process of formation of antibodies. J Am Chem Soc 1940; 62: 2643-57.
21 Pauling L, Delbrück M. The nature of the intermediate forces operative in biological processes. Science 1940; 92: 77-9.
22 Kay LE. Molecular biology and Pauling's immunochemistry: a neglected dimension. Hist Phil Life Sci 1989; 11: 211-9.
23 Burnet FM. The production of antibodies. 1st ed. New York: Macmillan; 1941.
24 Cohn EJ. The history of plasma fractionation. In: Andrus EC, Ed. Advances in military medicine, vol. 1. Boston: Little, Brown and Company; 1948. p. 364-443.
25 Edsall JT. A historical sketch of the Department of Physical Chemistry, Harvard Medical School, 1920–1950. Am Sci 1950; 38: 580-93.
26 Creager ANH. Biotechnology and blood: Edwin Cohn's plasma fractionation project, 1940–1953. In: Thackray A, Ed. Private science: biotechnology and the rise of the molecular sciences. Philadelphia: University of Pennsylvania Press; 1998. p. 39-62.
27 Creager ANH. Producing molecular therapeutics from human blood: Edwin Cohn's wartime enterprise. In: de Chadarevian S, Kamminga H, Eds. Molecularizing biology and medicine: new practices and alliances, 1910s–1970s. Amsterdam: Overseas Publishers Association for Harwood Academic Publishers; 1998. p. 107-38.
28 Mudd S, Flosdorf EW, Eagle H, Stokes J Jr, McGuinness AC. The preservation and concentration of human serums for clinical use. JAMA 1936; 107: 956-9.
29 Stokes J Jr, Wolman IJ, Carpenter HC, Margolis J. Prophylactic use of parents' whole blood in anterior poliomyelitis: Philadelphia epidemic in 1932. Am J Dis Child 1935; 50: 581-95.
30 Stokes J Jr, Maris EP, Gellis SS. Use of concentrated normal human serum gamma globulin for the prevention and attenuation of measles. J Clin Invest 1944; 23: 531-41.
31 Beeson PB. Jaundice occurring one to four months after transfusion of blood or plasma: report of 7 cases. J Am Med Assoc 1943; 121: 1332-44.
32 Stokes J Jr, Neefe JR. The prevention and attenuation of infectious hepatitis by gamma globulin. J Am Med Assoc 1945; 127: 144-5.
33 Gellis SS, Stokes J Jr, Brother GM, Hall WM, Gilmore HR, Beyer E, Morrissey RA. The use of human immune serum globulin (gamma globulin) in infectious (epidemic) hepatitis in the Mediterranean

Theater of Operations. I. Studies on prophylaxis in two epidemics of infectious hepatitis. J Am Med Assoc 1945; 128: 1062-3.
34 Creager ANH. 'What blood told Dr. Cohn': World War II, plasma fractionation, and the growth of human blood research. Stud Hist Phil Biol Biomed Sci 1999; 30: 377-405.
35 Smolens J, Stokes J Jr, Vogt AB. Human plasmapheresis and its effect on antibodies. J Immunol 1957; 79: 434-9.
36 Stokes J Jr, Smolens J. Repeated plasmapheresis in the same person – a rationale for modern bloodletting. Proc Am Phil Soc 1957 ; 101 : 330-5.
37 Benison S. Tom Rivers: reflections on a life in medicine and science. Cambridge, MA: MIT Press; 1967.
38 Bodian D. Neutralization of three immunological types of poliomyelitis virus by human gamma globulin. Proc Soc Exp Biol Med 1949; 72: 259-61.
39 Hammon WM, Pittsburgh PH, Coriell LL, Stokes J Jr. Evaluation of Red Cross gamma globulin as a prophylactic agent for poliomyelitis. 1. Plan of controlled field tests and results of 1951 pilot study in Utah. JAMA 1952; 150: 739-49.
40 Hammon WM, Coriell LL, Stokes J Jr. Evaluation of Red Cross gamma globulin as a prophylactic agent for poliomyelitis. 2. Conduct and early follow-up of 1952 Texas and Iowa-Nebraska studies. J Am Med Assoc 1952; 150: 750-60.
41 Hammon WM, Coriell LL, Wehrle PF, Stokes J Jr. Evaluation of Red Cross gamma globulin as a prophylactic agent for poliomyelitis. 4. Final report of results based on clinical diagnoses. J Am Med Assoc 1953; 151: 1272-85.
42 National Advisory Committee for the Evaluation of Gamma Globulin in the Prophylaxis of Poliomyelitis. An evaluation of the efficacy of gamma globulin in the prophylaxis of paralytic poliomyelitis as used in the United States 1953. Public Health Monograph No. 20. Washington, DC: US Government Printing Office; 1954.
43 Burnet FM. Lecture III. The ecological approach to the common virus diseases of today. Bull Johns Hopk Hosp 1951; 88: 157-79.
44 Paul JR. A history of poliomyelitis. New Haven, CT: Yale University Press; 1971.
45 Bruton OC. Agammaglobulinemia. Pediatrics 1952; 9: 722-8.
46 Good RA. Immunoglobulin deficiency syndromes in man. In: Merler E, Ed. Immunoglobulins: biologic aspects and clinical uses. Washington, DC: National Academy of Sciences; 1970. p. 191-9.
47 Buckley RH, Schiff RI. The use of immune globulin in immunodeficiency diseases. N Engl J Med 1991; 325: 110-7.
48 Granberry M. Gamma globulin shortage growing acute. Los Angeles Times 1995 Nov 7: Sect. B: 4.
49 Eichenwald K. Committee recommends steps to ease shortage of a drug. The New York Times 1998; Apr 29: Sect. A: 20 (col. 1).
50 Janeway C. The development of clinical uses of immunoglobulins. In: Merler E, Ed. Immunoglobulins: biologic aspects and clinical uses. Washington, DC: National Academy of Sciences; 1970. p. 3-14.
51 Enders JF. Antibodies in human gamma globulin. In: Tullis JL, Ed. Blood cells and plasma proteins: their state in nature. New York: Academic Press; 1953. p. 174-9.
52 Burnet FM, Fenner F. The production of antibodies. 2nd ed. London: Macmillan; 1949.
53 Porter RR. Separation and isolation of fractions of rabbit gamma globulin containing the antibody and antigenic combining sites. Nature 1958; 182: 670-1.
54 Porter RR. The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain. Biochem J 1959; 73: 119-27.
55 Edelman GM. Dissociation of γ-globulin. J Am Chem Soc 1959; 81: 3155-6.
56 Edelman GM, Poulik MD. Studies on structural units of the γ-globulins. J Exp Med 1961; 116: 861-84.
57 A note on the nomenclature of human immunoglobulins. In: Merler E, Ed. Immunoglobulins: biologic aspects and clinical uses. Washington, DC: National Academy of Sciences; 1970. p vi.
58 Kohler RE. The enzyme theory and the origin of biochemistry. Isis 1973; 64: 181-96.
59 Cambrosio A, Keating P. Exquisite specificity: the monoclonal antibody revolution. Oxford: Oxford University Press; 1995.
60 Löwy I. The strength of loose concepts – boundary concepts, federative experimental strategies and disciplinary growth: the case of immunology. Hist Sci 1992; 30: 371-96.
61 Jerne N. Waiting for the end. Cold Spring Harb Symp Quant Biol 1967; 32: 591-603.
62 Cambrosio A, Keating P, Tauber AI, eds. Immunology as a historical object. Special issue. J Hist Biol 1994; 27: 375-594.
63 Drake ME, Barondess JA, Bashe WJ Jr, Henle G, Henley W, Stokes J Jr. Failure of convalescent gamma globulin to protect against homologous serum hepatitis. J Am Med Assoc 1953; 152: 690-3.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Between scepticism and wild enthusiasm: the chequered history of allergen immunotherapy in Britain

Mark Jackson

In October 1986, the Committee on the Safety of Medicines issued a warning in the British medical press about the dangers of using desensitising vaccines for the treatment of allergic disorders. Responding to concerns about the standardisation of allergens and treatment protocols, about the efficacy of treatment, and, more critically, about the risk of death associated with desensitisation or immunotherapy (particularly in patients with asthma), the Committee recommended that desensitising vaccines 'should be used only where facilities for full cardiorespiratory resuscitation are immediately available' and that 'patients should be kept under medical observation for at least two hours after treatment'[1]. As many commentators noted with dismay, the Committee's recommendations and its subsequent guidelines "effectively curtailed" the use of allergen immunotherapy by general practitioners. 'In the absence of a system of hospital-based allergy clinics, allergen immunotherapy in the United Kingdom was effectively abolished overnight' [2].

The intervention of the Committee on the Safety of Medicines in 1986 initiated a series of intense arguments not only between clinical allergists and the Committee but also between different groups of both orthodox and complementary practitioners. In articles and letters in the medical press and in "position papers" issued by major national and international allergy societies, clinicians and scientists vigorously debated the safety and efficacy of immunotherapy in allergy. In this paper, I want to throw some light on the dynamics of this recent debate by reviewing briefly the history of immunotherapy since its introduction in 1911. In doing so, I want to make two points. Firstly, I want to argue that controversy surrounding the use of immunotherapy in allergic diseases is not new. The procedure has been plagued by doubts at a number of levels throughout the century. Secondly, I want to suggest that these controversies have not been entirely determined either by concerns about technical issues (about protocols, testing, standardisation, and so on) or by substantive alterations in the experience of adverse reactions by patients and doctors. In addition to these factors, the variable reception of immunotherapy by the established medical profession may have been influenced by broader professional and socio-economic concerns, notably by the ambiguous professional status of allergists and allergy studies in the United Kingdom.

The origins of immunotherapy

Immunotherapy (also variably referred to as desensitisation, hyposensitisation, prophylactic inoculation, or vaccine therapy) was formally introduced to the British medical

world in two articles in *The Lancet* by Leonard Noon and John Freeman in 1911 [3, 4]. Working together in the Department of Therapeutic Inoculation at St Mary's Hospital in London, Noon and Freeman had attempted to desensitise patients suffering from hay fever by injecting increasing doses of pollen extracts subcutaneously over a period of time. The results of their experiments led Noon to conclude that 'the sensibility of hay fever patients may be decreased, by properly directed dosage, at least a hundredfold' [3].

Noon and Freeman's strategy drew on two distinct traditions. In the first instance, their approach was heavily influenced by developments in bacteriology. The emergence of the germ theory of disease in the last half of the nineteenth century, and in particular the belief that specific diseases had specific causes, had led many investigators (and pharmaceutical companies) to search not only for antitoxins but also for bacterial vaccines to prevent a variety of infectious diseases. Working under the directorship of Almroth Wright, Noon and Freeman were well-acquainted with the development of vaccines designed to prevent and to treat both bacterial and non-infectious diseases [5]. Indeed, Freeman's research had originally concentrated not only on the use of the opsonic index as a test for immunity but also on the development of therapeutic immunisation [6]. Noon and Freeman's research on hay fever constituted a rational extension of Wright's preoccupation with the prophylactic and therapeutic power of the 'immunizator,' bringing 'pollen inoculation in line with the bacterial inoculation work of Wright and his school' [4].

Noon and Freeman's studies were also strongly informed by earlier work on hay fever. In the late nineteenth century, Charles Blackley, a Manchester physician, had employed a detailed microscopic analysis of numerous plants and pollens together with a number of clinical experiments to substantiate John Elliotson's suspicions (voiced in 1831) that hay fever was caused by the 'effluvia of grass and probably the pollen' [7, 8]. Significantly, noticing that farmers rarely suffered from the disorder, Blackley suggested that patients might be 'rendered insusceptible to the action of pollen by continued exposure to its influence' [9, 10].

Although subjected to some criticism by contemporaries [10], Blackley's aetiological explanation was confirmed in 1903 by William Dunbar, an American physician working in Hamburg. Having established pollen (rather than bacteria) as the causative agent, Dunbar suggested that the symptoms of hay fever were produced by the action of a pollen 'poison' or 'toxin' on the tissues of susceptible individuals. Working from this assumption, Dunbar then demonstrated the efficacy of passive immunisation using a 'pollen antitoxin' derived from horses and suggested that such techniques might be effective in the treatment of both hay fever and asthma [11].

In introducing the use of subcutaneous injections of pollen extracts, Noon and Freeman duly acknowledged the pivotal contributions made by Blackley and Dunbar to the study and treatment of hay fever [3, 12]. However, noting the limits of passive immunisation with specific serum such as 'Pollantin,' they combined these earlier studies of hay fever with their bacteriological knowledge and the Inoculation Department's captivation with developing vaccines to suggest that 'on general grounds a much more satisfactory result would be expected from the induction of an active immunity' [3]. In this way, the nascent bacteriological tradition was integrated with research on pollen toxins to generate a novel form of treatment for hay fever.

Although the theoretical basis on which Noon and Freeman had initiated trials of desensitisation was subsequently rejected (as hay fever was rapidly understood to be an allergic reaction of the type described by Clemens von Pirquet in 1906), the treatment of hay fever by inoculations of pollen vaccine was enthusiastically adopted by researchers and clinicians throughout the world. In the process, the techniques and treatment protocols were substantially amended and the range of both specific and non-specific vaccine preparations extended. In addition, in line with evidence suggesting aetiological and pathogenetic similarities between hay fever, asthma, food intolerance, and bee stings, collectively referred to as 'allergic disorders' or 'toxic idiopathies' [12, 13], immunotherapy was used both to treat these conditions and also to desensitise patients against serum sickness associated with the use of antisera. Active immunisation became particularly popular amongst North American allergists, who for many years considered immunotherapy the treatment of choice for hay fever and asthma and who continue to advocate its use in a wide range of conditions [2, 14, 15]. In Britain and some parts of Europe, immunotherapy was received more cautiously. Nevertheless, largely as the result of the continuing interest shown by researchers at St Mary's, desensitisation received extensive coverage in the medical press and (at least until the 1980s) was regarded by many European allergists as 'the cornerstone of allergy practice'[16, 17], and as the most effective treatment for allergic disorders even after the introduction of antihistamines, bronchodilators, and corticosteroids.

Debates about immunotherapy

From the start, both the practice and theory of immunotherapy were plagued by vigorous debates about procedures, efficacy, and safety and about the most appropriate practitioners to diagnose allergic conditions and administer the proper treatment. In the first place, it is clear that even advocates of immunotherapy have frequently disagreed about many practical aspects of the treatment. Shortly after Noon and Freeman's initial communications, B.P. Sormani (a lecturer in serology at the University of Amsterdam) manipulated some of Almroth Wright's principles as well as citing his own experimental experiences both to discredit 'Noon's dosage method' and to suggest an alternative method for preparing the pollen extract to be used in the treatment of hay fever [18]. Sormani's article in *The Lancet* prompted a swift and dismissive letter from John Freeman who, while claiming to be in 'considerable agreement' with Sormani's general advice, nevertheless rejected his particular methodological points emphatically and continued to employ a pragmatic (but vague) approach to dosage based upon clinical examination [19, 20]. Significantly, the question of appropriate dosage has continued to plague allergists throughout the century [21].

Debates about dosage were compounded by differences in the standardisation of specific pollen extracts. For many years, Freeman and his protégés employed the Noon Unit, 'equal to the amount of extract that can be obtained from one-millionth of a gramme of *Phleuem pratense* pollen' [20, 21]. Without entirely condemning the British method, American allergists, such as Robert Cooke, preferred to standardise extracts according to their nitrogen content, an approach that appeared 'to give preparations of equal and regular toxicity' [22]. The advent of commercially prepared allergen extracts did not obviate this problem but may have accentuated it. In 1930, Freeman complained that 'sets of protein

reagents' provided by 'several high-class firms' for diagnostic tests could become denatured and 'give only a poor reaction or none at all' [23]. More recently, the variable standard of commercial allergen extracts for diagnosis and treatment has not only been a source of considerable anxiety amongst allergists [24–26], but also one of the central concerns of the Committee on the Safety of Medicines, which pointed out that a 'confusing number of different units are used to express the allergen content of the products currently marketed' and warned that the 'absence of a standard unit means that products containing the same allergens are not interchangeable' [1].

The absence of standardised allergen extracts for diagnostic purposes (as well as treatment) coincided with inconsistencies in the diagnostic process itself. While most authors have routinely agreed on the need for the accurate identification of specific allergic sensitivity, in practice a variety of diagnostic tests have been employed by various researchers. According to Freeman, surveying the field in 1930, diagnosis could be achieved by applying allergens to the eye, the lips, and the skin (either by a scratch or prick test or by intradermal injection), or by demonstrating the presence of specific antibodies by passive transfer, a procedure introduced by Carl Prausnitz and Heinz Küstner in 1921 [27, 28]. The relative merits of these diagnostic procedures, and the precise clinical implications of the results, have only rarely been assessed, and when they have been considered their clinical value has frequently been questioned [29]. Although Freeman advocated skin tests both as a diagnostic tool and as a means of monitoring the progress and efficacy of treatment, he acknowledged that they had 'had too much medicinal virtue ascribed to them; in their clinical use they have almost developed into one of the pathological rituals designed towards magical healing' [27]. Significantly, the identification of IgE as the antibody responsible for many allergic reactions in the 1960s, and the subsequent development of diagnostic tests measuring specific IgE levels, have failed to resolve concerns about the correlation between clinical history, the results of diagnostic tests, and the efficacy of treatment.

Inconsistencies in diagnostic procedures and allergen preparations have been mirrored by vast differences in treatment protocols and by the absence of standard tests of efficacy. When Noon and Freeman first introduced prophylactic inoculations for hay fever in 1911, they employed what Freeman later referred to as 'leisurely desensitisation,' that is, a series of injections spread over a few weeks or months prior to the pollen season [30]. A few years later, Freeman adopted a procedure referred to as 'intensive desensitisation,' comprising daily injections of gradually increasing doses, particularly for patients sensitive to animals. Significantly, Freeman offered no conceptual rationale for this modification, pointing out simply that he 'fell into the way of inoculating these patients every day' because such patients 'were usually in a great hurry to go and hunt, or look after their dogs, or retrieve their cats from quarantine with the veterinary surgeon' [30]. The apparent success of intensive courses of immunotherapy in animal allergies encouraged Freeman to extend this approach to hay fever sufferers.

Freeman's sensitivity to the dictates of his patients is further evidenced both in his elaboration of an even more rapid treatment protocol, referred to as 'rush desensitisation' [30] and by his support for self-inoculation. Although Freeman indicated that the idea of speeding up the treatment process had been suggested by Besredka's work on anti-anaphylaxis, the advantages of rush desensitisation were clearly located in the practical interests

of the patient (and indeed the doctor) rather than in any concern for conceptual elegance [30]. More critically, the development of self-inoculation programmes for appropriately selected patients [12, 31] was driven both by concerns about the cumbersome protocol demanded by leisurely desensitisation and by the interests of pharmaceutical firms such as Parke, Davis and Co., which financed much of Freeman's work at St Mary's and produced self-inoculation kits for purchase and distribution by general practitioners [31, 32].

Interestingly, Freeman did not insist that it was necessary to 'adhere rigidly to the leisurely, intensive, or "rush" methods,' but accepted that 'an intelligent blend may serve one's turn better' [30]. Clinicians treating allergies followed Freeman's advice, devising their own protocols and preparing their own allergen extracts to suit the demands of their time and patients. Freeman's flexibility and his professed preference for clinical experience over statistics in judging dosage and efficacy may have suited the idiosyncrasies of his own character and the particular demands of the Inoculation Department at St Mary's in the inter-war years. However, it was, ironically, this flexibility that became a burden to the next generation of allergists. The multiplicity of approaches to immunotherapy made comparison of results extremely difficult. As the need for establishing the therapeutic efficacy and safety of pharmaceutical products became more urgent in the post-war period, allergists who failed to standardise their extracts or to perform randomised control trials were accused increasingly of relying on anecdotal, rather than scientific, evidence. Although the first controlled trial of immunotherapy was carried out by Frankland and Augustin in the Department of Allergic Disorders at St Mary's in 1954 [33], the persistent reluctance of allergists to test their theories and practices in this way became a source of increasing anxiety amongst both proponents and critics of allergen immunotherapy concerned about the efficacy of this form of treatment [1, 16, 25, 34].

Wide variations in diagnostic and treatment programmes highlighted, and in part stemmed from, a further fatal flaw in the immunotherapists' position, namely the absence of a coherent explanation of precisely how desensitisation might work. Noon and Freeman's understanding of the pathogenesis of hay fever (as the product of a toxin) was soon discarded and replaced by explanations that aligned both hay fever and a number of related complaints with the mechanism of experimental anaphylaxis [35, 36]. Hay fever, asthma, urticaria, and eczema were increasingly construed as allergic disorders in which the clinical symptoms and tissue damage were thought to be caused by a reaction between allergen and tissue-fixed antibody (or reagin) with the subsequent release of inflammatory mediators such as histamine [37].

While more sophisticated explanations of the pathogenesis of allergic disorders gave the emerging field of allergy some degree of intellectual and experimental focus, their incompatibility with the toxin hypothesis proved a persistent irritation for allergists attempting to establish the immunological mechanism of allergen immunotherapy. One of the first, and the most enduring, explanations of the mechanism of desensitisation was provided in 1935 by Robert Cooke and his colleagues in New York, who suggested that injections of allergen could promote the production of 'a peculiar blocking or inhibiting type of immune antibody that prevented the action of allergen on the sensitizing antibody' [38]. This promising hypothesis was subsequently explored by many allergists but it remained (and still remains) contested. In particular, reports of the lack of correlation between the levels of 'blocking antibody' (later identified as IgG) and clinical improvement prompted

allergists to postulate a variety of alternative (or additional) mechanisms for desensitisation [39–42]. The absence of a clear explanation of the mode of action of immunotherapy, together with the plurality of approaches to its execution, resulted in extensive criticism of this method of treatment for allergic disorders as being entirely empirical and arbitrary [41, 43].

Since 1911, proponents of prophylactic desensitisation have also been plagued by fears about its safety, especially in asthmatics. In 1915, Robert Cooke warned that 'liberal use' of desensitisation in patients with hay-asthma 'could conceivably induce death by anaphylactic shock' [22]. Cooke's caution was repeated regularly in the British medical press throughout the next few decades and was increasingly accompanied by case reports of serious adverse reactions and death during treatment and by suggestions for preventing fatalities, such as the inclusion of adrenalin in the syringe. In 1933, for example, David Harley warned that desensitisation was difficult and that severe reactions and fatalities had occurred [44]. The following year, an editorial in *The Lancet* highlighted the tendency of asthmatics in particular to suffer potentially fatal attacks of asthma during desensitisation [45]. In 1942, at a meeting of the Association of Clinical Pathologists, D.N. Nabarro described 'alarming and almost fatal anaphylaxis' after intradermal injection of mixed antigen into an asthmatic patient, a reaction that was successfully treated only after multiple doses of adrenalin and oxygen inhalation [46]. And in 1954, the death of a patient undergoing desensitisation for asthma at Guy's Hospital prompted both an informal coroner's inquiry and discussion in parliament [47].

Concerns about the safety of immunotherapy were linked to questions about who was qualified to vaccinate against allergic disorders. In 1914, Freeman had warned of the problems that could be caused by inexperienced local doctors continuing the treatment that had been started in his clinics [48]. Some years later, David Harley, one of Freeman's students at St. Mary's, pointedly referred to 'advanced allergists' producing 'brilliant cures' in certain cases [49]. In drawing a boundary between themselves and other practitioners, Freeman and Harley were asserting their jurisdictional rights over the diagnosis and treatment of patients with allergic disorders. In addition to establishing their own specialty, however, Freeman and Harley may also have been protecting their economic interests. The treatment of allergic disorders was becoming an increasingly profitable business both for pharmaceutical companies producing diagnostic sets and vaccine kits (and later a vast array of anti-allergy drugs) and for clinical allergists, such as Freeman, who had extensive private practices [6]. The extent to which the prophylactic treatment of allergies was influenced by commercial considerations did not pass unnoticed. In 1938, A.J. Cronin allowed Andrew Manson, the central character in his novel 'The Citadel', to desensitise a rich patient even though he considered the procedure to be useless [50].

For much of this century, advocates of allergen immunotherapy have been plagued by concerns about the standardisation of allergens, about diagnostic and treatment protocols, about efficacy and safety, about the supposed mechanism of action of desensitisation, and about the morality of links with the pharmaceutical business. Their failure to resolve these concerns may in part explain why allergy has been construed by some as a "Cinderella subject", occupying an ambiguous and uncomfortable space between orthodox and alternative medicine, a situation perhaps exacerbated by the fact that membership of the British Society for Allergy and Clinical Immunology has not been limited to clinicians, unlike

many other specialist professional medical organisations established in the middle decades of this century.

Significantly, the taint of the alternative has been accentuated by other factors. In the first place, allergists' classification of a wide array of non-specific symptoms and diseases as allergic disorders, and their persistent claims to be able to diagnose those conditions with skin tests and treat them with immunotherapy, have prompted suspicions of quackery [23]. More recently, these suspicions have been raised further by the elaboration of bizarre and apparently unsubstantiated diagnostic tests and treatment protocols by some clinical allergists (such as enzyme-potentiated desensitisation), and by the emergence of the 'total allergy syndrome' or "twentieth century disease," a phenomenon that resonates with earlier constructions of hay fever as a consequence of "higher civilization" [10, 11]. These developments have damaged the reputation of clinical allergists (at least in the eyes of orthodox practitioners) by linking them to the radical anti-modernist stance of alternative approaches to disease and the environment, exemplified by clinical ecology.

The fall of immunotherapy

It is tempting to read the intervention of the Committee on the Safety of Medicines in 1986 entirely as the product of accumulating concerns about technical aspects of immunotherapy, that is about protocols, standardisation, safety, and efficacy. There is certainly evidence to justify this interpretation. Like the update published by the Committee, most papers commenting on immunotherapy throughout the 1980s rehearsed many of the technical debates that have preoccupied clinical allergists since the introduction of prophylactic desensitisation in 1911. However, it is likely that the Committee's intervention was not determined solely by technical issues but also by contingent concerns about the professional status of allergists and by anxieties about competition between alternative and orthodox practitioners.

In 1980, a general practitioner from Southampton reported the death of a patient undergoing desensitisation for hay fever and asthma [51]. Three weeks later, in a letter to the *British Medical Journal*, Pamela Ewan responded to the report by suggesting that impressions of benefit from immunotherapy were 'based on anecdote' and that desensitisation was 'potentially dangerous and often ineffective' [52]. Ewan's disparaging remarks were immediately rebutted by Bill Frankland, an enthusiastic advocate of immunotherapy, who asserted that there was 'no doubt' that 'specific immunotherapy does give benefit' and that it was inappropriate to 'damn all hyposensitisation injections as dangerous' [53].

The tension evident in this exchange was expressed more overtly elsewhere. In 1986, prior to the Committee's intervention, *Clinical Allergy* solicited the opinions of physicians with differing views on desensitisation in the treatment of asthma [24]. In addition to rehearsing prominent reservations about efficacy and safety, I.W.B. Grant, a chest physician from Edinburgh, comprehensively maligned the character of many clinical allergists. 'Senior physicians,' he wrote, 'may recall with cynicism the financially profitable cult of hyposensitisation with haphazard mixtures of numerous allergens practised by unscrupulous self-styled allergists in the years before β_2-agonists, sodium cromoglycate and corticosteroid aerosols became available' [24]. In contrast, although they acknowledged the need for 'more systematic and controlled studies,' H. Mosbech and B. Weeke, from the Allergy Unit at the State University Hospital in Copenhagen, carefully distanced them-

selves from the 'mysticism, scepticism or even wild enthusiasm' associated with immunotherapy by soberly surveying the evidence in support of the procedure [24].

This debate prompted several letters from allergists criticising Grant's cynicism and pointing out the disadvantages of modern drug treatments [24]. However, the correspondence illustrates the extent to which allergists acknowledged the potency of Grant's scepticism and were concerned about their reputation. As S. Dreborg warned, immunotherapy 'cannot be undertaken in an uncritical, homeopathic fashion by unenlightened enthusiasts, but in an ordered supervised way by specialists, based on strict diagnostic criteria and with proper supervision of treatment' [24].

While such disputes about the professional status of allergists were clearly not new, they were given a particular impetus in the 1980s both by the re-emergence of market forces in medicine, which encouraged competition rather than collaboration between doctors, and by the realisation amongst orthodox practitioners that patients were increasingly consulting alternative practitioners for the treatment of allergies [25]. The recommendations of the Committee on the Safety of Medicines should be understood therefore not only as a response to long-standing concerns about efficacy and safety but also as a contribution to a highly-charged and specific debate about the legitimacy and professional status of clinical allergy, and about the nature of modern medicine.

In recent years, allergists have responded vigorously to the Committee's restrictions. In a stream of position papers and articles, allergists in Britain and elsewhere have attempted to counter prejudices against immunotherapy and to rehabilitate clinical allergy within the realm of modern scientific medicine by distancing themselves from 'unconventional and unproved forms of "allergy treatment"' [2], by stressing that diagnosis and treatment should only be performed by experienced allergists, by conducting extensive controlled trials [54], by reasserting the value of immunotherapy in the treatment of allergies to stinging insects, and by emphasising the need to balance 'science, theory and current practice' when assessing the validity of immunotherapy [17, 55–57]. It remains to be seen whether clinical allergists will succeed in their endeavours to establish the scientific basis of the traditional cornerstone of their practice.

Bibliography

1 CSM Update: desensitising vaccines. Br Med J 1986; 293: 948.
2 Frew AJ. Injection immunotherapy: British Society for Allergy and Clinical Immunology Working Party. Br Med J 1993; 307: 919-23.
3 Noon L. Prophylactic inoculation against hay fever. Lancet 1911; 10 June: 1572-3.
4 Freeman J. Further observations on the treatment of hay fever by hypodermic inoculations of pollen vaccine. Lancet 1911; 16 September: 814-7.
5 Worboys M. Vaccine therapy and laboratory medicine in Edwardian Britain. In: Pickstone JV, Ed. Innovations in medicine. Basingstoke: Macmillan; 1992. pp. 84-103.
6 Obituary: John Freeman. Lancet 1962; 27 January: 224-5.
7 Elliotson J. London Medical Gazette 1831; 8: 411-3.
8 Elliotson J. Hay fever. Lancet 1830-1831; 370-3.
9 Blackley CH. Experimental researches on the causes and nature of *Catarrhus Aestivus* (hay fever or hay asthma). London: Balliére, Tindall and Cox; 1873.
10 Waite K. Blackley and the development of hay fever as a disease of civilization in the nineteenth century. Med Hist 1995; 39: 186-96.

11 Dunbar W. Etiology and specific therapy of hay fever. Ann Otol Rhinol Laryngol 1903; 12. Reprinted in Ann Allergy 1962; 20: 752-64.
12 Freeman J. Hay-fever: a key to the allergic disorders. London: William Heinemann; 1950.
13 Freeman J. An address on toxic idiopathies. Lancet 1920; 31 July: 229-35.
14 Anon. The treatment of hay fever. Lancet 1922; 23 September: 678-9
15 Anon. Prevention of hay fever. Lancet 1929; 2 February: 248.
16 Aas K. Adequate clinical trials of immunotherapy. Allergy 1982; 37: 1-14.
17 Malling HJ, Weeke B, Eds. European Academy of Allergy and Clinical Immunology position paper: immunotherapy. Allergy 1993; 48 (Suppl 14): 1-35.
18 Sormani BP. Prophylactic vaccination against hay fever. Lancet 1916; 16 February: 348-50.
19 Freeman J. Prophylactic vaccination against hay fever. Lancet 1916; 4 March: 532.
20 Freeman J. Treatment of hay-fever. Lancet 1927; 30 April: 940-1.
21 Frankland AW. High and low dosage pollen extract treatment in summer hay fever and asthma. Acta Allergologica 1955; 9: 183-7.
22 Cooke RA. The treatment of hay fever by active immunization. Laryngoscope 1915; 25: 108-12.
23 Freeman J. The significance of idiopathic skin reactions. Lancet 1930; 31 May: 1197-9.
24 Does immunotherapy have a role to play in the treatment of asthma? Clin Allergy 1986; 16: 7-16, 179-80, 269-77.
25 Royal College of Physicians. Allergy: Conventional and Alternative Concepts. London: Royal College of Physicians; 1992.
26 Dreborg S, Frew A, Eds. European Academy of Allergology and Clinical Immunology position paper: allergen standardization and skin tests. Allergy 1993; 48 (Suppl.) : 14.
27 Freeman J. The significance of idiopathic skin reactions. Lancet 1930; 24 May: 1141-2.
28 Cohen SG, Samter M, Eds. Excerpts from classics in allergy. California: Symposia Foundation; 1992.
29 Nelson T, Porter A. Protein in asthma. Lancet 1931; 19 December: 1342-4.
30 Freeman J. 'Rush' inoculation, with special reference to hay-fever treatment. Lancet 1930; 5 April: 744-7.
31 Parke, Davis and Co. Vaccine and serum therapy. London: Parke, Davis and Co; 1935.
32 Anon. Hay fever reaction outfit. Lancet 1912; 25 May: 1448.
33 Frankland AW, Augustin R. Prophylaxis of summer hay-fever: a controlled trial comparing crude grass-pollen with the isolated main protein component. Lancet 1954; 22 May: 1055-7.
34 Editorial: a reevaluation of immunotherapy for asthma. Am Rev Respir Dis 1984; 129: 657-9.
35 Meltzer SJ. Bronchial asthma as a phenomenon of anaphylaxis. J Am Med Assoc 1910; 55: 1021-4.
36 Silverstein A. A history of immunology. San Diego: Academic Press; 1989.
37 Harley D. Hay fever: its immunological mechanism, diagnosis and treatment. Br Med J 1935; 13 April : 754-6.
38 Cooke R, Barnard JH, Hebald S, Stull A. Serological evidence of immunity with coexisting sensitization in a type of human allergy (hay fever). J Exper Med 1935; 62: 733-51.
39 Fitzgerald JD, Sherman WB. The specificity of blocking antibody induced by grass pollen extracts. J Allergy 1949; 20: 286-91.
40 Burdon KL. On possible mechanisms of hyposensitization: some pertinent laboratory findings. Ann Allergy 1967; 25: 483-95.
41 Editorial: hayfever. Lancet 1975; 5 April: 786-7.
42 Current status of allergen immunotherapy (hyposensitization): memorandum from a WHO/IUIS meeting. Bull WHO 1989; 67: 263-72.
43 Lancet 1973; 15 December: 1364.
44 Harley D. Reactions of the human skin to foreign sera. Lancet 1933; 1 April: 690-2.
45 Prophylaxis and anaphylaxis. Lancet 1934; 13 October: 817-8.
46 Lancet 1942; 21 March: 354.
47 Lancet 1954; 24 July: 188.
48 Freeman J. Vaccination against hay fever: report of results during the last three years. Lancet 1914; 25 April: 1178-80.
49 Harley D. Asthma: immunological mechanism, diagnosis, and treatment. Lancet 1936; 15 August: 367-70.
50 Cronin AJ. The citadel. London: Victor Gollanz; 1938.
51 Rands D. Anaphylactic reaction to desensitisation for allergic rhinitis and asthma. Br Med J 1980; 281: 854.
52 Ewan PW. Anaphylactic reaction to desensitisation. Br Med J 1980; 281: 1069.
53 Frankland AW. Anaphylactic reaction to desensitisation. Br Med J 1980; 281: 1429.

54 Varney VA, Gaga M, Frew AJ, Kay AB, Durham SR. Usefulness of immunotherapy in patients with severe summer hay fever uncontrolled by antiallergic drugs. Br Med J 1991; 302: 265-9.
55 Malling HJ, Ed. Immunotherapy: position paper. Allergy 1988; 43 (Suppl 6): 1-33.
56 Norman PS, Van Metre TE. The safety of allergenic immunotherapy. J Allergy Clin Immunol 1990; 85: 522-5.
57 Bousquet J, Hejjaoui A, Michel FB. Specific immunotherapy in asthma. J Allergy Clin Immunol 1990; 86: 292-305.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Vaccine viability: the shifting fortunes of hepatitis B immunization

Jennifer Stanton

Introduction: the broad background and key questions

When we think of major changes in illness over the past century, we generally see a reduction in the burden of infectious disease and a rise in chronic disease. Translated into causes of mortality, this shows as a decrease in infant diarrhoea, for example, and a rise in cancer and heart disease. Vaccination has played an important part in the decline of infectious diseases, although we know that the rise in the standard of living (in which we include nutrition, housing, and more public aspects like sanitation) also played a part. Clearly, if we take the case of smallpox, despite all the historic controversies, there has been some sort of triumph, since the disease has been eradicated worldwide. Whether we define the role of vaccination as a triumph of "medicine" or a triumph of "public health" depends on where we place the emphasis: on the ability of the vaccine to render an individual immune to a disease, or the ability of societies to deliver the vaccine to enough individuals to make the disease of negligible importance.

We are also aware of a distinction between "developed" countries mainly in the northern hemisphere, and "underdeveloped" countries which we think of as lagging behind, with a heavy burden of infectious disease still taking its toll especially on young children. Since the 1970s, the WHO/UNICEF policy of promoting immunization for six diseases that are big killers of children has been credited with saving millions of lives. It has also been criticized as a 'technical fix' imposed from above which does nothing to rescue children and their parents from a cycle of hunger, poverty, and disease [1]. In rich countries, in contrast, childhood immunizations were an adjunct of an all-embracing web of welfare provisions operating in an environment of relative abundance.

How does hepatitis B fit into this overall picture of Western success and Third World underdevelopment? This paper gives a fairly detailed account of hepatitis B vaccination in the United Kingdom, as an exemplar of policy in a wealthy country, and then moves on to consider the 'Third World' pattern. As a preliminary, however, it is necessary to sketch in the position of hepatitis B before a vaccine became available: what sort of understanding of the disease was held by medical/public health 'authorities'; where and to what extent was it seen as a problem?

Hepatitis B was in some ways a 'new' postwar disease in the sense that a distinction between two forms of viral hepatitis, A and B, was only established during and after the Second World War [2]. In Europe and America during the 1950s and 1960s, hepatitis had a

limited profile even within the medical profession: apart from experts in liver disease, few doctors knew much about it. Among British doctors interviewed by the author, several had heard of outbreaks during the war, connected with yellow fever vaccination; and several recalled that in the 1950s, they regarded hepatitis B as a disease that was prevalent among drug users [3]. One informant, a senior surgeon, recalled that in the early days of his career, hepatitis B was thought of as a 'dirty' disease associated with drug injecting, so that surgeons used to be nervous of catching it from patients who were suspected of being drug addicts [4]. In the blood transfusion service, hepatitis had a much higher profile as the most worrying side-effect of transfusion [5]. Then in the late 1960s, it appeared in a new setting: outbreaks of viral hepatitis caused havoc in kidney dialysis units throughout the United Kingdom, Europe, and America [6].

Following an 'accidental' discovery of the hepatitis B antigen by American biochemist and geneticist Baruch Blumberg, a test for hepatitis B was developed around 1968, with two very different effects [7]. The test enabled some countries to achieve near-elimination of hepatitis B from the blood supply. (This revealed the existence of other types of blood-borne viral hepatitis, initially known as 'non-A non-B'; the major hazard was later defined as hepatitis C.) The test also made epidemiological studies possible, shifting the perception of the disease, and providing a new context for the vaccine whilst it was being developed. These prevalence studies revealed that "acute" hepatitis B – the only form readily accessible to the medical 'gaze' up to this point – was merely the tip of the iceberg. Below the surface were much larger numbers of long-term carriers, with more or less "chronic" disease and more or less continued infectivity [8]. Even in countries with low prevalence, such as Britain, perhaps one in a thousand of the population were infective carriers of hepatitis B (although experts argued over carrier prevalence), while a far greater proportion of the population, having been exposed at some time, had developed immunity.

The carrier problem had a personal and a public health edge [9]. For the individual carrier, there was a chance of the chronic infection leading to severe liver damage (cirrhosis or liver cancer) two or three decades after initial infection. For the wider public there was the risk that carriers, often unaware of their status, might infect other people through any of the sorts of contact we are now familiar with thanks to AIDS – from sex to drugs to blood. In policy terms, an important difference was that health workers were seen from the outset as a risk group for hepatitis B. And although it is not so fatal, hepatitis B is much more infectious than HIV. A small proportion of people thus infected would develop acute hepatitis, for which there was (and still is) no cure, and which has a high case fatality rate.

Poor countries had a different public health problem with hepatitis B. Not surprisingly (although it had not been observable prior to the test), there was a much higher prevalence, with somewhere around 10 or 15% of the population in many Asian and African countries being carriers. A very high proportion of the population encountered the infection as newborns (maternal transmission) or infants, at which age there was little acute disease: the majority developed immunity. Since nearly everyone was exposed to the disease, transmission between adults was a lesser problem. However, early childhood infection resulted in a huge amount of liver disease among adult carriers. Primary liver cancer, rare in the West, was one of the commonest kinds of cancer in Asia and Africa, killing adults in the

economically productive age-band. Estimates as high as 200 to 300 million carriers placed hepatitis B ahead of AIDS in terms of global pandemic through the 1980s and into the 1990s [10].

To summarize thus far, a series of new scenarios for hepatitis B were created by medical science in the postwar period:

1) discriminating between previously unitary disease entities (hepatitis A and B, later C and D or δ);

2) creating new contexts for transmission, for example, blood transfusion, inoculations, dialysis;

3) revealing 'carriers' and new 'natural histories', especially a link with liver cancer;

4) offering hope of a vaccine.

The pattern of hepatitis B revealed by the test differed between developed countries and underdeveloped countries. In developed countries like Britain, with low prevalence, transmission between adults in specific settings or through particular behaviours was seen as the main public health problem. In poor countries, with high prevalence and transmission in infancy, long-term mortality from liver cancer was the main problem, but one that was scarcely recognised by the health authorities in such countries before the 1980s.

The advent of the vaccine, after trials in 1980, offered a means of prevention [11]. Against this background, we can frame questions about the viability of the vaccine. Did the vaccine solve the public health problems associated with hepatitis B? This will be illustrated here principally through the history of hepatitis B in the United Kingdom. How did patterns of uptake differ in high and low prevalence countries, and how much was this due to government health policies or other factors? How far does the viability of the vaccine depend on planned policy that follows from scientific experts giving health officials accurate information? The shortcomings of this 'rational' model of policy-making, in which science feeds logically into policy, will be discussed in the conclusion.

Hepatitis B vaccine in the UK

Despite intense debate in the pages of the *British Medical Journal* during the 1980s over wider use of hepatitis B vaccine for doctors and nurses – a debate in which cost featured prominently [12] – there were few cost-benefit analyses. Perhaps the most comprehensive cost-benefit study for any part of the United Kingdom during this period was a 1988 public health dissertation, which analysed costs and benefits of immunizing health care workers in Northern Ireland [13]. On first reading, two things were striking about this document. One was the recommendation "not" to vaccinate health care workers, on the grounds that it was not sufficiently cost-effective (compared, for example, to hip replacement operations). Instead, the writer argued for further improvements in hygienic precautions such as careful handling and disposal of 'sharps'. Sharp injuries – skin puncture with used needles or blades – were recognised as a major route of infection for health workers. But the hygiene precautions were so multifarious, so arcane in many ways (to the uninitiated reader), that it might appear simpler to offer these workers the vaccine that would safeguard them against hepatitis B. The catch – the factor counting against the simple option – was the high cost of the vaccine, at this time around £60 for a course of three shots.

The second thing that was striking was the extremely low prevalence and incidence of hepatitis B in Northern Ireland. This was the epidemiological 'fact' on which the whole

calculation of cost-benefit rested, but it was a puzzle. According to oral information, the politics of the 'Irish Question' entered here: due to vigilante control of petty crime by sectarian paramilitaries, there was a low rate of intravenous drug abuse in Northern Ireland, compared with other regions of the United Kingdom and with the Republic in the south. Since the use of shared needles by drug takers was a major route of transmission for hepatitis B, this control contributed to the low prevalence of the disease in 'the province'. Low overall prevalence meant low prevalence also among health workers, whose rates tended to be above average [14].

A number of points arise from this story. On a methodological issue, oral sources are a useful, though necessarily somewhat haphazard, supplement to the written record (mainly published papers) for this very recent or contemporary history. In terms of policy history, the primary focus was on health workers, despite epidemiological evidence that the highest rates of hepatitis B were among drug users and homosexual men. This was equally the case in England and Wales, as well as Northern Ireland. Why did policy fail to follow epidemiology – why did it focus on health workers rather than drug users? I would argue there were two main reasons. There is the managerial aspect: vaccine policy was channeled within the National Health Service (NHS), which held responsibility as the employer for health workers. There is also a personal aspect, which should not be ignored. Those most concerned with vaccine policy were medical civil servants in the Department of Health at the centre, and public health doctors, virologists, and infection control specialists advising regional health authorities and hospitals at local level. They probably felt especial sympathy for doctors and other health staff infected in the course of their duties.

A third point illustrated by the Northern Ireland story is that hygiene measures were, to a large extent, the currency of debate between policy makers and health and allied workers in the United Kingdom, from the introduction of a test for hepatitis B, that is, around 1970, until the early 1990s. Elsewhere I have argued that health workers and the Department of Health and Social Security (DHSS) worked out a tacit agreement in the 1970s (before the vaccine). Health workers were not subject to compulsory testing for hepatitis B, so long as they undertook stringent safety precautions when handling blood or, in the case of surgeons and dentists, operating on patients [15]. Compulsory testing might have resulted in hundreds of health staff being laid off, with compensation and recruitment implications. In addition, health staff might have demanded that patients be tested too, since they were probably the source of infection in the first place. Thus, screening of health workers would carry the risk of substantial financial consequences which the DHSS was unwilling to incur.

Thus when the vaccine became available in 1982 in the United Kingdom, the DHSS was in a cleft stick. On the one hand, here was the solution to the dilemma: a safe preventive measure that could clear up the public health hazard in the arena of most direct concern, the NHS. On the other hand, compulsory vaccination for hepatitis B would act as a surrogate test, since failure to respond to the vaccine (which involved three doses and a check on whether it had 'taken') would indicate possible carrier status. This could reopen the whole set of issues thrown up in the 1970s by the test. Moreover, the vaccine was extremely expensive, which may have been the prime reason the government was not keen on promoting widespread uptake [16].

Besides, from the 'consumer' viewpoint, this vaccine had a serious image problem. Its raw material was human serum containing large amounts of hepatitis B antigen – that is, serum from carriers' blood – and there rapidly grew up a folklore that it might be contaminated with HIV. The manufacturers had the misfortune to market the product in the United Kingdom around the time that news of the AIDS epidemic broke in the mass media. Health workers reasoned that hepatitis B carriers might also be AIDS carriers; there is some evidence that they boycotted the vaccine on these grounds even when it was offered to them (which it often was not) [17]. It took an enormous amount of campaigning effort by vaccine manufacturers Merck Sharp and Dohme (MSD) to convert even a fraction of the potential consumer market among health care workers to the view that the vaccine was completely safe [18]. The health workers union ASTMS voiced deep anxieties over the vaccine, having consistently thrown their full weight behind the 'precautionary' approach to handling samples possibly contaminated with hepatitis B in the laboratory [19].

There were, however, strong allies in favour of the vaccine. Many, though not all, of the individual experts involved in hepatitis research, together with bodies such as the Blood Products Laboratory and regional transfusion laboratories, favoured wider promotion of the vaccine. Among professional groups, dentists were most enthusiastic. Health and safety officers of the Royal College of Nursing rapidly called for the vaccine to be made available to more of their members, and supported meetings and conferences designed to convince the broader membership as well as the DHSS to expand vaccine availability [20]. Doctors seem to have offered a fragmented response, in which anxieties over the safety of the vaccine were sometimes evident, alongside resistance to being screened for the virus.

Throughout the 1980s we can trace the growth of an alliance of experts and professional groups pressing for broader guidelines on hepatitis B vaccine use. Debates continued, but it seems that resistance 'from below' was a lesser constraint compared with the restrictive guidelines from the DHSS. These guidelines focussed on the health care setting where, as we have seen, they aroused controversy; but they avoided 'lifestyle' groups outside the health care setting. Thus, we see the paradox of a rising rate of acute hepatitis B in the United Kingdom (England and Wales) during the 1980s, when the vaccine was theoretically available. Incidence nearly doubled from 1,000 to almost 2,000 per annum between 1980 and 1984 [21]. The rise probably occurred chiefly among intravenous drug users [22]. For every acute case, there would probably have been dozens of subclinical cases, some resulting in chronic infection, mainly unnoticed for decades. Notifications of acute hepatitis B decreased from the mid-1980s. This could be due to a number of causes, including patients not coming forward for fear of being diagnosed as HIV-positive; but it seems most likely to be associated with safe sex among homosexual men, and safer practices among drug users, in connection with AIDS education [23]. Hepatitis B vaccine was not at this time reaching significant numbers of homosexual men or drug users.

In the late 1980s a number of changes occurred that altered the picture. Hepatitis transmission among drug users decreased, when needle exchange schemes for drug users were initiated to reduce the risk of HIV spread. There was technological innovation: Smith Kline Beecham introduced a new genetically engineered vaccine, slightly cheaper than the earlier vaccine and crucially with a 'cleaner' image, more acceptable to the groups that had reservations about the serum-based vaccine. In 1988, the DHSS issued expanded guide-

lines on use of hepatitis B vaccine, this time including 'lifestyle' groups such as homosexual men and drug users, although again these were not immediately offered the vaccine [24] Indeed, a 1990 study showed that two-thirds of drug addiction treatment centres did not offer screening or vaccination for hepatitis B [25]

Lastly but perhaps with most impact on policy, Crown immunity was removed from NHS properties and personnel as part of a wider effort to make the law more even-handed As never before, it was now possible to hold hospitals and doctors liable for 'accidents' such as hepatitis B transmission A study from the Communicable Disease Surveillance Centre revealed a longstanding problem of transmission to patients [26] In place of the previous tendency to hush up such incidents, from 1991 on there was greater openness Whenever practitioner-to-patient transmission of hepatitis B became evident, patients were notified and called in for testing, and the media were involved in the process [27] Vaccine was rapidly made available, and strongly promoted, among groups such as nurses in mental handicap institutions In 1994, a surgeon who had hidden his carrier status was prosecuted and jailed [28] Medical students were to be subjected to compulsory hepatitis B testing and vaccination, with carriers barred from entry to medical school, but this move aroused strong opposition among the medical profession [29] There was similar opposition to a move to screen all hospital doctors [30] We can see a definite move towards active targeting of health professionals for both testing and vaccination in the 1990s

Clearly, the broadening hepatitis B vaccine policy of the late 1980s and early 1990s reached out to some groups more energetically than others Pregnant women were actively targeted leaflets handed out at ante-natal clinics strongly advised vaccination for babies of carrier mothers, and urged such mothers to carry out hygiene precautions such as burning sanitary towels [31] The Confederation of Health Service Employees [COHSE] produced their own leaflet in the late 1980s, giving very full coverage to safety measures [32] The vaccine was mentioned as effective but not widely available to health staff The Terrence Higgins Trust and Group B, a self-help homosexual men's group, produced a leaflet advising homosexual men to seek hepatitis B vaccination, while also describing safe sex techniques [33] More vaccine was reaching more clients now but in a fragmented way

In 1992 the Department of Health commissioned research into the costs and benefits of different hepatitis B vaccine policies that it might pursue in the future Options ranged from continuation of the current 'targeted' approach, to expansion toward universal childhood vaccination at about age 12 years, a policy already instituted in countries like Italy with a slightly higher prevalence [34] This 'universal adolescent' vaccination policy was publicly heralded as the coming thing as early as 1991, but was not introduced in the United Kingdom for many years [35]

In the course of my own research in the 1990s, I witnessed a shift in the fortunes of hepatitis B immunization in this country, with the promise of more to come To a greater extent than was apparent in the local, United Kingdom sources, these shifts were part of a global picture, in which the Europe regional office of WHO was influenced by worldwide developments Influenced, but not determined, apart from any other factors, the epidemiology of hepatitis B was different in most Third World countries from affluent countries such as the United Kingdom, as indicated in the earlier part of this paper [36]

Hepatitis B immunization in Asia and Africa

The link between hepatitis B and liver cancer is a recent construct that arose from the epidemiological studies of the 1970s, combined with work on the natural history of the disease and immune responses Colonial doctors in earlier periods allocated blame for the high rate of liver cancer in African countries to use of indigenous herbal medicine, or to the dietary deficiencies that also produced kwashiorkor in infancy The high incidence of this rapidly fatal cancer in parts of Africa and Asia became linked with hepatitis B once the antigen test allowed epidemiological studies to be undertaken It would seem logical that immunization on a wide scale would follow as soon as a vaccine for hepatitis B became available

China, Taiwan, and Japan did initiate such policies Japan is of course one of the most developed countries, with Taiwan and China at an intermediate stage, and all three had a very high prevalence of hepatitis B and liver cancer and an awareness of the problem Each initiated local production quite early after the development of the first human-serum vaccine, reducing the cost, and in China costs were partly recouped by charging recipients But most high prevalence countries did not – why not? Why the difference? And how did this shift?

There were two main reasons for a lack of enthusiasm about hepatitis B immunization in the countries which suffered highest prevalence of the disease One was simply cost if cost had been an inhibitor in British vaccine policy, it was a hundred times more so for poor countries Cost reduction is a vital element in the history of promotion of hepatitis B immunization in these countries Its complex recent history is untangled in an account by American historian William Muraskin, detailing technology transfer and local production in Korea and elsewhere [37] Even when technology transfer and local production cut the cost of a course of vaccine from $100 to $1, this still represented an impossibly heavy burden for the poorest countries with annual per capita expenditure on health as low as $3 The spread of the vaccine to sub-Saharan Africa was markedly slower than the Southeast Asian experience

The other reason for a reluctance to prioritize hepatitis B immunization in underdeveloped countries is more subtle, and recalls the complexities of the ways that the disease was seen and handled by different sectors and policy-makers in wealthier countries The epidemiological evidence might indicate a certain course of action but this was only clear when looking at that problem in isolation For health ministers, officials, scientists, and health care workers in Third World countries, there appeared to be an array of more pressing problems There was no established 'liver cancer' lobby as there was, for example, an established maternal and child health lobby Reflecting this, there was a lack of urgency at the international centre, the World Health Organization The WHO Expert Advisory Panel on Virus Diseases recognized that universal immunization against hepatitis B at birth would be the ideal aim for high prevalence countries, but this policy did not receive official WHO endorsement until 1992

In Muraskin's account, all these obstacles – indifference at WHO, governmental reluctance, the high cost of the vaccine – were tackled by the International Task Force on Hepatitis B Immunization set up in 1986 This US-based group, a sort of scientific Magnificent Seven (actually, nine when fully formed), included Alfred Prince, renowned virologist at New York Blood Center, who devised a cheap plasma-derived vaccine He called the

Merck vaccine a 'rich man's vaccine' for a 'poor man's disease ' By offering assistance with technology, and promoting competition between interested firms and agencies, in 1987 the Task Force succeeded in securing a bid from the Korean Green Cross Corporation to supply the vaccine at just below $1 a dose [38]

Even at this price it was difficult to convince many governments in poor countries of the need for hepatitis B immunization The Task Force promoted two pilot projects, carried out 'on the ground' by the Program for Appropriate Technology in Health, an American non-profit organization whose head, Richard Mahoney, was a founding member of the Task Force In Indonesia, acceptance of the project was assisted by backing from President Suharto whose Foreign Minister had died of liver cancer In Thailand, local production of vaccine and continuation of immunization beyond the pilot area and time span were agreed beforehand, as long-term goals In both countries, hepatitis B immunization was integrated into the existing Expanded Programme on Immunization (EPI) for six childhood diseases But the hepatitis B element was more ambitious, since it involved locating all newborn babies in order to administer the first dose of vaccine within 7 days of birth, which was not necessary for the other EPI vaccinations (polio, diphtheria, pertussis, tetanus, tuberculosis, measles) Both the pilot projects were judged very successful

Meanwhile other initiatives were afoot According to Muraskin, when Japanese leaders of the Western Pacific Regional Office (a WHO agency) helped 'open up' China to hepatitis B immunization, the Task Force stepped in with a model project to demonstrate feasibility in a poor rural area However the Task Force could not dissuade the Chinese government from seeking to convert their own vaccine production to the prestigious but expensive genetically engineered (recombinant DNA) technology Blumberg, the discoverer of the hepatitis B antigen, gives a rather different account of the development of vaccine policy and production in China, emphasizing initiatives from Chinese public health experts and scientists seeking technology transfer from the West at each stage of vaccine development [39]

International and UK dimensions in the 1990s

By 1991 when a joint WHO/Task Force conference on hepatitis B immunization was held at Yaounde in Cameroon, universal infant hepatitis B immunization had been adopted in 30 countries The 'Yaounde Declaration' issued by that conference stated that hepatitis B exceeded in severity diphtheria, pertussis, polio, cholera, rotavirus diarrhoea, and AIDS, called for integration of hepatitis B vaccine into the EPI, and for global funding for purchase and delivery of the vaccine This call was followed by an EPI advisory group meeting in Turkey which set a target date of 1995 for universal hepatitis B vaccination in countries with carrier rates above 8% and 1997 for all other countries These goals were further endorsed by the WHO World Health Assembly in May 1992, calling for universal hepatitis B immunization by 1997, although countries with rates below 2%, like the United Kingdom, could consider adolescent vaccination as an alternative

This partly explains the change in attitude of the British government the debate over targeted versus universal immunization was no longer a matter between the Department of Health [40] and its advisors, but now had an unavoidable international profile [41] A rethink was forced on policy makers by the international health community Local circumstances could still be pleaded as a reason for local solutions, with more or less

expensive options considered. But the choice increasingly appeared to be between universal infant vaccination, or universal adolescent vaccination. The targeted route was falling out of favour.

However, by the end of 1997, Britain had not met the WHO deadline, along with the Scandinavian countries, the Netherlands, and Ireland: most other West European countries had done so. These recalcitrant non-vaccinators shared a relatively low prevalence of hepatitis B, and perhaps a scepticism about the philosophy of universal provision. This is reflected in a comment by a representative of the UK Public Health Laboratory Service:

> This disease is not a priority for UK people. It is still relatively rare and the vaccine is expensive. We have to think how best to spend disease prevention cash. It is not necessarily the correct thing to vaccinate millions of people to stop the annual appearance of only a few hundred cases, of which only a fraction will result in serious illness [42].

Nevertheless, during the 1990s the British authorities pursued the policy of targeted hepatitis B vaccination with zeal, notably in relation to the infants of carrier mothers, as illustrated by the case of a family whose child was forcibly vaccinated against their wishes [43]. Mothers and infants have historically acted as favoured targets for public health interventions in both developed and developing countries. They have long formed the category where the tension between individual liberty and public interest sways most readily towards public interest – that is, in the eyes of policy makers [44].

Conclusion

The leading contention of this paper is simply that vaccine viability does not depend only on the quality and effectiveness of the vaccine, combined with the size of the problem, and qualified by delivery issues such as the cold chain. If that were the case, we could write a sort of equation:

Large infectious disease problem + good vaccine = universal immunization

This 'rational model' of the relationship between science and policy clearly fails to meet the historical facts. For a wealthy country like Britain, the introduction of an effective but expensive vaccine for hepatitis B was predicated onto covert negotiations between certain groups – mainly health workers – and the Department of Health, over screening issues. For many years, hygiene measures continued to be the preferred avenue of defense, as illustrated by the Northern Ireland study referred to above. Policy initially failed to target the groups most likely to contract hepatitis B; but in any case, the implementation of the 'targeted' policy remained very feeble throughout the 1980s.

For Third World countries, where the vaccine might have been much more urgently needed, barriers of cost together with prioritization of more established problems meant that only in a few countries was the vaccine energetically promoted during the 1980s. For most countries it took the concerted effort of outside agencies to pilot viable delivery programmes. Apparently these were so successful, and the fall in price of the vaccine so dramatic, that a definite swing occurred in the early 1990s towards global promotion of universal vaccination. This does not represent a triumph for global pharmaceutical companies, as the bigger Western companies have given place to producers in the Far East, while a number of poor countries aim to make their own vaccine (although their ability to do so remains constrained).

To understand these sequences, we have to follow how problems and solutions were seen by those involved. This paper has traced two main strands to this 'social construction' of vaccine viability. One is the intertwining of alliances, evident both in my account of moves in the United Kingdom and in Muraskin's account of shuttle diplomacy and brokering by the International Task Force. The other is the longer history of the framing of the disease, including previous negotiations, before the advent of the vaccine, which set the terms within which its use could be considered. Paradoxically this process, based in Western science, produced a sharper image of the disease and more varied solutions in countries with low prevalence like Britain; while in poorer countries, there was very little indigenous framing of the disease. Only in those Asian countries with developed or planned economies plus a high prevalence of hepatitis B, did the push for a viable vaccine come from within.

So in essence I am arguing that 'vaccine viability' is to be understood in social, economic, and political terms rather than biological, scientific, and medical terms: even where the technical efficacy of a vaccine was generally agreed (and, as we have seen, this was not always the case) the use of vaccines depended on a complex of social, economic, and political factors – some explicit, some implicit. To understand the history of vaccines, we have to recover these factors, otherwise our history remains a sketch. Stories of 'success' and 'failure' of vaccines in purely scientific terms explain little, and are correspondingly inefficient in illuminating our present and our future.

Acknowledgements

Acknowledgements to the Wallcome Trust for funding the research of which this is one outcome. Responsability for interpretation remains with the author.

Bibliography

1 Goodfield J. The planned miracle. London: Cardinal; 1991. Goodfield gives a mainly positive account of immunization programmes in poor countries but shows awareness of problems of sustainability; she also describes transmission accidents.
2 Stanton J. Health policy and medical research: hepatitis B in the UK since the 1940s [thesis]. London: University of London; 1995. chapter 2.
3 Interviews conducted between 1991 and 1993, discussed here anonymously.
4 Interview with liver disease consultant, London, 1992.
5 Interview with virologist working in blood transfusion centre, London, 1992. See also, Mollison P. Blood transfusion in clinical medicine. 5th ed. Oxford: Blackwell Scientific Publications; 1972.
6 Stanton. Hepatitis B thesis. chapter 4.
7 Blumberg B. The Australia antigen story. In: Millman I, Eisenstein T, Blumberg B, Eds. Hepatitis B: the virus, the disease, the vaccine. New York: Plenum Publishing Corp.; 1984. p. 5-31.
8 London W, Blumberg B. Comments on the role of epidemiology in the investigation of hepatitis B virus. Epidemiol Rev 1985; 7: 59-79.
9 There is a wide medical literature on this question but the best introduction for historians is: Muraskin W. The silent epidemic: the social, ethical and medical problems surrounding the fight against hepatitis B. J Soc Hist 1988; 22: 277-98. My own grasp on the subject was greatly assisted by talking with a medical sociologist and a genitourinary medicine (GUM) specialist at a London GUM clinic, 1991.
10 Jilg W, Deinhardt F. Hepatitis B: eradicable? World Health July 1988; 10-2.
11 Szmuness W, et al. Hepatitis B vaccine: demonstration of efficacy in a controlled clinical trial in a high-risk population in the United States. N Engl J Med 1980; 303: 833-41.

12 See corr. following articles by leading hepatitis B expert Zuckerman A. Priorities for immunisation against hepatitis B. Br Med J 1982; 284: 686-8; Zuckerman A. Who should be immunised against hepatitis B? Br Med J 1984; 289: 1243-4.
13 McKee CM. Hepatitis B in Northern Ireland – who should be immunised? [Dissertation for part 2 of Membership of Faculty of Community Medicine]. 1988.
14 Interview with epidemiologist, London, 1991.
15 Stanton J. What shapes vaccine policy? The case of hepatitis B in the UK. Soc Hist Med 1994; 7: 427-46, esp. 433-5.
16 Restricted guidelines set out in DHSS circular letter CMO[81]11, 31 December 1981 and circular CMO[82]13/CNO[82]11, 15 October 1982; focussed on limited categories of NHS staff and patients.
17 Shanson D. Attitudes of staff [to] vaccination in a London hospital. In: Short R, Jones G, Eds. Hepatitis B in the UK. Proceedings of conference at Royal Society of Medicine, 14 October 1986. p. 42-4.
18 Besides sponsoring conferences, and publication of conference proceedings, MSD employed a public relations company to liaise with medical experts: interview, public relations officer, London, 1991. Subsequently, Smith Kline Beecham, the company producing the genetically engineered vaccine, employed similar tactics but went further, setting up a 'Viral Hepatitis Prevention Board.'
19 ASTMS Health and Safety Office Special Report. The risk of hepatitis to laboratory workers: the case against the attempt to downgrade safety standards in laboratories testing hepatitis B virus specimens. London: Association of Scientific, Technical and Managerial Staffs; 1980. An ASTMS health and safety representative spoke out against the vaccine at conferences in the 1980s: interview with epidemiologist, London, 1991.
20 Royal College of Nursing, Safety Representatives Conference Co-ordinating Committee. Hepatitis B and nursing in the UK: report from Wembley Conference. Newspaper format, April 1987. RCN archives. Front page headline reads: 'RCN urges DHSS to revise vaccine guidelines.'
21 Polakoff S. Acute viral hepatitis B reported to the Public Health Laboratory Service. J Infect 1990; 20: 163-8. These figures for England and Wales reflect a similar pattern to that suggested for Scotland (see below).
22 There were 'epidemics' of both drug use and hepatitis B in several inner cities in the early 1980s. For Edinburgh, see Bath G, Carson RA. Hepatitis B notifications in Edinburgh – a study by Edinburgh District Council and Lothian Health Board. Typescript, February 1986.
23 Interviews: group of homosexual men with hepatitis B, London, 1991; Department of Health drug policy adviser, London, 1992.
24 DHSS EL[88]P/125, July 1988.
25 66% of NHS drug treatment centres did not screen or offer hepatitis B vaccine, according to Farrell M, Battersby M, Strang J. Screening for hepatitis B and vaccination of injecting drug users in NHS drug treatment services. Br J Addict 1990; 85: 1657-9.
26 Heptonstall J. Outbreaks of hepatitis B virus infection associated with infected surgical staff. Commun Dis Rep 1991; 1 (8): R81-5.
27 For example, four contract hepatitis B in rented ward. Guardian, 11 May 1991; Women seek checks in hepatitis scare: risk from surgeon's infection minimal, says hospital. Guardian, 2 February 1993; Patients recalled as surgeon is told he carries hepatitis B virus. Guardian, 5 November 1994.
28 Hepatitis doctor jailed for 'terrible' deception. Times, 30 September 1994.
29 NHS staff to be tested for hepatitis B: Doctors' concern over compulsory testing for HIV. Independent, 11 January 1993.
30 Hepatitis B screening queried. Times Higher Ed Suppl, 8 July 1994.
31 Hepatitis B: does it affect you? Some advice for women. Leaflet by Ruth King for King's College Hospital Liver Unit, October 1989. At foot of front page, in bold type: 'There is now a vaccine which can protect new-born babies and other people from the virus.'
32 COHSE fact-sheet on hepatitis B, no date but from internal evidence late 1980s.
33 Group B. Hepatitis B and you. 26-page booklet, November 1990.
34 Mangtani P, Hall AJ, Normand CE. Hepatitis B vaccination: the cost effectiveness of alternative strategies in England and Wales. J Epidemiol Community Health 1995; 49 (3): 238-44.
35 Liver disease jabs 'for all at 12'. Guardian, 14 October 1991.
36 West European countries had lowest prevalence; some southern and eastern European countries had 'intermediate' prevalence, similar to that in Latin America.
37 Muraskin W. The war against hepatitis B. A history of the International Task Force on hepatitis B immunization. Philadelphia: University of Pennsylvania Press; 1995.
38 In time for first meeting of Task Force in Nairobi, March 1987. See: Muraskin. War against hepatitis B, p. 95.

39 Interview, Oxford, 1991.
40 The DHSS split in 1988, the Department of Health separating from Social Security.
41 One of the government's chief advisors on hepatitis, AJ Zuckerman, who headed a WHO Reference Centre on hepatitis in London, and sat on WHO Europe hepatitis committees over many years, actively advocated wider hepatitis B immunization for the United Kingdom.
42 Quoted in: McKie R. Killer in the blood that Britain ignores. Observer, 23 November 1997.
43 McCrystal Cal. Snatch and jab. Observer Life, 4 August 1996.
44 Stanton J, Berridge V. Vertical ancestries and horizontal risk: hepatitis B and AIDS. In: Löwy I, Gaudillière JP, Eds. Transmission: between heredity and infection. Harwood Press; forthcoming.

PUBLIC IMAGES OF IMMUNOLOGY

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Figuring immunity: towards the genealogy of a metaphor

Ed Cohen

"The choice of explanations in medicine is always a choice of values...."
Lawrence Kirmayer [1]
"... in this culture, medical thought is fully engaged in the philosophical status of man."
Michel Foucault [2]
"Science projects are civics projects; they remake citizens."
Donna Haraway [3]

A metaphorical introduction

From antiquity until the beginning of the nineteenth century, the Western imagination recognized that nature exercised a curative power on the organism, a power which medicine sought at best to emulate or to enhance. Within the prevailing Hippocratic tradition, this force was known as "Vis Medicatrix Naturae," the healing power of nature. According to this doctrine, healing was imagined as a natural manifestation of the organism's latent elasticity; it was thus understood to embody the organism's more expansive relations to the world, embracing the forces that animated the cosmos as a whole. Traditionally, natural healing expressed the enmeshment of living beings in the universe and affirmed their fundamental connection to the matrix from which they arose and to which they would one day return [4].

By the end of the nineteenth century, however, the "Vis Medicatrix Naturae" had fallen out of favor among Western bioscientists. Deemed unduly vitalistic for the emerging reductionist paradigms proposed by scientific medicine (which sought exclusively biochemical explanations for biological processes) healing became an increasingly anachronistic notion. Instead, Western medical rationality embraced a new image, adopted from a very old legal doctrine, for how organisms survive illness: "immunity." This metaphoric innovation fundamentally reconceived the amelioration of illness. No longer represented as evidence of the organism's essential connection to the world in which it lived, immunity refigured healing as an effect of the organism's necessary struggle to defend itself against the world.

We now take this reinterpretation to be definitive. Most of us who rely on biomedical treatments accept the idea that our "immune systems" ought to "defend" us against disease (even as we are also increasingly aware that they do not always live up to this promise). Nevertheless, it is not entirely self-evident why a term that had served almost exclusively as a political and legal concept for more than two millennia could have been so rapidly

assimilated into modern biological thinking as one of its central metaphors. Indeed, given the seeming impropriety of using an explicitly juridico-political term to elucidate a biological phenomenon, we might want to know how this metaphoric translation came to make sense. We might also want to ask what the consequences of this translation were both for the practice of medicine and for the experience of healing.

The word "immunity" was transplanted into bioscience a little more than 100 years ago in order to describe the ways in which multicellular organisms systematically mediate relations with the environments in which they exist – environments replete with other living beings. Needless to say, there is no natural or necessary connection that grounds the analogy between the legal and biological uses of the word. Moreover, as with all transformative metaphors, its rhetorical success depends upon the way it both retains and resolves the tensions that exist between the categories which it brings into relation. When immunity first began to appear in medical discourses during the middle of the nineteenth century, its metaphorical efficacy was constrained by the unresolved incongruity of the terms it put into play. At that time, immunity did not significantly increase medical knowledge because the force of the legal concept, which underscores exemption and distinction, overtly clashed with prevailing humoral and environmental explanations, which stressed continuity and connection among organisms and their contexts. Hence, immunity appeared in these early biological uses as a paradoxical metaphor at best and not a particularly helpful paradox at that. By the end of the century, however, political, philosophical, and scientific developments had significantly changed the context for biological understanding. So much so that when the concept was reinvoked in the early 1880s and used to explain the miraculous results of Pasteur's vaccination experiments, it rapidly transformed the basic understanding of how organisms – and especially human organisms – endure.

During the last 100 years, then, a term which first appeared in Roman Law has come to serve as the metaphoric crux of numerous technoscientific endeavors in human biology. Its naturalization has provided, and continues to provide, a principle of articulation for the practice of scientific medicine on and within the bodies of the citizen-subjects of the industrialized West. Indeed, by facilitating the application of biochemical practice to organismic function, immunology has successfully modeled the "objects" of twentieth-century biomedicine (in the double sense of its experimental domains and its therapeutic aims) [5]. Today immunity literally underwrites the project of scientific medicine as it has been incorporated in our culture since the end of the nineteenth century. It is a trope which seems to make biochemical reductionism seem to make sense as the best, if not the only, way of understanding the complex, and at times contradictory, articulations of organism and environment. Immunity as metaphor underlies the truth of immunity as a bioscientific concept precisely in so far as we forget the metaphoricity which makes this "truth" seem true [6].

This essay explores the bio-political effects that produced and were reproduced by the medical investment in "immunity" as a scientific trope. What I want to try to understand is: How did an ancient juridical concept pass over into the human interior at the end of the nineteenth century in order to scientifically elaborate the body as a knowable object? And how did the acceptance of the metaphor immunity as an organic truth shape the goals of modern bioscience? By recovering the metaphoric migration of immunity from the political to the biological, then, I hope to unravel some of the complex meanings, practices,

events, relations, and histories that congeal in and through the incorporation of immunity as principle of articulation within contemporary Western medicine. For if, as Haraway reminds us, "... a research program is virtually always also a very mobile metaphor," then we can also welcome the significance of her admonition that, "[m]etaphors are tools and tropes. The point is to learn to remember that we might have been otherwise, and might yet be, as a matter of embodied fact" [7].

Embodied metaphor

Standard medical histories usually represent the "discovery" of the immune response as a turning point in the "progress" of medical science. Before the advent of immunization in the 1880s and 1890s – arguably among the first effective biotechnological treatments – scientific medicine had done little more in the way of ameliorating illness than either its pre-scientific predecessors or its contemporary competitors. Other than Jenner's famous experiments with small pox vaccinations at the end of the eighteenth century (a practice derived in part from the folk knowledge of dairy maids) the advances in biomedical knowledge until the 1870s were primarily anatomical and physiological. In the absence of new treatment options, the emerging science of medicine focused on pathology and diagnosis as practices of classification that defined the epistemological superiority of the modern physician over and against all other health practitioners. The focus was on elaborating the "natural histories" of specific diseases in order to fix diagnosable pathologies as privileged objects for "rational" medical inquiry. Even through the first decades of the twentieth century, diagnosis, and not treatment, still defined the physician's métier [8].

From this pathogenic perspective, the physician's attention was directed away from the event of illness as a passage in the flux of human embodiment towards the search for an essentially "foreign" agent which could now be conceptually localized as the "cause" of disease [9]. With the introduction of vaccines and anti-toxins as viable medical technologies in the last decades of the nineteenth century, doctors arrived for the first time at an apparently reliable, scientifically-derived, industrially-produced, and commercially profitable means to intervene in the process of disease transmission. Needless to say, this technology played a critical role in establishing both the epistemological and the economic success of modern medical practice. Yet even after Pasteur spectacularly introduced an effective vaccine produced from attenuated forms of the bacteria that "caused" chicken cholera and anthrax, thereby providing a technique that could be applied to other infectious diseases, there was no cogent explanation for why or how his technique worked. Not, that is, until the scientific elaboration of "immunity" as the organismic activity of "self-defense" provided an effective remedy to this plaguing theoretical wound.

In order to begin to understand the difference immunity has made to modern biomedicine, it is necessary to reconsider the conceptual developments which led to its incorporation in medical theory and practice, developments which made it possible for immunity to begin to function as a productive category for bioscience in the first place. While medical historians have noted the sporadic use of the term "immunity" over the last 2,000 years to designate the decreased vulnerability to the effects of epidemics found in those who survived previous incidents of the same kind of illness, this usage was at best highly impressionistic [10]. According to Anne-Marie Moulin, the first modern biological use of the term "immunity" does not appear until 1775 when Van Sweiten, a Dutch physician,

used the word *immunitas* in Latin to describe the effects induced by an early attempt at "variolization" (inoculation for small pox) by his compatriot Boerhaave:

> It would be desirable thanks to a medical artifice to treat the body in such a manner that everything happens as it does in those who have been attacked once by a disease or those who, thanks to their own idiosyncrasy, were not touched by the contagion and as a consequence possess immunity (immunitas) with regard to this illness [11].

Here the metaphor "immunity" serves to mark a derivative quality "possessed" by certain exemplary individuals who remain unmarked by those symptoms which signify the presence of a "contagion" among the general population. Hence, the backdrop against which the singular exemptions represented as "immunity" become intelligible as marked exceptions is that of a generalized illness which affects not just the atomized, individual body but more significantly the aggregation of bodies that constitute the visible social field. In so far as the most devastating effects of epidemics during the eighteenth and nineteenth centuries primarily occurred in urban areas, we might say that biological immunity emerged as a way of designating a quality that inheres in individuals who remain unmarked by those afflictions that primarily appear as municipal diseases.

Moulin also notes that the first systematic treatise published on immunity appeared in 1852 in the monograph of a French doctor, André Threse Chrestien bearing the title "On Immunity and Morbid Susceptibility" (*De l'immunité et de la susceptibilité morbide, au point de vue de la clinique médicale*) [11]. Here Chrestien's juxtaposition of "immunity" to "morbid susceptibility" is explicitly located within the "clinical" model that Michel Foucault has elucidated as emerging during the last decades of the eighteenth century and the first decades of the nineteenth century. As Foucault cogently explained in *The Birth of the Clinic*, "morbidity" develops as a medical concept at the moment when bioscience begins to "open up a few corpses" in order to distinguish the somatic signs inscribed by diseases, signs that define disease as the "author" of pathogenesis. In Foucault's account, the emergence of clinical medicine marks the movement away from the grand metaphysical concepts embedded in humoral medicine towards the concrete empirical practices of scientific medicine, represented most graphically by the autopsies which underwrote pathological anatomy [12]. As the cadaver became the privileged locus from which the truth of disease could be read off, "morbidity" became the teleological context in which disease comes to make sense. As Foucault succinctly comments: "It is not because he falls ill that man dies; fundamentally, it is because he dies that he falls ill" [13]. For clinical medicine, the corpse provided a corporeal map that rendered disease both visible and intelligible as an object of human understanding, thereby locating the concept of disease itself within the penumbra of death.

The appearance of biological "immunity" in the shadow of the morbid turns our attention to the particularizing function which nineteenth-century theories of disease adopt. By localizing disease in the tissues of the body, clinical medicine began to restrict the consideration of how diseases appear by defining them as forms of individuality. For Chrestien, *immunité morbide* specified "an effect, in itself spontaneous, which only recedes upon becoming a cadaver and which manifests itself in the most positive manner by the faculty it has of frequently resisting the causes of illness which surround [*environment*] men on all sides" [14]. Derived from a vitalist understanding that he opposes an emerging biochem-

ical reductionism, Chrestien's immunity gestures towards a "force of resistance more or less inherent in life," a living force that counteracts the deleterious effects of a social environment within which the organism may "contract certain illnesses" [*contracter certaines maladies*] [15]. As Chrestien's politico-economic metaphor suggests, immunity manifests a contractual exemption – or, more pertinently, an exemption from having a contract imposed upon one – where illness appears as a biological manifestation inscribed within a social relation. However, the juxtaposition of "morbid susceptibility" to "immunity" in Chretien's text also demonstrates why "immunity" does not emerge from this context as a robust bioscientific concept: for if morbidity appears under the shadow of mortality and if mortality is a destiny to which all flesh is heir, then how does one speak of "immunity" in relation to "morbid susceptibility"? Since this contradiction undermines the metaphor's coherence, it unwittingly disrupts the very meanings which it seeks to incorporate in the first place. Imported from a juridico-political domain, immunity disrupts the intelligibility of the logic that, at this point in time, internally links morbidity to mortality and therefore it remains an eccentric formulation with respect to the prevailing accounts of pathogenesis. Thus, while Chrestien's usage might seem closer to a contemporary bioscientific concept, immunity still functions in 1852 primarily as an analogy that gestures towards a kind of singularity marked out by exemption from contagion, where contagion continues to impinge on the political as one of its most devastating "facts."

Grounding the fortress body

The difficulty with establishing the biomedical meaning of the term "immunity" before the end of the nineteenth century derives from its ambiguous status as a mixed metaphor within the prevailing systems of biological explanation. Until the development of biochemical reductionism as a dominant paradigm for allopathic medicine during the mid to late nineteenth century, the reigning mode for interpreting disease processes descended from the categories of ancient Greek medicine. In the tradition inscribed under the names Hippocrates and Galen, diseases were understood to signify disturbances in the four humors (blood, phlegm, yellow bile, and black bile), which themselves derived from the four elements (water, fire, earth, and air) as modified by the four cardinal qualities (wet, dry, hot, and cold). Mapped by this constellation, pre-scientific Western medicine defined the human body as a localization of fundamental elements that were coextensive with the known universe. Since bodies, healthy or diseased, could not be known apart from the ways that they fleshed out their extension in the world, diagnosticians directed their attention not to an atomized, individuated body as the host for a symptom producing disease entity, but rather to a deeply embedded body understood as a relation among humors, elements, and qualities.

Given the inclusive and fluid metaphoric system underlying pre-modern medicine, to speak of "immunity" with respect to embodied states would not only be improper but nonsensical. If disease signified a relation among elemental qualities and humors that were materially constitutive of both the living organism and its life context, then "exemption from" them on the model of juridico-political immunity would be a non sequitur at best. In so far as bodies are radically fluid for the pre-modern imagination, explanations for susceptibility or resistance to disease do not partake of the closed political economies that are the hallmarks of modern immune discourse. It is only after the figure of the "fortress

body" emerges in Western political philosophy during the course of the seventeenth and eighteenth centuries that the groundwork was laid for "immunity" to begin to function as a productive image for medical understanding.

An early conceptual elaboration of this figure appeared in the writings of that paradigmatic ideologue of American virtue, Cotton Mather. Here we begin to discern a formulation that anticipates biological immunity's emergence at the intersection of juridico-political and biomedical thinking. In the chapter entitled "Variolae trimphatae, or The Small-Pox Encountered" of his manuscript, "The Angel of Bethesda" (1724), probably the first medical text written in the thirteen colonies, Mather characteristically juxtaposes bombastic Calvinist interpretations of small pox as the scourge of God to a truncated version of an earlier humoralist expulsion theory before providing a detailed account of the "New and Right Method of treating the Small-Pox, and reclaiming People from the Madness of Killing one another with Kindness, and praeposterous Proceedings" [16]. While not surprising, Mather's exposition of this prophylactic "method" primarily concerns the "Way of Proceeding in the practice," he nonetheless begins his discussion with an extended analogy that attempts to elucidate the effects of the technique he takes pains to describe:

> Behold, the Enemy at once gott into the very Center of the Citadel: And the Invaded party must be very Strong indeed, if it can struggle with him, and after all Entirely Expel and Conquer him. Whereas, the Miasmas of the Small-Pox being admitted in the Way of Inoculation, their Approaches are made only by the Outerworks of the Citadel, and at Considerable Distance from the Center of it. The Enemy, tis true, getts in so far as to make Some Spoil, yea, so much as to satisfy him, and leaves no Prey in the Body of the Patient, for him ever afterwards to seize upon. But the Vital Powers kept so clear from his Assaults, that they can manage the Combats bravely and, tho' not without a Surrender of those Humours in the Blood, which the Invader makes a Seizure on, they oblige him to march out the same way he came in, and are sure of never being troubled with him any more. [17]

Here Mather unites a humoral depletion theory with the emerging notion of the body as fortress to construct his story of how variolization works. Likening the body to a "citadel" at whose front gate the "venomous Miasmas" of small pox normally lay siege, Mather portrays inoculation as a pre-emptive strike. If the physician can introduce the seeds of contagion by way of a secondary fortification whose maintenance is not critical to the integrity of the whole organismic structure, Mather suggests, then the appropriate gusto of the enemy will exhaust itself by depleting "the prey" that it covets, while only causing a minor loss of humoral wealth. By relinquishing that part of its property which makes it a desirable target for "the Invader [to] make a Seizure on," the inoculated body frees itself from future life threatening attacks on its fortifications though a project of strategic impoverishment. Needless to say, both the logic and the metaphors in this account are somewhat fuzzy. Nevertheless, what seems very clear is that Mather begins to forge a biological analogy grounded in a proprietary understanding of individuality, an understanding founded on the image of the body as a citadel whose vital properties are strategically alienable (in the same way, for example, that wage labor is beginning to be understood as the contractual alienation of the body's exertions at exactly this historical moment). Thus, Mather presciently characterizes modern embodiment as an essentially defensive posture

through which the fortress body must fend off the marauding tendencies of hostile "enemies" which would attempt to "seize" its vital properties for their own ends.

Incorporating biopolitics

Coinciding with the "scientificization" of medicine itself, the transition to a distinctly modern conceptual formation of immunity begins to crystallize at the cusp of the nineteenth century. From a technological perspective, Edward Jenner's experiments in the 1790s demonstrated that resistance to small pox could be induced by prophylactically inoculating humans with the relatively harmless infectious matter of cow-pox (hence, his coinage of the term "vaccination" from the Latin *vacca* for cow). However since Jenner, a practicing physician, was primarily an empiricist and not a theorist, he was almost entirely concerned with documenting rather than with explaining the efficacy of his technique. In his *Vaccination Against Smallpox: An Inquiry into the Causes and Effects of the Variolae Vaccinae, or Cow-Pox* (1798) [18], he offers a compendium of cases, updated over each of the next several years, that describe the effects of his experimental protocol. Largely abjuring analytic inference from the material he presents, Jenner tentatively broaches a conclusion in his supplement of 1799:

> The results of all my trials with the virus on human subjects has been uniform. In every instance the patient who has felt its influence, has completely lost the susceptibility for the variolous contagion..." [18].

In this passage, which represents one of his very few synthetic statements, Jenner invokes the nebulous efficacy of vaccination's "influence" to account for decreased "susceptibility" to "variolous contagion." Not surprisingly, given his belief that small pox derives from an undue "familiarization... with a great number of animals [e.g., dogs, cats, cows, hogs, sheep, and horses] which may not have been intended for his associates" [19], Jenner deploys an environmentalist terminology to cautiously explain the results of his experiments. Vaccination, thus, capitalizes on the disease inducing proximity between humans and animals (a condition of boundary confusion that signals "the deviation of man from the stage in which he was originally placed by nature") by controlling the direction in which the infectious matter circulates. Yet, since Jenner made no attempt to elucidate the rationale for his, albeit successful, technique, vaccination did not provide a model either for further scientific experimentation or for theoretical elaboration until almost a century later when Pasteur would retroactively appropriate both the concept and the name.

If at the end of the eighteenth century, Jenner was unwilling or unable to articulate a more general theory concerning the complex biosocial nexus within which a vaccinated body manifests its non-susceptibility to small pox, or conversely within which a non-vaccinated body manifests its susceptibility to small pox, this is not to suggest, however, that such a theory was unimaginable at this time. Indeed, we could argue that this was precisely the moment when the theory most radically incorporating such a body was first effectively articulated within the very liberal political philosophy which underwrote both the French and American revolutions. As the liberal individual's self relation to his [sic] own body became the ground upon which political rights could be demanded, over and against the sovereign claims previously made in the name of God and souls by European monarchs, the biopolitical artifact called "the body," by the very fact of its birth as such,

came to define both the possibilities for and the limits on social relations. In Thomas Paine's famous formulation in "The Rights of Man" (1791) [20], the notion of equality of life is predicated on the generality of birth as the necessary organismic condition of human embodiment. Birth equalizes insofar as it equally incorporates citizen-subjects both biologically and politically, the former becoming the inescapable condition of the latter [20]. Inscribed as the originary moment of the self's constitution as a political subject, biological birth comes to mark the human body as both the site and the source of political relations. Insofar as human beings – or in Paine's literal formulation, "men" – are born, they are born equal. Equalized by the birth that gives them life, humans are equally political and biological creatures. As Giorgio Agamben has astutely noted in his brilliant political meditation "Homo Sacer":

> ... [T]he concept of the 'body' is always already caught in the a deployment of power. The 'body' is always already a biopolitical body and bare life.... In its extreme form, the biopolitical body of the West appears as the threshold of absolute indistinction between law and fact, juridical rule and biological life.... [W]e are not only, in Foucault's words, animals whose life as living beings is at issue in their politics, but also – inversely – citizens whose very politics is at issue in their natural bodies [21].

Emerging into the world as nodes of biosocial relations, human bodies came to constitute, at the end of the eighteenth century, the threshold across which life and politics commune. Concomitantly, this moment also marks the formulation of "life itself," of "bare life," of "animal... life as living beings," as a political determination, where the facts of birth, death, and illness, are reimagined as constituents of the health of the nation as a whole. Thus, the incorporation of the body as the substrate of political life animates the conditions within which the body and its health, or illness, become indices of the well-being of Western nations.

Perhaps the most striking example of this indexical relation appears in Thomas Malthus's "Essay on the Principle of Population" (1798/1803) [22], written at exactly the same moment that Jenner was undertaking his experiments. Whereas Jenner had judiciously restricted his consideration of the play between health and illness to finding empirical justification for his prophylactic technique, his contemporary Malthus seemed to have no qualms about globalizing their implications. Indeed, Malthus famously – albeit pessimistically – adumbrated the ways that the human organism qua "susceptible" organism becomes implicated in and as the frontier of social relations. He thereby crystallized contemporary ways of imagining the human body and inscribed them within a social theory that envisioned this body as a critical nexus for the nation.

The first edition of the "Essay", appearing in 1798, developed Malthus's remarkably dismal premise that if otherwise "unchecked" by "misery and vice," the "fixed laws of our nature" (which demanded both the nourishment of food and the gratification of sex) would produce a national population that would outstrip the national means of subsistence [22]. As George Canguilhem succinctly puts it:

> Malthus's problem was the following: How could a tendency be made compatible with a limit? How could two aspects of nature – the multiplication of living things and the limited amount of available space and food – be reconciled? [23].

Yet, according to Malthus, the "wants of the body" are also the "stimulants" that spur "the brain of infant man into sentient activity" so that they necessarily underlie all forms of social productivity [24]. This formulation, as Catherine Gallagher has aptly noted [25], simultaneously positions the biological body as absolutely central to and absolutely problematic for the social formation. In other words, the body with its need for daily nutriment and its "necessity" for "passion between the sexes," becomes the crux of an analysis that seeks to explain why the organismic capacity for economic and biological production and reproduction can simultaneously give rise to abject poverty and starvation. Since the analysis works only by attributing this paradox to the "laws of nature" inhering in human bodies themselves, Malthus imagines "the body" at and as the intersection of economic and biological processes. He thereby fleshes out earlier proprietary notions of the body as "property" and conversely grounds problems of social welfare – or as it would soon come to be known, "public health" – within a somatized "natural history." By drawing together these eighteenth-century developments into a coherent and persuasive explanation for social problems such as poverty, malnutrition, disease, and unemployment, Malthus's work establishes the body as a nexus that both naturalizes the unequal distribution of resources within the nation and inscribes the attendant human suffering as the effect of biological processes. As a consequence, given its explicit imbrication of economic and biological causalities, Malthus's formulation simultaneously anticipates and frames the practices through which health care became a form of social policing during the course of the nineteenth century.

Construing public health

While it had long been the case that during epidemics governments would actively intercede to attempt to contain the spread of deadly diseases by establishing protocols such as curfews, quarantines, inspections, or disposing of corpses, it was not until the nineteenth century that the function of promoting public health came to be understood as an ongoing obligation of the state. In part, this new perspective emerged from the confluence of economic and political factors that resulted in the development of the great European urban centers, where the coincidence of destitution and disease rendered increasingly graphic displays of morbidity and mortality among the citizenry. As repeated waves of cholera, typhus, typhoid, and influenza broke across Europe during the first half of the century, claiming hundreds of thousands of lives, they left in their wake rampant fears concerning the ability to defend social life against such pestilential tides. In this context, Malthus provided a theoretical frame within which to comprehend the bio-social significance of these virulent epidemics. Moreover, the Malthusian definition of poverty became so imbedded in the social imagination of epidemic disease that the one increasingly became an index for the other. As the foremost French hygienic investigator of the early nineteenth century, Louis René Villermé, put it in 1833: the frequent recurrence of epidemics is a sign of poverty, or, what comes to the same thing, an excess of population with respect to the means of subsistance" [26]. Through the interpretive strategies deployed to make sense of these deadly epidemics, the political significance of biological existence came to be imagined as implicating the "health" of both individual and national "bodies."

This coincidence of clinical and social interpretations served as the basis upon which "public hygiene" emerged as a prevailing medical ideology during this period [27]. As medical theorists and practitioners began to conceive of the human organism as a critical biosocial nexus whose somatic vulnerability defined it as the contradictory site of both socioeconomic reproduction and devastation, they engendered the possibility for an explicit articulation of political and biomedical thinking [28]. Indeed, as Bruno Latour has wryly commented, the hygienists' practice was predicated on the elaboration of this very contradiction: "The conflict between health and wealth reached such a breaking point in the mid [nineteenth] century that wealth was threatened by bad health" [29]. In so far as the aggregation of human organisms in cities foregrounded the coincidence of biological and sociological vulnerabilities, public health increasingly came to define the horizon within which the most effective practice of medicine took place. Indeed, as many medical historians have noted, the social improvements that resulted from the agenda of public hygiene movement (including improvements in sewage, water supplies, housing, working conditions, ventilation, quality of basic food stuffs, etc.) provided a remarkable amelioration of urban morbidity and mortality prior to the development of the bio-medical technologies that followed in the wake of Pasteur.

The medical responses to the numerous epidemics that marked the lives of European urban dwellers during the nineteenth century were political, then, in several senses. First, they attempted to ameliorate the suffering and death of the citizens, especially in so far these disrupted the productive forces of the nation. Second, they sought to reestablish the sanctity of the nation's geopolitical boundaries which offered little obstacle to onslaughts of disease. Third, they tried to mitigate tensions within the city exacerbated by the disproportionate toll that epidemics took upon the poor and laboring classes. Fourth, they legitimated the responsibility for the medical policing of the populous through both legislative and extra-judicial authority. Fifth, they established medical practitioners as mediators between citizen-subjects and the state. Through this complex ensemble of relations, hygienic practice shifted the primary locus of medical intervention from the illnesses of particular individuals (where it was fairly useless) to the aggregate phenomenon of endemic and epidemic disease among vulnerable cohorts of national citizens. As Anthony Wohl has remarked apropos the English responses to the wave of epidemics that crested during the 1830s:

> These diseases called for more than the doctor's healing art or the research chemist's endeavors. They called for the state to inspect and ultimately control the excesses of unregulated urban growth and rural neglect: in short, they projected the state into public health and placed it in the position of the guardian over the environment [30].

Through these moments of biosocial crisis, medicine was effectively transformed from an ameliorative healing practice directed towards redressing the micro- and macrocosmic imbalances which manifest themselves as the bodily suffering of individuals into a regulatory governmental practice aimed at safe guarding the geopolitical domain as a whole.

Given the increasing political, economic, and human costs entailed by the epidemics of infectious disease that plagued Europe during the first half of the nineteenth century, it makes sense that the public hygiene movement would gather momentum during these years. Hence, Foucault argues that public hygiene served as the instrument-effect of a new

regime of biopower precisely in so far as it was able to "install mechanisms of security around this risk which is inherent in a population of living beings" [31]. Yet what was surprising about public health initiatives, as Bruno Latour has noted in his remarkable study "The Pasteurization of France" [29], was the patent incoherence of the doctrine to which the hygienists adhered in framing their social practice. Lacking a uni-causal understanding of infection and disease and drawing instead on the legacy of humoral and environmental theories that located diseases in their complex life contexts, the public health movement saw the constituents of illness everywhere:

> If anything can cause illness, nothing can be ignored; it is necessary to be able to act everywhere and on everything at once. The style reflects the action planned by the hygienists.... Since anything might cause illness, it was necessary to act upon everything at once, but to act everywhere is to act nowhere. Sometimes the hygienists give a definition of their science that is coextensive with reality. They claim to be acting on food, urbanism, sexuality, education, the army. Nothing that is human is alien to them. Even the human being is too narrow a field; they must also concern themselves with air, light, heat, water, and the soil. But to understand everything is to understand nothing [32].

For the hygienists, diseases arose out of the extensive nexus of a life world and hence could only be located throughout the nexus as a whole. Disease was proximate to the individual but not localized within the individual per se. To extend Latour's syntax: for the hygienists, widespread disease was paradoxically everywhere and nowhere at once, precisely because such disease made the contradictions incorporated in, and as, "the individual" palpable in the very place where that abstract individual lived, i.e., the biosocial world. Thus, despite how it might appear to our post-Pasteurian eyes, the hygienists' style of reasoning about disease did not demonstrate their ignorance or naiveté but rather evinced their commitment to a thoroughly political interpretation of disease [33]. Taking disease, and especially epidemic disease, as a manifestation of what following Donna Haraway could call "artifactual social nature" [34], the nineteenth-century public health movement located the vectors of its causality along a continuum that circumscribed human bodies within the entire municipal domain. Infectious diseases appeared not simply as internalized struggles for individual survival in a hostile and life-threatening world, but rather as collective disturbances of the social field. Addressing the complex implications of disease therefore necessitated widening the medical purview to include the multiple ecological factors that coincided in and as the life of the human organism since this political ecology defined the situation within which diseases manifest themselves.

Unfortunately, while this expansive purview did lead to projects which generally improved the lives of many nineteenth-century urban dwellers – probably including declines in overall morbidity and mortality – it was not specifically effective against the waves of infection that came to serve as the index of their success [35]. Thus arose what Latour calls "the paradox of the hygienist movement":

> [O]n the one hand, it was a social movement of gigantic proportions that declared itself ready to take charge of everything, and on the other, it was a succession of measures that were being quietly undermined by unknown and erratic agents. As a result, the period showed keen interest in identifying the corrupting forces, the double agents, the miasmas and contagions, and accorded immediate trust to those

> who might, in identifying them, be able to take measures against them. It was at this precise point that the microbe and the revealers of the microbes appear [36].

It is no small irony that at precisely the point when public medicine seemed to thoroughly understand the social dimensions of living as the context within which diseases emerge that it should have to have recourse to biological theories predicated on localizing microscopic pathogens within isolated individuals in order to maintain its political purchase. Seemingly, the amelioration of endemic disease by multifarious means was not enough to establish the success of public hygiene as an effective medical technology in the face of frequent epidemics of untreatable illnesses. In so far as social theories of disease relied upon diffuse notions of causality, they provided limited means for comprehending the ways in which particular bodies manifest their vulnerability to contagion. Hence, while waves of infectious diseases that threatened nineteenth-century Europe were patently social phenomena in the aggregate, the inability to assuage the symptoms embodied by particular individuals rendered the public hygiene movement vulnerable to the conceptual contagion known as the "germ theory" of disease.

Bringing the outside in: Bernard's "*milieu intérieur*"

Needless to say, "germ theory" has long since become synonymous with the spectacular triumphs of Louis Pasteur, despite the fact that these Pasteurian "triumphs" only emerged out of a complex matrix of laboratories, hospitals, academic institutions, scientific publications, industrial concerns, farms, animals, technicians, political ideologies, and popular representations of experimental science. By and large, most accounts of modern medical history still unproblematically define the legendary successes of Pasteur's vaccines against chicken cholera (1879) and anthrax (1881) as the theoretical and commercial threshold of modern biotechnologies. However, while Pasteur's fortuitous experiments undoubtedly provided the model for what would soon come to be widely known as immunization, both their intellectual and technological conditions of possibility must be located in the path-breaking work of the famous French experimental physiologist, Claude Bernard [37].

Forsaking the Hippocratic tradition of healing, which had relied on contemplative observation and description, in favor of an interventionist regime aimed at the "scientific domination of living nature," Bernard made it possible to reimagine both the object and the aim of nineteenth-century medical practice [38]. With his simple yet labile concept, *milieu intérieur*, he arguably transformed nineteenth-century bioscience more radically than anyone who had preceded him [39]. Indeed, the experimental imaginary of Bernard's *milieu intérieur* provides the incipient logic that underlies and underwrites the corporeal atomism of most post-Pasteurian medicine, especially in so far as it has been naturalized through the organismic function described as "immune response." In the wake of Bernard's critical innovation, biomedical interest could begin to be "scientifically" directed away from the complex ecological interpretations of disease and healing – upheld in different ways by both the Hippocratic tradition and the hygienic movement – towards the artifactually-reproduced "independent" and "free" body of the experimental subject. It is from Bernard's biosocial construction: the individual animal/artifact, that the "defended" organism corporalized within bioscientific paradigms of immunity descends.

In its earliest presentation, Bernard's *milieu intérieur* adapted the residual effects of a humoralist understanding of blood as the essential medium of life in order to articulate his

break with humoralism [40]. Since the internal elements of any complex organism no longer directly contact the environment from which they must derive sustenance, Bernard reasoned, an "artifice" was needed to mediate the separation: "That artifice is circulation; the blood is the environment" [41]. This initial formulation provided Bernard with a rough theory which allowed him to justify his claims for the preeminence of physiological determinants over environmental variables in experiments on animals. If the living body necessarily relies upon the surrounding environment, it is nevertheless effectively self-contained within its own internal milieu:

> That sort of independence which the organism has in the external environment derives from the fact that in the living being the tissues are in fact withdrawn from direct external influences, and are protected by a veritable internal environment, which is constituted in particular by the fluids circulating in the body. That independence, moreover, becomes greater the more elevated the organism is in the scale of organization, that is the more completely protective an internal environment it possesses [42].

The focus on the "elevated" animal's "independence" from the environment marks a critical – and political – turn in biomedical discourse. For as Sagen and Margulis remind us, "Independence is a political, not a scientific term" [43]. By transplanting this tenet of liberal political philosophy into his "scientific" theory of the organism and then ignoring the imaginary work it performs, Bernard directly overturns the hygienists' presumption that the social environment constitutes the human organism's life context. Instead, by way of a paradoxical formulation, *milieu intérieur*, Bernard inverts the topological relations of inside/outside, such that the organism's interior comes to serve as a determining context that effectively isolates it from the life world which now only secondarily environs it. Hence, in Bernard's formative depiction, the highest (a.k.a., human) animal's "independence" is fundamentally predicated on the "more completely protective an internal environment it possesses."

Prior to Bernard's reformulation, *milieu* denominated the extensive context within which an organism purdures, the domain of which is in the middle, or the medium of its existence. Indeed, as George Canguilhem demonstrated, *milieu* emerged as an imbricated biological (Lamark) and sociological (Comte) concept at the beginning of the nineteenth century precisely because it foregrounded the complex determinations of human organisms in their life contexts [44]. In the wake of Bernard's conceptual innovation, however, *milieu intérieur* comes to figure as not only the means of separation from, but a "complete protective mechanism" [45] against material environment which enables higher organisms to become "free and independent ":

> Constant or free life belongs to the most highly organized animals. In it, life is not suspended in any circumstance, it unrolls along a constant course, apparently indifferent to the variations in the cosmic environment, or the changes in the material conditions that surround the animal. Organs, apparatus, and tissues function in an apparently uniform manner, without their activity undergoing those considerable variations exhibited by animals with an oscillating life. This is because in reality the *milieu intérieur* that envelops the organs, the tissues, and the elements of the tissues does not change; the variations in the atmosphere stop there. So that it is true to say that physical conditions of the environment are constant in higher animals; it is

> enveloped in an invariable medium, which acts as an atmosphere of its own in the constantly changing cosmic environment. It is an organism that has placed itself in a hothouse. Thus, the perpetual changes in the cosmic environment do not touch it; it is not chained to them, it is free and independent [46].

This formulation which appears in the posthumously published *Leçons sur les phénomènes de la vie commune aux animaux et aux végétaux* [Lectures on the Phenomena of Life Common to Animals and Plants] (1878) [46] summarizes Bernard's thinking about the implications of the *milieu intérieur* after more than 20 years of deliberation. Linking the highest values of liberal political thought – freedom and independence – to the "internal" biological conditions of individual, "higher" organisms, Bernard implicitly circumscribes the sociality that the public hygiene movement located within the urban *milieu* in and as the biological context of the human organism itself, albeit a sociality among "lower" beings: "A complex organism must be considered as an association of simple beings, which are the anatomical elements, and which live in the fluid internal environment [*milieu intérieur*]" [47]. An aggregation enveloped in its own context, the higher organism specifies its own relation to the world as a self-relation among its elements. Moreover, the biosocial context of life only impinges on such a self-relating organism in a mediated fashion. Hence, all efforts to assuage conditions that afflict the complex organism must be addressed towards the domain in which it really lives, that is, "the fluid internal environment," and not to the epiphenomenal social world which merely touches it indirectly. Not surprisingly, Bernard made the policy implications of this "scientific" perception explicit:

> Medicine must act on individuals. It is not destined to act on collectivities or people.... In reality, one only acts on individuals. Collectives are entrained in currents upon which we can have no effect. These are general actions which are beyond us. It is the same with epidemics and epizootics. One can act on the individual who presents oidium, plague or cholera; but one cannot act on the general cause of plague, cholera, etc., etc. [48].

By insisting on the experimentally derived validity of the "higher" organism's relative autonomy (an autonomy which reciprocally defines the organism's "elevation" in the first place), Bernard explicitly contradicts the hygienic movement's focus on the human organism as part of a political ecology. Instead, he advances a putatively non-political, "scientific" practice that identifies the "free and independent" individual as the necessary locus of all biomedical inquiry and action. He argues (successfully as it turns out) for the circumscription of medical experimentation within the context he names the *milieu intérieur* as if this metaphor designated a manifestly non-political or extra-political domain. Ironically, however, the explanatory power that this metaphoric gesture obtains derives primarily from its paradoxical formulation; indeed, we might argue that the reason the *milieu intérieur* performs such profound imaginary work at and as the threshold of modern biomedicine is that it names a contradiction [49]. By folding the environment back into the organism, Bernard transforms the domain within which biomedicine operates. Even as he insists that organisms are necessarily environmentally determined, he incorporates the liberal abstraction that (human) individuals are "free and independent" as a biological explanation for organismic function and thereby produces a theory that "naturally" consol-

idates an atomistic understanding of "the organism." Yet, rather than undermining the force of his concept, the antithetical possibilities recognized (or re-cognized) by this biopolitical formulation are manifestly productive ones: they imagine the paradoxical situation of the organism in its life context as one of determinant indetermination. The difficulty with Bernard's formulation then is not that it is contradictory, which is not a scandal, but rather that it fails to admit this contradiction as such and instead resolves this tension into an ideological affirmation of the higher organism's "independence." In so doing, it clearly privileges individuation as the "highest" of all possible modes of organismic differentiation and thereby ideologically forecloses the possibility for a more nuanced scientific understanding of the organism-milieu dyad [50]. Moreover, by naturalizing autonomy and freedom as biological rather than political values and then defining them as the aggregate effects of the *milieu intérieur*, Bernard emphasizes the defining functions of protection over nutrition, separation over contiguity, independence over dependence, all of which foreshadow the triumph of biological immunity over political community as a medical world view.

Immune from incriminating disease, or the law of non-recidivism

Following from this biopolitically decisive and defensive formulation, it is fairly easy to discern the penumbra of immunity emerging from Bernard's assertion that: "External influences, therefore, bring about changes and disturbances in the intensity of organic functions only in so far as the protective system of the organism's *milieu intérieur* becomes insufficient in given conditions" [51]. While he himself did not adopt the position that subsequently became identified with Pasteur's germ theory, preferring until the end of his life to comprehend disease as a general imbalance in the internal environment rather than as the specific effect of particular microorganisms, Bernard's work nevertheless provided the philosophical and political bedrock upon which Pasteur's theoretical and institutional edifice was built [52]. For, whatever its many scientific merits, germ theory also emerges in the middle of the nineteenth century as a defensive formulation specifically designed to impugn the credibility of the transformative multifactoral causality that underwrote both the promises of spontaneous generation and the premises of public hygiene [53]. Needless to say, this theory opened the way for a radical reconceptualization both of infectious diseases and of the living hosts in which these germs of pathogenesis flourished.

A chemist who trained his professional sights on the biochemical effects of microorganisms, Pasteur appears not to have had much respect for or interest in the host organism, except abstractly as the medium of bacterial growth [54]. Instead of focusing attention on the host organism as the locus of disease, Pasteur and the Pasteurians fetishize the microbe as the agent or "*auteur*" of disease and then seek to manipulate its pathogenic qualities by "culturing" it in less virulent forms. Taking the diseased organism as a "*milieu de culture*" (an appropriation from Bernard's *milieu intérieur*), Pasteur imagines disease as the ecological disturbances effected by a pathogenic agent within an organic "culture." However, Pasteur was not much concerned with the phenomenon of disease per se as it affected the animals who served as his experimental subjects (the flocks of chickens and sheep who lost their lives to the cause), but rather with its propagation and transmission by microbiotic vectors. Thus, Pasteur's project focused almost exclusively on manipulating strains of bacteria that had been correlated with recognizable patterns of symptoms (a.k.a.,

"diseases") in the hope of mitigating the pathogenic effects produced and reproduced by their movements in the "external" world.

To the extent that Pasteur did consider disease, then, he did so only in terms of how the activity of microorganisms transformed the enclosed economy of the organism. In his 1880 discussion of chicken cholera, for example, he attempts to account for the effects of microorganisms on the dying chicken as a struggle for scarce resources in the *milieu intérieur.*

> By the acts of its nutrition, the microbe produces the gravity of illness and leads to death. One can easily comprehend it. The microbe, for example, is aerobic; it absorbs great quantities of oxygen and it burns many of the elements [principes] of its culture medium [*milieu de culture*] (of which it is easy to assure oneself by comparing the extracts of chicken bouillon before and after the culturing of the little organisms). Everything indicates that it takes the oxygen that is necessary for its life from the globules of blood, across the blood vessels; and the proof of this is that during its life, and often even as it approaches death, one sees the comb of the sick chicken turning purple, while the microbe no longer exists in its blood, or only exists in such a minuscule quantity that it escapes observation. This genre of asphyxia would be one of the most curious traits of the disease that occupies us, if it was proved that it can not be attributed to a difficulty in circulation brought on by the disease itself [55].

Pasteur interprets the disease he has induced in his laboratory subject by identifying the animal with the equipment of the laboratory itself. For Pasteur not only is the chicken's living metabolism reduced to a feathered *milieu de culture*, but the chicken's living per se is reduced to its status as experimental subject. In this restricted context, it makes complete sense to understand disease as the consequence of the activity of the asphyxiating microbes, since only the microbe – and Pasteur, of course – have any agency at all in this scenario. Yet by taking account of the mise-en-scène, we can begin to understand what is so ostentatiously bracketed in Pasteur's thinking, that is, the sociohistorical frame in which he has created the very illness that he seeks to describe by subjecting a chicken to inoculation with a virulent strain of bacteria which he has cultured precisely in order to kill it. The scarce resources (which in the case of the chicken could actually have been renewable resources) that the microbes supposedly deplete in their nutritive frenzy, thus appear only as artifacts of the experiment in which the living processes of the disease's host are rendered insignificant. In other words, disease denotes for Pasteur the depletion of the living capital ("*les principes*") that sustains the organism and for which it is necessarily in competition with all other organisms in a world in which it will receive no sustenance from outside itself.

Following this logic, Pasteur understands the effects induced in animals inoculated with attenuated strains of "disease-causing" bacteria, whose virulence he had manipulated in the laboratory, as due to pre-emptive modifications in the *milieu de culture* which make it less hospitable to the more virulent, naturally occurring strains (much as Cotton Mather had postulated 150 years earlier).

> ... The condition of the existence and multiplication of microbes, the causes of virulent illnesses, is that they find in the inoculated organism or the culture medium [*milieu de culture*] where one introduces them elements for their nutrition; the proof is that if one filters the chicken bouillon which served to cultivate the microbe of

> chicken cholera, this bouillon becomes improper for a new culture of the same organism, whereas it can still serve to cultivate other microbes, anthrax bacteria for example. Why? Because, in all probability, the first culture exhausted the elements necessary for the life, the multiplication of the microbe of chicken cholera, and not those necessary for the bacteria.
>
> If this happened in my test tubes, couldn't it equally happen in the animal organism, in the human body? [56]

This formulation provides one of the most explicit statements of Pasteur's understanding of why the practice of inoculation might work. Based on the evidence of bacteria in a culture medium, whose exponential growth is cut short at the point where the population has depleted necessary nutrients – itself a microcosmic version of Malthus's tendency/limit problem – Pasteur reduced the domain of the host organism to that of a giant flask. If the growth of bacteria in a laboratory is imagined as the same as the growth of bacteria in a living organism, then the problem of disease lies entirely on the side of the microorganism. Moreover, to conceive of the organism as a passive "*milieu de culture*" suppresses the fact that the "*milieu de culture*" is itself necessarily a social artifact, produced within a laboratory through the intersection of human agents, equipment, and microbes, not to mention the flows of capital and information within which all these are experimentally articulated in the first place. However, as the terms he used to characterize the desired effect of inoculation suggest, Pasteur also relied on the slippage between the multiple senses of "culture," as medium, as activity, as human nexus [57], to make this understanding plausible.

> As for the cause of non-recidivism [a term Pasteur uses interchangeably with immunity], one cannot refrain from the idea that the microbe, author of the malady, finds in the body of the animal a *milieu de culture* and that in order to satisfy its own life, it alters or destroys, which amounts to the same thing, certain materials whether it prepares them for its profit or burns them with the oxygen it borrows from the blood. When complete immunity is attained, one can inoculate the most virulent microbe in whatever muscles without producing the least effect, which is to say that all culture becomes impossible in these muscles. They do not contain any food for the microbe [58].

The interchangeable terms "non-recidivism" and "immunity" stud Pasteur's many texts on chicken cholera and anthrax, marking the overdetermined "cultural" implications of Pasteur's analytic scenario: for Pasteur, the disease-authoring bacteria wrongfully "profit" from the exploitable resources of other living bodies which they appropriate without remuneration. He reiterates and proliferates these terms, along with a third less favored metaphor, "refractory," in his attempt to represent the effects induced in organisms inoculated with his attenuated forms of microorganisms. Besides implying that disease is tantamount to a "natural" crime (i.e., if it recurs it is a repeat offense, *une récidive*) these metaphors underscore the inevitable coincidence of Pasteur's bio-scientific and biosocial interests: through his interpretive efforts to make the results of his laboratory practices intelligible, Pasteur produces and reproduces a new socially pathogenic element, "the germ," whose domain of predation is the living body. Indeed, throughout the explanations he gave of his famous experiments, Pasteur consistently prefers the now scientifically forgotten notion "non-recidivism" over the still current and highly prized scientific concept "immunity" as a way of elucidating the effects that he thought he had induced by inoculating his exper-

imental subjects. So much so, in fact, that by the Spring of 1881 he was referring to "the general law of the non-recidivism of virulent illnesses" (*la loi générale de la non-récidive des maladies virulentes*) [59]. If Pasteur initially comprehends diseases in terms of whether or not they were recidivists or non-recidivists, what difference does this make to our understanding of the biological immunity which both derives from and supersedes it? In other words, how do Pasteur's metaphors shape the scientific and political effects induced by the experimental imaginary that underwrites the inoculation-vaccination-immunization projects which he set in motion?

As Anne-Marie Moulin points out: "Pasteurian medicine doesn't rely upon theoretical hypotheses concerning immunity, but on an empirical program of immunization which appeals to the attenuation of micro-organisms" [60]. Nevertheless, the theory implied by Pasteur's choice of metaphors suggests that he interprets the effects which pathogenic bacteria produced in the experimental subjects upon which they were inflicted as an analogue of repeated criminal behaviors. Certainly ascribing the repetition of socially and legally reprehensible activities to microorganisms, along with the attendant implications of patterned criminality whether innate or conditioned, is not what Pasteur had in mind. Yet by invoking the category of recidivism as a privileged rubric within which he sought to comprehend the mechanism of his empirical program, Pasteur arrogated the putatively "natural" activity of the bio-scientifically engineered micro-organisms which he injected into his laboratory subjects for the social project of which he was a part. In other words, Pasteur's analysis of his laboratory results isolated the social effects of disease in the bacteria which were defined as its cause (or "author") and not in the larger biopolitical ecology in which they produced and reproduced their pathogenic symptoms – an ecology which contained not only these microorganisms and the multi-cellular organisms into which they were introduced, but also the laboratories in which the experiments were conducted, the scientists who conducted them, and the public who applauded their efforts. By recognizing the local effects of inoculation through the manifestly juridical concepts recidivism and immunity, Pasteur simultaneously ascribed the political dimensions of disease to the activity of microorganisms and localized the geopolitical domain of disease within the *milieu de culture* constituted by the host organism. Immunity functions only vaguely here since Pasteur is only marginally concerned with living qualities of the host; immunity thus serves Pasteur as a non-specific quality derived from the non-recurrence of disease. Instead, recidivism, a category associated from the first half of the nineteenth century on with the statistical analyses of prison populations, delimits the social relevance of infectious diseases to the activities of micro-organisms within the diseased body.

Not one to shy away from expanding the domain of his influence, Pasteur did not hesitate to expound the overtly political implications of his paradigm. For example, in testimony given before the Counseil d'Hygiène Publique et de Salubrité on March 9, 1888, concerning plans for the construction of a new sewage treatment facility, Pasteur unabashedly declaims:

> A new science has been born whose progress has been such that it has imposed its superior teachings in all the universities of the world. Under its momentum, surgery and medicine have transformed their therapeutic methods. It has revolutionized our knowledge of virulent and contagious illnesses and these illnesses comprise the majority of pathologies (if one excepts nervous or hereditary illnesses). Now the

> principle which dominates all microbiology is the following: virulent and contagious diseases are never spontaneous; they have for their origin a living ferment of disease, living a proper life, a microbe, and the spontaneity of the life of these microbes is as chimerical as the spontaneity of the lives of large animals or plants. Suppress the microbe of typhoid fever, of diptheria, of scarlatina, of measles, of glanders, of anthrax, of cholera, etc., or place them in conditions in which they can do no harm, and you will never find a single case of these diseases. Whatever the conditions of life, of the physiological misery of the individual, never by its own nature, never can it create the illnesses of which I have spoken nor be affected by them whatsoever.... All virulent and contagious diseases arise from the presence and the development of microscopic beings. These are ineluctable facts [61].

While the "ineluctable facts" that Pasteur evokes were indeed persuasive, there was a major flaw in his not so ineluctable logic. For even if "all virulent and contagious diseases arise from the presence and the development of microscopic beings," it is not the case that all those exposed to the same "living ferment of disease" will actually succumb to the same deleterious effects. Outside of laboratories or science fiction scenarios, it is a very rare microbe indeed that can produce one hundred percent morbidity in all those who are exposed to it. Given variable susceptibility of organisms outside the controlled conditions of his laboratory, Pasteur was a bit premature in asserting that "[w]hatever the conditions of life, of the physiological misery of the individual, never by its own nature, never can it create the illnesses of which I have spoken nor be affected by them whatsoever." Yet, by emphatically negating the "conditions of life" and "physiological misery" as possible causal factors, Pasteur effectively asserts the superiority of his uni-causal explanation for disease over and against all multi-factorial theories. He seeks to restrict the domain of medical and bio-scientific interest to the microbiotic domain because this is the locus within which he has the potential to maximize control. Thus by restricting his attention to what was at the time the smallest possible scale of intervention, he is able to produce global results, results lauded "in all the universities of the world." The germ becomes the privileged object for Pasteur because it allows him to articulate multiple levels of determination defined across varying scales ranging from the micro- to the macro-cosmic. Both the extremity of Pasteur's political claims and the fallibility of his logic, then, derive from his assumption that microorganisms are the sole agents of pathogenesis to the exclusion of any consideration of how "infected" organisms might themselves enter into complex ecological interactions with microscopic beings.

As Bruno Latour so deftly illustrates, the events inscribed under the name of Pasteur introduce a new political agent, the germ, into the social realm. Given the triumphal displays that heralded its arrival, the germ's significance quickly percolated through the biological imaginary – both scientific and popular. Its infectious power was enthusiastically acknowledged even, or even especially, among the advocates of public hygiene, who had long sought a convincing environmental scenario for contagious diseases, despite the fact they had the most to lose by doing so. Needless to say, the germ rapidly displaced the multiple causal factors through which the public hygiene movement sought to take into account the ecologies of diseases, so that the living conditions of the infected organism came to seem irrelevant except in so far as they facilitated or inhibited the transmission of pathogenic microbes. More specifically, it was on the basis of this new agent, which Latour dubs "the cultivated-microbe-whose-virulence-they varied," that a new biosocial

articulation of bodies, subjects, diseases, economies, and interests could be incorporated. Yet more than just an efficacious re-articulation of multiple levels of biosocial flows, germ theory also defines the passive interiority of the organism as the primary space within which the political and economic struggle for survival takes place. Like Malthus before him, Pasteur rejects the notion that socioeconomic factors significantly impinge on the health of populations and instead isolates the "true" experience of infectious disease within the epidermal envelope of the singular organism. In so doing, he re-articulates the domains of the natural and the social, so that the former seems to saturate the passive interiority of the organism, while the latter envelops and contains the organism from without. By bracketing the biosocial circumstances of the organism into whose life the experience called infectious disease has emerged, Pasteur privileges the microbes over and against both those circumstances that may render living beings vulnerable to infection and the vital exchanges which may enable those infected to heal.

Immunity in action

Even at the time of his greatest triumphs, however, there were indications that there might be difficulties with asserting the uncontested validity of Pasteur's analysis: there was neither an account of how infected organisms participate in disease processes, nor an understanding of how afflicted organisms survive an infectious illness, nor a way of explaining the enduring resistance to disease that inoculated, or recuperated, subjects maintained. Moreover, from the mid-1880s on, other researchers were also challenging the ineluctable facts of Pasteur's practice by demonstrating, for example, that effective vaccination is possible with dead bacteria (Salmon and Smith) or even in the absence of bacteria altogether (von Behring and Kitasato). These were plaguing problems indeed for the advocates the new Pasteurian bio-technologies. While the Pasteurian project was, as Moulin notes, primarily an empirical one, nevertheless, its generalization as both a biomedical and biotechnical enterprise depended upon the possibility of extending its range of applicability. Yet in Pasteur's hands, immunity had not sufficiently crystallized as a robust trope that could serve as the basis for the expansive research program he envisioned. Or, to put it more concretely, the "poetic borrowing" [62] that immunity constituted within the Pasteurian project did not provide a sufficient foundation for the magnificent edifice that Pasteur was in the process of erecting for himself: namely, the Institut Pasteur. Thus, Pasteur-the-empiricist needed a theory that would allow him to account for the specific effects induced by the bacterial attenuation techniques which he had introduced in order to capitalize upon the resulting "immunity" that he claimed to have induced.

To his great good fortune, Pasteur was able to redress this plaguing problem by importing and then incorporating the leading figures of the field that would soon become known as immunology: Elie Metchnikoff and his phagocytes. A Russian zoologist with an abiding interest in the comparative embryology of invertebrates, Metchnikoff became internationally renowned during the 1880s for proposing the first account of immunity as an organismic activity, which he understood as a cellular struggle for survival and among whose primary functions he included not only organismic protection but also with the disposal of cells that "degenerate" in the course of organismic function. Drawing on his interest in a group of leukocytes (macrophages) that he called "phagocytes," or "eating cells," which were amebic cells capable of encompassing and degrading both microorgan-

isms and cellular detritus found in the *milieu intérieur*, Metchnikoff introduced his theory of immunity in 1883. He identified the effects of intracellular digestion in meta-cellular organisms as the paradigmatic example of "host defense," proffering a concept that radically transformed scientific medicine by redrawing its epistemological maps of the organism's bio-social ecology. Thanks to Metchnikoff, the "host" organism came to be imagined as a materially localized entity, inscribed within a recognizable frontier, whose "immunity" from the biosocial context of infection appeared within the furthest limit of its ability to marshal a defense against the invasive forces of marauding parasites. Needless to say, Pasteur enthusiastically placed his imprimatur upon Metchnikoff and his phagocytes: he published an essay by Metchnikoff in the inaugural issue of the *Annals of the Institut Pasteur* and provided a French home for the Russian embryologist under the auspices of his newly constructed research facility (where Metchnikoff remained ensconced for the next 30 years), thereby annexing Metchnikoff's insights for his own inoculation-vaccination-immunization regime – a win-win situation if there ever was one.

By taking inflammation as a particular manifestation of a more general organismic response, which was itself seen as an evolutionary holdover from the intracellular digestion of single-celled organisms, Metchnikoff redefined biological and bio-medical interpretations both of what an organism was and of how it maintained its integrity within its life context. Bringing his keen appreciation for organismic processes to bear on the insights of bacteriology (for which, as noted earlier, the diseased organism had heretofore functioned primarily as a passive context for the propagation of microbiotic agents) Metchnikoff transformed the concept of biological immunity from a vague metaphoric gesture towards the inhibition of microbiotic recidivism and into the robust basis for a "theoretically articulated experimental research program," that is, immunology [63]. What Metchnikoff offered to biomedicine, then, was no less than a way to account for what Bernard had deemed the organism's "freedom and independence" from its biosocial environment through an activity that at once produced and reproduced its integrity and its boundary. Finessing the problematic Pasteurian basis for microbiotic non-recidivism, Metchnikoff instead redirected the question by framing it as the effect of a process whose telos was manifesting the autonomy of the organism itself, as Latour notes: "The microbes were becoming particular cases of a general problem: the integrity of the organism" [64]. In the wake of Metchnikoff, immunity ceased to function as an ad hoc appropriation of an ancient juridico-political concept; instead, in an amazingly short period of time, it came to operate within bio-medical discourse as a robust metaphor for an active principal of organismic existence whose province was no less than championing the life of the organism itself against the deleterious effects of a hostile environment.

Bibliography

1 Kirmayer L. Mind and body as metaphors: hidden values in biomedicine. In: Lock M, Gordon DR, Eds. Biomedicine Examined. Boston: Kluwer; 1988. p. 82.
2 Foucault M. The birth of the clinic: an archaeology of medical perception. Sheridan Smith AM, trans. New York: Vintage; 1973. p. 198.
3 Haraway DJ. Modest witness @ second millennium. Femaleman meets oncomouse: feminism and technoscience. New York: Routledge; 1997. p. 175.
4 Neuberger M. Doctrine of the healing power of nature throughout the course of time. Boyd LJ, trans. New York; 1932 [1926]. p. 115-80.
5 Tauber A. The immune self: theory or metaphor? Cambridge: Cambridge University Press; 1994. p. 68.

6 Nietzsche F. Truth and lie in an extra-moral sense. In: Levy O, Ed. The complete works of Freidrich Nietzsche. New York: Russell and Russell; 1964. ii. p. 173-85.
7 Haraway DJ. Modest witness. p. 83, 39.
8 Golub E. The limits of medicine: how science shapes our hope for the cure. Chicago: University of Chicago Press; 1994. p. 178-9.
9 Figlio K. The historiography of scientific medicine: an invitation to the human sciences. Comparative Studies in Society and History 1977; 19: 278.
10 Silverstein A. A history of immunology. San Diego: Academic Press; 1989. p. 1-3.
11 Moulin AM. Le dernier langage de la médecine: histoire de l'immunologie de Pasteur au Sida. Paris: Presses Universitaires de France; 1991. p. 24.
12 Foucault M. The birth of the clinic. p. 159.
13 Foucault M. ibid, p. 155.
14 Chrestien A. De l'immunité et de la susceptibilité morbide, au point de vue de la clinique médicale. Montpellier; 1852. p. 2.
15 Chrestien A. ibid, p. 7.
16 Mather C. The angel of Bethesda. Barre, MA: American Antiquarian Society; 1972 [1724]. p. 98.
17 Mather C. ibid, p. 112.
18 Jenner E. Vaccination against smallpox. An inquiry into the causes and effects of the variolae vaccinae, or cow-pox. Amherst, NY: Prometheus Books; 1996 [1798]. p. 73-4.
19 Jenner E. ibid, p. 13.
20 Paine T. The rights of man. Garden City, NY: Anchor Press; 1973 [1791]. p. 304.
21 Agamben G. Homo sacer: sovereign power and bare life. Stanford: Stanford University Press; 1998. p. 187-8.
22 Malthus T. An essay on the principle of population, or, a view of its past and present effects on human happiness; with an inquiry into our prospects respecting the future removal or mitigation of the evils which it occasions. In: Himmelfarb G, Ed. On population. New York: Modern Library; 1960. p. 8.
23 Canguilhem G. The development of the concept of biological regulation in the eighteenth and nineteenth centuries. In: Goldhammer A, trans. Ideology and rationality in the history of the life sciences. Cambridge, MA: Harvard University Press; 1988. p. 93.
24 Malthus T. An essay. p. 129.
25 Gallagher C. The body versus the social body in the works of Thomas Malthus and Henry Mayhew. Representations 1986 ; 14 : 85.
26 Louis-René Villermé. Des épidémies sous les rapports de l'hygiène publique, de la statistique médicale et de l'économie politique. In: Delaporte F, Ed.; Goldhammer, A. trans. Disease and civilization: the cholera in Paris, 1832. Cambridge: MIT Press; 1986. p. 53.
27 Coleman W. Death is a social disease: public health and political economy in early industrial France. Madison: University of Wisconsin Press; 1982.
28 Foucault M. Faire vivre et laisser mourir. Les Temps Modernes 1991; 535: 37-61.
29 Latour B. The pasteurization of France. Sheridan A, Law J, trans. Cambridge, MA: Harvard University Press; 1988. p. 18.
30 Wohl A. Endangered lives: public health in Victorian Britain. Cambridge, MA: Harvard University Press; 1983. p. 118.
31 Foucault M. Faire vivre et laisser mourir. p. 44.
32 Latour B. The pasteurization of France. p. 21.
33 Coleman W. Death is a social disease. p. 202.
34 Haraway D. The promise of monsters. In: Nelson C, Treichler P, Eds. Cultural studies. New York: Routledge; 1992. p. 311.
35 La Berge A. Mission and method: the early nineteenth-century French public health movement. Cambridge, MA: Harvard University Press; 1992.
36 Latour B. The pasteurization of France. p. 33.
37 On Pasteur's pedagogical debt to Bernard, see Debré P. Louis Pasteur. Forster E, trans. Baltimore: Johns Hopkins University Press; 1998. p. 348-61.
38 Canguilhem G. Études d'histoire et de philosophie des sciences. Paris: J. Vrin; 1968. p. 132.
39 Canguilhem G. Études. p. 150. See also: Wasserstein A. Death and the internal milieu: Claude Bernard and the origins of experimental medicine. Perspectives in Biology and Medicine 1996; 39: 319.
40 Holmes FL. Claude Bernard and the milieu intérieur. Archives Internationales d'Histoire des Sciences 1963 ; 65 : 369-76; Milieu intérieur and cell theory, Bulletin of the History of Medicine 1963; 3 : 315-35; Origins of the concept of *milieu intérieur*, In: Grande F, Visscher, MR, Eds. Claude Bernard and experimental medicine. New York: Harper and Row; 1967. p. 179-91.

41 Bernard C. Cours de physiologie générale de la Faculté des sciences. [1854] In: Holmes FL, Ed. Origins. p. 184.
42 Bernard C. Leçons sur les propriétés physiologiques et les altérations pathologiques des liquides de l'organisme. (1859) In: Holmes FL. Origins. p. 186.
43 Margulis L, Sagen D. What is life? New York; 1995, p. 26.
44 Canguilhem G. Le vivant et son milieu. In: La connaissance de la vie. Paris: J Vrin; 1992. p. 129-54.
45 Bernard C. An introduction to the study of experimental medicine. Greene HCG, trans. New York: Henry Schuman; 1949. p. 62.
46 Bernard C. Lectures on the phenomena of life common to animals and plants. Hoff H, Guillemin R, Guillemin M, trans. Springfield, IL: Charles Thomas; 1974. I. p. 83.
47 Bernard C. Lectures. p. 84.
48 Bernard C. Pensées: notes détachées. Paris: Librairies JB Baillière & Fils; 1937. p. 76-7.
49 Holmes FL. Origins. p. 187.
50 Simondon G. L'individu et sa genèse physico-biologique. Grenoble: Jérôme Million; 1995. p. 21-34.
51 Bernard C. Introduction. p. 62.
52 On the Pasteur-Bernard dynamic: Moulin AM. La rivalité des pères de la médecine moderne reste bien vivante. La Recherche 1995 ; 26 : 906-11; Geison G. The private science of Louis Pasteur. Princeton: Princeton University Press; 1998. p. 18-21.
53 Farley J. The spontaneous generation controversy from Descartes to Operin. Baltimore: Johns Hopkins University Press; 1974. Also, Geison G. Private science. p. 110-42.
54 Latour B. The pasteurization of France. p. 103.
55 Pasteur L. Oeuvres de Pasteur. Vallery-Radot P, Ed. Paris; 1922–1939. VI. p. 310.
56 Pasteur L. ibid. p. 295.
57 Williams R. Keywords: a vocabulary of culture and society. New York: Oxford University Press; 1985. p. 43-5.
58 Pasteur L. Oeuvres. VI. p. 305.
59 Pasteur L. ibid, p. 339.
60 Moulin AM. Le dernier langage. p. 47.
61 Pasteur L. Oeuvres. VII. p. 125.
62 Moulin AM. Le dernier langage. p. 36.
63 Tauber and Chernyak. Metchnikoff and the origins of immunology: from metaphor to theory. New York: Oxford University Press; 1991. p.148.
64 Latour B. The pasteurization of France. p. 107.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

The Lübeck catastrophe and its consequences for anti-tuberculosis BCG vaccination

Philippe Menut

Introduction

The anti-tuberculosis vaccine, BCG ("Bacille de Calmette Guérin", from the name of its co-discoverers) is the product of the attenuation of bovine Koch's bacillus [1]. Benjamin Weill-Hallé and Raymond Turpin first used the vaccine on humans at the Charity Hospital in Paris in 1921. There then followed a long series of controversies on the safety and effectiveness of the vaccine [2] punctuated by favorable judgements: in 1928, the Commission on Hygiene of the Society of Nations meeting in Paris officially declared that the vaccination was inoffensive (the vote was unanimous excepting the Austrian professor Nobel) [3]. In 1931, the National Academy of Medicine in France reaffirmed the safety of the French vaccine [4].

In May of 1930, the international opinion was alerted by a serious BCG accident in Lübeck, Germany, following a large vaccination campaign. Until then, Germany had remained suspicious of the French vaccine and experiments with human vaccination had remained local and limited. Only H. Buschmann of Bleialf in a rural area near Bonn, C. Prausnitz of Breslaw and I. Zadek of the Berlin Neuköln had used the vaccine in human medicine [5]. The accident resulted in a long trial during which BCG's qualities and the vaccination question were discussed against the background of rising nationalism.

This catastrophe has left an indelible mark on the history of vaccination using BCG. By returning to this historical episode we hope to show that it can teach us something about the functioning of medicine and its relationship to bacteriology. The Lübeck vaccination debate is, in itself, an interesting example of the history of vaccination in general.

The prelude to the catastrophe: the introduction of anti-tuberculosis vaccination using BCG

In July of 1929, spurred on by Councillor Bielefeld, the Lübeck Parliament decided to introduce BCG vaccination. Conferences were organized in order to convince physicians and midwives. Preparation of the vaccine was delegated to Deycke (1865–1937), a former military physician and well-known bacteriologist. Alsteadt, a young physician and student of Deycke, and who had just taken charge of the local hygiene services, requested a BCG culture from the Pasteur Institute in Paris [6]. Vaccination began on 26 February 1930 following a propaganda campaign in the local press. Parental accord was required for the vaccination, even though the propaganda had carefully avoided using the term "vaccina-

tion." Parents who accepted the vaccination for their children received money as did midwives who decided to vaccinate, with the vaccination being administered orally [7]

On 6 March, the first child died at the hospital in Lübeck, although a link to vaccination was not established [8]. It was necessary to wait until a third death on 26 April before the physicians decided to stop the vaccination. Up until May 2 a "remedy" put together by Deycke in order to "avoid at all costs that other children be vaccinated" was used [8, 09/25/30]. The public was not notified until 13 May 1930 by the Senate of Lübeck. On 28 May 1930, Deycke and Klotz (the Director of the Lübeck Hospital), Alsteadt, and a nurse implicated in the preparation of the vaccine were accused of negligence [9]. By the end of 1932, of the 256 children vaccinated, 77 were dead and many were ill [7, p. 67].

The trial that followed lasted 76 days and called upon numerous experts, two of which were sent by the government: Bruno Lange of the Robert Koch Institute in Berlin and Ludwig Lang of the Ministry of Health. The other experts, convoked by the plaintiffs and the prosecutors, included Schmincke, a hygienist, Hans Much (1880–1933), Paul Ulhenhuth (1870–1957) of the Institute of Hygiene of Freiburg University, Martin Hahn (1865–1934) of the Hygiene Institute of the University of Berlin, Wilhelm Kolle (1868–1935) of the State Institute for Experimental Therapy in Frankfurt, Rudolph Abel (1868–1942) of the Hygiene Institute of the University of Jena, all bacteriologists, Schürmann (1895–1941), an anatomo-pathologist, and Poll, a biologist. All, with the exception of the biologist, were recognized as specialists in bacteriology and, in particular, as applied to tuberculosis. All had conducted experiments using Cox bacteria, some had even worked with BCG. On the other hand, none had practiced vaccination with BCG in human medicine. Moreover, with the exception of Schmincke, a communist, all recognized the high quality of Deycke's work and his merits as a person "who had devoted his life to medicine" [8, 12/19/31]. Although the defenders of vaccination were in the minority, the opponents of the procedure were divided: either the BCG was not safe, in which case, the question of its efficacy was superfluous (Ulhenhuth) [8, 10/14/31] or it was safe, but useless (Kolle) [8, 10/14/31].

The problems of expertise

The tribunal asked the experts two principal questions: what was the causal agent of the tuberculosis that appeared in the vaccinated children? Had there been willful negligence involved in the preparation of the vaccine?

In order to prepare their expert testimony, the scientists had very little material: The bacterial strains that had been used in the vaccination had been destroyed just after the accident. All that remained was a culture labeled "BCG-143 Deycke" that was taken to be a sample of the local BCG. In addition to this were cultures that had been obtained from the vaccinated children who were either dead or ill. Finally, the material was completed by a strain of virulent bacteria from Kiel, which was supposed to have been the only human strain present in the Lübeck laboratory [8, 11/24/31].

Schürmann was in charge of the autopsies and established the existence of a tubercular process of an intestinal origin based on the presence of intestinal inoculation cankers [10]. This latter argument did not garner unanimity amongst the experts. In particular, Much opposed this interpretation and quoted a letter from Joseph Koch of the Robert Koch Institute in Berlin in support of his opinion [8, 11/41/31].

Based on epidemiological observations, Ludwig Lange distinguished four different periods characterized by different profiles of morbidity and mortality [11]. He concluded that the children had in fact been vaccinated by different products in the course of the campaign, thus dragging the debate back into the bacteriological terrain. From this, a series of hypotheses emerged to explain the accident. Either it really was the BCG that was at the origin of the accident and its virulence had evolved in the course of the campaign (return to virulence hypothesis) or there had been a mix of BCG and a virulent strain in variable proportions in the course of the campaign (the contamination hypothesis) or, finally, the children had been vaccinated with something other than BCG (the substitution hypothesis). As for the hypothesis that the product had arrived contaminated from Paris, this was rejected at the very beginning of the trial. Sister cultures of the Lübeck culture had, in effect, been used in Riga by Kirschenstein, in Mexico by Castrejon, and in France without any accident whatsoever [12]. This useful information had been rapidly sent by Calmette, which had enabled the German experts to avoid a painful and costly voyage to Paris to the BCG Laboratories and had preserved the local character of the catastrophe!

The experts positioned themselves with respect to these three hypotheses according to their bacteriological practices. Two styles of thought confronted each other in the expert testimony. Either the expert was convinced of the existence of bacterial species or he was not.

For the defenders of the idea of fixed species of bacteria, experiments that increased the virulence of BCG were not considered conclusive and were mainly regarded as results obtained by isolated and marginal researchers. The defenders accorded greater importance to Calmette's results at the Institute Pasteur in Paris or the Reich's Ministry of Health. The two experts delegated by the Reich, Lange and Lange, belonged to this line of thought. From the outset, they oriented their expertise towards the comparison of the Kiel virulent human strains with those that had been isolated from the organs of victims. They categorically exonerated BCG from any responsibility. Moreover, since bacteriology typed the different forms of the Koch bacillus on the basis of differential virulence, BCG, which was reputed to be completely attenuated, escaped any form of determination! What is more, their expertise showed that in the majority of the cases studied, it was difficult to determine which bacteria was present, the virulence being variable and labile [13].

Confronted with the impossibility of identifying precisely the agent responsible, the experts proposed three different methods that gave convergent results. The first, proposed by Bruno Lange, was based on a syllogism. Bacteria isolated from the victims had a labile virulence. Now, it is rare to encounter germs presenting such lability. The Kiel strain possessed this characteristic and the strains taken from the vaccinated children did too. The expert therefore deduced the identity of the Keel strains and those taken from the vaccinated children with a probability approaching certitude [8, 12/16/31]. The second method proposed by Ludwig Lange was based on the study of the singularities of the Kiel strain. During the trial he showed the jury a strain of the Kiel bacterium that presented a green coloration. The strains isolated from wet nurses in Lübeck possessed this same characteristic [8, 12/16/31]. How had the experts ignored such a property during the long months of their inquiry? The last approach, biochemical, was a result of studies done by the soon to be famous biochemist, Erwin Chargaff, a collaborator of Han. Both approached the question by comparing the lipoid content of the different Kiel, cow, human and

BCG strains. They considered the strain taken from sick or dead children to belong biochemically to the Kiel type [8, 12/22/31]. This new technique had, in fact, been developed for the trial.

Their opponents, that is, those who opposed the idea of fixed bacterial species, used different material and often cited work done by BCG's detractors whether they were well known or not. The latter advanced as proof of the possibility of bacteria regaining virulence cases of BCG. They considered the introduction of vaccination a grave error. These experts were aided in their study by work done by Freidmann, who had supported the publication of a brochure on the 150 opponents to Calmette's method [8, 10/15/31]. Freidmann, in fact, was a promoter of his own vaccine, the para-tubercular bacillus of the sea tortoise (BBM). He used the trial to promote his vaccine in order to gain a market share in Germany and elsewhere [14]. The most important opponent of the idea of fixed bacterial species was Much, a former collaborator of the Nobel Prize winner Emil Von Behring, who used intervention to launch an attack against "scholastics and dogma in medicine" [8, 11/14/31]. He was opposed to dogma, and particularly the dogma of the stability of bacterial species which he considered little more than an "*idée fixe*" [8, 12/14/31]. What counts, he claimed, is "the particular constitution of the organisms infected" or in other words, the terrain [8, 12/14/31]. The research he had undertaken for the Lübeck trial enabled him to produce virulent strains of BCG in organisms that had been experimentally prepared. He obtained such results by jointly injecting lactic acid and BCG [8, 12/14/31]

He found a momentary ally in Poll, a biologist. The latter remarked that there exists modifications "that we do not know how to measure exactly" in a living world. For Poll, a culture only represented a mean that never gave the value of the extremes. Modifications, therefore, could be produced when the conditions were favorable. The material situation in Lübeck, for example, and, in particular, the recourse to milieux that had not been recommended could have created such a modification. He acknowledged nonetheless that science could not recognize this modification [8, 12/15/31].

Through the confrontation of these two types of expertise it was in fact bacteriology and its medical applications that were directly threatened. If germs could vary, then their specification would become an illusion and it would therefore be impossible to do any good through the use of attenuated bacteria. For the more orthodox bacteriologists, the calling into question of this dogma inverted the order that bacteriology had brought to medicine [15]. How, in other words, could one have predicted the characteristics of a BCG strain that had become once again virulent?

The crime site and the weapon

If there had been substitution or contamination, it remained to be determined where and when (one or some) pathogenic bacteria had been able to penetrate the vaccine ampoules. The first anomaly noticed by the experts was that the BCG in Lübeck had not been produced according to the rules set down by the Pasteur Institute in Paris and which accompanied every strain of bacteria sent out from Paris. They had not used the same milieu for the culture, the glassware was not that which had been recommended, the spatial separation of the activities involved in the vaccine preparation and other laboratory activities had not been respected [8, 12/11/31]. The experts also noted that Deycke had not closely followed his nurse's preparation of the emulsions.

Nonetheless, an on-site laboratory visit by the Tribunal showed that the material conditions were sufficient. A military-style order reigned in the laboratories and Ludwig Lange himself admitted that he would have happily assumed the responsibility of producing the vaccine in such conditions [8, 12/12/31].

On 28 October 1931, a tribunal demanded that Nurse Schutze repeat the preparation of the partial antigens of the Koch bacillus in the presence of experts who criticized many of her gestures. "She held the balloons vertically making contamination possible" declared Bruno Lange during the trial [8, 12/11/31]. Moreover she used etiquette that easily fell from the tubes and her writing was difficult to read.

All these details referred to possible sources of error but offered little information of any great precision with regards to the place and the means of the contamination, or the substitution. The experts then brought their attention to bear on the use of the drying stoves. The issue had emerged from the testimony of the Director of the Lübeck Dispensary: Deycke apparently had given him two tubes coming from the same drying stove, one containing BCG, the other containing human culture, in order to show him how difficult it was to confuse the two cultures and that the source of the accidents in Lübeck was to be sought for elsewhere [8, 12/4/31]. According to the experts, the fact that the two cultures had been so close together made the contamination hypothesis at the laboratory quite reasonable. Lange and Lang took advantage of these observations in order to request that a law controlling the production of vaccines be passed and that henceforth, production and analysis be disassociated. This is the first time that such a law came to control the practices in order to assure the use of vaccines in the best conditions possible.

Was it necessary then to introduce the vaccine into Lübeck?

Although no law made the action illegal, the vaccinators should have been sufficiently well informed and remained vigilant with regard to criticisms. The plaintiffs accused the physicians of having underestimated the controversy surrounding BCG [8, 01/21/32]. They recalled Arthur Schlossmann's warning of March 1927, who had said that the time had not yet come for the introduction of the vaccine [8, 01/11/32]. The vaccinators defended themselves using statistics that were particularly favorable to Calmette and the decision of the Society of Nations of 1928 which made notification obsolete [16]. The accused also replied to economic arguments. The lack of anti-tuberculosis equipment at Lübeck and the high tuberculosis morbidity had forced them to cut costs. The experts agreed with the accused on this last point but nonetheless regretted that test inoculations on animals had not been undertaken with the vaccine emulsions before moving on to humans.

When this debate was taken up by the lawyers and the public prosecutors, the temperature rose. The entire organization of German medicine was denounced. Deycke was described as an autocrat who refused all criticism. Alsteadt was described as a physician who was "young and who had arrived too quickly at a position of responsibility" that he had obtained "in publishing vast quantities to the detriment of quality" which was "the disease of our time" [8, 01/18/32]. They alone had decided to undertake such an enterprise for which they were not prepared. The prosecutor, Von Beust, spoke of "a failure of the system" of which the accused were the representatives, a system where physicians were all-powerful and masters of the destinies of others [8, 01/18/32].

The debate was taken up again by Julius Moses, a member of parliament, socialist and physician to the Reichstag tribunal. He criticized the overriding power of the physicians who had hidden from the parents the reality of what had been done to their children [6, pp. 35-7]. The accused defended themselves saying that the word vaccination was not popular and that to hide it did not in itself constitute a crime. Julius Moses, on the other hand, saw an infraction against individual liberties in the transformation of a vaccination campaign into an experiment on humans [6, p. 39]. He called for a new ethical attitude on behalf of physicians whom the system had not sufficiently controlled. He referred to the physicians as "servants of their patients" [17] who expected from the physicians miracles which had lead the latter into excess. Finally, he severely criticized the hegemony of bacteriology which dominated scientific medicine and which treated the individual more like an object than a patient [6, pp. 26-39]. He held that measures of social hygiene were the only means of democratic struggle and would only accept a recourse to vaccination when the children grew up in a contaminated milieu [6, p. 15].

An unsurprising judgment?

In his judgement, the President of the Tribunal blamed negligence in the preparation of the cultures, excluding de facto the possibility that BCG had regained virulence and thus saved bacteriology that had stood on the brink of ruin. He based his decision primarily on the depositions of experts representing the large German research institutes and consequently the followers of order, namely those who accepted the idea of the constancy of species. The efficiency of their demonstration was greatly enhanced through the production of visible demonstrations. All could see the beautiful green color of the cultures and there had been a very impressive pedagogical display of posters and charts [8, 12/16/31]. This had reinforced the administration of justice as the experts had shown not only bacteria but also a wide variety of evidence assembled in the course of a police search of the laboratories. At this point the judicial and the scientific discourses became one and the judge found himself in the classic case of the administration of justice: the weapon had been found, the crime scene established, the guilty recognized.

Much and Poll's arguments were not revisited. Much had little credibility. From the beginning his commercial collaboration with Deycke in the production of the pathogens had been raised in order to discredit him. His angry temperament, his nationalist aspirations and the marginality of his theories isolated him still further [8, 12/21/31]. As for the biologist, Poll, he had only offered theoretical arguments. While the bacteriologists readily admitted biological evolution, they had no proof of its existence with bacteria [8, 12/18/31].

Deycke was sentenced to 24 months in prison (a sentence that he did not serve [7, p. 61]) and Dr. Alsteadt to 18 months. The Supreme Court Leipzig confirmed this judgement on appeal, thus ending the trial on 1 June 1933 [18]. The relative clemency of the sentence reflected the traditional solidarity amongst *élites* increasingly threatened by the reforms of the Weimar Republic characteristic of this period [19].

The consequences of the Lübeck "vaccination"

The international press gave detailed coverage of the trial. In France and Germany, nationalist feelings were excited. For example, the French Daily *Candide*, created by Arthème Fayard and which took up the nationalist themes of Action Française, spoke of a trial

where "an atmosphere of passion, bad faith and hatred against French savants" reigned [20]. In Germany, the nationalist press deplored the absence of the truly guilty party, Professor Calmette. The bacteria were represented as an arm of modern warfare and a vector of French hostility towards Germany. Hitler himself had threatened to come to Lübeck to demonstrate in order to demand reparation from France [21].

German Socialists mobilized against those responsible for the vaccinations in Lübeck and criticized foreigners only to the extent that they expressed the same perversions of medicine as had been expressed in Germany. The Health Council of the Reich voted on 20 February 1931 a directive to control human experimentation and to protect patients which responded point by point to the issues raised in the course of the trial. Although this ethical code was supposed to be signed by all practicing physicians, it seems that nobody actually ever signed it and that when the NSDAP took power, the code was rendered obsolete [22]. In any event the code shows the extent to which the German judicial apparatus was ready to counter the medical atrocities of the following years which so cruelly illustrate.

On the other side of the Rhine, opinion seem more preoccupied with the fact that BCG had been declared innocent than by the possible ethical implications. The judgement renewed confidence in the Pasteur Institute where the man in the street felt that: "maximum precautions are taken" and where "no pathogenic ... bacteria can enter into contact with the vaccine." In Lübeck, on the other hand, "the criminal negligence was evident. It jumped out at you" [23].

Nonetheless, during the trial, many countries suspended BCG vaccination. Such was the case, for example, of Holland, Belgium, Poland, Switzerland and, of course, Germany. In France, to overcome the crisis, the Minister of Foreign Affairs put radio stations and the public agencies of information at Calmette's disposition. All communiqués repeated that BCG was innocent in the deaths in Lübeck and gave the latest figures favorable to vaccination in order to save France's image and the image of French science abroad. Via the prefectures, propaganda in favor of BCG attained new heights lending much necessary support to the construction of a 'temple' for BCG at the Pasteur Institute in Paris.

In addition to this state network, there emerged a second network composed of the friends of Calmette abroad. In particular, in Germany, Calmette maintained links with researchers who defended BCG in a local press and struggled against the political turn of the trial.

The Lübeck trial would continue to draw attention following the Liberation. Thus, Weill-Hallé, who was the first to practice BCG on a suckling child, wrote in 1946 in the *French Hospital Weekly* "a German Professor undertook a vaccination trial in 1930 and introduced perhaps for experimental purposes bacteria having maintained the virulence into the vaccine produced by the Pasteur Institute, causing the death of a number of suckling infants and whose death prior to a long and minute investigation was falsely attributed to BCG [24]. Similarly, Debré in his book "L'Honneur de vivre" described the Lübeck trial as a result of "a crazed criminal experiment" undertaken by a German vaccinator. One must evidently consider these words within the context of a post-war era [25, p. 178].

Finally, in 1950, in the course of vote on a law concerning obligatory BCG vaccination, the Ligue pour la Liberate de la Vaccination used the Lübeck incident as one of the

foundations for its anti-vaccinationist argument. Curiously, at many points the League's discourse resembled that of the German Socialists at the beginning of the 1930s.

In the course of the Lübeck catastrophe, scientists and physicians found themselves in a rather unusual situation: they were constrained by similar material and similar questions without being able to interact directly with the foundations of their research. Moreover, they made their depositions in public without mediation. The hesitations and the disagreements gave a precise image of research and medical organization at the time and showed us the weak points were practiced, that is insufficiently controlled. The final word was reserved for the judge whose judgment possessed a high value and who endorsed, to a certain extent, the scientific proof offered in the course of the trial. Had he concluded that BCG had been responsible for the accident, it would have been a hard blow for medical bacteriology. Indeed the avidity with which scientists received the decision and accorded it the value of a true scientific judgement evidently favorable to vaccination in bacteriology bears witness to this.

By thrusting such a medicine into the limelight, the trial acquired an exemplary value in so far as it highlights the hopes and the fears of the public with regards to medicine. A certain number of fears clearly appear: the fear of a vaccination acting as a means of control of the body which raises the question of individual liberty vis-à-vis medicine, the fear of an act which can be aggressive and which does not protect, the fear of a practice that costs little, but which is opposed to the improvement of hygiene, in itself costly, but which profits everyone. We should therefore not be surprised to see the opponents of vaccination turn towards more "natural" medicines, such as homeopathy, which are supposed to respect the body and the mind!

Finally, this catastrophe underscores the fact that vaccination demands the calculation of a double risk: that of encountering a contagious individual and falling ill and that of the consequences of vaccinating someone who is healthy.

Bibliography

1 Calmette A, Guérin C, Weill-Hallé B, Nègre L, Boquet A, Wilbert R, Turpin R. Revue de la Tuberculose 1924; 5: 481.
2 Menut P. Contribution à l'étude de la définition des standards de bactériologie médicale. Paris: Université de Paris VII DEA; 1995.
3 Bernard L. Rapport de la conférence technique pour l'étude de la vaccination antituberculeuse par le BCG. Presse médicale 1928; 36: 1410-1.
4 Procès verbal de la Commission du BCG, Paris, 07/07/31, Fonds BCG, Pasteur Institute Archives.
5 Calmette A. La vaccination préventive de la tuberculose par le BCG dans les pays étrangers. Ses effets sur la décroissance de la mortalité générale infantile. Ann Institut Pasteur 1930; 45: 525-46.
6 Moses J. Der Totentanz von Lübeck. Rabeudel b. Dresden: Dr. Madaus & Co.; 1930.
7 Hahn S. Der Lübecker Totentanz. Zur rechtlichen und ethischen Problematik der Katastrophe bei der Erprobung der Tuberkuloseimpfung 1930 in Deutschland. Medizinhistorisches J 1995; 30; 61-79.
8 Lübecker Trial transcript. Lübecker General Anzeiger. 11/13/31.
9 Dépêche de l'Ambassade de France à Berlin, 05/28/30, Fonds BCG. Pasteur Institute Archives.
10 Schürmann P. Tuberkulosevorträge auf der Pathologentagung in Munschen vom 9-11 April 1931. Zeischr Tuberkulose 1931; 61: 299-302.
11 Lange L. Zu den Tuberkuloseschutzimpfung in Lübeck 1930; 57: 305-10.
12 Calmette A. Le matin. 05/21/30.
13 Lange B. Untersuchungen zur Klärung der Ursachen der im Anschluss an die Calmette-Impfung aufgetretenen Saülingserkrankungen in Lübeck 1931; 59: 1-18.
14 Letter from Dr. Specklin of Mulhouse to Albert Calmette. 11/01/31. Fonds BCG. Archives of the Pasteur Institute, Paris.

15 Amsterdamska O. Medical and biological constraints: early research on variation in bacteriology. Soc Stud Sci 1987; 17: 657-87.
16 Bernard L. Rapport de la conférence technique pour l'étude de la vaccination antituberculeuse par le BCG. Presse médicale 1928; 36: 1410-1.
17 Moses J. Der Kampf um die Kurier-Freiheit. Rabeudeul b. Dresden: Dr. Madaus & Co.; 1930.
18 Jugement du procès de Lübeck à la cour suprême de Leipzig. 06/01/33. Fonds BCG. Pasteur Institute Archives, Paris.
19 Peukert DJK. La république de Weimer. P. Kessler, trans. Paris: Aubier; 1995.
20 BCG. Candide 02/05/31.
21 Letter from H. Buschmann to A. Calmette. 01/31/32. Fonds BCG. Pasteur Institute Archives, Paris.
22 Ambroselli C. L'éthique médicale. Paris: Presses Universitaires de France; 1988.
23 Propos d'un parisien. Le matin. 02/09/32.
24 Weill-Hallé B. La semaine des hôpitaux de Paris. 03/07/46.
25 Debré R. L'honneur de vivre : témoignages. Paris: Hermann & Stock; 1974.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Immunology in developing countries. A distorted image

Dominique Frommel

How long ago did immunology, defined as "a scientific study of protection against or resistance to disease" [1], spread into tropical countries? It was introduced one century ago, deriving from the overseas establishment of the Pasteur Institutes. These centers, more overtly perhaps than endeavors of British medicine, were emblematic of the faith in vaccination for averting infectious disease [2]. Hereinafter, immunization campaigns, sero-epidemiological studies, and the production of vaccines represented the entrance of applied immunology in the South. These issues, however, remained associated with white power and its moral and medical connotations.

Practical guides for native health workers who performed immunizations became progressively available after World War II. However, they offered little, if any, explanation on the nature of immune response and on the mode of action of vaccines. During the time of independence, the creation of new public health colleges and medical schools promoted the training of health officers and physicians in their home countries. Their cursus included a few lectures on immunology, usually proffered in the programs of the microbiology and parasitology departments. As in other disciplines, the teaching material was of European or North American origin, the most popular treatises being Roitt's "Essential Immunology," Talwar's "Immunology for Medical Students" in India, and the fascicles "*Cours d'Immunologie de l'Institut Pasteur de Paris*" in French-speaking countries. The mention of ancient knowledge generated in non-western cultures, such as pox inoculation and leishmanization in China and in the Ottoman Empire did not appear in these basic textbooks. Furthermore, expatriate tutors, who in the 1960s and 1970s frequently taught immunology overseas, seldom alluded to such historical events capable of bridging different concepts of preventive care.

The question as to whether instructors or leaders of immunization teams – today chiefly indigenous health professionals – were capable of imparting the principles of acquired immunity has been largely neglected, although the medical literature abounds with papers dealing with the operational difficulties of vaccinations [3, 4].

Having been one of these foreigners who tried to share his conversance with immunobiology with various groups of students, an enquiry about the image that immunology health professionals had exuded appeared relevant. Two questionnaires were submitted to the trainees who attended sessions on mother and child health care. [1]

1. The survey took place during courses of mother and child health care attended in Yaoundé and Paris by physicians, midwives, nurses and health officers from Africa, South America, and the Far-East. Forty-two responses were considered.

The first questionnaire dealt with the functions related to immunology. In *table I*, the five most frequent examples are given in alphabetical order. Replies conformed to a fairly universal acceptance of the tasks assigned to the immune system.

Table I. Functions related to immunology.

Immunology is related to representation of:
Defense
Killing of germs
Protection
Survival
Vaccine

The second questionnaire was to list ten terms and expressions which, in their experience as promoters and partakers of preventive medicine, evoked immunization procedures (*table II*). The least that can be said is that this florilegium of bellicose and warlike wordings failed to comply with a Hippocratic lexicon. One may, however, discover similar trends in reports of "politically-correct" institutions such as WHO and UNICEF [5, 6]: "missed opportunities for immunization – targets being not only diseases but also pregnant women and children."

Table II. Terms associated with vaccination.

adverse (effect)	injection
anti-(body),- (gen)	invasion
anxiety	killed (vaccine)
blood	live (vaccine)
campaign	mandatory (immunization)
control	mass (immunization)
(avoidable) death	national
disease	pain
elimination	population
eradication	prevention
exclusion	strategy
failure	(immuno)suppression
fear	survival
fight	transmission
foreign(er)	triumph (over smallpox)
immunization	virulence
infection	(magic) weapon

Cartoons, usually routed by pharmaceutical companies, illustrating successes and victories of the fighter cells mobilized by vaccines, popularized martial approaches of immunology.

Beyond any doubt, there is a semantic gap between the North and the South concerning the expression of the essence, the achievements as well as the goals of immunology; more importantly, in developing countries the scope of immunology remains chiefly reduced to operational aspects. Socio-anthropological aspects of health and disease may, in part, underlie such discrepancies. Rather than digressions on this topic, notes taken after discussions with two Bengali physicians offer another view of this dissonance.

> Why is immunology the mirror of the cultural corpus of the Occident, a genuine 'self' for you, rather than a 'non-self' for us? Indeed, immunology you are exporting to our countries – to our great benefit, we would add – is merely confined to vaccinology. Even in our remote places, we understand that, with its penetration into molecular-based analyses, immunology is becoming the paradigm of the science of communication. Communication between cells and not between individuals!
>
> Immunology has uncovered signals of our inner as well as of our outer worlds; it has also deciphered the mechanisms governing the subtle interactions taking place in the immune system producing some kind of multidisciplinary cybernetics. For illustrating facts or speculations, immunologists often resort to words and terms more familiar to anthropologists, sociologists, and philosophers than to basic scientists. Since you adopted a large palette for describing immunological phenomena, we are surprised that the applied side of immunology is reaching us enwrapped in a fairly inflexible rationalism that is still exotic to us. Indeed, immunization practices you recommend tend to shake our social cohesion.
>
> Well before the West, our forefathers practized measures aimed at preventing diseases. Today's immunization modes are based on the fiction of collective willingness, whereas the past inoculations and scarifications, the latter conveying symbolic meanings of little value against most diseases, involved strong social relations and took place in a dialogue between healers and individuals seeking their assistance.
>
> Thus, how is it that in their immeasurable repertoire, immunologists face difficulties in finding room, or receptors, addressing another diversity, the cultural one?

Our colleagues wanted to be provocative as one often is in Calcutta. They overlooked one fact, that in New Delhi the National Institute of Immunology, a center of excellence, had no need to envy the best ones in industrialized countries. In India, emancipation from the North has come into being. Perhaps unconsciously, their reaction might also have been rooted in a more profound split between two worlds, the scientifico-technical one with an eroded vision of spiritual mastery and the still traditional one in which mankind remains the mediator between the revealed and the un-revealed. Is it not the immune system that mediates many of man's relations with his environment?

Nonetheless, accepting their somewhat contradictory viewpoints, we tried to formulate proposals which made it possible to curve this asymmetry. We agreed that basic immunology, centered on recognition and regulatory mechanisms, should be part of the cursus of physiology. Applied aspects of immunology should be integrated in the module of preventive medicine and public health and that immunology, related to pathology or diseases, should be taught in medicine and other clinical disciplines. Writers and editors of the North would be invited to revise the presentation of their books and reminded that the incorporation of too many recent discoveries obscure the major concepts conveyed by modern immunology. Moreover, they would be reminded that experiments carried out on mice are quite unappealing to audiences for which these small mammals personify predators. Finally, we recognized that we all awaited treatises on immunology framed according to Asian images of modern science.

On a different level, African colleagues suggested that campaigns or days of immunization should be preceded or accompanied by festive events so as to accommodate modern measures with local and living traditions. Scarifications are visible events, injections are

not. Only a community fair or ceremony can be linked to protection, related to an initiation on how to cope harmoniously with the environment and, occasionally, with maleficent influences. (One may argue that it would be difficult in Europe to implement similar measures on the occasion of traditional events, such as the pilgrimage to Lourdes). Can such a proposal, although the young Federation of African Immunological Societies (FAIS) [7] would endorse a problem-solving one, remain undebated?

Dialogue and disputes on immunology between health care providers are before all of us and are a fruitful challenge.

Acknowledgements

D. Frommel is a member of the Institut National de la Santé et de la Recherche Médicale (Inserm), Paris, seconded to the International Children's Center.

Bibliography

1 Oxford advanced learner's dictionary. 5th ed. Oxford: Oxford University Press.
2 Delaunay A. L'Institut Pasteur, des origines à aujourd'hui. Paris: Éditions France-Empire; 1962.
3 Eng E, Naimoli J, Naimoli G, Parker KA, Lowenthal N. The acceptability of childhood immunization to Togolese mothers: a sociobehavioral perspective. Health Educ Quart 1991; 18: 97-110.
4 Lévy-Bruhl D, Cook J, Legonou B, Jaffré Y, Amévigbe B, Sanou G, Guérin N. Approches méthodologiques de la vaccination: exemple de trois enquêtes menées en Afrique de l'Ouest. Sci Soc Santé 1993; 11: 9-25.
5 Expanded Programme on Immunization. Sick children: targets for immunization – United Republic of Cameroon. Weekly Epidemiological Record WHO 1983; 58: 29-30.
6 UNICEF. The state of world's children. New York: UNICEF; 1995.
7 Mshana R. Immunology in Africa. Immunologist 1995; 3: 3-4.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Contemporary and prototypic figures of immunology in the medical press

Daniel Jacobi

Immunotherapeutic drugs market

This text sets out to study a corpus of graphic representations used by contemporary immunology. The research focuses specifically on the visual devices marshaled in the cause of communication in the sphere of immunological therapy. In other words, the purpose is to draw attention to the devices used for presenting the various families of immunotherapeutic drugs which have been marketed over the last few years (excluding the more commonly used serums and vaccines), to comment on the use to which they are put, and to explain their mode of action. The distinction made here between presentation, comment, and explanation is clearly an arbitrary one. In our discussion of the notion of pain, we have shown that the passage from one discursive genre to another is a constant feature in scientific communication[1].

We have gathered these graphic representations into a fairly wide-ranging body of documents. Although we have consulted the primary scientific reviews (those in which specialists publish the results of their research for their peers) and the scientific press aimed at the general public, our corpus is not in fact centered on either of these categories[2]. Instead, we have chosen publications which have in common the fact that they are published or produced for a specific category of readers, namely professionals working in the health field and who, theoretically, have had the benefit of a university education and may therefore be supposed to possess a basic culture in immunology.

It is common knowledge that there is a branch of the specialized press aimed specifically at nursing staff and, of particular interest to us here, pharmacists. The corpus which we have built up consists exclusively of texts accompanied by visual sequences (illustrations, photographs, curves, tables, etc.). In addition to press articles written by practicing scientists or specialized journalists, the corpus consists of documents published (sometimes at great expense, judging by the four-color printing and glossy paper used) by pharmaceutical companies for reasons of both information and publicity: technical notes, reports of trials, articles written by researchers and resembling a primary scientific publication but

1. Jacobi D. Expliquer et faire comprendre la douleur; formes et ressources des discours explicatifs. Recherches en soins infirmiers 1998 ; 53 : 5-16.
2. On this distinction, see Agostini F. Science en bibliothèque. Paris: Cercle de la librairie; 1994.

distributed by the company itself as an off-print inserted in a publicity folder, presentation sheets and instructions aimed at general hospital pharmacies, etc.[3]

In view of the hybrid character of these publications, our analysis makes provision for imagery not normally considered as scientific, that is to say analogical reproductions of medicinal products and their packaging as displayed in the publicity and logos of companies marketing active products for health disorders supposed to be most closely linked with dysfunctions of the immune system.

In order to grasp the specificity of the figures appearing in this kind of publication, we have compared them first to a limited number of primary texts which pharmacists might be expected to consult on occasion, and second to immunology manuals used in higher education and to texts aimed at a general public which, once again, pharmacists occasionally consult or have consulted[4]. With regard to the last of these categories (texts aimed at a general public) and with a view to ensuring that our comparison remains relevant, we have limited our inquiry to texts published in reviews aimed at cultivated and scientifically reasonably well informed readers (reviews such as *Pour la Science*, *La Recherche* and *Biofutur*) – in other words readers who are culturally and professionally comparable to pharmacists.

Indeed, one of the goals of this research was to study the characteristics of immunological imagery when such imagery is intended for professional pharmacists working in pharmacies or hospitals and supplying (or selling) active immunological principles to members of the public for their treatment[5]. Is the imagery used in this literature similar to that to be found elsewhere or is it of a specific nature?

In disseminating immunological theories, does the professional medical and pharmaceutical press, ie, the press exclusively read and consulted by scientifically educated and informed health personnel, have frequent recourse to imagery? And if so, what are the characteristics of this imagery?

This research forms part of a longstanding discussion of the role and place of imagery in scientific communication, not only in terms of the history of immunology but also with respect to the way this imagery is used in scientific manuals and publications aimed at a more general public[6]. The visual models at work in this discipline belong to a dual tradi-

---

3. From a communication point of view, it would no doubt be illuminating to submit the special economy of the medical and paramedical press to thorough analysis. This lavish specialist press, relying heavily on industrial advertising or sponsorship for finance, was the first manifestation of what has now become a common phenomenon: the rise of a specialized press aimed at a restricted and targeted readership accompanied by the relative eclipse of a generalized all-purpose press seeking financial health by circulating to as many readers as possible. Miège B. La société conquise par la communication. 2. Grenoble: Presses de l'Université de Grenoble; 1997.

4. Most of the examples cited in this article are taken from a research project conducted by a pharmacist working in our laboratory while preparing her Ph.D. It is clearly most unusual for a specialist to subject the imagery of specialized publications to exhaustive analysis. Bruneton C. Les représentations graphiques du médicament; analyse comparative des procédés de figurabilité utilisés pour représenter des médicaments immunologiques dans la presse professionnelle. Pharmacy thesis under the supervision of Jacobi D. Dijon, France, 1997. I wish to congratulate Claire for her excellent work and to thank her for help in writing this article.

5 Bruneton C. Les représentations graphiques du médicament; analyse comparative des procédés de figurabilité utilisés pour représenter des médicaments immunologiques dans la presse professionnelle. Pharmacy thesis under the supervision of Jacobi D. Dijon, France, 1997.

6. See our textxs :Cambrosio A, Jacobi D, Keating P Ehrlich's "beautiful pictures" and the controversial beginnings of immunological imagery. Isis 1993 ; 84 : 662-99. Jacobi D. La communication scientifique; discours, figures, modèles. Grenoble: Presses de l'Université de Grenoble; 1999.

tion: the tradition governing the methodological and experimental paradigms materializing the research, and that governing communication with an educational intent with recourse to images and illustrations[7].

The publications are considered as a particular discursive production designed to prove, to argue and to convince rather than to illustrate. Our decision to focus on a specific category of health care professionals, in this instance the cultivated, intermediate category constituted by pharmacists, for the purpose of research on scientific popularization, is therefore inspired by a desire to test what strikes us as being a relatively credible hypothesis. This hypothesis is to the effect that the dissemination of scientific information to as wide an audience as possible (in this case the patients with whom the health care professionals are in daily contact) is less likely to succeed when it is characterized by direct communications from the manufacturers of pharmaceutical products calculated above all to convince the prescribers of these products, ie, the pharmacists[8]. In these situations, the ability of the pharmaceutical companies to explain, argue and convince will be a fairly accurate reflection of the communicational strategies which pharmacists are likely to adopt in their relations with their patients.

Figurability resources of immunoactive pharmaceutical substances

The comparatively new discipline of immunology has already given rise to numerous therapeutic applications[9] of which the best known are vaccination and the prevention and treatment of organ or bone-marrow transplant rejection. Today, however, attention is increasingly shifting towards allergies, cancer, autoimmune diseases and retroviruses. In many cases, treatments are already in preparation while in others the prospects for development look promising.

The purpose of this therapeutic armory is to control the immune reactions or to deflect them in a particular direction. The drug's role is to generate an immunomodulation of the response of the patient's organism. Four classes of drug are distinguished: immunostimulants (bacteria extracts supposedly capable of preventing recurring respiratory infections), immunosuppressants, e.g., cyclosporine or monoclonal antibodies, immunosubstitutes (serums and immunoglobulins) and lastly enigmatic cytokines (interferons, interleukins) whose effects appear to be both beneficial and harmful[10].

Any attempt to present these active substances, to pinpoint their role or explain how they intervene, almost inevitably involves situating them in their antigen-antibody complex. We should bear in mind that immunology is a scientific discipline that does not lend itself easily to visualization, partly because it is a branch of research concerning phenomena which are immaterial (regulations) or imperceptible without recourse to complex

7. On scientific imagery, see Latour B, Ed. Les vues de l'esprit; visualisation et connaissance scientifique. Culture technique. 1985 ; 14. On the superiority effect of images, see Reid DJ. The role of pictures in learning biology. J Biol Educ 1990 ; 24 : 3-4; 161-72; 251-8.

8. See Jacobi D. Textes et images de la vulgarisation scientifique [new edition]. Bern: P. Lang; 1999.

9. Moulin AM. Le dernier langage de la médecine; histoire de l'immunologie, de Pasteur au Sida. Paris: Presses Universitaires de France; 1991.

10. Perrin LF, Laurent,PE. Immunopathologie clinique. Paris: Masson; 1990. For an example of a highly colored popularization of interleukin 2, see Jacobi La communication scientifique (op. cit).

appliances and techniques, and partly because it increasingly involves the relationship between persons, substances and a series of complex reaction mechanisms[11].

In a previous work, we analyzed this difficulty in connection with the different visualizations of the complement. The situation is somewhat different where the medicinal product is concerned. Here, writer and illustrator have at their disposal items of information corresponding to the principal characteristics of the pharmaceutical specialty, namely its international nonproprietary name, its composition and in particular the chemical structure of its active principle, its pharmacological properties, therapeutic indications, dosage, administration route, interactions with certain physiological functions and possibly with other substances, adverse secondary effects (if any), and the conditions governing its issue and market authorization.

To these intangible characteristics must be added two other special features: firstly, the commercial name and form of presentation chosen by the laboratory (powder, tablet, drops, capsule, etc.), and secondly, the shape and color of the packaging in which the product is sold to the public. Finally, depending on the publication platform, it may be desirable to display the product's mode of action and to give some indication of the workings of the active principle, if only to explain why it is effective.

Each of the points mentioned above is a candidate for visual treatment, though clearly the figurability resources differ widely and the visual potential marshaled will very according to the type of publication. In her dissertation, Claire Bruneton identified the same resources described and listed by Daniel Jacobi (along with other researchers) and corresponding to two distinct imagery categories used in the socio-dissemination of life sciences.

The first category consists of the scientific imagery specifically related to research and to the production of knowledge (results tables, curves and columns, graphic semiology, semiotic codes peculiar to the various disciplines, readings produced by apparatuses of varying complexity, micrographs, etc.).

The second category is made up of imagery of didactic intent, more concerned to explain, and to help the reader grasp and retain the essential points. Here, the principal registers adopted are diagrams, plans, simplified drawings, and the partial reuse of scientific imagery.

Bruneton also demonstrated that, in the case of publications aiming to inform health specialists of new specialties or of those produced by pharmaceutical companies, these two categories were accompanied by a register closer to that of commercial advertising.

In this case, a first figurability resource may be identified: the visualization of the drug's active substance. In point of fact, two entirely different approaches are adopted – the reproduction of the molecule and the reproduction of its packaging.

With regard to the former, the illustrator typically uses the chemical formula (letter symbols and figures), its so-called structured representation and its likeness in the form of a line drawing (*figure 1*).

At a more mundane level, the illustrator can make do with the presentational drawing (capsule, pastille, tablet, etc.), the elegantly typed commercial name under which the

11. Cf. Cambrosio A & Keating P.

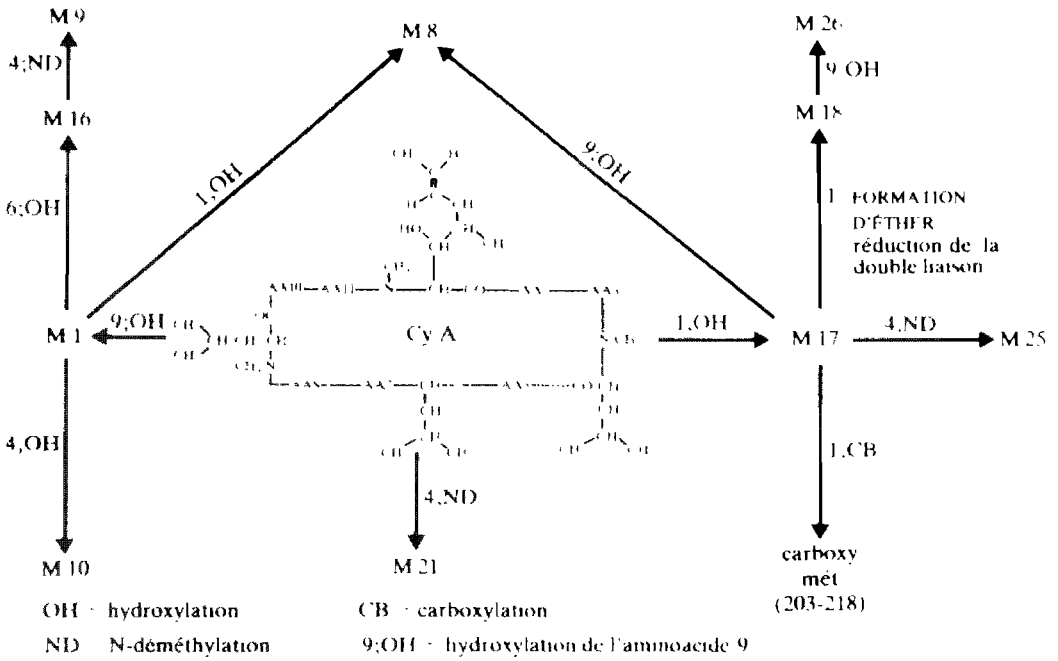


Figure 1. Developed chemical formula of cyclosporine [in Sandimmune® technical file, 4th ed. 1993. p. 10]. This technical file includes the flat structural chemical formula of cyclosporine

active substance is marketed, analogical reproductions of the packaging or wrapping (of varying degrees of sophistication) chosen by the manufacturer (box, tube, blister, etc.), the colors and graphic display adopted for the packaging and serving to distinguish the drug from those produced by its competitors (once the patent has expired) (*figure 2*).

Less frequently, and only in certain kinds of publications, we encounter diagrams depicting directions for use, the localization of the treatment, and lastly, in the case of promotional inserts, images of prescribers and their patients, that is, children.

Recourse to visualization processes in function of the type of publication

When disseminating scientific information via recognized scientific journals, scientists give preference to curves, tables of results, and inscriptions. Virtually the same graphic representations are to be found in the technical dossiers intended specifically for hospital pharmacists. The inference is that for this serious-minded and captive audience, scientific arguments count for more than anything else.

For example, the curve illustrating the effectiveness of mycophenolate mofetil in preventing the rejection of a renal allograft and originally published in *The Lancet* was reproduced virtually without change in the Cellcept® technical file distributed to hospital pharmacists[12].

12. Bruneton C. Les représentations graphiques du médicament; analyse comparative des procédés de figurabilité utilisés pour représenter des médicaments immunologiques dans la presse professionnelle. Pharmacy thesis under the supervision of Jacobi D. Dijon, France, 1997.

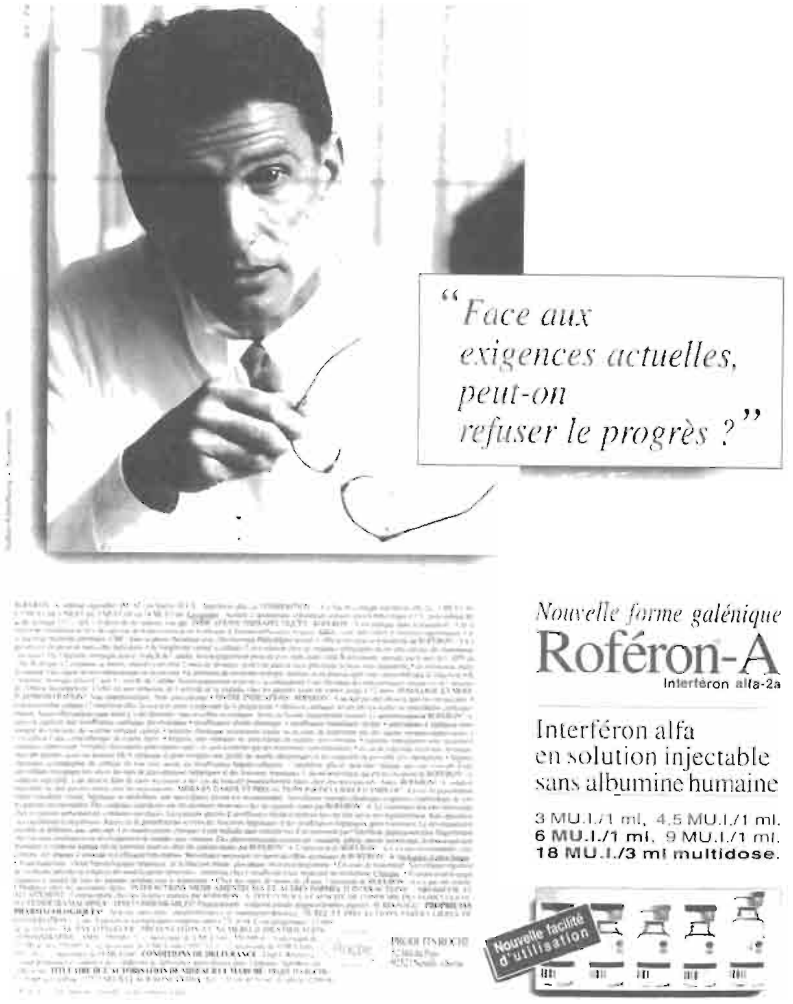


Figure 2. A full-page advertisement by Roche Laboratories for Roféron-A® [in *La Presse Médicale* 1997 ; 26 : 6, outside back cover].
This standard, lavishly designed advertisement provides two levels of discourse. The first (towards the bottom of the page) consists of a classic, regulatory presentation of the drug, giving full details and instructions for use. The text concludes with a discreet reminder of the company logo inscribed in an oblate hexagon; the right part of the advertisement, gives the name of the drug, using a sober and austere graphic display, and analogical representations of five bottles of injectable solution (at the right dose).
The second discourse component, quite different from the first and notably covering the upper part of the page, attracts the prescriber's attention with a typical slogan. This slogan – "Confronted with today's demands, can we afford to turn our backs on progress?" – is displayed in large italic print inside a box partly covering a photograph of a smartly attired man whom we may suppose to be a doctor (his tie is clearly visible under his white coat); the man has removed his fine-rimmed glasses as though to address the reader with the words ".... can we afford to turn our backs on progress?". Naturally, this direct question refers us back to two seemingly insignificant phrases in the other more academic and conventional body of text: "New galenic form" (above the name of the drug) and "Improved ease of use" – written in video inverse (white on black) – in another smaller, oblique but highly visible strip barring the photographs of the bottles.
This is a classic advertising ploy: the photograph and the name of the vaunted product have been superimposed by a likeness of the "target," in this case an elegantly dressed, youthful-looking doctor who has gone to the trouble to take off his glasses so that he can address his audience more directly.

However, another source is also used for illustrating documents aimed at pharmacists: imagery produced for educational purposes. in such cases, of course, constant recourse is made to university manuals. It is known that immunological regulation calls into play a series of intercellular communication mediators each of which derives its name from the type of cellular actors which it produces: "lymphokines" for lymphocytes, "monokines" by monocytes or macrophages, etc. The more transversal term "interleukine" has also been created to emphasize their role in intercellular relations. Finally, since many more cells produce mediators for exchanging information with the others, the generic term of "cytokines" has been devised to designate all the mediators produced by the cells and agents of immunological regulation. Quite clearly, this terminological profusion is more than a little unsettling and can give rise to confusion at a phonetic level (*figure 3*).

It is very interesting to see how different types of publications, while favoring what in theory is their own register, do not scruple to resort to others: the advertising pages or off-prints sent by companies to professional practitioners for promotional purposes, frequently contain curves, columns, tables of results and micrographs next door to such typical advertising features as pictures of patients and the commercial product[13].

The mobilization and exploitation of visual patterns

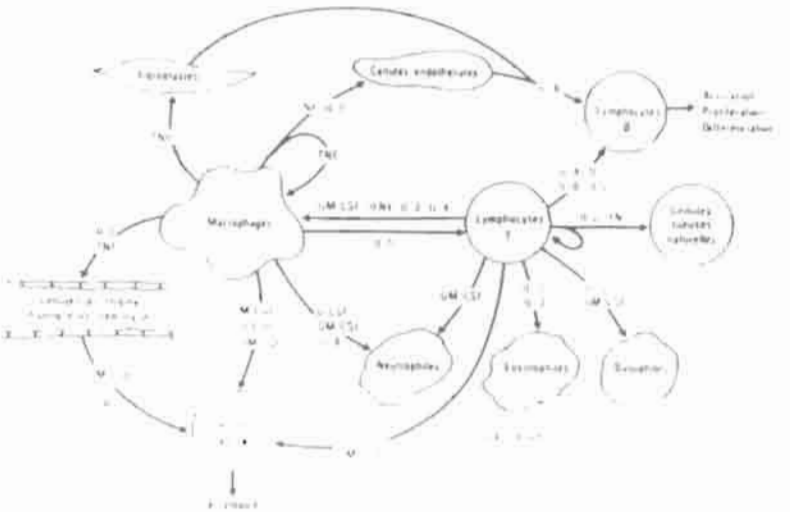
The fact that a limited number of very similar figures turn up time and again in a wide range of documents is a clear indication of a scientific consensus concerning the figuration of immunology. Concerning these prototypes, mention may be made of the three-dimensional shapes of macro-proteins as imagined and drawn with extreme precision by Irving Geiss and others[14]. There are, too, many flowcharts visualizing the complexity of regulation and cooperation between agents and mediators required to produce a reaction. We should also take note of the highly-colored sequences of micrographs, invariably inscribed on a bold black background and depicting the struggle or fight between the body's defender and invader[15]. But the undisputed champion of immunological figurability is the Y representation of immunoglobulins. The Y has become the pattern (or model) of the antibodies. Scientifically solid arguments have suggested that the biochemical structure of these protein macromolecules has a shape which approximates to that of the letter Y (upper case). An examination of the contemporary immunological literature shows that this model is able to be figured in different forms: as an austere, spare outline in which the rigid structure of the Y is built up through rectilinear fragments, each section of which is identified by its biochemical structure; as a simplified, supple shape, of a more animal-like aspect; as an extremely sober drawing showing how the extremities of the Y's arms fit

13. Bruneton has worked out the percentages for each of these categories in each type of support. On the whole a certain regularity emerges but it is interesting to note that no category is, as it were, out of bounds for a support. Thus, the figuration of the cellular actors of immunology (macrophages, polymorphonuclear leukocytes, etc.) is transversal. Bruneton C. Les représentations graphiques du médicament; analyse comparative des procédés de figurabilité utilisés pour représenter des médicaments immunologiques dans la presse professionnelle (op. cit).

14. Alberto Cambrosio has gathered a mass of information and evidence on this remarkable producer of original figures for scientific publications. In his early career, he established a reputation for his work aimed at a general readership, for example the covers of *Scientific American*. Later, he worked with research teams to create and trace visual representations capable of rendering the remarkable properties of macromolecules.

15. See the analysis of these sequences in Jacobi D. La communication scientifique ; discours, figures, modèles (op. cit.).

a



b

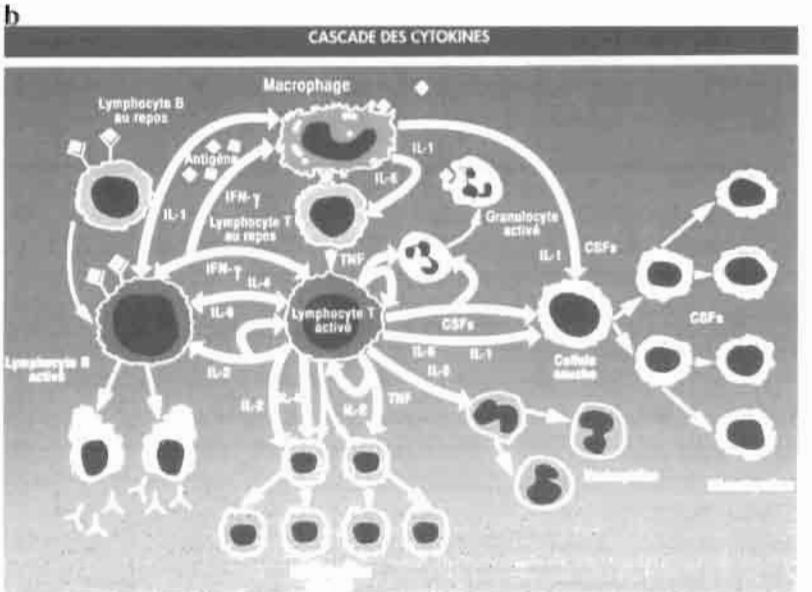


Figure 3. Two flowcharts presenting the cytokine "network." Top: Bach JF. Traité d'immunologie. Flammarion; 1993. p. 444; bottom: Dossier Technique du Roferon-A®, 1995. p. 13.

The connection between the two figures is obvious. Various cells of the immune system are presented in diagram form and linked by means of numerous arrows, the purpose being to indicate the names of the cytokines and the transformations or effects which they generate. These are complex systemic flowcharts. The diagram in the manual is a line drawing and the schematization of the cells is deliberately almost abstract, consisting as it does of somewhat forbidding geometric circles or forms (spindle for the stylized fibroblast) (a). In marked contrast, the figure intended for the health care professionals is printed in bright colors (b). The thick yellow arrows are clearly visible against the blue background, while the differently colored cells (blue, pink, indigo, etc.) are sharply differentiated according to their histological characteristics. Antigen and antibody are also presented (in Y) in the left part (dormant and activated lymphocyte).

together with the diverticula of the surface of the antigens; a detail of an extremity of one arm of the Y fitted on a geometrized site of the antigen's membrane; as a complex flow-chart reconstructing an immunological sequence or mechanism in which the antibody is an actor identified by its simple Y form, drawn with a full line which may be followed through the various stages; or as a spectacular three-dimensional synthesis image built up by little angles and reticulate segments forming sorts of colored clouds branching off against a black background, etc (*figure 4*).

The emergence of these Y shapes marked a turning point in both immunological and pharmacological research since it very soon became clear that the laboratory production of monoclonal antibodies was an interesting therapeutic line of investigation, and by the middle years of the 1980s advertisements for this type of product could be seen in specialized reviews.[16]

The construction of a singular and striking visual identity with reference to immunology

It goes without saying that commercial considerations weigh heavily with the industrial manufacturers of active synthetic substances sold on a large scale. The process of research, development ,and testing required before a product can be launched is lengthy and difficult, and the costs involved are considerable. On the other hand, of course, if a new drug can conquer a large captive "market" of patients, the potential rewards can be exorbitant.

Naturally there is fierce competition, nowadays at the international level, between industrial pharmaceutical companies, to win markets in the industrialized countries. Since the results of biomedical research are readily available, the pharmaceutical groups obviously seek to produce the same classes of drug or to update the form of substances which have come into the public domain and are no longer protected by patents. The result is an intensely competitive climate with considerable resources being devoted to building up a commercial (and thus visual) identify for a given specialty.

It might be supposed that marketing and advertising imperatives would serve to trivialize the figurability of immunological substances but in actual fact this not the case. While marketing budgets for pharmaceutical products are quite clearly far from tight (if the glossy, lavish nature of advertising documents is any guide), managers nonetheless prefer to exercise a certain caution as though mindful of the specific nature of their audience. The advertisements and indeed all other documents disseminated by laboratories for health care professionals are truly fascinating from a visual point of view: sophisticated, carefully produced with genuine plastic qualities, for all the world as though those in charge had commissioned the services of particularly demanding and creative graphic artists.

In particular, they incorporate either a sort of graphic inventiveness, or attractive and carefully produced analogical images, or again logotypes which are the result of research and design carried out by communication specialists or professional graphic artists. It is fascinating to observe how the shape of immunoglobulins and hence of antibodies has become a veritable, almost obsessional leitmotiv or rhyme (*figure 5*).

Nowadays all the major laboratories make use of a company or product logo (a sort of idiographic visual symbol) as a means of consolidating or even improving their commer-

16. See the pages of advertisements inserted in the dossier, Les défenses du corps humain. La Recherche 1986 ; 177.

Figure 4. Advertisement (full page) for Sandoglobuline® inserted in the Revue Médicale de la Suisse Romande 1990 ; 3 : 251.

This advertisement for the Sandoz pharmaceutical company, is a sophisticated and elegant variation in the Y shape of immunoglobulins. The caption states that the drug is a preparation of "lyophilized human immunoglobulins, intended for intravenous administration, used for the treatment of congenital forms of immunodeficiency...."

At the top of the illustration, the graphic artist has chosen to place a thin white Y in a black square, paving the way for the drawing developed in the middle: it constitutes the basic module used by the artist. The model-pattern is positioned above a slogan proclaiming the drug's purpose: Naturally reinforces the overworked immune system. We may note in passing the thrust of the modalization wrought by the adverb "naturally" and the anthropomorphism introduced by the choice of the word "overworked" which has the effect of transforming the immune system into a familiar character.

The large line drawing at the center of the page is a graphic conceit in the style of Escher: a geometrical network with links made up of thin squares in which is inserted either an entire Y or a Y without its left upper oblique arm. The network as a whole has the shape of a square deformed on its right side, and the tangle created by the inextricable superimposition of squares at the center gives way, at the top and bottom, to Y shapes which are detached from the rest of the network and therefore clearly readable.

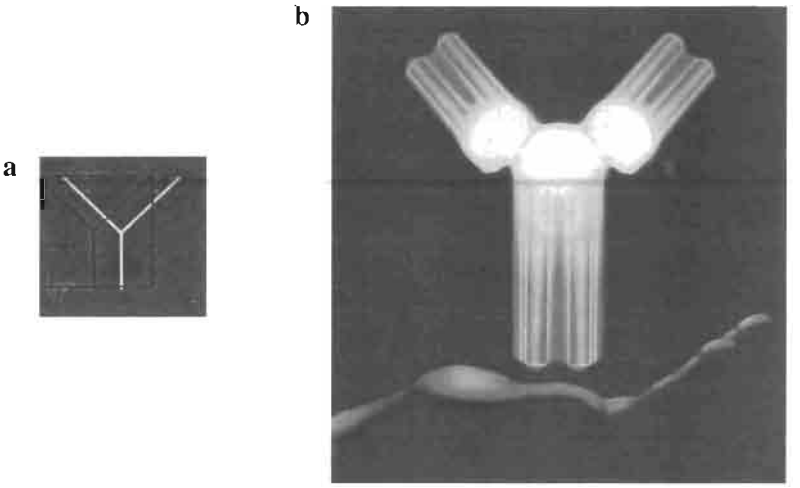


Figure 5. Sandoglobuline®: the old and new logotype (Sandoz Laboratories).
The original logo (a) for this drug was already a graphic work in its own right, bringing the qualities of simplicity, sobriety, balance, and stability to the task of exploiting the visual rhyme of the antibodies' Y shape. We have already seen how it was used and harnessed in the advertisement reproduced in figure 4. The new logo (b), while remaining faithful to the «brand image,» resorts to different semiotic means: black and white has given way to colors, with the Y (red-orange on a purple background) resembling a fluted tube of light drawing attention to the presence of three luminous white spots. These spots are placed at the center, that is to say at the point where the three right segments of the Y congregate. There is also a band or ribbon of the same color hovering at the bottom, combining with the luminous points to confer a certain dynamic to the logo.

cial relations. The logo is a graphic device carefully designed to be at one and the same time highly individualized (calling attention to the company or product), strongly structured, sober and elegant (graphic qualities) and easily memorized (so that it can be recognized by the target audience without any possibility of confusion). It owes more to marketing and advertising requirements than to the discipline of communication. Indeed, it would be an exaggeration to equate the logo with an ideogram in the sense of a super sign capable of encapsulating a large body of information in condensed form. It is neither an icon (a virtually analogical reproduction of the given element) nor an index (causally related to the referent) but rather – to adopt Peirce's distinction in a rough and ready manner – a symbol, an entirely arbitrary sign. In the wake of the commercial model, non-commercial and public institutions have been emboldened to produce their own logos. In his commentary of Peirce's trilogy, Eco has shown that the icon is a sort of mental image, capable on its own of communicating a concept without any other intermediary[17]. The logo makes use of identity and repetition in creating colors and shapes to characterize and singularize a company or product. By familiarizing users with the image which the company strives to project, it helps to create an identity and is an integral part not only of

17. Eco U.Le signe. Paris: Le livre de poche; 1988. Bertin J. Sémiologie graphique. Paris: Mouton; 1967.

all commercial initiatives but also all aspects of external communication, from business correspondence to the designs displayed on the walls of buildings.

In their search at virtually any cost for a reference incorporating an almost obsessional rhyme of contemporary immunology, publishers and laboratories alike do not scruple to seek out the Y shape in other visual fields. An article published in *Tempo médical*, for example, featured the Eiffel Tower, while LFB Laboratories (for its Tégéline® immunoglobulin) and Fujisawa (Prograf® Tacrolimus) both used in the same year (1996) a photograph of the Tancarville suspension bridge, a civil engineering feat whose anchor piling and cable supports are in the form of the letter Y.

Persuading patients through their prescribers

Lastly, French pharmaceutical laboratories use a final register, but only in the advertising sections of documents intended for practitioners. Although this register resembles a classic advertising device it is worth describing here.

The advertisements incorporate two actors: the product to be promoted (always reproduced in a strictly referential manner) and an image of the target consumer presented in a gratifying or sometimes critical situation. In stark contrast to the previously described advertising pages or inserts which are remarkable for their very real visual and plastic qualities and for the care lavished on their production, these images are banal in the extreme and visually far less striking. Typically, they consist of a full-page photograph or analogical design depicting the doctor (or pharmacist) and his/her patient, whereas the drug itself and its packaging are presented on the opposite page, always accompanied by the full text of its data sheet (doubtless to satisfy regulatory requirements which in France make it illegal to advertise drugs in the same way as ordinary consumer goods).

A variation on this theme shows, for example, a little boy thumbing his nose at a large sphere full of spikes which may be supposed to represent a virus. Another image shows a man in whose body a red circle of little Ys seems to be battling against a darker, less clearly defined spot.

Other images borrow from popular iconic tradition in depicting animals intended to conjure up such concepts as gentleness, aggression/defense (sea urchin).[18] (on the interpretation of the advertising image, see also[19]. How to explain these banal, mediocre, and, in truth, conventional or even trivial advertisements? It may be noted that these advertisements tend to involve so-called comfort drugs whose efficacy is sometimes open to doubt (immunostimulants supposedly providing protection against recurrent infections of the respiratory system). The frequent use of the little boy and of children's favorite animal (bears) gives a clue to the answer. The purpose is not so much to convince the prescriber directly (although, to the advertiser's way of thinking, the conventional style serves to reassure the client and to denote the seriousness of the product in question) as to provide

18 Durand G. Les structures anthropologiques de l'imaginaire Paris: Bordas; 1979. On the interpretation of the advertising images, see also Peninou G. Sémiologie et publicité. In: Fraenkel B, Ed. Entreprise et sémiologie. Paris: Dunod; 1999. p. 23-37 Fresnault-Deruelle P. L'image placardée; pragmatique et rhétorique de l'affiche Paris: Nathan; 1997. Adam JM. L'argumentation publicitaire, rhétorique de l'éloge et de la persuasion Paris. Nathan; 1997. Chamboredon JC. Le parfum sémantique de l'image Ethnologie Française 1994; 2 : 308-9.

19 Arnheim R. La pensée visuelle. Paris: Flammarion; 1976.

him with a series of images and stereotypes to help him advise parents and to prescribe these substances to their accompanying children.

Images of knowledge and images of the product

At first sight, it may seem strange to take an interest in the drawings, illustrations and photographs – in short the published image – in pharmaceutical documents of information. What is the justification for such an interest? The amount of care and space accorded by publishers to these visual zones is a measure of the significance attached to them. After all there would be little point in devoting so much energy to a device of purely marginal importance. What is the figurability value of this imagery? Studies conducted on visual communication suggest that their effectiveness may differ from that of linguistic communication. In this scheme of things, looking at a visual entity and reading a linguistic utterance constitute different intellectual skills authorizing distinct accesses to information [19]. The techniques of visual communication would thus constitute a means of exploiting these differences, of completing the information or reformulating it in another register. To show is to display, to describe, to figure in the literal sense of the word, to spatialize, to synopsize, to treat information.

This incursion into the sphere of specialized professional publications is extremely instructive and brings to light, perhaps more rapidly and clearly, the contradictions inherent in the communication used by manufacturers of medicinal products with health care personnel with regard to objective information or discussions between specialists on immunotherapy. Given a finite number of figurability resources and in view of the characteristics of the audiences concerned, the choices open to publishers are limited: it is difficult, and in any case scarcely acceptable, to stray from the rules and charter governing the scientific figurability of immunology as laid down in primary reviews and higher education text books.

As soon as publishers move away from these canons, even in the case of glossy, sophisticated advertising which clearly endeavors to conform with a deontology supposedly at once with the ideal archetypal vision of the cultivated health professional, we are confronted with the most commonplace visual models used for marketing purposes. When advertisements adopt a more trivial register, it is a sign that these discursive procedures are less designed to convince readers than to supply them, more or less consciously, with arguments to be advanced with their patients.

Acknowledgements

I wish to thank N. Autissier, Faculty of Pharmacy, University of Burgundy, Dijon, the Burgundy Regional Council Research Fund and especially Claire Bruneton for her excellent work and for the help she has given me in writing this article.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Multiple splendor. The one and many versions of the immune system

Anne Marie Moulin

At first view, an interdisciplinary meeting on immunology would seem to be facilitated by the increasing circulation of immunological concepts in daily life, from the "Self" to "T4." The notions of immune system and immunity have spread far beyond the borders of the scientific community and generated a great wave of interest in a broad audience. This circulation reveals shared views between professionals and non-professionals of immunology on the meanings of defense and survival, identity and foreignness, recognition and rejection.

These popular views may be abhorrent to immunologists who prefer sophisticated[1] exchanges about refined theories of their own, when, for example, they strive to replace the antiquated Self/NonSelf framework with other theories such as the "danger" one, and rely on fastidious experimental procedures to test their hypotheses. These laboratory experiments are nevertheless human constructs. There is no doubt that the popularity of allegedly simplistic views on immunity bears witness to a profound resonance between specialised immunological topics and everyman's interest in self-defense and the curing powers of nature. This suggests the existence of an anthropological background[2] that articulates a network of meanings between scientists and their clients[3].

My chapter stands at the junction between the public and private (or scientific and cultural) uses of immunology. I suggest that common reveries underlie immunological concepts, and that the metaphysical character of the immune system turns it into an anthropological tool available for the interpretation of various cultural patterns. In other words, the immune system not only functions as a cultural frame inside the Western sphere, but may also be used as a transcultural key. The immune system behaves as a coded language[4] that can be reformatted, glossed, translated, distorted, but unquestionably is a tool for communication.

---

1. "In the original sense of the word", as Rolk Zinkernagel remarks in his presentation of his model of the immune system. Localization, dose and time of antigens determine immune reactivity. Semin Immunol – Self-NonSelf revisited, 2000; 12: 169.
2. This chapter is a completely reformulated version of my earlier paper: Un objet scientifique à la charnière des sciences biologiques et sociales : le système immunitaire. Rio de Janeiro: Historia, Ciências, Saude, Manguinhos; 1996; 2. 300-18.
3. Kleinman A. Medicine's symbolic reality. On a central problem in the philosophy of medicine. Inquiry 1973; 16: 206-13.
4. Moulin AM. Text and context in biology. Poetics Today 1988, 9: 145-61.

Mythologies

Immunologists in their popular books often make reference to mythology:

> Myth is constituted by the loss of the historical quality of things, things lose the memory that they once were made. The world... comes out of myth as a harmonious display of essences[5].

The immune system is supposed to say something about the doings of Nature. In referring to myths, immunologists implicitly acknowledge the continuity between common sense and immunological concerns – what I call the anthropological background.

In his book "La sculpture du vivant,"[6] the immunologist Jean-Claude Ameisen refers to the myths of antique Greece on immortality. He recalls how mythology expressed dreams of immortalizing the human species and picks up Ariadne's thread that leads to contemporary attempts to postpone untimely death through manipulation of the immune system, defined as an ever-changing balance between destructive and constructive forces, embedded in its structure. He locates in mythology remote analogs of standardized immunological techniques. The magician Medea, for example, in order to rejuvenate his father-in-law Eson, opened up his veins and replaced his extenuated blood with plant juices and extracts from animals known for their longevity. In a similar way, the goddess Demeter anointed Eleusis king's son with ambrosia in order to turn him into an immortal. However, when she wanted to test his immunity by passing him through the flame, the glow of fire woke up the mother who wrought the child from his divine nurse[7]. It would be since this time that children cry when they are vaccinated...![8] In another tale, the story of Achillus's heel might also be quoted as a fair example of immunological defect.

But let us turn to harder texts.

Biology

Nowadays the immune system is currently conceived of as a network of cells and molecules endowed with a function in the body called immunity. Immunity, once pointing to a specific property, intervening only in very special cases for adaptive purposes to changes in the environment, has been turned into a general property incessantly involved in the survival of the organism. From an exceptional event linked to the onset of an epidemic or the meeting with a vaccinator's lancet, immunity has become a permanent function of the body, designed to sustain its ongoing combat against disease and death.

The immunological viewpoint is now accepted as one fundamental facet of bodily activities, of physiology and pathology, in the same way as evolution, development, or reproduction. Immunity has become a basic function, ranking beside the other great ones such as nutrition, circulation, respiration. The notion of immune system has led to convergence between scientists operating in distinct fields such as transplantation,

5 Barthes R. Mythologies. transl. by Lavers A. New York: Hill and Wang; 1972.
6. Paris: Seuil; 1999.
7. Chuvin P. La mythologie grecque. Paris: Fayard; 1992.
8. Moulin AM. Immunologie. In: Dictionnaire culturel des sciences. Paris: Editions du Regard; (in press).

zoology, genetics and favored the explosion of immunology[9] in the 1970s, which, in terms of grants and academic prestige, ranked high among the leading disciplines in biomedical sciences[10].

The definition of the immune system as an adaptive system evolving over time both raised and solved the issue of autoimmunity. Nineteenth century biologists had treated the possible existence of "autoimmunity" (immunity directed toward the components of the body) as a major taboo. Immunologists postulated the presence of ingenious mechanisms that would ensure the observation of the taboo by the cellular machinery[11]. Paul Ehrlich invented a regulatory mechanism consisting of antibodies able to knock out any emerging autoantibodies in the wake of immunization[12]. But the permanent destruction of elements belonging to the organism was contradictory per se[13]. This was a paradox: in order to maintain its integrity, the organism had to anticipate and destroy potentially obnoxious elements belonging to the fabric of the body.

This criticism led researchers to abandon Ehrlich's taboo of autotoxicity and to acknowledge a physiological level of silent autoimmunity. In the 1960s, it was hypothesized that this autoimmunity, once amplified, gave birth to autoimmune diseases, a framework which, firstly restricted to some rare anemias and experimental syndromes, soon included affections formerly recognized in other nosological frameworks and that were poorly understood, such as diabetes and multiple sclerosis. The idea that pathological autoimmunity was a mere amplification of a low-rate physiological reaction revived the so-called Broussais problem, i.e., the debate on the existence of a sharp divide between the normal and the pathological, subsequently discussed by Georges Canguilhem[14].

The theory of generalized autoimmunity acquired even more strength with the vogue of idiotypic regulatory networks[15]. Antonio Coutinho gave to this idiotypic regulation a literary formulation, when he tried to emphasize the conviviality of networks permanently regulating the mounting of an immune response. He was pictured in the French daily newspaper *Libération*, with two mice in his hand, one black and one white, and the following humorous caption: "Immunology, a school of tolerance!"[16] But Coutinho carefully refrained from saying what tolerance exactly was about: the enjoyment or the stoic acceptance of diversity, or the unawareness of differences, obviously corresponding to very different patterns of tolerance.

---

9. Moulin AM. Le dernier langage de la médecine. De Pasteur au Sida. Paris: Presses universitaires de France; 1994.
10. Nossal GB. Trials and triumphs of immunology in the 1980s. Immunology Today 1988; 9: 286-91.
11. Silverstein AM. Cell Immunol 1986; 97: 173-88.
12. Ehrlich P. Zur Theorie der Lysinwirkung (Uber Haemolysine). Berl Klinische Wochenschr 1901, 38. 569.
13. See comments in Jacquemart F, Coutinho A. Observer, immune system and their respective objects. In: Sercarz EE, et al., eds. The semiotics of cellular communication in the immune system. NATO ASI Series H. Berlin: Springer Verlag; 1988; 23: 173.
14. Canguilhem G. Le normal et le pathologique. Paris: Presses universitaires de France; 1988. Moulin AM. La médecine moderne selon Georges Canguilhem. Concepts en attente. In: Georges Canguilhem, philosophe, historien des sciences. Paris: Albin Michel; 1993. p. 121-34.
15. Jerne NK. Towards a network theory of the immune system. Ann Immunol 1974; 125C. 373. Urbain J, et al. Idiotypic regulation of the immune system by the induction of antibodies against anti-idiotypic antibodies. Proc Natl Acad Sci USA 1977; 74: 5126. Jerne NK, Rolland J, Cazenave PA. Recurrent idiotypes and internal images. EMBO J 1982; 1: 243. Jerne NK. Idiotypic networks and other preconceived ideas. Immunol Rev 1984; 79: 5-24.
16. Liberation 22 May 1991, p. 20.

As autoimmunity no longer pointed to pathological reactions, it became available for more general functions hereafter: self-recognition and internal communication. Coutinho speculated about the existence of a distinction between a peripheral immune system and a central one featuring high "connectivity" and multireactivity[17]. He characterized the core system by a high level of regulation and oscillatory mode of work. In case of disease, the immune system tends to function according to an unusual clonal mode, including expansion and hyperproduction of deleterious molecules and antibodies. If one accepted Coutinho's suggestion, then one might consider as a form of therapeutics the introduction of autoantibodies, allowing the body to reconquer its usual balance and modify its microenvironment. The immune system would up- or downregulate the activity of antibodies directed against physiological targets belonging to other systems of the body such as the cardio-vascular, endocrine, or nervous systems.

The relationship between germs and their hosts covers a multiplicity of situations, from mutual assistance to irreconcilable fighting through oscillating phases. Parasites such as *Plasmodium*, the agent of malaria, offer fascinating models of this relationship. Adults, in endemic areas, survive with an enormous *Plasmodium* burden (multiple parasites in their blood cells) that would kill naive people (or young children) exposed for the first time to the infection.

Even if few therapeutics of this kind have passed into current use, some have suggested manipulating the immune system by introducing molecules directed against the self-components. Even if all diseases are not necessarily conceivable in immunological terms, they could be immunologically manipulated. Monoclonal antibodies offering the promise of "exquisite specificity," the Eldorado of immunologists, are candidate therapeutic agents to control hypertension. They might, in addition, neutralize the heat shock proteins released in acute infections in intensive care patients. Immunorestitution, a new password, points to the recovery of a putatively lost immune integrity. Here we may be forced to take into consideration Georges Canguilhem's caveat, "the cure never means the return to the cellular innocence."

Self/NonSelf in danger

A chapter of this book recalls in detail how Elie Metchnikoff elaborated his theory of immunity[18]. Metchnikoff's construct underwent sharp criticism by those who felt that it smelt of teleology and was heavily dependent on a philosophical agenda that included notions such as the pursuit of harmony and of a peaceful and timely mode of passing away, in short, natural death rather than immortality. Phagocytosis had for many years been relegated to a minor rank before being rehabilitated as innate immunity which is considered as a first line of defense against the hazards of life during evolution.

Following Burnet's work in the 1960s, the Self and NonSelf approach was adopted for several decades as a satisfactory background theory although from the very beginning, the

17. Coutinho A. Beyond clonal selection and network. Immunol Rev 1989; 110: 66-87.
18. Metchnikoff E. Leçons sur la pathologie comparée de l'inflammation. Paris: Masson; 1891. Tauber AI. Metchnikoff and the origins of immunology. From metaphor to theory. New York: Oxford University Press; 1991.

biological nature of the Self seemed elusive[19]. After some attempts at its characterization, its promoter Frank Macfarlane Burnet gave up trying to define it precisely. Yet many biologists resented the almost metaphysical assumptions underlying it and sought to drop it.

The so-called danger model was introduced by Polly Matzinger in the early 1990s[20] and reprised Melvin Cohn's proposal that whereas two cellular signals are required for the mounting of a specific immune response, one signal, if isolated, tolerizes the cell. Matzinger's idea is that one of these signals occurs when a cell is damaged (by a pathogen, for example) and when some inflammatory cytokines are released in the medium. A myriad of molecules, from NO oxides to heat-shock proteins act as chaperones and can thus trigger an immune response. The danger model, by interweaving non-specificity (damage) and specificity (recognition effected by antigen-presenting cells), follows the tradition of associating non specific and specific components, a "form" and a "matter," in the description of the immune response. The danger model epitomizes the famous statement that ontogeny reproduces phylogeny, the non-specific component being the more primitive.

But the phrasing of the "danger model" does no more to explain the explosion of an immune response than the Self-NonSelf model; it is not more natural for a cell to sense danger than to preserve its integrity. In other words, to feel danger is very analogous to looking after its own Self. As David Napier argues[21], sensing danger refers to the sense of the integrity and this integrity in turn refers to something or to a substance that might be a kind of Self.

Janeway, approved by Matzinger, has gone further by suggesting that all cells, when prompted to die before the due time for apoptosis, can deliver signals of the first kind[22]. For Janeway, this represents the most archaic line of defense. All cells, at the beginning of our era, were probably able to deliver such a pathetic cry for help. If we accept Janeway's proposal, the immune system would no longer stand apart in the body with its anatomical basis, its central and peripheral organs, reflex and regulatory pathways. Immunity would tend to become a general property of all cells and living tissues, as indeed was the case in organisms before the division of labor: this is the case of protozoa such as amoebas, a model favored by Cohn and Langman in their version of immunity[23]. In other words, the immune system would stand for the body.

Even if we question this way of animating the cellular dialogue and endowing the machinery of molecular biology with human affects and behavior, we must acknowledge the influence of this immune system talk on medical language.

---

19. Moulin AM. La métaphore du soi et le tabou de l'autoimmunité In: Bessis M, Bernard J, Debru C, Eds Le soi et le non-soi. Paris: Seuil; 1990. p. 55-68 The so-called model of the peptidic self led to distinguish a somatic self and an immunological self: Kourilsky P, Claverie JM. Le modèle du soi peptidique Med/Sci 1988; 4: 177-83. Löwy I The immunological construction of the Self. In: Tauber AI, Ed. Organism and the origins of Self. Boston: Kluwer, 1991 p. 43-75.

20. Matzinger P. Tolerance, danger and the extended family. Ann Rev Immunol 1994; 12: 991-1045

21. Napier D. The age of immunology (forthcoming).

22. Janeway CA Jr Innate immunity: the virtues of a non clonal system of recognition. Cell 1997, 91: 295-98. Matzinger P. An innate sense of danger. Semin Immunol 1998; 10: 399-415.

23. Langman RE The immune system. San Diego: Academic Press; 1989.

Medical

Following the direction sketched above, the complexity of the immune system has considerably increased. If the definition of the immune system as a network of cells and molecules makes the link with the past, the identification of these cells and molecules has considerably progressed, and the tools of molecular biology have permitted the dissection of numerous receptors and cytokines[24] which foster great expectations from these new molecules. The popular metaphor of the immune system as an orchestra, once invented by the immunologist Fred Gershon, implies a variety of instruments and musicians, or possibly the replacement of the director by many secondary choir leaders. In keeping with the music metaphor, one might say that immunity is phrased as identification of signals and their integration into the cellular symphony. Genetic orientations are built into the fabric of the immune system and determine its power to mount a response against ordinary and extraordinary pathogens. Whether an imperfect response against an antigen is due to ineffective antibodies, and/or a genetic incapacity to mobilize appropriate categories of cells and/or the propensity to develop illness in a given organ, in any case, the global response for all these events has to be found in the immune system and its activity. The immune system has thus emerged as a cause, in the fullest, "Aristotelian" sense of the word: a formal, material, final cause of pathogenesis and recovery[25]. It has integrated the three time dimensions by including infections and traumas of the past, present encounters with environmental hazards, and prospective risks embodied in genes involved in susceptibility to disease. Among other issues, the immune system makes it possible to resume the traditional debate on "soil and germ".

But an epistemological difficulty is linked to the fluidity of the framework articulating medical discourse. So many reaction cascades are available for explaining a disorder that it is hard to choose between various explanatory pathways. Let us illustrate the case by telling a story about a patient who suffered from an intriguing infection. An intestinal worm, usually restricted to the guts, had honed in on the most bizarre locations; it had colonized the stomach in spite of the acidity of its content and passed the meningeal barrier, reputed to be impermeable to macroparasites. It was hard to decide whether some defect in immunity had caused this dissemination or whether this parasitic burden had induced an impairment of immunity. Furthermore, it was impossible, from a holistic point of view, to determine the sequence of the events (infections, misery, sexually-transmitted diseases...) that had led to the present condition. The discovery, years later, that the patient was infected by human T-cell leukemia virus (HTLV), introduced a further additional cause for immunosuppression, without solving the sequence of pathological phenomena.

The immune system conjures up a graph where it is possible to plot the onset, the development and the outcome of pathological processes. It provides a response to inquiries on pathogenesis for poorly understood disorders, somatic or even psychological. In Western countries, practitioners currently mention modified immunity to account for various

24. Moulin AM, Silverstein AM. History of immunophysiology. In: Oppenheim JJ, Schevach E, Eds. Immunophysiology. The role of cells and cytokines in immunity and inflammation. Oxford: Oxford University Press; 1990. p. 3-13.
25. Moulin AM. The dilemma of medical causality and the issue of biological individuality. In: Deleskamp-Hayes C, Gardell Cutter MA, Eds. Science, technology and the art of medicine: European-American dialogues. Dordrecht: Reidel; 1994. p. 153-64.

health problems: flu, common cold, chronic affections, or conversely incriminate pregnancy, sorrows, travelling, depression in a hypothetic immune dysfunction.

We have seen this before with the expansion of the notion of "stress" into a growing number of areas[26]. In the same way the immune system is now being presented as crucial for survival as well as quality of life, and immunomodulation is advertised as the most effective and softest management of disease, similarly, stress research in the past fostered the idea that the understanding of stress would yield clues for a healthier and happier life.

The concept of stress was invented by Hans Selye in the 1930s to account for the shock vascular syndrome in animal experimentation. It consisted of inflammatory reactions and corticoid secretion after aggression, and was soon extended to the response to various kinds of traumas. In the 1950s, in the context of cold war and messianic expectations of economic prosperity on the Western side, stress was endowed with new and extended meanings and became synonymous with mobilizing the energy of the organism and possibly responsible for deleterious side-effects. How Red Riding Hood meets the wolf and successfully overcomes her panic and runs away, has been told as a successful story of stress. According to Selye who greatly contributed to popularizing the notion[27], stress ultimately meant the hardship of life: unemployment, racial vexations, sorrow, mourning and suffering in general[28]. Stress was thereafter included in the factors influencing the level of immune defense and lately, stress has been incorporated into the domain of psycho-neuroimmunology[29].

Like stress, once primarily a scientific object modeled and studied in the laboratory, the immune system has become available as a framework for narratives of subjective experiences of illness and recovery. Solicited once by a journalist working for the tabloid press, I was surprised to be confronted with what I misunderstood as an appetite for the latest news about the immune system. I did my best to explain about apoptosis and lymphocyte stimulation, autoimmunity and the promise of new vaccines. The journalist listened to me earnestly without making notes and claimed that she understood everything. Imagine my surprise when upon reading the digest of my lectures in the next issue of the journal, I discovered how I had taught the reader and subsequently the readers how to manipulate his or her own immune system by feeding it with spinach and yolk egg and adopting the appropriate lifestyle. In short, a fashionable taste for immunology had crept into patients' complaints, taking its place beside the French *crise de foie* or the Japanese *katakori*[30], when immunity literally exploded onto the scene of popular culture with the AIDS epidemics[31].

A good example of the availability of the immune system as an explanatory framework in this period is the convergence of observations of an unknown syndrome that brought

26. Selye H. The development of the stress concept. In: Parvez H, et al , Eds. Advances in experimental medicine. A centenary tribute to C Bernard. North Holland Biomedical Press – Elsevier; 1980. p 43-69. See also Young A. The discourse on stress and the reproduction of conventional knowledge. Soc Sci Med 1980 ; 14B : 133-46. Moulin AM. Une devise pour l'organisme. In: Résister. Autrement, March 1994; 22-9.

27. Selye H. The stress of life. New York: McGraw Hill; 1956; Stress without distress. New York: Lippincott; 1974; The stress of my life. New York: Van Nostrand Reinhold; 1979.

28. Cooper EL, Ed. Stress, immunity and aging. New York: Marcel Dekker; 1984.

29. Locke A, Ader R, et al. Foundations of psychoneuroimmunology. New York: Aldine, 1985 Corson SA. Historical and philosophical background of immunoneuromodulation. Intern J Neurosci 1988; 39: 283-7.

30. Kuriyama S. The historical origins of katakori. Jpn Rev 1997; 9: 122-49.

31. Seytre B Sida, les secrets d'une polémique. Paris: Presses universitaires de France, 1993.

together infections by so-called opportunistic germs who displayed their pathogenicity in organisms of altered resistance (due sometimes to the iatrogenesis once vilified by Ivan Illich). The immune system suddenly appeared in everybody's intellectual armamentarium. The AIDS virus emerged as a virus specifically damaging for the immune system, adhering electively to receptors carried by those cells whose defensive action is crucial for survival.

The role played by T4 helpers in the development of the disease and the delayed failure of defense mechanisms has since become commonsense[32]. The meaning of the immune defeat went far beyond the tissular damage induced by retroviruses. Immune diseases strike not only those whose defenses are biologically compromised but those who have them socially altered (such as drug addicts, for example) and whom the collapse of social solidarity leaves helpless – groups most at risk in industrialized societies or poor populations of the developing world[33].

When in 2000, seizing the opportunity of the big AIDS meeting in Durban, President Mbeki of South Africa contested the current interpretation of AIDS as a contagious disease, he raised an uproar of protest. At first sight, it seemed that President Mbeki concurred with the American scholar Peter Duesberg who, at an early stage of AIDS research, had contested the scientific focus on HIV viruses[34] and had suggested scientists investigate other factors contributing to the emergence of the disease[35]. But was President Mbeki expressing a biological opinion when he said that political men would do better to address poverty than stick to the scientific issue of virus receptors? He pointed to factors important in contamination and the development of the disease such as the lack of education and the absence of decent means of living. He sent the political message that these factors and all factors alike might be actually more relevant than the virus itself and in short pleaded that the construction of immunity was as much culturally and socially as naturally grounded.

An object which so successfully links together disparate things and people is likely to involve a remarkable degree of semiotic flexibility or adaptability to successive contexts and variable purposes in users' circles. To the complexity of the construct corresponds a "fluid epistemology," a software with which are built up understandings and *faux sens*. But is it possible to develop a fruitful concept without a multiplicity of meanings which stimulate creative minds and potentially leads to falsification of the original contents?

Over centuries, doctors have elaborated a theoretical model of health and disease around a central tenet. The Hippocratic paradigm pivoted around ideas of depletion and plethora.

---

32 Many novels have focused on the tribulations of the immune system in AIDS or cancer-stricken protagonists, such as Navarre Y. Ce sont amis que vent emporte; Camus R. Elégies pour quelques-uns; Barbedette G. Mémoires d'un jeune homme devenu vieux. Hervé Guibert described the clinical stages of the immune drama in: À l'ami qui ne m'a pas sauvé la vie; the opportunistic infections in: Cytomegalovirus: the clinical trials. in: Le protocole compassionnel (1991), before the ultimate Paradis of 1992. See also Lévy J, Nouss A. Sida-fiction. Essais d'anthropologie romanesque. Lyon: Presses universitaires de Lyon; 1994. Rémission is the title chosen by Alain Roger.

33. Brandt A. AIDS and metaphor. Toward the social meaning of social diseases. Soc Res 1988; 55: 413-32.

34. Duesberg P. HIV is not the cause of AIDS. Science 1988; 241: 514-6; Human immunodeficiency virus and acquired immunodeficiency syndrome. Correlation but not causation. Proc Natl Acad USA 1989 86: 755-64.

35. The Barbara McClintock Project (1993), supported by the association Act-Up, was named after the American biologist McClintock, whose non-conventional approach to genetics was crowned by the Nobel Prize. It aims at stimulating alternative hypotheses to the role of the virus.

Based obsessively on the paradigm of female physiology, the management of the disease consisted of dietetic guidance and regular bloodlettings aiming at maintaining or restoring humoral balance.

This paradigm has survived and has co-evolved from the end of the nineteenth century with a new vision centered around the fight between host and pathogen. At that point, doctors declared war on germs and aimed at the eradication of the enemy[36]. They speculated on the sterilization of the living medium, purified of all possible bacteria, not only the skin but the mazes of the intestines, the pulmonary passages... The idea of stamping out microbes and consequently disease in individuals by all possible means held sway on the physicians' minds and paved the way for the craze of eradication which seized international organizations in the second half of the twentieth century.

Today, a more sober discourse has prevailed, admitting the necessity of compromises. The awareness of resistances has led to strategies other than scorched earth, lest like in the Gospel, fiercer demons replace the former ones, previously chased from the home. From the idea of diseases emerging sporadically in an immunocompromised host, one can easily move to the idea of a compromise between every organism and the host of microorganisms besieging it both inside and outside. This idea that the immune system is constantly challenged and challenging is supported by the observations of the heretofore unnoticed immunological daily life. In Senegal, the village of Dielmo has been the seat of an epidemiological watch for the last 15 years. All inhabitants have been investigated around the clock for fevers or other symptoms. The unrelenting surveillance system has allowed researchers to detect transient drops in immunity, most notably in the post-partum period, leading to malaria development in mothers exposed to contamination, during the first term after delivery[37]. Casual encounters with pathogen agents incessantly shift the equilibrium in the body. Vaccination long viewed as unequivocally protective has been reconceptualized as the activation of multiple pathways in the body, with complex consequences difficult to assess which do not exclude a degree of rearrangement in immune receptivity.

Public/private images

We have surmised, from a rapid review of transitional forms of discourse between scientists and non-scientists, that immunological language can transmit general human concerns: this is probably what Melvin Cohn means when he mentions the "loaded"[38] character of the Self and NonSelf issue. It is not so much because scientists consciously attempt to popularize their esoteric knowledge than because scientists first address and convince themselves: they need, in the first place, to make nature intelligible to themselves. The danger theory illustrates how immunologists wrap (and need to wrap) in simple words complex biochemical reactions in a way that make sense for them and incidentally for their human fellows.

---

36. Gradmann C. Invisible enemies: bacteriology and the language of politics in imperial Germany. Science in Context 2000; 13. 9-30.

37. Trape JF Criteria for diagnosing clinial malaria among a semi-immune population exposed to intense and perennial transmission. Trans R Soc Trop Med Hyg 1983. 435-42; Increased susceptibility to malaria during the early post-partum period. N Engl J Med 2000; 343: 598-603.

38. Langman RE, Cohn M. Editorial introduction. Semin Immunol 2000; 12: 159.

This transition between private and public science can be illustrated by a beautiful video-clip, produced in the 1990s by a senior immunologist[39]. He sketched the main lines of the immune system in terms of a fictive correspondence between a scientist and his mother. The choice of the protagonists is no mere chance. The mother, once an actor in the immunological drama of pregnancy, plays the *confident* in the tragedy. She allows the hero to express his beliefs and doubts. Mother and son were once fusional partners. Quite naturally, it is she who shelters and fosters the early version of a dialogue between Self and 'former' Self and mediates it between Self and the Other.

"How can I tell you how happy I am when you disclose to me your secret garden". The Garden of Eden once harbored universal knowledge. And it is again the metaphor of the garden that comes to the mother's pen during her initiation into the "field" of research: "I cannot help of thinking of the immune system as a big park with its trees, groves and alleys."

But a second metaphor competes with the former one: "I see the immune system as the big book where Grand Pa used to record the main family events: but it is difficult to single out changes in an ever-changing context." The difficulty of defining a changing bodily identity is echoed by the anthropologist Sarah Richardson in her article on the "End of the Self"[40]: "The Self is constantly being defined anew". But Richardson goes further by adding: "which is another way of saying that it doesn't really exist at all." The mother does not go to this extreme. She quietly addresses the Heraclitean dimension of life, the constant dissolution of elements that nevertheless maintain a form: "the system is our guardian, on a permanent watch and fulfills its duty by confronting continuously what exists today with what existed once."

The permeability of the border between the outside world and the interior of the organism[41] is also grasped as obvious: " We are open wide to the outside world: we are the milieu for a host of microbes: is it not to acknowledge that we are not alien to the milieu that surrounds us?" This is the intuition that beings originating from the outside (such as viruses, for instance) may have been part of our inner selves. The mother takes a similar intuitive approach to autoimmunity, which stands in symmetrical relation to the invasion from the outside.

Gaston Bachelard used to invite scientists and philosophers to explore and exploit jointly the resources of the unconscious, in the name of creative rationality. This poetical correspondence about the immune system, an official creation of contemporary biology, illustrates the kinship between biological and metaphysical thinking. The development of immunology as a broad biological science opened a new semantic field. It can be considered as a language into which it has become possible to express and potentially to solve general questions raised by the interaction between man and his environment, Self and the Other, seen as another kind of Self[42]. "*Aime ton prochain comme toi-même,*" says the

39. Daëron M. Cellular and clinical immunology. Paris: INSERM U 255, Institut Curie. The clip was produced by the INSERM for a broad audience and school students.
40. Richardson S. The end of the Self. Discover 1996; 17: 80. I thank David Napier for having attracted my attention to this quotation.
41. Moulin AM. Le dernier langage de la médecine. De Pasteur au Sida. Paris: Presses universitaires de France; 1991.
42. Daëron M. Le système immunitaire ou « Connais-toi toi même ». Colloque de Cérizy Praxis et Cognition, 1988.

Gospel, suggesting a strange and profound kinship between the various selves. Many cosmogonies commonly associate love and hatred as the principles of being.

Anthropology: the Self and the Other

The immune system is one of the rational constructs of our contemporary biomedicine: the identification of the body as a network of cells and molecules, strong and fragile like the web of life. This powerful representation challenges the imagination: any event registered in the Book of the system is determining for survival, until we reach, with our back turned to the future, the term fixed for the end. This system might contain part of the secret of biological individuality and the key to decay and destruction. Whoever detains the password of the Self and NonSelf and is able of manipulating the immune system becomes the master of life and survival.

While questioning the divide between the Self and the NonSelf, immunology first of all meets the philosophical queries such as in Plato's Meno: Is to know a form of remembrance, in which case NonSelf is identified with Self (see Descartes's innate ideas), or invention, and then how is recognition possible? Christ says to the mystic: "Thou would not look for Me if thou had not already found Me."

Anthropologists were eager to point to the convergence between biological systems and other semiotic products of our culture. Donna Haraway, in the wake of her work on cyborgs and post-modern beings, described the immune system as a typical post-modern icon: she views the body, either healthy or ailing, as a robot ready to be assessed and manipulated by the medical profession[43]. Following the adventure of smallpox inoculation[44] immunization campaigns against other pestilential diseases have tried to synchronize populations' immune reactions to pathogens. The state increasingly controls bodies, exploiting the "governmentality of life." The politics of public health have dramatically modified demographic trends. Donna Haraway conceives the manipulations of both individual and collective bodies as typical of modern science, obsessed with warfare and accounting, unable to turn away from dichotomic and mechanistic thinking stigmatised by Evelyn Fox-Keller[45]. Only in recent years have epidemiologists organized the careful registration of "vaccination-related adverse events," and discovered a long-neglected potential source for original clues to pending questions on individual pathways of immunity. Moreover, while acknowledging the output of public health measures on the modern rise of populations, historians point to the authoritarian character of such measures, the break in traditional lifestyle and social customs. They point to the risk of exclusion for some social groups targeted as "vectors" of diseases[46].

---

43. Haraway D. The biopolitics of post-modern bodies: determination of Self in the immune system discourse. Journal of Feminist Cultural Studies 1989; 11: 3-43.

44. Moulin AM, Ed. L'aventure de la vaccination. Paris: Fayard; 1996.

45. Fox-Keller E. Refiguring life: metaphors of twentieth-century biology. New York: Columbia University Press; 1995 Napier D. Penser 'vaccinologiquement'. une sélection qui n'est pas vraiment naturelle, ou les modèles sociaux du monde microbien. In. Moulin AM, Ed. L'aventure humaine de la vaccination. Paris: Fayard; p. 409-22.

46 Rivet D. Hygiénisme colonial et médicalisation de la société marocaine au temps du protectorat français: 1912–1956. In: Longuenesse E, Ed. Santé, médecine et société dans le monde arabe. Paris: L'Harmattan; 1995. p. 105-28. Peter JP. Dimensions mythiques des épidémies et Sida. Action et recherches sociales 1989; 3: 15-29. Weindling P. Medicine and holocaust. The case of typhus In: Löwy I, Ed Medicine and change. historical and sociological studies of medical innovation. Montrouge: Editions INSERM/John Libbey; 1993. p. 447-64.

Another American anthropologist, Emily Martin, has analyzed the vision of the immune system, on the basis of research conducted, in a hospital, on the reception of immunological knowledge[47]. In the 1960s, the immune system was seen as composed of a central organ such as the thymus and peripheral effectors such as the lymphatic glands, and displayed centrifugal and centripetal pathways. The latest version of the system is different and exhibits the characteristics of economic life in advanced societies: mobility, polyvalence, and plasticity[48]. Hermann Wolf Fridman chose to call his popular essay on immunity "The mobile brain" (*Le cerveau mobile*)[49]. The immune system integrates the passwords of the globalized world, in the age of liberal economic thinking. Rolf Zinkernagel sees in the interplay of time, antigen dose and location in the body the secret of flexible autoimmunity regulation[1]. In some of the "danger models," we have seen that any damaged cells can deliver a signal and trigger an immune response.

The traits of contemporary American society are reflected as well in immunology as in the postulates of interactionist sociology of the Chicago School. In the market place, everything is negotiable, and standards are unstable. In the immune system, not only lymphocytes, but most tissues can be recruited for immunological tricks, in the "ball of cells," according to Lewis Thomas's title.

Ethnography

The immune system can be interpreted as a post-modern icon, influenced by the contemporary currents. However, although immunological language is a Western code rooted in twentieth century biomedical knowledge, its use is far from being restricted to industrialized societies[50]. It can function as a reference in comparative anthropological studies. From a cross-cultural perspective, the immune system can be used to explore various cultural views of illness, judged at first sight as incommensurable[51]. The immune system has created a semiotic field where one can project distinct cultural patterns of disease, belonging to remote pasts or exotic countries, and made it possible to detect hidden links between apparently distant narratives of illness.

The anthropologist Dominique Buchillet has detected a variant of king Mythridate's myth of immunization among the Desana of Amazonia[52]. The destitute Indians resent the apparent good health of their former invaders and their insolent immunity to epidemics

47. Martin E. Flexible bodies: tracking immunity in American culture from the days of polio to the age of AIDS. Boston: Beacon Press; 1994.

48. Martin E. The cultural construction of gendered bodies: biology and metaphors of production and destruction. Ethnos 1989; 54: 143-60. Toward an anthropology of immunology: the body as nation state. Med Anthrop Q 1990; 4: 410-26.

49. Fridman WH.Le cerveau mobile. Paris: Hermann; 1991.

50. For other examples, see Kleinman A. Concepts as a model for the comparison of medical systems and cultural systems. Soc Sci Med 1978; 12: 85-93.

51. It does not imply to create a continuity between various traditions or medical systems: see Zimmermann F. Généalogie des médecines douces. Paris: Presses universitaires de France; 1995. From classic texts to learned practice: methodological remarks on the study of Indian medicine. Soc Sci Med 1978; 12: 97-103. For translation matters related to the commensurability issue, see also Zimmermann F. Terminological problems in the process of editing and translating in Sanskrit medical texts. In: Unschuld PO, Ed. Approaches to traditional Chinese medical literature. Dordrecht: Kluwer; 1989. p. 141-50.

52. Buchillet D. Perles de verre, parures de blancs et « pots de paludisme ». Epidémiologie et représentations Desana de la maladie infectieuse (haut Rio Negro, Brésil). Journal de la Société des Américanistes 1995; 81: 181-206.

that are fatal to the tribes. They tell that the Whites did not flinch from drinking immortality with coca, while Indians were more squeamish. Since this time they pay the price for not having been brave.

Anthropologists have been attentive to the way populations adjust or distort in their myths and rituals scientific innovations from elsewhere. In the Meiji period, the Japanese reinvested the *Kamis*, geniuses that protected against smallpox (*Kami* is sometimes translated by guardian angel), with new meanings and linked their ancient rituals to the practice of Western vaccination[53]. Today, historians no longer consider popular resistance to scientific progress as a simple proof of backwardness but undertake a careful in-depth analysis of acceptance and rejection[54].

The American anthropologist David Napier has promoted a study of reactions to Western knowledge about AIDS in various cultures: how do Balinese, for example, integrate the western version of immunity into their religious beliefs and rituals?[55] He notes that they seem to understand and integrate current assumptions about the immune system and its role in HIV infection: the idea that a tiny invisible being has the power of investing and harming the body fits easily with their current demonology. Only those historians who think that bacteriology has been an absolute departure from the former beliefs into occult forces and invisible miasmas can deny the analogy between scientific beliefs in the multiple entities of the immune system[56] and such cultural demonology[57].

The immune system recapitulates the questions on the uncertain status of the body in Nature. It solicits queries on the boundaries between the body and other living beings, constantly renegotiated during the course of life[58]. On the one hand, ecologists complain that modern practices have inadvertently ignored the natural barriers between species (e.g., the bovine spongiform encephalopathy affair). However, 40 years ago grafting had already bypassed the dogma of the "uniqueness of the individual." The success of transplantation with living or deceased donors has pointed to the fluid relationship between the living and the dead or within the living community itself, again an open space for negotiation[59]. On the other hand, looking to the inside of the organism, no absolute guarantee against autoimmunity can be given, and the emergence of immunity may remain a matter of context, as Rolf Zinkernagel has indicated.

If the immune system is suggestive of the issue of limits between nature and culture, the Self and the Other, the living and the dead, it is possible to identify analogs in other cultures.

---

53. Rotermund HO. Hôsôgami ou la petite vérole aisément. Paris. Maisonneuve et Larose: 1991.
54. Moulin AM. Premiers vaccins, premières réticences. Pour la science 1999; 264: 12-5.
55. Napier AD. Foreign bodies: performance, art and symbolic anthropology. Berkeley: California University Press; 1992 Also see Jacquemart F Préliminaires à une théorie générale anthropocentriste des objets mous [doctoral thesis in immunology]. Paris: University of Paris 6; 1990
56. Cambrosio A, Keating P. A matter of FACS: constituting novel entities in immunology. Med Anthrop Q 1992; 6: 362-84.
57. Napier D. The age of immunology (forthcoming).
58. Moulin AM. A science "dans le siècle": immunology or the science of boundaries. In: Krige J, Pestre D, Eds. Science in the twentieth century. Amsterdam: Harwood Academic Publishers; 1998. p. 475-95.
59. Moulin AM. Body parts: the modern dilemma Transplant Rev 1993; 95: 33-55. La crise éthique de la transplantation d'organes. À la recherche de la "compatibilité" culturelle. Diogène 1995; 172: 76-96; Postface. In. Carvais R, Sasportes M, Eds. La greffe humaine. Paris: Presses universitaires de France; 2000; p 749-64.

Among the Ngbaka villagers of Central Africa, along the Lobaye River, the genesis of disease works along two interpretive modes. Either disease is attributed to aggressive agents from the outside – flies, mosquitoes, melipones (wild honey flies which colonize the natural openings of the body in the rain forest) – or it is attributed to the doings of a wild beast crouched in the body, responsible for antagonistic effects. Alternatively, it stimulates and chastises, replenishes and starves the body, induces or restricts bleeding, stops or impedes the humoral flow[60]. This devouring beast, alternatively supportive or destructive, is a perfect illustration of the equivocacy of the Self to self-relationship, a strange relation associating fusion and autonomy, love and hatred.

Among the Avikam in Ivory Coast, a wound is perceived as a major hazard, putting the body at risk by disturbing the gentle flow in and out of the organism and canceling the concentration gradient, leading to homogenization and death[61]. The Avikam live on a lagoon, with log cabins built on woodpiles in a marine environment. Their daily life is an unrelenting battle against the invasion of salty water that destroys homes and drowns bodies, but that also provides the main resources for this population of fishermen. The story of the Avikam is similar to the epic story of Hans, the little Dutch hero, who kept his finger in a hole in the dike all night, waiting for help.

Surgery was, in the last century, celebrated as one of the greatest medical achievements and received almost unanimous applause. With the growing awareness of so-called nosocomial affections, a more sober view now prevails which considers surgical decisions as a potential harm whose consequences must be carefully weighted. As with the Avikam, any surgical intervention figures as a trauma with a risk of infection and pathological cascades. Practices such as the resuscitation of comatose patients, or the use of automated instruments to replace failing basic vital functions, need also to be reconsidered in this respect.

Any biological equilibrium is a negotiation with the fluctuating composition of the external and internal milieus, involving the up- and down-regulation of cells' capacity to proliferate, differentiate, secrete, adhere, or aggregate. Such was the message of the English immunologist Gell when he warned that the immune system is at risk of "bleeding to death"[62] when one of its branches is cut off or conversely is under the permanent threat of inflation by proliferation of cells or antibodies. Even immunization programs, unanimously praised for their major contribution to the decline of infant mortality, are not easily interpreted in a unilateral way. Vaccination activates multiple pathways in the body, varying with the type of antigen, leading to possibly contradictory effects. Some recent data suggest that some vaccines could, on the one hand, stimulate immunity and decrease infant mortality in an unspecific way, while other vaccines might increase atopy and raise

60. Amenorrhea or metrorrhagias are thus easy to explain. The amenorrhea corresponds to a period of nourishment for the parasite. This parasitic beast is necessary to the transmission of life;. Moulin AM, Lévi-Strauss à Kaka (République centrafricaine), CNRS typescript, Bangui 1983; Pagézy H, Couillot MF, Moulin AM. Enquête ethnographique sur la grossesse et l'allaitement, examens médicaux et considérations ethnomédicales. CNRS typescript, Paris, 1984.
61. Gely M. Le corps fragile: écologie du corps et syncrétisme médical chez les Avikam lagunaires de Côte d'Ivoire. Sci Soc Santé 1991; 9: 5.
62. Gell PG. Network concepts in science and the arts. In: Lefkovits I, Ed. The immune system; a Festschrift in honor of Niels Kaj Jerne. Basel: S Karger; 1980. p. 58.

mortality[63]. Moreover, the clarification of the immunological consequences of repeated inoculation of vaccine antigens remains to be worked out. The rise of lymphomas or autoimmune diseases, during the last two decades, has remained a controversial issue.

Metaphysical

Immunity emerges thus as a topic for reverie on the instability of biological borders and frontiers between beings. Literature elaborates in this mood the theme of a citadel where the enemy silently lurks inside or haunts, invisible, outside. The heroes of Dino Buzzati's "Desert of Tartars," wait on the battlements for a mysterious invader without precise indications of time and space. But who sets the limit between the Self and NonSelf, the citizen and the foreigner?

The time-space location suggested by Rolf Zinkernagel plays with thresholds and doses, light and darkness, hidden and exposed, latent and manifest antigens in the body. Ultimately, immunology illustrates a paradoxical element in recognition: the mixture of closeness and distance in the perception of the enemy, nonselves as other selves, the impossibility of identification in the absence of kinship, even inimical. Recognition is a gradual and complex phenomenon where the enemy is more easily recognized, different in that from the absolute foreigner whose name and essence are beyond all knowledge. This is the sense of the Greek distinction, the sociologist Georg Simmel used to recall, between the "meteque," alien in the city but belonging to it and depending on its laws, and the barbarian, who has no human language (the original meaning of barbarian) and no human bond.

A South African novel by Coetzee, "Waiting for Barbarians," depicts the odyssey of a judge settled on the frontier that protects the country from wild natives' invasion. A female prisoner is one day taken to him. The story of this judge who makes love with his prisoner and progressively identifies her as a human being, before being banned from his own society, is not only a political fable written in the apartheid country before the time came for political change, it is also a metaphysical and poetical tale about immunity with underpinnings on the arbitrary character of identities and the reversibility of categories.

But the making of identities is not necessarily the royal pathway to order and morals. Dealing with the case of the former Soviet Asian republics, the political scientist Olivier Roy breaks with the tradition that opposed "good" identities (cultural, linguistic, social) and "bad" identities (racial or pseudo-biological founded on skin, hair, blood...) and shows that the making of cultural identities can also be a source of violence and fantasy[64].

Conclusion

The concept of immune system has made its way both inside and outside the scientific community because it is the product of imagination focusing on the status of the body in the world and among the other bodies. The Western world distanced itself from cultures where disease is attributed to a divine or human maleficent volition. But with their helper, suppressor and dendritic cells, immunologists also have invented all kinds of entities

63. Aaby P, et al. Early BCG vaccination and reduction in atopy in Guinea-Bissau. Clin Exp Allergy 2000 ; 30 : 644-50. Non-specific beneficial effect of measles immunization: analysis of mortality studies from developing countries. Br Med J 1995; 311: 481-5.
64. Roy O. La fabrication des identités. Paris: Le Seuil: 1998.

endowed with intention. Dendritic cells may be reminiscent of some satanic monsters "quaerens quem devoret" (seeking to devour somebody).

The controversies where immunologists confront their views on the mechanisms of memory, recognition and defense, the connotations of the words they use or the affects to which they refer, manifest what I have called the anthropological background of a perhaps too human science. The awareness of this background might help – this is the gamble of interdisciplinary exchange – to discriminate more aptly between the contingent and the essential in the arguments and the underlying issues.

No immunologist would accept the reduction of his work to queries on the identification of the Self and boundaries between Self and the other selves, or to the medieval question of micro/macrocosm. But no philosopher would accept to deny that metaphysical queries underlie the experimental work of biology. The dialogue between both, at the extreme of their minds, harbors not only uncontested, at least I hope, pleasure, but a promise of mutual enrichment if not decisive progress toward a common truth.

In this chapter, I have not claimed that immunologists should refrain from using metaphors and attributing volitions and concerns to the immune system. I have not tried to stigmatize their transient use of non-scientific categories, but, instead, have tried to show that, while doing science, they interfere with categories profoundly relevant to human life and experience.

The Self and NonSelf debate, whether it is or not reformulated as the danger model, turns on an issue of the utmost importance: the overlapping and elusive character of identities. The immune system can be considered as an illustration of the ambivalent and reciprocal status of the Same and the Other[65]. If, in some versions of the immune system, the Other is viewed as the germ or the pathogen, the ideal target for eradication, other versions offer a more flexible, changing and adaptive picture. In this perspective, the immune system illustrates the cognitive and political problems of exploring the differences without abolishing or reifying identities. We suggested that the immune system may represent the central issue of anthropology or the encounter with the Other[66] how to structure this encounter and elude the danger of absorbing or being absorbed?

Definitions of Self and selves, crucial to guide one's way into chaos, are another possible source of closure and schizophrenia, hostility and fighting. Interdisciplinary exchange should serve the purpose of challenging the boundaries in which the scientists could be tempted to enclose themselves.

The immune system illustrates the importance of metaphor in science, one of the foundations that makes interdisciplinarity possible and fruitful, politically correct and epistemologically legitimate, a useful basis for a meeting and hopefully many others such as the one that took place in Saint-Julien, in Claude Bernard's home.

---

65. Pouillon J. Malade et médecin: le même et/ou l'autre? (Remarques ethnologiques). Nouvelle Revue de Psychanalyse 1970; 1: 78-98. Segalen M, Ed. L'Autre et le Semblable. Regards sur l'ethnologie des sociétés contemporaines. Paris: Presses du CNRS; 1989.

66. Fabian J. Time and the Other. How anthropology makes its object. New York: Columbia University Press; 1983. Presence and representing the Other and anthropological inquiry. Crit Inquiry 1990; 16: 753-72.

THE HISTORIOGRAPHY OF IMMUNOLOGY

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Tales of neglected (orphaned?) historiographies

Alfred I. Tauber

Diversity of historiographies

History of science today reflects the same historiographic pluralism found in other historical disciplines. In this post-structural age, there are seemingly no organizing principles or modes of discourse that dominate in ordering the phenomena in question, or the narrative styles used to describe them. So it is no surprise that in a collection of studies grouped under the rubric of "historical issues" in immunology, we witness highly diverse approaches in the attempt to capture the pertinent issues and the lens by which we might view them. This diversity is quite typical. For instance, in November, 1996, I participated in a conference at The Dibner Institute (Cambridge, Massachusetts), titled "Pasteur, Germs, and the Bacteriological Laboratory." Although considerably more focused thematically, as in our own conference, there was a sorting process that reflected quite different orientations to the subject at hand. Let me just briefly sketch the approaches represented at the American meeting, for I believe, when compared to the papers collected herein, we can readily appreciate that the multifaceted perspectives is not a function of the *scope* of the topic, but rather the eclectic styles of historiography.

Consider, how, even in the well-circumscribed Dibner historical topic, the disparate issues described: the growth of the discipline of microbiology and the influence of thought styles formed the matrix for William Bynum, Patricia Gossel, Victoria Harden, Claire Salomon-Bayet, and Michael Worboys; the socioeconomic context was examined by Wolfgang Eckart, Christoph Gradmann and Gerald Geison; the methodological strategies that reflect cognitive models or standards of objectivity was assessed by Thomas Schlich; the role of technology and its uses served as a subordinated theme for Geison, Harden, and Gossel; the problem of defining ontological entities was deciphered by Bynum, Schlich, and Worboys; assessments of the origins of this field for our own time served as a critical nexus of discussion for John Farley, Geison, and Anne-Marie Moulin; and when the history of ideas was utilized by Bernardino Fantini, Andrew Mendelsohn, Jan Sapp, and Arthur Silverstein, we seemingly had exhausted the current exemplars of historiographic modalities.

There were, however, other matrices of discussion, which, while remaining implicit, functioned as the backgrounds for each of these papers. Already mentioned is the interplay of technology with scientific methodology and the modes by which this factor directs and even drives the science both in theory construction and practice. The second is perhaps more subtle, and concerns how medical and industrial interests operate to focus the impetus of microbiological research. One might easily forgive omission of this particular

question considering the well-discussed contributions of Bruno Latour [1] in this regard, but interestingly, there were at least two other kinds of history not represented at the Dibner meeting and our own, an omission, I believe, which is deliberate. My paper is an attempt to consider these "orphaned" histories in order to reflect on those historiographic strategies that we currently embrace with more enthusiasm.

The demise of the Great Man

The first of the neglected historiographies today is history oriented around the "Great Man." In our meeting, Arthur Silverstein and Paul Weindling offered readings of Ehrlich's career that pointed to his central role in early immunology, but this kind of historical narrative that places the leader's throne in context is generally neglected for other pursuits. I believe the current status of this general perspective is better illustrated by Gerald Geison's *The Private Science of Louis Pasteur* [2], a vivid testament of "dethroning." This work reflects a direct challenge to Great Man history, within which in Geison's honest attempt to present a balanced account of Pasteur's success, the critical gaze appears to have polarized those stalwart champions of the unique virtuoso from chroniclers who would seek a more nuanced view (see e.g., Max Perutz's attack [3] and the ensuing debate engaged by Geison [4] and Perutz [5], and then by Summers [6] and Perutz [7]). In the field of immunology there are several "heroes" we might consider in some detail – Pasteur, Metchnikoff, Ehrlich, Bordet, Koch, Landsteiner, Burnet, Jerne – but each, even while recognizing their extraordinary geniuses, are at the same time placed under the lens of a "normalizing" refraction, where the creative contribution is either not specifically addressed or when considered, strenuous efforts are made to show the human limits of such heroes. I admittedly generalize, but by and large, in the effort to remain "detached" and "objective," the biographer who embraces his subject too enthusiastically, is prone to be accused of "hero worship." Thomas Soderqvist's biography of Jerne is an exemplar of this cautious attitude. Soderqvist, while fascinated, and perhaps even awed by Jerne, scrupulously records his obsessions, excesses, and psychological weaknesses in the effort to offer a full account of an existential personality. But there is, indeed, no neutrality, and we are left with a general *Gestalt* that can hardly support the ethos of a hero.

In short, the implications of the leadership roles are minimized and to the extent the respective accomplishments of Great Men are acknowledged, we are warned to balance their merit with a judicious, if not jaundiced view of their foibles. So when Koch's fallibilities were demonstrated by Eckart and Gradmann in the Dibner meeting, we gleaned a multidimensional portrait of Koch's career and better appreciate the conflicts and disappointments that were at play. From another perspective, one I would label as the "voice of silence," Harden and Gossel described the adoption of bacteriological methods in America as falling to the bureaucracies concerned with public health, but hardly advanced by a leading scientific personality. America produced no scientist comparable to Pasteur during this period, and thus George Sternberg, arguably the greatest American microbiologist of the period, can hardly be compared to the eminent European leaders he emulated [8]. To what extent the missing supportive political and economic culture versus the absence of the scientific commander accounts for the slow start of American bacteriology is left moot, when in fact this leadership gap remains a significant factor in the growth of the discipline.

Silences concerning psychology and the operative factors of becoming a successful scientist may simply reflect an aversion to dipping into a problem too complex for a short paper; but on the other hand, when we witness any discussion in this regard we see in both the lacunae left by Harden and Gossel, as well as in the critiques offered by Eckart and Gradmann, a tendency to move away from that older genre of a history of adoration. After all, we would never expect to read an assessment like William Osler's deification of Pasteur: "He was the most perfect man who has ever entered the kingdom of science" [9, p. xvi]. In short, a biography of Pasteur such as that of Rene Dubos [10, 11] is simply out of fashion today and we hear no echoes of its underlying ethos. I would not argue for a return to an earlier "innocence," even were that possible, but the fate and lessons of Great Man history has an interesting parallel in the second missing historiographic approach, namely that of exploring science's metaphysical context, indeed, how the epistemology of science emerges from its underlying metaphysical foundations. This topic is difficult to tackle for various reasons, but most obviously the metaphysical/epistemological boundary is ever-shifting, and thus this reciprocal relationship is oftentimes difficult to tease apart. Historians of science might easily dismiss the issue as one for philosophers or historians of philosophy to pursue, but I maintain that this topic, elusive by contemporary sociologically-influenced practice-oriented histories (and thus "dangerously" philosophical), is forfeited with a high cost to the history of science. This thesis is pursued here in greater detail.

Factoring in metaphysics

Charles Rosenberg refers to the creative dialectic of diverse historiographies as a split between "macro" and "micro" orientations, and when one attitude dominates over another, as for instance in the general case here of "micro" narratives, there is a tendency to lose track of the laboratory's orientation in the world. The rebuttal takes several forms, but my agenda is to reinforce the claim for the dialectic of micro and macro, of part and whole, of individual and society, with more attention placed on the "macro" issues than is currently fashionable. I refer here to an older generation of philosophical historians who sought a more complex intellectual intercontextualization for explaining the deeper conceptual foundations of science. The key issues were to what extent "science in effect creates philosophy" [12, p. 3]; how metaphysics might "lead" philosophy; and the basis by which a synthesis of these perspectives might arise.

My proclivities rely on the identification and synthesis of the underlying philosophies, or more specifically, metaphysics of the scientists in the consideration of how these often undeclared cognitive predilections play formative and directive roles in practice. Prominent examples of this genre include Alfred North Whitehead's *Science and the Modern World* [13], Edwin A. Burtt's *The Metaphysical Foundations of Modern Science* [14], and Alexandre Koyré's *From the Closed World to the Infinite Universe* [15]. Putting aside the issue of whether science has progressed rationally, in these studies an endeavor was made to place particular scientific discoveries and theory construction within the dominant philosophical *Weltanschauung* of the period. Convinced that the scientist was in a sense shackled to his world view, these metaphysicians drew vivid philosophical canvases that portrayed science as part of those forces that framed philosophy, and concomitantly became another expression of a philosophical orientation. Science as an intellectual enter-

prise was thus situated and even constrained by the metaphysical assumptions of its practitioners. In sociological terms, this orientation was adopted to assess a particular research community in terms of the construction of knowledge by Ludwig Fleck, who referred to a "thought style" [16], to reflect the social bases and manifestations of this infrastructure. Fleck's originality was to expand the "metaphysical" (in the sense of the above) with other social contextualization factors. Thomas Kuhn acknowledged and built on this foundation, radicalizing these two perspectives; from *The Structure of Scientific Revolutions* [17] we may easily trace the genealogy of current science studies in their commitment to consider the diverse social and intellectual factors at play in constructing scientific knowledge.

Although Kuhn was still concerned with metaphysical questions, the eclipse of "metaphysical histories" was in large measure due to this new emphasis on sociological factors (i.e., laboratory practice, cultural influences, political and economic factors). However, we should bear in mind that although generally out of current favor, such metaphysical histories indeed represent the earlier effort to contextualize science in its intellectual culture and became instrumental in preparing for our contemporary pluralism. In many respects, the search for hidden metaphysical influences challenged the dominant rational philosophy of history, whereby such practitioners as George Sarton saw science as advancing in a retrospectively appreciated order. On that view, progress was documented by the reasoned assessment of one theory supplanting another by the sheer force of its superior intellectual content. Interestingly, and not coincidentally, such a normative historiographic approach was influenced and strengthened by logical positivism, which sought rational and universal criteria for truth and used science to illustrate and warrant that exercise. Granted that Rudolf Carnap, Philipp Frank, and Hans Reichenbach differentiated the "rationality of explanation" from the "logic of discovery," science was anointed as the exemplar for offering a method to substantiate truth claims, and its rationality was to be universalized. The positivist's philosophical examination became, in some sense, myopic to its narrow concerns, and the larger contextualization issues raised by the "metaphysicians" were counted among the casualties of this approach. The tides, of course, shifted, and the constraints of "internal" examination have largely yielded to "externalized histories."

Despite the current dominance of sociologically-oriented contextualized histories, current historiographic strategies have generally skirted the grand metaphysical assessments made by earlier generations of historians and philosophers. In the search for identifying the microprocesses and effects empowering science, we have by and large opted for a more detailed, and some would argue, careful consideration of those sociological factors that might be described without resorting to the highly abstract and finally elusive categories of the metaphysician. Certainly a commitment to a positivism is at play here, but we may also be witnessing manifestations of an anti-metaphysical sentiment, where in the very discipline it might be supported, namely philosophy, there is strong movement pledged to purge any vestiges of such contamination! Nevertheless, the deeper philosophical assumptions of science are critical factors in theory construction and adoption, for in an important sense we live with a dialectic, where empirical evidence and theory interplay with metaphysics to create the very fabric of our world views. In eschewing the deeper philosophies of scientists, we omit a crucial component of their endeavor. I will first illus-

trate this thesis with an example drawn from my own research that is directly relevant to the topic at hand.

A case study: Metchnikoff

The celebrated polemics between Elie Metchnikoff and the German school of microbiologists guided by Robert Koch, and later the immunochemists led by Koch's protégé, Paul Ehrlich, has been subject to much scrutiny, and elsewhere in this volume Eileen Crist and I explore this question again. As we discuss, there were, to be sure, rather obvious divergent theoretical and methodological perspectives embraced by the protagonists; and while due attention must be paid to those issues, I have long argued that there were even more fundamental differences in the deeper philosophical understanding of biology that ultimately would allow no resolution [18-20]. Metchnikoff was an embryologist, who sought to discover genealogical relationships in the context of Darwinian problematics [21]. He was intrigued with the problem of how divergent cell lineages were integrated into a coherent, functioning organism. He was thus preoccupied with the problems of development as process, which he regarded as Darwinian: cell lineages were inherently in conflict to establish their own hegemony, and he thus hypothesized that a police system was required to impose order, or what he called "harmony," on the disharmonious elements of the animal. He found such a system in the phagocyte (a mobile amoeboid cell), which retained its ancient phylogenetic eating function: devouring effete, dead, or injured cells that violated the phagocyte's "sense of identity." In other words, the phagocyte became the sentinel of the organism's individuality and, according to this theory, determined what was to be preserved as "self" and destroyed as foreign.

Thus the phagocyte was initially viewed as a purveyor of identity, and when Metchnikoff became engaged in the nascent field of infectious diseases at the beginning of the 1880s, he was poised to apply his phagocytes to the duty of protecting the organism from pathogens (i.e., maintaining integrity [22, p. 20; p. 62-3]. It was a grand scheme, which he presented in a series of public lectures in Paris in the spring of 1891, later published as *Lectures on the Comparative Pathology of Inflammation* [23]. There Metchnikoff argued that the phagocyte had preserved its most ancient phylogenetic function: in simple organisms such cells functioned as the nutritive organ (eating resident microbes), and in animals with a gut, phagocytes continued to eat, but now for defense. In Metchnikoff's theory, therefore, immunity was a particular case of what he called "physiological inflammation," a normal process of animal economy.

But there was a more subtle message: 1) immunity was an active process with the phagocyte's response seemingly mounted with a sense of independent arbitration (viz. agency); and 2) organismal identity was a problem bequeathed from a Darwinian perspective that placed all life in an evolutionary context, which Metchnikoff extended to the physiology of the individual animal. The agency quality of his argument and the radical sense of self-definition reflected major Nietzschean themes, a parallel I have attempted to make explicit [22, 24, 25]. On this reading, Metchnikoff possessed a particular understanding of biology that reflected a universe in disharmony, where struggle and self-aggrandizement were the very substance of life. Existentially, Metchnikoff was seemingly preoccupied with self-realization (or self-actualization), a Darwinian mode characteristic of his time; religiously, he placed his abundant spirituality in the service of science's ration-

ality; psychologically, he was highly volatile, swinging between optimism and pessimism; and philosophically, he was committed to a unique blend of teleology and vitalism whose roots in earlier nineteenth century descriptive biology brought him into dissonant controversy with the dominating reductionism of the period (discussed more fully in the paper with Crist).

Metchnikoff was brushed aside by his German detractors as a hopeless Romantic, with outdated teleological precepts, and his phagocytes were caricatured as possessing volition and intention, and thus vitalist independence. Falling beyond the boundaries of the reductionist ethos of the time, this dismissal was in effect a rejection of his entire underlying Darwinian philosophy. To be sure, Metchnikoff's polemics with the Germans was complicated by both political and personal issues [26], but the conceptual differences dominate: on one level the conflict is expressed as a stance against the strong reductionist program of contemporary immunochemists, but at a more fundamental level, we observe different conceptions of the organism. I would not make that claim solely on the basis of examining the Nobel archives, weighing public testimony, nor by contrasting his specific scientific views with other scientists who were involved in similar research [8, 22]. Rather, the case ultimately rests upon *interpreting* how his scientific posture employed emergent and dynamical thinking appropriate to an organismic orientation of a biologist keenly aware of the problem of identity in a post-Darwinian age [19, 20].

On that reading, Metchnikoff deeply comprehended the Darwinian revolution. He maintained that throughout the life experience of the organism there are changing environments, new insults, encounters with novel challenges, and it is the immune adaptability and versatility that determines overall success. From this point of view, the primary lesson of evolutionary biology is a radically different conception of the organism from that of the pre-Darwinian era. The contrast is between a view of the organism as "given," essentially unchanging and stable in distinction to the dynamic image implicit in Metchnikoff's formulation, where the organism is in a dialectical relationship with its world. In an ever-changing set of relationships, at many different levels of engagement, the organism lives in response to its environment; and it was this component of 'active' reaction that was so revolutionary and drew the ire of his opponents. They saw it as a vitalistic or teleological formulation, because they did not either understand or share the metaphysical foundations upon which Metchnikoff based his theory.

So in the public debates surrounding what was called the "phagocytosis theory," Metchnikoff focused upon the reaction of the host organism invaded by pathogens, and in that struggle he postulated an active response to infection. The philosophical issues framing his position never explicitly surfaced, which in itself is a curious fact considering Metchnikoff's deep intellectualism. But the playing ground was defined by the Germans, and so the "cellular-humoral" polemics of the 1890s were argued on the basis of how to interpret particular experimental findings. Metchnikoff's opponents would not engage him on the deeper conceptual issues, perhaps because their own theoretical formulations were so ill-formed. In any case, the proposal, even narrowly construed as a model of host defense, was radical in itself: the phagocytosis theory was the first that posited active immune mechanisms as opposed to passive ones to deal with infection.

Pasteur proposed his own passive notion of immunity as an extrapolation of the test tube scenario shortly before Metchnikoff published his immunity hypothesis. On Pasteur's

view, bacteria simply exhausted critical nutrients in the host during initial infection which then deprived the micro-organisms of necessary food. Either the pathogens starved during their initial foray or failed to grasp a landing on second exposure because of the deprived nutritional state of their host. By 1885, when Salmon and Smith demonstrated immunity with dead bacteria, the passive theories were laid to rest. Although there were profound effects of Metchnikoff's theory detected in the development of later immunology [22], and perhaps for medicine in general [27, 28], from our vantage point, Metchnikoff's dogged defense of the phagocytosis theory was soon overwhelmed by the immunochemical program of his opponents, and for our purposes we may leave him here.

Another case: Pasteur

To recognize Metchnikoff's own holistic view of the organism, the integrated and comprehensive approach to his biology as theory-driven, and the neo-Romanticism of his extrapolated biological thought to humankind and himself, is to place him within certain broader intellectual currents of his time. And more particularly, I have maintained that the appreciation of an active host response originated with Metchnikoff, because during the nascent years in which infectious diseases emerged as a scientific discipline, he, alone, appreciated a fecund implication of the Darwinian perspective, albeit derived from a particular metaphysical orientation. In short, I have argued that to understand Metchnikoff's science one must take account of his broader philosophical posture. Let us now consider Pasteur as a second case where similar concern to comprehend his underlying philosophy becomes an important factor in assessing the scientist. I do so with some hesitation only because I have not directly examined primary sources, but instead must rely on the narratives of other historians who share my propensity to consider such factors as operative. Nevertheless, the illustration is so compelling I must briefly summarize it.

My initial interest in Pasteur arose from the friendship that he and Metchnikoff shared, whose basis is not obvious at first glance considering their differing scientific interests and backgrounds. But based on the warm welcome the refugee Metchnikoff received by Pasteur in 1888, who offered the Russian a laboratory at the new Pasteur Institute (and a salary of one franc a year!), there seemed to be some fundamental affinities in their philosophical outlooks. After all, according to Metchnikoff, when he and Pasteur first met, Pasteur proclaimed, "I at once placed myself at your side, for I have for many years been struck by the struggle between the diverse microorganisms which I have occasion to observe. I believe you are on the right road" [26, p. 22].

Pasteur indeed embraced Metchnikoff, and the two scientists formed a close alliance. As Metchnikoff wrote, "Pasteur, like every human being had his weaknesses, but even without speaking of the enormous contributions which he heaped upon mankind, he combined with his scientific genius a vibrant soul, a profound kindness, and a character of exceptional stature" [26, p.25].

Part of their relationship was based upon a nationalistic alliance of the Russian and Frenchman against their German competitors. But their kinship had deeper roots than the geopolitical conflicts of the time.

At one obvious level they were intellectually compatible, inasmuch as they each sought the broadest biological principles to guide the interpretation of their research, and their respective theories were not in conflict. Although their particular visions of biology

differed, each revolved about a broad vision of biological unity. Their respective scientific inquiries were formulated in the endeavors to define the integrity of the organism for Metchnikoff and the essential biochemical unity of life for Pasteur. In this sense, each sought theories to govern the functional harmonization of diverse behavior. But there is a deeper and perhaps more elusive affinity between Pasteur and Metchnikoff, which pertains to the non-positivistic elements of their respective philosophies. Neither Metchnikoff nor Pasteur were comfortable with positivism; and I maintain that their scientific philosophies were intimately connected to this discontent, specifically their deep humility before the altar of life that resisted a fully positivist program. In these brief remarks I wish to explore this factor in more detail.

Consider the following quotation from Pasteur's 1882 inauguration lecture to the Académie Française. In this excerpt, Pasteur is dealing with positivism as a method applied to history, sociology, and literature, basically all the humane sciences. He assumes a most critical stance of knowledge in that arena based on such methods: Auguste Comte's fundamental principle is to eliminate all metaphysical questions concerning the first and final causes, to attempt to account for all ideas and theories in terms of concrete facts, and to consider as valid and established only that which has been shown by experience... "'[E]xperience' has a meaning very different from that of the word 'experiment' in scientific language. In the former case, experience is merely the simple observation of things with the induction which concludes, more or less legitimately, from what has been to what could be. In contrast, the true experimental method aims at reaching a level of proof immune to any objection" [Quoted by Dubos, 10, p. 387-8].

Pasteur was fully confident of his scientific method applied to his own sphere of investigation, where scientific facts would emerge by vigorous experimentation. But he was highly doubtful of any social science that attempted to ape the methods of natural science to answer queries of a different kind, specifically those not amenable to experimentation. But he goes further: "The conditions and the daily results of the scientist's work lead his mind to identify the idea of progress with that of invention. In order to evaluate positivist philosophy, therefore, my first thought was to search it for the evidence of invention, and I did not find it...Positivism, offering me no new idea, leaves me reserved and suspicious" [10, p. 388-9].

Positivism was too restricted a mode of inquiry! Scientific advance in Pasteur's view required a creative component, for its success transcended the mere observation and description of a naked positivism. At the heart of the Pasteur studies by Bruno Latour [1] and Gerald Geison [2] is the appreciation of how Pasteur at times abandoned strict adherence to Scientific Method, even employing shrewd rhetorical skills to make his case against opponents. Pasteur's scientific accomplishments could not have been made without rigorous techniques, executed with "irreproachable exactitude." But it is not accomplished by the routine application of some mechanical scientific method. It is more than that. It is a gift, a talent, a skill, an art – and Pasteur was most decidedly an artist of the invisible world [2, p. 133].

I will not dwell on his interpretation of positivism and its application, but rather let us see how this critique serves as revealing the limits that Pasteur himself recognized for his science.

I am referring to Pasteur's religious convictions. He maintained that the positivists could not account for the deep sources of inspiration, the appreciation of the infinite, which he believed served as the very basis of man's humanity. This was translated into his scientific philosophy quite explicitly and should not be regarded as either an inconsistency in his role as scientist or a divorce of faith and reason. For Pasteur, there was a fundamental mystery at the heart of the organic, and I believe this sheds important light on his scientific persona.

At the time of the controversy on spontaneous generation, Pasteur quite explicitly described his opinion regarding the ability of experimental science to decipher the riddle of life. Concurring with Claude Bernard, he regarded the mystery of life as residing not in the manifestations of vital processes, which are amenable to scrutiny by ordinary physico-chemical assessments and characterizations, but in the predetermined specific characters of the organisms which are transmitted through the ovum, through the so-called germ. In what was to become an often repeated refrain, Pasteur wrote

> Life is the germ with its becoming, and the germ is life... Once the germ exists, it needs only inanimate substances and proper conditions of temperature to obey the laws of its development...it will then grow and manifest all the phenomena that we call 'vital,' but these are only physical and chemical phenomena.....
>
> This is why the problem of spontaneous generation is all-absorbing and all-important. It is the very problem of life and of its origin. To bring about spontaneous generation would be to create a germ. It would be creating life; it would be to solve the problem of its origin. It would mean to go from matter to life through conditions of environment and of matter.
>
> God as author of life would then no longer be needed. Matter would replace Him" [Quoted by Dubos, 10, p. 395-6].

Pasteur acknowledged that his work had not proved that spontaneous generation was impossible, only the fallacy of all known claims to that effect. He recognized that he could not prove the negative. But by the same token he protested that spontaneous generation had been the origin of life in the universe. In unpublished remarks written in 1878, Pasteur asks a remarkable question: "I have been looking for spontaneous generation during 20 years without discovering it. No, I do not judge it impossible. But what allows you to make it the origin of life? You place matter before life, and you decide that matter has existed for all eternity. How do you know that the incessant progress of science will not compel scientists...to consider that life has existed during eternity and not matter? You pass from matter to life because your intelligence of today...cannot conceive things otherwise. How do you know that in 10,000 years one will not consider it more likely that matter emerged from life...?" [10, p. 396-7].

Geison has admirably drawn the contours of Pasteur's political and philosophical orientation in regards to spontaneous generation [2, p. 121ff.], and the richness of his study in large measure results from the sensitivity shown to these "macro" issues. On one plane, considering Pasteur's political and orthodox Catholic religious associations, he would naturally align himself against the positivists and materialists who were associated with the opposition to Church and State. In this regard, the spontaneous generation debate was a touchstone that defined underlying attitudes towards science, philosophy and religious beliefs. Pasteur, at least in this respect, was committed by his general political and reli-

gious convictions, and Geison suggests that Pasteur's highly selective use of experiments was thus prejudiced scientific intuition. Quoting Pasteur: "I did not publish these experiments for the consequences it was necessary to draw from them were too grave for me not to suspect some hidden cause of error in spite of the care I had taken to make them irreproachable" [Quoted by Geison, 2, p. 130].

The general practice of selectively using experimental results to bolster "intuition" or "theory" is well-appreciated, and Pasteur is no exception. But rather than decry the fall of an unblemished scientific method, we might more profitably seek to understand how Pasteur's philosophy of biology was built on a complex array of factors that led him to the orientation he adopted. Beyond his religious belief system, there was a strong scientific component built from a speculative understanding of the world. From his earliest chemical research in the optical polarity of compounds, Pasteur saw distinctions that he surmised were characteristic of the organic [2, p. 135 ff.]. At the very core of his chemical presuppositions regarding the nature of symmetric (inorganic) and asymmetric (vital) compounds, Pasteur brought a strict duality that in a fundamental sense framed his view of the world. This is a metaphysical posture that both directed his interpretation of his science, but also was reinforced by the results of his research. In my view, this is not a circular reinforcement, but rather a dialectical process.

One final note. Pasteur's philosophy of science distinguished quite clearly the limits of thought and experimentation. "Cause" was reserved for the primary divine impulse which gave birth to the universe and life. On his view, we can observe nothing but correlations, recognizing that cause and effect relationships are inferred when one phenomenon always follows another in time; we discern them as relational, ie, we believe they must occur in a certain order. But this is not Cause. As Dubos observed, Pasteur, like Claude Bernard, was satisfied with relation or conditions in lieu of cause, which was reserved for First, or divine, Cause [10, p. 398]. At a profound spiritual level, Pasteur invoked determinism as a vehicle for our technical mastery of nature, but not as a reference for a fundamental account of Life. I would go as far as to suggest that on this view, Pasteur recognized a profound mystery that framed and motivated his work: biology was not the science of life, but rather the means by which we comprehend vital processes in physico-chemical terms.

But I would not close with a neat demarcation of science and religion in such a complex personality. On the one hand, Pasteur viewed the spontaneous generation issue as a religious question, whose resolution re-affirmed his own beliefs in the divine. At the same time, there was apparently a bit of Dr. Faust in Pasteur, since throughout his life he contemplated how to capture that elusive divine principle in the laboratory [2, pp. 138-142]. It seems that at the same time as recognizing the limits of deciphering cause and applying method in his research, he still dreamt of apprehending "a new and fundamental force in nature – a 'cosmic asymmetric force' that was ultimately responsible for life itself" [2, p. 139].

Conclusion

My two protagonists lived with metaphysical tensions. Metchnikoff's immunity theory was plagued with accusations of teleology and vitalism, because the phagocyte was regarded as somehow possessing intention, which had no scientific basis. But he, like Pasteur, believed that there were fundamental aspects of biology that could not be

addressed by reductionist analysis. For Metchnikoff it was the establishment of organismal identity, and for Pasteur, the origin of life. In both cases, the most basic nature of life was a black hole, one indefinable by our rational tools, scientific or otherwise. As heretics of the ascendant reductionism, where chemistry, physics, and soon genetics would reign, they assumed a stalwart circumspection in regards to science's explanatory power. This characterization is largely dwarfed, however, by their respective roles as champions of science. After all, Pasteur was an early, key architect of this new molecular biology, and unlike Metchnikoff, was firmly committed to the application of a reductionist strategy to analyze vital processes, albeit with certain implicit restrictions. And Metchnikoff, despite reliance on an older descriptive biology, was nevertheless a vigorous prophet of science, the creed he believed that would save humankind from its own folly and biological limitations [18, p. 23]. In short, if one is to have a full appreciation of these figures, and perhaps more importantly, their philosophies of nature and science, divergent aspects of their intellectual personae must be factored in and accounted for.

Dazzled by the awe of life, both Metchnikoff and Pasteur were diligent scientific examiners who always seemed aware of the limits of their craft. This affinity between Pasteur and Metchnikoff formed a critical bond that bypassed generational differences, scientific goals, and sociological categories. It allied them against a common foe in the guise of Robert Koch and his school, who were blind to their respective larger philosophical-religious postures. There is a vast literature on the growth of materialism and reductionism in nineteenth century biology, and I make no attempt to go beyond the implicit contrasts with that movement exhibited by my two protagonists. Metchnikoff's staunch defense of the agency of the phagocyte and Pasteur's experiments on spontaneous generation were each driven by deeper metaphysical presuppositions and the desire to defend them. These beliefs placed them in a fundamentally different orientation with respect to their opponents. We need not be shy to consider including these factors in our assessments, for to ignore them is to forfeit a full characterization of the practicing scientist. In short, to entertain the import of metaphysical concerns is to portray formative, and even directive influences in science as exercised by very human agents. Unlike an older generation of "metaphysical historians," who might have sought a dominance in their philosophical inquiries, I would be content with simply adding such discussions to the rich pluralism adopted by other historiographic strategies, recognizing that while one may contextualize metaphysical assumptions, such ideas are worthy of study in their own right inasmuch as they may have independent effects of their own.

Acknowledgements

I wish to thank Robert S. Cohen, Eileen Crist, and Gerald Geison for their generous counsel, whose informed criticism improved earlier versions of this manuscript.

Bibliography

1 Latour B. The Pasteurization of France. Sheridan A, Law J, transl. Cambridge, MA: Harvard University Press; 1988.
2 Geison GL. The private science of Louis Pasteur. Princeton: Princeton University Press; 1995.
3 Perutz MF. The pioneer defended. [Book review of Geison GL. The private science of Louis Pasteur]. New York Review of Books 21 December 1995; 42: 54-8.
4 Geisen GL. Pasteur and the culture wars: an exchange. New York Review of Books 4 April 1996; 43: 68-9.

5 Perutz MF. Pasteur and the culture wars: an exchange. New York Review of Books 4 April 1996; 43: 69.
6 Summers WC. Pasteur's private science. New York Review of Books 6 February 1997; 44: 41.
7 Perutz MF. Pasteur's private science; reply. New York Review of Books 6 February 1997; 44: 41-2.
8 Tauber AI. The birth of immunology: III. The fate of the phagocytosis theory. Cell Immunol 1992; 139: 505-30.
9 Osler W. Introduction. In: Vallery-Radot R, Ed. The life of Pasteur. Devonshire RL, transl. Garden City, NY: Garden City Publishing Co.; 1919.
10 Dubos R. Louis Pasteur, free lance of science. Boston: Little, Brown, and Co.; 1950.
11 Dubos R. Pasteur and modern science. Garden City, NY: Anchor Books; 1960.
12 Bachelard G. The new scientific spirit. Boston: Beacon Press; 1984.
13 Whitehead AN. Science and the modern world. New York: Macmillan; 1925.
14 Burtt EA. The metaphysical foundations of modern science. New York: The Humanities Press; 1952.
15 Koyré A. From the closed world to the infinite universe. Baltimore and London: Johns Hopkins University Press; 1957.
16 Fleck L. Genesis and development of a scientific fact. Chicago: The University of Chicago Press; 1979.
17 Kuhn TS. The structure of scientific revolutions. Chicago: The University of Chicago Press; 1962.
18 Tauber AI, Chernyak L. Metchnikoff and the origins of immunology. New York and Oxford: Oxford University Press; 1991.
19 Tauber AI. The immunological self: a centenary perspective. Perspect Biol Med 1991; 35: 74-86.
20 Tauber AI. Introduction: speculations concerning the origins of self. In: Tauber AI, ed. Organism and the origins of self. Dordrecht: Kluwer Academic Publishers; 1991. p. 1-39.
21 Gourko H, Williamson DI, Tauber AI. The evolutionary biology papers of Elie Metchnikoff. Dordrecht: Kluwer Academic Publishers; 2000.
22 Tauber AI. The immune self: theory or metaphor? New York and Cambridge: Cambridge University Press; 1994.
23 Metchnikoff E. Lectures on the comparative pathology of inflammation. Starling FA, Starling EH, transl. New York: Dover Publications; 1968.
24 Tauber AI. The organismal self: it's philosophical context. In: Rouner L, Ed. Selves, people, and persons. Vol. 13. Boston University Studies in Philosophy and Religion. Notre Dame: Notre Dame University Press; 1992. p. 149-167.
25 Tauber AI. A typology of Nietzsche's biology. Biol Philos 1994; 9: 25-44.
26 Tauber AI. A case of defense: Metchnikoff at the Pasteur Institute. In: Cazenave PA, Talwar GP, Eds. L'immunologie: l'héritage de Pasteur. New Delhi: Wiley Eastern Limited; 1991. p. 21-36.
27 Tauber AI, Chernyak L. Metchnikoff and a theory of medicine. J Roy Soc Med 1989; 82: 699-701.
28 Tauber AI. Darwinian aftershocks: repercussions in late twentieth century medicine. J Roy Soc Med 1994; 87: 27-31.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Georges Canguilhem's "On the normal and the pathological": a restatement and a commentary

Peter Keating

Introduction

Georges Canguilhem is one of the best-known exponents of historical epistemology. His most important book, "On the Normal and the Pathological" [1], has been translated into a number of languages and the impact of this book on historians of medicine has been similar to the impact of Thomas Kuhn's "The Structure of Scientific Revolutions" on historians of science [2]. Canguilhem is also well known in the other social sciences. Indeed, according to Mike Gane, "Canguilhem's work is perhaps best known in the Anglophone world in connection with the debate on epistemology and the social sciences introduced by Althusser and Foucault in the 1960s" [3, p. 312]. Canguilhem's most celebrated student, Michel Foucault, offered a similar analysis 20 years earlier in the introduction to the first English edition of "On the Normal and the Pathological" when he wrote, "[t]ake away Canguilhem and you will no longer understand much about Althusser, Althusserism and a whole series of discussions which have taken place among French Marxists (...)" [4, p. ix]. While interest in Althusser has since subsided considerably, there are still good reasons to read Canguilhem today. In this paper, I would like to restate the central thesis of "On the Normal and the Pathological" and offer a further illustration of its continuing relevance to historians of medicine by giving several historical examples taken from bacteriology and immunology. I then raise two questions: 1) Is Canguilhem's distinction between the pathological and the *ab*normal still relevant? 2) Is pathology still independent of physiology? To understand the pertinence of these questions, let us turn to Canguilhem's thesis.

The reduction of the pathological to the normal

According to Canguilhem, there is a recurrent theme or, more precisely, ideology of medicine and, in particular, of the relationship of physiology or biology to pathology that dates back to the beginning of the nineteenth century. While, as we shall see, the theme takes various forms, it may be stated simply as the thesis that there is no difference between the normal and the pathological. This belief can be understood as an expression of the triumph of eighteenth century rationalism over other forms of magical and religious causation. For physicians and researchers from the nineteenth century on, to admit that there was something special about pathology, that it was something other than a particular application of

the laws of physiology, was to admit that there was something special about disease. To say that opened an unwanted door to the past where disease represented, for example, punishment for sin.

Canguilhem criticizes three kinds of reduction of the pathological to the normal. Let us begin with the first, the idea that the pathological is a quantitative extension of the normal. Stated somewhat differently, the thesis claims that there is a quantitative continuum running between the two and that there is, therefore, a qualitative identity. This thesis finds clear expression in the celebrated work of Claude Bernard. Indeed, according to Canguilhem, one of Bernard's major preoccupations was the demonstration of such a thesis.

The best example of this articulation of the normal and the pathological in Bernard is diabetes. According to Bernard, each of the cardinal symptoms of diabetes can be understood as merely a quantitative extension of a normal physiological process. Thus, he claimed that the intense thirst, frequent urination, wasting and sugar in the urine could all be observed in a normal human being. While thirst, urination, and weight loss are common human experiences, Bernard had difficulty showing that the last symptom, glycosurea, was present in normal subjects. Failing to make the required demonstration, Bernard simply asserted that there must be something wrong with the equipment (i.e., that it was not sensitive enough) used to measure glucose in urine [1, pp. 34-5]. Canguilhem's critique in this case is quite clear. Diabetes is not simply too much sugar in the urine; there is no physiological equivalent to glycosurea.

The second form of reduction criticized by Canguilhem consists of the reduction of the disease to part of its mechanism. To use the same example, Bernard took the mechanism of glycosurea (which, to repeat, he considered a normal phenomenon) to be the mechanism of diabetes. According to Bernard, the liver had simply put too much sugar in the blood and the kidney was flooded; the consequent overflow leaked into the urine. If one were to ask Bernard the cause of the increase in glycogenesis, he would simply defer to the nervous system. Canguilhem's criticism in this case centers on what Bernard could not have known, namely that the sugar in the blood is not a mere quantitative increase in glycogenesis but the result of an absolute (qualitative) lack of insulin. Notice here that the argument against Bernard could also have been stated as a classical confusion between symptom and disease.

The third form of reduction is one that Canguilhem calls "same cause but different effects." Two examples are offered to explicate this form of reasoning and while the second is no longer convincing, it seems to me that the first could never have been so. Let us consider the first. According to Canguilhem, the movement from arterial tension to hypertension is generally seen as a quantitative move in the sense that, as the prefix "hyper" implies, the same underlying mechanisms are believed to be at work in the production of both forms of tension. There is no fixed threshold, in other words, beyond which arterial pressure may be said to be pathological. This, at least, is the view of pathologists in the 1930s. Canguilhem points out, however, that there is indeed a qualitative difference in the effects of normal blood pressure and hypertension even though the same cause has only varied quantitatively [1, p. 55]. Hence, although the identity of the normal and the pathological can be maintained in terms of a common mechanism – blood pressure

–, it cannot in terms of the effects of the quantitative differences in the performance of the mechanism.

This is not a very strong argument and there is, I believe, a better one that could have just as easily been made at the time. Specifically, one could ask: Is it really, as Canguilhem suggests, the same cause that has varied quantitatively? As the etiological cause of hypertension, then as now, is simply unknown, one could more directly conclude that the mechanism of hypertension is unknown, or, at least, incomplete. In this sense, the third kind of reductionism is simply another example of the second.

Canguilhem offers a second example of the "same cause, different effects" argument by contrasting immunity and anaphylaxis. Presumably both forms of the immune response proceed from the same cause: penetration of a foreign substance into the body. The effects, of course, are completely different. Canguilhem tells us, "The presence of antibodies in the blood serum is therefore always normal, the organism having reacted to an aggression by its milieu by modifying its constants (...) but in one case the normality is physiological (immunity) and in the other case pathological (anaphylaxis)" [1, p. 138]. Once again, there are better reasons to agree with Canguilhem. This time, however, it is the newer knowledge of immunology that allows us to say that that we do not have one cause varying quantitatively, but two different mechanisms and, in fact, two different kinds of antibody in the blood.

Canguilhem raises another kind of objection against the identification of the normal and the pathological namely, that such an identification presupposes an objective definition of the normal [1, p. 26]. Without an objective definition of normal, there is no starting point for the quantitative extension of the normal to the pathological. Such numbers do not, according to Canguilhem, exist and the argument is as follows: It is only with regards to some external norm or standard that it is possible to speak of more or less. In physiology, the norm is often taken to be the mean of a series of observations. This leaves out, however, the fact that the number – the mean – is not itself the norm but the numerical expression of the norm. Moreover, the physiological process under investigation has a positive (qualitative) value prior to having a quantitative expression. The positive value derives from the fact that, assuming that there are not any useless (neutral) physiological processes, then all processes must contribute in some way to the maintenance of the life of the organism and must have been the result of natural selection. In other words, all physiological processes had at some point, and retain so far, some positive selective value. Similarly, negative physiological values, like the body temperature of a dinosaur, must already be extinct (and presumably cannot be measured anyway). We shall see later that diseases such as diabetes create mixed cases.

The idea here is fairly obvious to students of natural selection. If a trait or characteristic has a negative selective value, then it will, by definition, be eliminated in the course of evolution. This value, however, is only ever retrospective. It is not predictive and therefore can be neither judgmental nor normative. This Darwinian theme, however, can seemingly become a platform for the re-enchantment of nature as when Sinding tells us, for example, that, according to Canguilhem: "A living thing institutes a value judgment on itself in so far as the perpetuation of the species implies that in the course of evolution some individuals and some species have been valorized whereas others have been condemned to disappear" [5, p. 23]. I say seemingly because Sinding is, of course, well aware that, in

fact, no judgment has been made. Indeed, she goes on to say that "Now, no change, no anomaly, no mutation can be considered a priori negative or positive in reference to a pre-established type: the validity of the new organization can only be referred to its eventual success" [5, p. 23]. In other words, as previously stated, the "value judgment" is always retrospective. So, why, we may ask, is Sinding drawn into what appears to be a contradiction? Sinding sees a practical side (other than promoting physiology) to Bernard's attempt to identify the normal and the pathological and it is this: If the mechanisms of pathology can be reduced to those of physiology, then medical researchers will have considerably widened the sphere of preventive medicine. It seems to Sinding that we are in fact on this road to prevention. Whereas Bernard lacked the concept of gene and the technology permitting its manipulation, now that we have it, the reduction is possible. In particular, were physicians and researchers to genetically modify non-viable embryos, then artificial selection will have replaced natural selection. In this case Sinding's fictitious value judgments become possible. In the absence of artificial selection, however, and thus for the present, physiological norms are ultimately dependent on the environment and thus cannot be taken as stand-alone, external standards.

The primacy of the clinic

If the pathological is not reducible to the normal, then, it might be asked, what, besides "nature," gives pathology its autonomous status vis-à-vis physiology? According to Canguilhem, the history of medicine shows that, despite textbook descriptions, research has, more often than not, led from the pathological to the normal. It is usually the case that some clinical phenomena has drawn the attention of researchers to some organ or process. The discovery of the normal function of the islets of Langerhams, for example, resulted from research on diabetes. In general, as Canguilhem puts it: "Almost every time it is said in human physiology, 'We now know that...,' one will find, if one looks hard enough – and without reducing the role of experimentation – that the problem was raised and the solution outlined in the clinic and through therapeutic intervention, and often at the expense, biologically that is, of the patient" [1, p. 142]. This assumes, of course, that the primary aim of medicine is the cure of human beings. While this may be so for clinicians, it is not necessarily the immediate case for pathologists: understanding disease often precedes cure and the model for understanding is not necessarily human patients. In their recent textbook on pathology "dedicated to all patients in all times and places and to all those who helped us to understand the primal patient – the cell", Majno and Joris, for example, say that "We chose to focus on the cell as the elementary patient (...)" [6, p. xii]. Nonetheless, Canguilhem's claim is clear: pathology is not only autonomous vis-à-vis physiology, but, as far as clinical medicine is concerned, it is primary. Some commentators have interpreted the latter to mean that Canguilhem is asserting the "primacy of the patient." Stuart Spicker, for example, seems to believe that one of the consequences of Canguilhem's work should be to "stress the uniqueness of each patient (organism) and the implications of this thesis in producing and prescribing complex, yet individualized, drug regimens (...)" [7, p. 407].

While it is doubtful that Canguilhem himself would have gone that far, there is no doubt that he considers the patient to be the historical starting point of pathological investigations. Canguilhem develops this theme through a fictitious example. Suppose, he says, a

murder victim is discovered to have cancer that had not been diagnosed or even suspected by the victim who had remained symptom-free until his death. Did the murder victim have cancer? According to Canguilhem, the answer is no. The reason offered at first appears psychological, namely "there is nothing in science that has not first appeared in consciousness" [1, p. 53]. In other words, prior to a conscious complaint by the patient or some other form of physician/patient contact, there is no clinical phenomena.

It has been suggested that since individual complaints vary from culture to culture, Canguilhem ends up in a kind of relativism. Christopher Lawrence, for example, tells us, "What then defines an event as pathological in any particular context? Canguilhem's answer is in one sense both logical and yet at the same time richly unexpected. It is the sick man" [8, p. 96]. Similarly, François Dagognet begins a commentary on Canguilhem with an illustration of the cultural variability involved in the interpretation of scar tissue [9]. According to Dagognet, whether and to what extent a scar is unwanted depends on a number of factors, including where it is and whether or not it was the product of ritual scarification and therefore wanted (at least initially). Dagognet then proceeds to explain that, in a similar fashion, pathology is not objective because the definition of illness depends on the social and cultural context within which the symptoms emerge. Unsurprisingly, Dagognet takes both of his examples from the field of mental diseases, which leads to the obvious objection that harder cases abound. Death by smallpox, for example, would seem to transcend cultural relativism and stand out in any cultural context as definitely unwanted.

One might further object here that, in any event, many interesting, important and real diseases do not affect humans and that such plant and animal diseases do not require complaints by their sufferers in order to enter science [10]. This objection can, of course, be easily sidestepped simply by deferring the complaint, the consequent prior entry into consciousness, and the relativism entailed to the owner or steward of the plant or animal in question. Such a manoeuvre, however, is insufficient. The replacement of the doctor/patient relationship by a triad composed of, say, the veterinarian, the animal owner, and the animal does not account for cases where animals and plants are routinely diagnosed as ill and destroyed without any complaint having been made by the owner or steward. Moreover, ecologists recognize diseases as real and important sources of stability and diversity in nature [11, pp. 272-93]. Pathology, in other words, does not always require pathos or a conscious sufferer. A sense of order will suffice.

Nonetheless, Canguilhem's argument concerning the primacy of the clinic rides on more than the conscious presentation of symptoms by a human subject. As Canguilhem explains, the primacy rests on a genealogy of clinical facts and techniques. Even though clinical scientists use laboratory techniques to detect diseases that individuals may not be aware of, these techniques are, in the final assizes, based on previous clinical encounters. In other words, "If today a physician's knowledge of disease allows him to foresee the experience of the disease by the patient, it is because in the past the patient called upon the physician." In history, then, pathology proceeds physiology. "Cells of the renal, pulmonary or splenic parenchyma can only be said to be sick, and sick from some disease, today, by an anatomo-pathologist who never sets foot in a hospital or clinic because they were removed or resemble those that were removed yesterday or 100 years ago, it doesn't

matter which, by a practicing clinician from the body or amputated organ of a man whose behavior he had observed" [1, p. 151].

True. But origin is not basis and one might easily claim that it is not because the knowledge of a cell as diseased originates in the clinic (i.e., is determined by how we have come to know about the status of that cell) that the basis for a cell being diagnosed as diseased is clinical. We can distinguish between how we come to know something and what that thing is. While it is often true that diseased cells presuppose a human being with symptoms – we ignore, here, plants and animals –, there is not necessarily a continuous chain – historical or logical – between clinical observation and a clinician's or a pathologist's judgment. Clinicians originally diagnosed AIDS, for example, on the basis of patients presenting with *Pneumocystis carinii* pneumonia (PCP) or Kaposi's sarcoma. While this is the origin of the AIDS diagnosis, it is no longer the basis of the diagnosis of HIV disease; that diagnosis depends entirely on laboratory tests. There is no clinical encounter or observation of behavior necessary. Moreover, the behavior or symptoms observed in the past are not necessary for the present diagnosis. Indeed, antibiotic therapy for PCP, the discovery that Kaposi's sarcoma requires prior herpes infection and, above all, the discovery of the HIV virus have broken the links in the chain. Laboratory investigations have, in fact, inverted the chain of events between the original clinical observations and the diagnosis. Although in the course of this paragraph, I have deliberately switched from the older term – AIDS – to the modern term – HIV disease, I have not done so in order to suppress the fact that HIV disease is a laboratory construct whereas AIDS is a clinical designation. Rather, I have done so in order to stress that both diagnoses require laboratory confirmation (in the Western world.) In particular, since the mid-1990s "The transition to a diagnosis of AIDS is now based on a laboratory parameter, e.g., a CD4 T lymphocyte count below 200/mm^3 (normal range approximately 500–1,000/mm^3) rather than the prior clinical definition of the life-threatening opportunistic infection or cancer" [12, p. 162]. In either case, as insurance carriers know, the present virological diagnosis can be made without the patient's knowledge, symptoms, or consent. So, let us return to the original question: Did the murder victim have cancer? Since it is fiction, let us suppose that upon examination the coroner discovered that the victim suffered from HIV disease and that unbeknownst to the victim, his faithful wife discovered, during a routine blood examination, that that she and her future child also suffered from the same disease. Does the county coroner have a possible murder suspect or just a philosophical problem?

Two examples

While the above remarks may seem overly critical, I do believe that Canguilhem is right to insist that attempts to reduce the pathological to the normal or to describe disease without any notion of lesion or cognate notions are most often doomed to failure. I would like to give two further examples, not discussed by Canguilhem, of attempts to reduce the pathological to the normal. I have taken these examples from my own work in the history of bacteriology and immunology in part because the bacteriological revolution has come to symbolize the triumph of ontological views of disease over physiological views such as those expressed by Bernard. Consequently, the continued attempts to reduce pathology to biology from within bacteriology have special significance and are not, I believe, motivated by the same concerns that directed earlier attempts to reduce pathology to physiol-

ogy. While my first example confirms what Canguilhem has said about Bernard, the second example raises a further issue that is somewhat more complicated, namely the status of genetic polymorphisms in the production of disease.

Let me begin with the first example, vaccine therapy [13]. Prior to antibiotics, therapeutic success in bacteriology was limited mainly to the antitoxins (notably diphtheria and tetanus antitoxins) and passive serotherapy. One might wish to add malariotherapy to the list, but that would seem to be a more empirical discovery if it can indeed be counted as one. The only other therapy directly derived from bacteriology was vaccine therapy, invented around the turn of the century by the British immunologist Almroth Wright. Wright's therapy was practiced all over Europe and North America until the advent of penicillin and then it disappeared. In its heyday, say between 1900 and 1920, practitioners applied vaccine therapy to virtually every infectious disease and even to some diseases that we would not today consider infectious, such as mental illness or cancer. By the middle of the 1930s, the therapy was reserved mainly for chronic, local infections. Nonetheless, as late as 1970, researchers at the Pasteur Institute in Paris still produced a limited number of therapeutic vaccines. In the case of some dermatological disorders, they still claimed a success rate of around 50%, although they admitted that it was impossible to say why the vaccines worked.

Vaccine therapy consisted of the subcutaneous injection of heat-killed bacteria. The species of the injected bacteria was supposed to be the same as that causing the disease. Vaguely reminiscent of the "hair of the dog that bit you" reasoning, the rationale for the therapy was based on a particular idea how immunity worked. According to Wright, infectious diseases could be divided into general and local. Local diseases – and this could include tuberculosis and pneumonia – sometimes evaded host defense systems and thus became chronic. Vaccine therapy supposedly recharged the immune system by provoking an increased production of serum substances known as opsonins.

In order to know if the immune system was depressed and thus to know if an individual was a suitable subject for vaccine therapy, the physician had to calculate the "opsonic index." The procedure for the calculation was more or less as follows: You took a blood sample from the patient, and separated the blood serum and the white cells. At the time investigators supposed that the serum contained the opsonins and that the white cells contained phagocytes. They furthermore supposed that the two entities worked together in the efficient killing of bacteria. In the meantime, you would have already cultivated the pathogenic bacteria from the patient. If you incubated the white cells, the bacteria and the serum together in a capillary tube, under a microscope you could calculate the number of bacteria ingested by the white cells. This gave you the numerator of the fraction. To get the denominator, you took the same white cells and the same bacteria and incubated the two of them with the serum from a normal, healthy person. You would do the count again. An index of less than 0.8 signified that the patient lacked a normal number of opsonins and thus required vaccine therapy to boost the number to a normal level and overcome the infection.

If we look at this calculation, I think we can see here something similar to Claude Bernard's dictum that there is only a quantitative difference between the normal and the pathological. That, at least, would seem to be the idea lying behind the comparison of normal and pathological samples on the basis of numbers alone. However, in addition to

this unstated thesis, there is a more surprising presupposition lurking behind the calculation. In order to substitute normal serum for the pathological serum, it must be assumed that normal serum is capable of withstanding infection. It must be assumed, more precisely, that the pathological event of importance is not infection, per se, but the failure to resist. Consequently, the ability to resist must function as a biological variable and this variable must be normally distributed. Infectious disease is thus not the result of infection but the result of a failure of resistance. As French authors sometimes said in this period "*Le terrain est tout.*" This is the constitutional doctrine.

Among the many constitutional doctrines that circulated between the two world wars, one known as constitutional serology applied specifically to bacteriology [14]. Ludwik Hirszfeld invented the doctrine of constitutional serology just after World War I. Prior to the war, Hirszfeld had worked with von Dungern in Heidelberg. History credits the two with having shown that the ABO blood groups are inherited in a Mendelian fashion. During the war, while stationed in Salonia, Hirszfeld and his wife worked with a multiethnic group of allied troops. They conducted a series of blood tests on the soldiers and showed that the distribution of blood types varied between ethnic groups.

After World War I, in a series of articles later collected into a monograph, Hirszfeld proposed a constitutional serology that he presented and justified as a criticism of bacteriology. According to Hirszfeld, bacteriology had so far been too concerned with clinical end-points; full-blown disease. Hirszfeld believed that, instead, one should look at the continuum running from health to disease and produce a biology of infectious disease. In order to do so, it had to be shown that the immune response to infection was merely a quantitative extension of a normal physiological process. To show this, two things had to be demonstrated: 1) that normal or physiological antibodies existed and 2) that these antibodies pre-existed any bacterial or viral infection. Blood group antibodies seemed to fit this description.

The problem with blood-group antibodies was that they did not seem to have any physiological function. Moreover, the other so-called normal antibodies seemed, at least during the inter-war period, to be a theoretical impossibility, even though they were occasionally found in the laboratory. This had not always been the case. Ehrlich, for instance, writing in the 1890s, had proposed that normal and immune antibodies were identical and suggested that normal antibodies participated in the physiological process of digestion. But Hirszfeld could not accept this. Throughout the 1920s, Karl Landsteiner had been producing antibodies against artificial haptens. The idea that pre-formed antibodies would be present and specific to substances that had no evolutionary existence seemed unlikely and indeed, until the 1960s, most immunologists considered Landsteiner's work the best evidence against pre-formed antibodies. That was the theory.

In practice, producers of diphtheria antitoxin knew that horses that had normal antibodies against diphtheria were the best producers of diphtheria antitoxin. That was taken as a sign that they were constitutionally primed to produce a vigorous reaction to diphtheria. Based on this practical knowledge, Hirszfeld elaborated the theory of "serogenesis." Simply stated, "Normal antibodies represent a 'spontaneous' cell function; antibody formation is the development and strengthening of a genotypically determined cell capacity. The response of the organism follows a predetermined path" [15, p. 489]. Notice, first of all, that this looks like a precursor to the clonal selection theory. Notice also, however,

that two different notions of the "same" are at work here: common structure (strengthening) or common origin (development). This remained the biologically plausible basis of constitutional serology.

The second proposition – that normal antibodies had some physiological function – was somewhat more difficult to establish. In fact, while it was supposed to be the basis of the doctrine, it functioned as the presupposition of the doctrine. Taking only the blood group antibodies, Hirszfeld proposed a correlation between blood groups and disease. However, a number of studies prior to Hirszfeld had shown that diseases were fairly well distributed amongst blood groups. To counter this, Hirszfeld argued that it was not 'incidence' that counted, but 'disposition'. Now, disposition was an invisible variable for clinical medicine. It included all those people that had had an infection without any clinical signs. Sick people were just the tip of the iceberg. The test for disposition was the Schick test for diphtheria and this formed the technical basis of Hirszfeld's theory.

Hirszfeld's doctrine generated numerous studies and then died out in World War II. I would like to give just one example of the kind of work that was done. In the early 1930s, Jungeblut in New York reported what he considered "direct experimental evidence" that blood type determined clinical outcome in polio. He had taken sera from convalescent polio patients that had been collected during the 1916 epidemic. Then he compared blood type of the sera with a random sample of blood types from the New York area. Using a χ^2 test, he found that there were significantly less type B than would be expected. So he concluded that Type B was resistant to polio. I would like to cite his conclusion as it is quite striking. Polio is "an extreme example of an infectious disease in the pathogenesis of which the significance of the etiological agent has become almost completely subordinate to the predominance of constitutional factors in the host" [16, p. 41]. Of course, critics were quick to point out, Esquimos had almost no B and yet no polio.

Both of the examples just given seem to point to the same mistake as if well into the twentieth century, fears of equating disease and sin still motivated researchers to reduce the pathological to the normal. As I said previously, I do not think that we have the same motives here. More importantly, I do not think we have the same kind of mistake. I think there is a difference between, say, Bernard's positivism and Hirszfeld's enthusiasm for a biology of disease. In both the immunological cases, the ambiguity between the normal and the pathological does not rely on a denial of the importance of the clinical origin of medical knowledge. In fact, it seems that the ambivalence arises from the fact that the objects under study have an intrinsically ambiguous status. To take the first case, unlike all other organ systems, the immune system does not have just a normal function. Inflammation is not a normal process even though it is frequent. Infection is not normal even though it is universal. The function of normal antibodies remains unknown and the status of Wright's opsonins can even today be differently described as either normal (the C3 portion of complement) or as pathological (IgG antibodies involved in phagocytosis). Concerning the second case, genetic polymorphisms (such as blood groups) are neither lethal mutations nor biological abnormalities. They can, however, attain something close to that status in certain instances. The classic example from HLA polymorphisms is the relation between HLA-B27 and ankylosing spondylitis (spinal inflammation) where bearers of B27 are 87 times more likely to develop the disease [17, p. 355]. In other words, what

appears to be an error motivated by ideological concerns can also be seen as a further naturalization of disease as a biological process.

Conclusion: What counts as pathology today?

In the era of "biomedicine" and molecular pathology we might be tempted to say that biology has supplanted pathology as the form of explanation of disease. In this respect, I would like to conclude with two questions, the first of which is raised by Canguilhem: Can there be a biological definition of disease that does not ultimately depend upon a clinical encounter? With the emergence of molecular biology and medical genetics, the question may be recast as follows: To what extent do congenital diseases escape the clinical encounter in so far as their diagnosis often precedes any form of clinical expression? Do we have, in other words, with the advent of molecular and preventive medicine an entirely new series of diseases similar to HIV disease in that they do not require a patient's complaint for their diagnosis?

Canguilhem dealt with this issue by distinguishing between anomaly and disease. According to Canguilhem, hemophilia, for example, is an anomaly. It can become a disease, but an anomaly by itself is not a disease. Although one might agree with the principle, the example and the subsequent reasoning require a bit of a stretch. Specifically, Canguilhem claims, "All the functions of a hemophiliac are carried out similarly to those of healthy individuals. (...) In sum, the life of a hemophiliac would be normal if animal life did not normally entail relations with a milieu, risky relations that may provoke lesions (...)" [1, p. 88]. As both we and Canguilhem know, it is really all but one of the functions (blood coagulation) that are normal. And that abnormality is indeed lethal (unless treated). On this basis alone, hemophilia could easily be classified as a disease. One might further object, why does Canguilhem classify alcaptonuria as a disease while hemophilia is deemed an anomaly? They are both genetic disorders and like the hemophiliac avoiding hemorrhaging, the same difficulties confront the sufferer of alcaptonuria who attempts to avoid tyrosine and phenylalanine in his or her diet [1, pp. 42; 206].

Canguilhem admits that the "problem of the distinction between an anomaly (...) and a pathological state is very obscure (...)" [1, p. 88]. Indeed, if we ignore the vagaries of the individual we can understand that Canguilhem wishes to raise here a very general issue. He wants to ask: Should hemophilia be classified as an instance of biological variation or should it be considered a (potentially) lethal mutation? In order to shed light on the issue, Canguilhem offers an example of biological variation that is seemingly lethal, yet is not. Fruit flies with vestigial wings do manage to compete successfully with winged flies in certain (windy) milieux thus transforming a potentially lethal liability into an asset. The idea of lethal, in this case, is thus relative to the environment. Canguilhem does not, of course, offer a milieu that is advantageous to hemophiliacs. But the die is cast, so to speak, and Canguilhem goes on to describe a number of physiological variations produced by diet and life-style. He is sufficiently impressed by this variation and by its persistence within certain ethnic groups that he is led to endorse the inheritance of acquired characteristics which, in a footnote added in a later edition, he ultimately rejects.

Whether or not Canguilhem flirted with neo-Lamarckianism is not the point here. The problem would be the same for a strict Mendelian and it is this: If the environment determines what will be considered pathological, just as it determines what will be considered

normal, then, just as there is no objective criteria for the normal, there is equally none for the pathological. We thus return here to a point raised earlier and that Paul Rabinow has put as follows: "Hence there is no purely objective pathology; rather, the basic unit is a living being in shifting relations with a changing environment" [18, p. 84]. At this point, one might wonder how Canguilhem can maintain that there is a qualitative difference between the normal and the pathological. Since many of our descriptions of an environment can be simply quantitative (more or less wind), then so would the determination of what is normal and what is pathological.

Certainly, recent trends in what are termed "diseases of civilization" are consistent with this line of thought as the language of biology tends to override the language of pathology. In an article discussing problems in developing objective measures of asthma, for example, Britton frames the problem as one of "defining the 'asthma phenotype'" [19, p. 2]. The objective of such an exercise was not necessarily to improve the diagnosis of asthma for clinical purposes. Rather, the short-term agenda sought to develop measures of the phenotype to serve those "studying the disease in general populations" since the "emergence of genetics as a new and especially active area of asthma research has added new demands for appropriate ways of defining the 'asthma phenotype'" [19, p. 2]. Notice here that while the requirements for isolating a phenotype and those for diagnosis certainly overlap, they are not necessarily the same. Symptoms such as wheezing and bronchial obstruction that are measured in the course of screening are continuous variables. Persons screened for the phenotype may thus present at a subclinical level in the sense that they themselves formulate no complaint. It is therefore certain that some members of the population that fall within the phenotype definition would not, under normal circumstances, be diagnosed with asthma. While it can therefore be said that this is certainly a form of medicine, it is not strictly oriented around the clinic and the physician/patient encounter. As in the case of plant and animal pathologies, the physician/patient/disease triad has been replaced, in this case, with a researcher/population/phenotype triad. It might be argued that, at the end of the day, the phenotype definition depends upon the initial physician/patient encounter and the specification of wheezing and bronchial obstructions as cardinal symptoms. Perhaps it does. However, the genotype will ultimately have been defined on the basis of population studies that are more than the sum of individual encounters between the doctor and the patient. When that day comes, moreover, when there is a genotype definition available, will not preventive therapy eliminate the need for the expression of the phenotype for medical intervention to occur?

The definition of a disease as a biological phenotype, moreover, need not necessarily arise from a clinical encounter. An anthropological encounter will do. This at least is what we are led to be believe by proponents of the "thrifty genotype" hypothesis. Concerning such diseases of civilizations as diabetes, obesity and hypertension, and their heightened expression in anthropological phenotypes, for example, Neel and his collaborators, suggest that "some positive selective value of the gene complex [underlying these diseases] in the past must be visualized" [20]. They propose moreover that the "ultimate genetic complexity of each of these diseases qualifies it for the term syndrome. Perhaps collectively we can speak of the 'syndromes of impaired genetic homeostasis' or, more colloquially, the 'civilization syndromes,' or the 'altered lifestyle syndromes,' to which other diseases may be added" [20, p. 61]. Notice that the altered lifestyle in a new envi-

ronment does not give rise to a disease, but to a syndrome: a configuration of diseases. The diseases themselves are not relativized, just their relative frequency. Do we nonetheless encounter here the relativism advanced by Canguilhem?

I do not think that the observation that biological variability has selective value should lead us to believe that disease is entirely relative or that pathology is completely subjective. Recall that pathological processes, mechanisms, or events are not the same as or even relative to normal processes and mechanisms. While the value (survival or health) of some capacities, dispositions, genes, or polymorphisms may indeed be relative to the environment, the events and mechanisms are not. Mutation is not transcription and proliferation is not cell division. Cell surface markers are not tumor antigens (if there are any). The vocabulary and the concepts of pathology are not reducible to physiology. In other words, I think Canguilhem's original statement could be reinforced: not only is the pathological qualitatively different from the normal, but that pathological events are not simply relative to an environment; they are absolutely different from physiological processes.

If pathology is not reducible to biology then does it follow – this is the second question – that pathology as a discipline is independent of biology? Recent events suggest convergence rather than independence. In July of 1995, *The American Journal of Pathology* added the following subtitle to the journal: "Cellular and molecular biology of disease". Founded in 1901 under the name *The Journal of Medical Research*, in 1925, following the merger of the American Society for Experimental Pathology and the American Association of Pathologists and Bacteriologists, it became the *American Journal of Pathology*. The most recent change in name had followed a change in name, 2 years previously, of the society publishing the journal. Formerly known as the American Association of Pathologists, in 1992 the society voted to change its name to the American Society of Investigative Pathology [21]. Whatever the reason for the change in name, the addition of the subtitle was prompted by the increasing importance of cellular and molecular biological methods in the study of disease mechanisms. As an editorial announcing the addition explained, "During the last decade, studies of pathogenesis and diagnosis have relied, to an even greater extent, on the methods of molecular and cell biology. These new developments sharpened the focus of pathology and, at the same time, greatly expanded its horizons by removing artificial barriers among disciplines" [21, p. 1].

Now a change in method is hardly cause for modifying the title of a journal. Methods change all the time. As stated, the object of the society is still the study of the pathogenesis of disease. As the editors point out, however, it is more than the addition or change of methods that the name-change is meant to mark. There is something of a historical movement at issue here, a movement forward from which there is, apparently, no turning back. The application of these new methods had, as a consequence, "placed the field into 'fast-forward' and made it leap from descriptive morphology to molecular pathogenesis" [21, p. 1]. While this in itself may be reason enough for the change in name, there is more.

Citing the proposition to change the name, we learn that the change in methods had, in fact, changed the nature of pathology. Or so it seems. If we look closely at the proposition however, there are at least two different possibilities. The proposition states that the renamed society "... takes pride in its members research contributions to the understanding of basic mechanisms of disease. Recognizing that this interest goes beyond traditional concepts of pathology, the society has chosen its new name to reflect the breadth

and interests of its members" [21, p. 1]. From the foregoing, it is the interest in the "understanding of basic mechanisms of disease" that "goes beyond traditional concepts of pathology." This could mean that either pathologists are now interested in more than disease, such as biology, insofar as it is the background for basic mechanisms, or that biologists are now interested in disease, insofar as it reveals, for example, pathological variations in biological pathways that point to basic biological mechanisms. Both are probably true. The pathologization of biology has proceeded apace with the biologization of pathology. The point is, nonetheless, that the formulation does not allow us to decide whether the "interests" or the "concepts" determine the subject (pathologists or members) of the sentence.

Has pathology therefore been reduced to biology? If we distinguish actors and events then so far it has not. Consider, for example, the following review article entitled "The use of biological variables to predict outcome in multiple myeloma" [22]. Given the above, we might ask what are these biological variables? According to the article, there are two: IL6 (interleukin 6) and a human herpes virus HHV8. In the first case, the IL6, rather than performing its normal function, was suspected of contributing (was a factor in) to uncontrolled proliferation. In the second case, a preliminary study showed evidence of HHV8 infection in 100% of cases of melanoma. It was moreover suspected that there was a relation between the two biological variables in that the viral genome contained a homologue (a homologous sequence) to IL6. Now it is true that viruses are biological entities and that IL6 is a normal biological substance. But can it truly be said that infection by herpes is normal or that promoting uncontrolled cellular proliferation is a normal function of IL6? Are infection and uncontrolled proliferation really biological variables or are they pathological events? Clearly, although the players are normal, biological entities, the events are pathological.

Bibliography

1 Canguilhem G. Le normal et le pathologique. 2nd ed. Paris: Presses universitaires de France; 1972.
2 Delaporte F. A vital rationalist: Selected writings from Georges Canguilhem. New York: Zone Books; 1994.
3 Gane M. Canguilhem and the problem of pathology. Econ Soc 1998; 27: 298-312.
4 Foucault M. Introduction. In: Canguilhem G, Ed.. On the normal and the pathological. Dordrecht: D.Reidel; 1978.
5 Sinding C. Relire Canguilhem: de la normativité à la normalité. Prospective et Santé 1986-87 ; 50 :21- 5.
6 Majno G, Joris I. Cells, tissue and disease: principles of pathology. Cambridge, MA: Blackwell; 1996.
7 Spicker SF. An introduction to the medical epistemology of Georges Canguilhem: moving beyond Michel Foucault. J Med Philos 1987; 12: 397-411.
8 Lawrence C. Review of Georges Canguilhem, On the normal and the pathological. Br J Hist Sci 1983; 16: 95-6.
9 Dagognet F. Le normal et le pathologique: prolégèmes. Prospective et santé 1986-87; 40: 7-10.
10 Swabe J. Animals, disease and human society: human-animal relations and the rise of veterinary medicine. London: Routledge; 1999.
11 Wills C. Yellow fever, black goddess: the coevolution of peoples and plagues. Reading, MA: Addison-Wesley; 1996.
12 Groopman J, Gill P, Sloand E. Pathogenesis and therapy of HIV disease. In: Hematology 1996. Washington, DC: American Society of Hematology; 1996.
13 Keating P. Vaccine therapy and the problem of opsonins. J Hist Med Allied Sci 1988 ; 43 : 275-98.
14 Keating P. Holistic bacteriology: Ludwik Hirszfeld's constitutional serology between the two world wars. In: Weisz G, Lawrence C, Eds. The holistic turn in western biomedicine: 1920-1950. Cambridge: Cambridge University Press; 1998. p. 283-302.

15 Hirszfeld L. Ueber die Konstitutionsserologie im Zusammenhang mit der Blutgruppenforschung. Ergebnisse der Hygiene, Bakteriologie, Immunitatsforschung und experimentellen Therapie 1928 ; 8 : 367-512.
16 Jungeblut CW, Smith LW. Blood grouping in poliomyelitis: its relation to susceptibility and the neutralizing property of convalescent sera. J Immunol 1932; 23: 35-47.
17 Roitt I. Essential immunology, 8th ed. Oxford: Blackwell; 1994; Ibid., p. 88.
18 Rabinow P. Georges Canguilhem. In: Rabinow P, Ed. Essays on the anthropology of reason. Princeton, NJ: Princeton University Press; 1996.
19 Britton J. Symptoms and objective measures to define the asthma phenotype. Clin Exp Allergy 1998 ; 28 (Suppl 1) : 2-7.
20 Neel JV, Weder AB, Julius S. Type II diabetes, essential hypertension, and obesity as "syndromes of impaired genetic homeostasis": the "thrifty genotype" hypothesis enters the 21st century. Perspect in Biol Med 1998; 42: 45-74.
21 Fausto N. What's in a name? The American Society for Investigative Pathology. Am J Pathol 1993; 142: 1.
22 Davies FE, Jack AS, Morgan CJ. The use of biological variables to predict outcome in multiple myeloma. Br J Haematol 1997; 99: 719-25.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

On the coexistence of multiple time frames in historical accounts of immunology

Alberto Cambrosio

À chacun son sablier, pour en finir avec le sablier [1]

Introduction: history and time

History, by definition, unfolds in time. In other words, in spite of the existence of different attitudes towards the coherence of historical narratives [2], time is an intrinsic dimension of all historical accounts. Lest one be tempted to argue that the same applies to 'society' in the case of sociological accounts, or 'culture' in the case of anthropological accounts, it should be noted that time, unlike 'society' and 'culture,' is not the actual object of inquiry but, rather, a constitutive element of the various, sometime widely heterogeneous events or phenomena (from battles to mentalities) undergoing historical analysis. As a result, historians, more often than not, take the temporal dimension of their narratives for granted, one obvious and notable exception – although, as we will see below, less of an exception than one would at first assume – being, of course, Fernand Braudel's notions of "*longue et courte durée*" [3]. For the moment, let us simply note that, implicitly or explicitly, the existence of a uniform, linear, unidirectional flow of time is generally perceived as a condition of possibility of any modern understanding of historical events, other notions of time, for instance circular or relativistic ones, being relegated to pre-modern or esoteric societies or to the no less esoteric debates at the interface between philosophy and modern physics [e.g., 4].

As long as alternative conceptions of time were confined to primitive (literally speaking, "pre-historical") societies, or to domains, such as the physical world, falling within the province of nature, as opposed to society or human affairs, historians did not need to worry. However, in recent years sociologists and anthropologists have turned their ethnographic eye to the temples of modern rationality, namely scientific laboratories and research institutions [e.g., 5], by the same token also questioning such time-honored great divides as those between primitive and modern thought, or nature and society. Several of these contributions have involved an explicit discussion of time parameters. For instance, Sharon Traweek [6] has examined how the activities of high-energy physicists take place in a web of entangled temporalities; Allan Young [7] has analyzed the architecture of time in psychiatric narratives; Hans-Jörg Rheinberger [8], drawing on Bachelard and Derrida, has cast his analysis of experimental systems in biochemistry and molecular biology in terms of recurrence and historiality; and Bruno Latour [9] has explored strategies of temporal attribution in order to claim that scientific objects can be simultaneously historical and real.

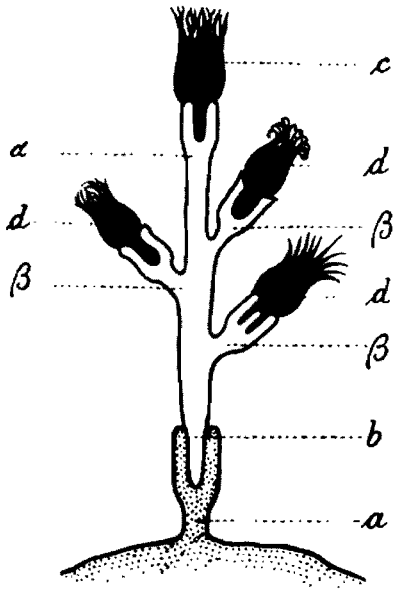
At this point, readers familiar with historiographic debates will argue that there is nothing new under the sun, citing as proof Braudel's previously mentioned famous distinction between "*courte durée*" and "*longue durée*" [3]. There have been slightly different versions of this distinction, but the basic idea is the same: in writing history, one must distinguish between different historical layers that do not obey the same temporality. In opposition to the "*longue durée*" – the almost immobile history of the relations between human beings and their environment – stands the "*courte durée*" – the highly rhythmical history of individuals (the "*histoire événementielle*") – a middle layer being occupied by the slow-moving history of social formations. Yet, one should not read into the distinction between different temporal layers an attempt to contrast incommensurable narratives since, in the end, these different temporalities must somehow fit together, as in an archeological or geological excavation where the different strata all cast light on the evolution of a same formation: "In fact, these different temporalities are interdependent: it is not time as such that is a creation of our mind but, rather, its fragmentation" [3, p. 76, my translation].

Braudel has applied his temporal framework to science and technology, by arguing, for instance, that the "*longue durée*" accounts for such long-standing formations as the Aristotelian world and, subsequently, the geometrical world of Galileo, Descartes, and Newton [3, p. 52]. Yet, lest one is tempted to identify the Braudelian "*longue durée*" with Kuhnian paradigms [10, p. 293-319], and, on that basis, to sociologize the whole thing, it should be noted that Braudel (like Kuhn; 'pace', among others, Barnes [11]) has consistently resisted any sociological interpretation of his approach. Casting Bachelard, and his admittedly unusual discussion of time and rhythmicity [12], into the somewhat surprising role of the sociological villain, Braudel has explicitly rejected any notion of a variable time, that is, of the existence of multiple times that would somehow be reducible to or engendered by a given social formation. What is of interest to historians, according to Braudel, are intersections and breaking points between temporal layers, and this presupposes the existence of a uniform time [3, pp. 77-78].

Today readers looking for a distinctly sociological account of time will probably turn to Norbert Elias [13] rather than Bachelard [12]. Elias's contribution amounts to a strong argument in favor of an instrumental, as opposed to a transcendental understanding of time. In spite of his outdated evolutionary grand narrative, Elias shares with recent work in the field of science studies an interest in the empirical tools that generate time, for he insists on the fact that time is neither a natural, physical object nor a psychological a priori but, rather, the result of an evolving set of contingent practices involving the production of standardized tools such as clocks, timetables, and so on. Time, according to Elias, blurs the lines between nature and society. But if this is so, two consequences follow: first, 'contra' Braudel, multiple temporalities, as produced by different tools and as embedded in different, parallel networks, can be separately articulated; and, second, these multiple temporalities can be reflexively injected into historical accounts.

The temporality of immunological imagery

As hinted in the previous section, scientific objects have the interesting property of de-stabilizing taken-for-granted notions of historical time: they are a source of temporal instability. For instance, previously inexistent scientific entities, once accepted as real, are granted by scientists a temporal privilege that turns them into a-historical objects (they have always



a) Receptor der Zelle. — b) Haptophore Gruppe des Amboceptors. — c) Dominantes Complement. — d) Nichtdominante Complemente.
Complementophile Gruppen des Amboceptors: α) für das dominante Complement, — β) für die nichtdominanten Complemente.

Figure 1a. 1902 representation of an antibody with receptors for dominant and non-dominant complements [19, p. 585]. Original in black and white.

been there), although they can later slide, once again, into inexistence, thereby regaining historical contingency [9, 14]. Another kind of temporal instability is exemplified by medical pedigrees, i.e., the family trees utilized in human genetics [15]. The use of pedigrees to track human hereditary diseases goes back to the nineteenth century and a comparison between early twentieth century and present-day pedigrees shows the existence of a continuity that lies less in their content than in their form [16]. In the course of the century, pedigrees have been associated with widely different genetic technologies [17], and although one could indeed argue that with each new association pedigrees have been translated into different tools, one could also maintain that in spite of and through all these transformations, pedigrees basically remained the same ruler, pen and paper technology. Castaneda [16], following Foucault [18], has argued that, in this particular case, the continuity in form can be ascribed to the fact that pedigrees embody the transition from the classical to the modern *épistémè*: form and content, in her account, thus obey different temporal rhythms.

In the present paper I will explore a somewhat similar argument, although cast in a less epochal form. As with the just mentioned example of pedigrees, my focus will be on scientific, more specifically: immunological, imagery. In spite of the epistemological privilege often granted to textual and verbal components of scientific practice, visual objects are also full-fledged ingredients, rather than occasional by-products, of scientific activities. Thus, immunological imagery is not to be conceived of simply as a pedagogical device,

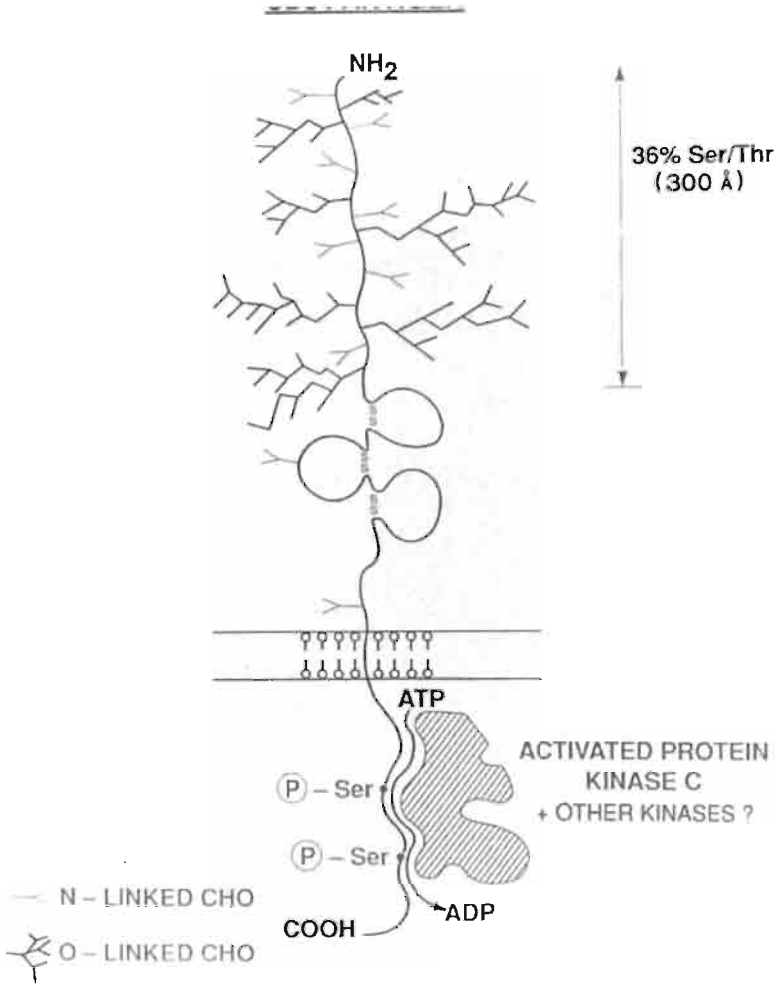


Figure 1b. 1992 schematic representation of the CD34 receptor [20, fig. 4, p. 120]. Original in black, red, and white. Illustration courtesy of Mary Ann Liebert Inc., Publishers. Original artwork kindly provided by DR Sutherland.

but, rather, as a constitutive element of the entities investigated by immunologists [19]. A detailed empirical analysis of both present-day and turn of the century immunological imagery [19, 20], conjugating semiotics with the history and sociology of science, can lead to an interesting subversion of traditional historical narratives by inaugurating the possibility of multiple temporalities. This is so because certain images can be accounted for by relating them to the material practices by which they were generated – they are, in other words, both context-bound and temporally located – while other images seem to escape narrow temporal straitjackets. This raises the interesting possibility of variable analytical temporal frames.

As a way of proceeding with the abstract argument, let me immediately turn to a

concrete example. I will contrast two sets of images:

The first set consists of three illustrations, a turn-of-the century drawing and two recent images. *Figure 1a* shows a 1902 drawing by Ehrlich and Marshall, representing an antibody (or, to use their terminology, an "amboceptor") with receptors for what the authors refer to as "dominant and non-dominant complements" [21]. *Figure 1b*, published in a 1992 article in the *Journal of Hematotherapy,* consists of a "schematic representation of the structural characteristics of the human CD34 antigen" [22, p. 120]. Finally, *Figure 1c* is an 'artistic,' computer-generated interpretation of *figure 1b* that was featured on the cover of an "educational chart" distributed by the biotech company Coulter/Immunotech in the mid-1990s.

The second set consists of two illustrations. *Figure 2a* shows a "summary diagram" from a 1908 book authored by one of Ehrlich's followers. It shows three types of cells (antigen cell, antibody cell, and complement-producing cell) with their various receptors, as well as free receptors [23]. *Figure 2b* was featured on the cover of the 1996–1997 Cytometry catalogue produced, once again, by Coulter/Immunotech, and although no explanation is given for that figure, it 'obviously' represents a cell with antibodies attached to the cell surface.

It is my claim that the illustrations in the first set do not inhabit the same temporal space, while the drawings of the second set can be construed as sharing a common temporal frame.

From a formal semiotic point of view, the figures within each set present different characteristics. For instance, some are in black and white (*figure 1a*), others in black and white with a few dashes of color (*figures 1b* and *2a*), while yet others are in full color (*figures 1c* and *2b*). Some figures have a flat, two-dimensional structure (*figures 1a, 1b* and *2a*), while others resort to perspective and three-dimensional effects (*figures 1c* and *2b*). Yet, the figures within each set also evoke a sense of similarity: a tree- or seaweed-like form in the case of the first set, and, to use an anachronistic analogy (at least for the 1908 image), a "module-landing-on-the-moon-surface" appearance in the case of the second set. In turn, this sense of similarity can be ascribed to some of the semiotic properties of the drawings. For instance, in both illustrations in the second set, a curved surface occupies the corners of the figure, and a boundary is used to separate different domains (the intracellular from the extracellular space) within each figure. Moreover, in both illustrations small entities appear to rest on or be approaching the boundary, or, to be more exact, are drawn in contact or close proximity to the boundary, the perception of movement towards the cell membrane being an effect produced by graphic conventions.

Admittedly, in *figure 2a* the bounded surface is empty, as in a portolan – those ancient maps of Africa listing only coastline towns, the continent's interior bearing only the inscription "hic sunt leones" – while in *figure 2b* the cell interior is filled with distinct shapes (we recognize them as a DNA molecule, mitochondria, and the endoplasmic reticulum). Even the boundary, rather than a reinforced but undifferentiated single line as in *figure 2a*, consists, in *figure 2b*, of a multi-layered structure. Yet, this does not detract from the fundamental similarity of the two figures, in the same way as both the African portolan and a modern African map will be recognized as maps of the same continent. In addition, notice another similarity: in both cases the relative scale of the various elements (hypothetical or real) is not respected. In *figure 2b*, for instance, the DNA molecule (recognizable to modern readers because of the conventional double-helix structure) undergoes a

progressive change of scale, a sort of perspectival magnification. As with the DNA molecule, in both *figures 2a* and *2b* the elements sitting on or close to the cell membrane (the antibodies) are 'obviously' magnified with respect to the cell. 'Obviously' because, in spite of the fact that in *figure 2a* the antibodies are represented as conventional symbols whereas in *figure 2b* they are featured by showing one of their possible molecular configurations, in both cases the artist would have known that cells are visible under a light microscope while antibodies are not.

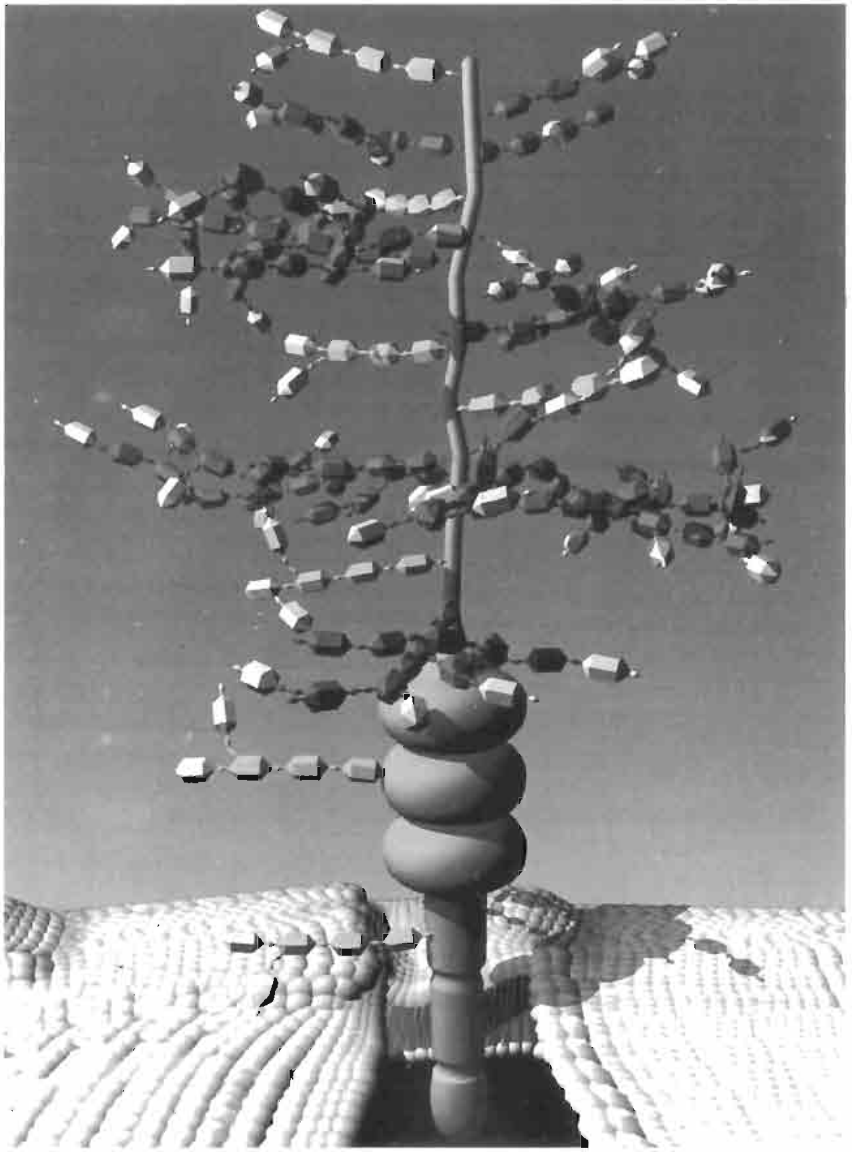


Figure 1c. Artistic interpretation of the structure of a CD34 receptor. Original in color. Illustration courtesy of Beckman Coulter, Inc. 2000.

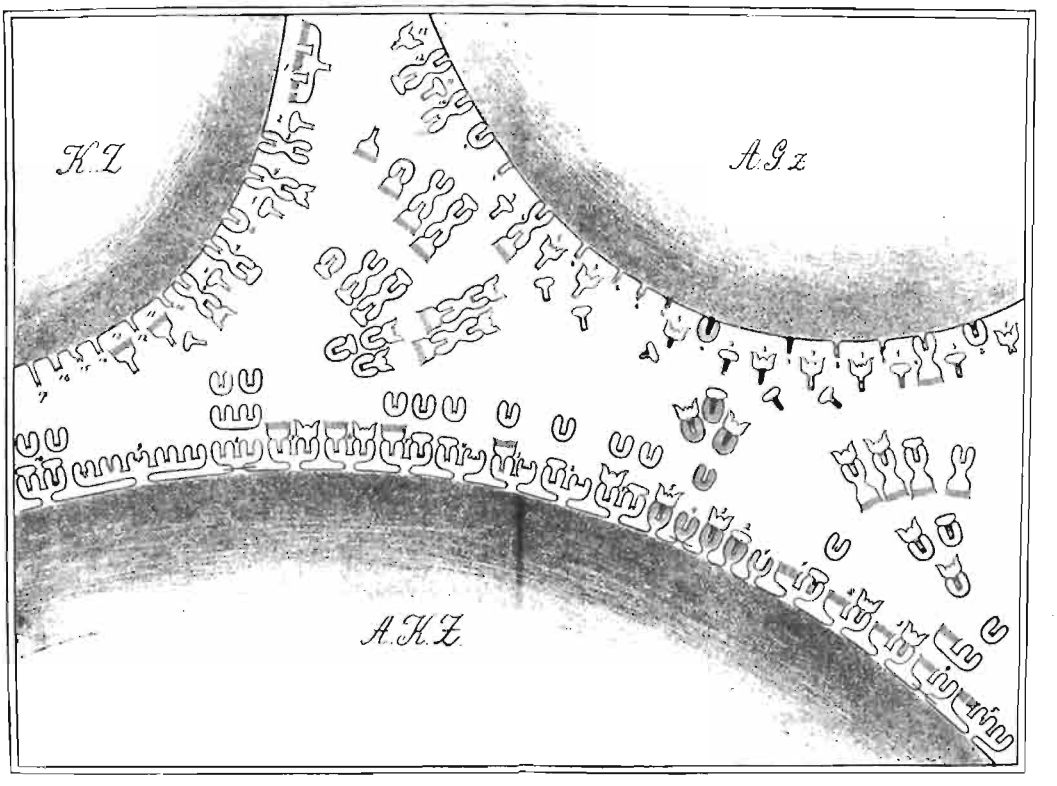


Figure 2a. 1908 "summary diagram" of the various receptors involved in immune reactions [21; unpaginated plate]. Original in black and white, with a few dashes of red, green, and blue.

What about *figures 1a, 1b* and *1c*? In all three cases the illustrations can be termed "artistic renditions," an expression that hedges the visual realism of the entities represented. But there are important distinctions. In *figure 1a*, we find ourselves in a self-conscious world of speculative analogy. In the absence of any information on the chemical structure of antibodies (whose existence as discrete chemical substances was in itself a matter of controversy), Ehrlich's turn of the century diagrams were more suggestive of marine animals than of chemical formulae. Contemporary observers noted that the figures had "something [of] the appearance of cells with pseudopodia extending out from their periphery," that they resembled "hungry pollywogs biting eagerly at inviting bits of protruding protoplasm of just the right size to make a mouthful," and Ehrlich himself spoke of "the 'tentacle' or grappling arm of the protoplasm" [19, p. 676]. In other words, nobody, especially not Ehrlich, would have argued that his drawings were realistic representations of antibodies and receptors.

In the case of *figure 1b*, things are murkier. The figure's original legend informs us that we are looking at a "schematic representation" of the structural characteristics of a receptor. Unlike Ehrlich's 1902 diagram, the 1992 schematic drawing features several

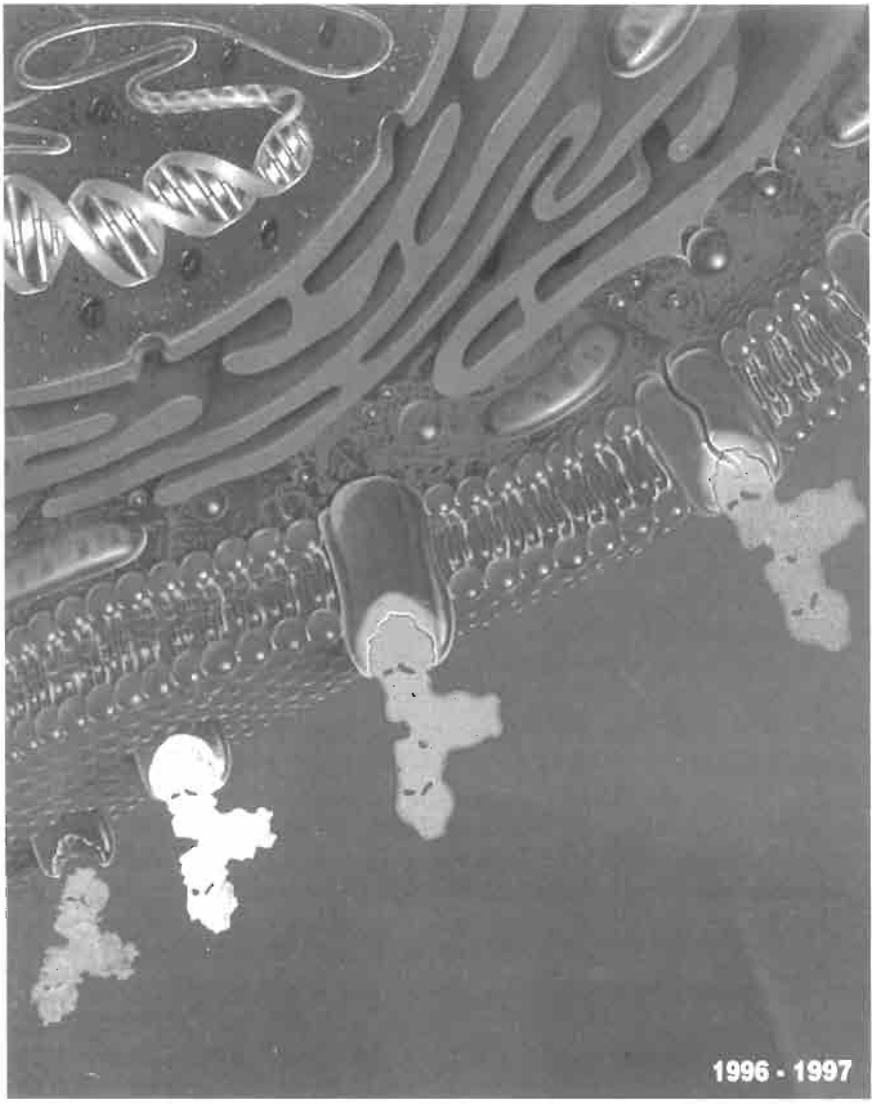


Figure 2b. Illustration on the cover of the 1996–1997 Cytometry catalogue of Coulter/Immunotech. Original in color. Illustration courtesy of Beckman Coulter, Inc. 2000.

subdomains of the molecule (O-linked glycans, amino terminals, serine residues, and so on) the occurrence of which has been secured by other researchers using various chemical and physical techniques. The drawing, then, is a composite picture obtained by piecing together various kinds of evidence and conventions. Yet, the figure still qualifies as "schematic." In the main body of the article we find additional qualifications: we learn, for example, that the "the native molecule is *anticipated* to contain" a certain structure, and that a given molecular domain "*may* be large and/or complex," "*may* induce the polypep-

tide to take on an extended configuration," "*can be anticipated* to protrude a considerable distance above the cell membrane," and "*probably* exhibits a globular conformation" [22, pp. 119-20; italics added]. In short, structural claims are accompanied by modifiers that signal their tentative status, so that, in conclusion, the wealth of immunochemical and cDNA data mobilized by the authors merely 'suggests,' as opposed to establishing, the receptor's shape.

As for *figure 1c*, it explicitly presents itself as an "artistic view" of *figure 1b*, a qualifier that while preserving some of the scientific legitimacy of the latter, adds yet another degree of representational freedom. In spite of the fact that the main author of *figure 1b* consulted with the authors of *figure 1c* in order to prevent any major misinterpretation of the original data, he claims to have had no real say in how the final product turned out and to be less than satisfied with the results, for, to him, the molecule in *figure 1c* looks very "top-heavy," whereas it should look (note the interesting analogy) like a flag-pole, "where the protein backbone is the pole and the flags are the different types of sugars decorating the molecule" [24].

The tentative status attributed to *figures 1b* and *1c* is not surprising, not simply because these illustrations can be construed as preliminary attempts at deciphering the molecular structure of the CD34 receptor, but, more decisively, because the practice of molecular modeling embodies a constant negotiation between a realist and a constructivist understanding [25]. As noted by a group of leading protein biophysicists, "for something smaller than the wavelength of visible light, there is no such thing as showing how it really looks on the molecular level" [26, p. 1186]. Yet, for the purpose at hand, models will be treated as truthful representations of molecular objects. The realist interpretation is nicely illustrated by the following anecdote. A few years ago, during a meeting on the history of immunology to which many famous immunologists, old and young, had been invited, a younger but already famous scientist who had been instrumental in producing three-dimensional computer representations of structures such as T-cell receptors, grew visibly annoyed by a theoretical debate launched by his senior peers. He stood up and briskly argued that there was no point in pursuing those old debates: why not simply settle the issues by inspecting the molecular spatial configuration of the relevant structures? This was not simply a rhetorical outburst during a conference. When a controversy erupted among immunologists over whether the Major Histocompatibility Complex I or II was involved in autoimmune diabetes, the production of a three-dimensional configuration of the relevant molecules settled the debate in a way other experimental results had not been able to do [27]. A similar sort of pragmatic realism is constitutive of the production of designer-drugs [28]. Scientists thus assume, for all practical purposes, that molecular representations such as *figure 1b* and, to a lesser degree, *figure 1c*, correspond to real structures. "For all practical purposes" here means: for the purpose of further experimental or clinical intervention.

So, what makes the difference between *figures 1a* and *1b/c*? A quick answer is: a thick layer of instrumental inscriptions that sustain the hybrid realist status of representations such as *figures 1b* and *1c*, thus grounding them (almost paradoxically) in a specific time and place. By speaking of instrumental inscriptions, we do not intend to reduce complex issues of representation to a question of mere technology narrowly conceived, for instrumentalities should be understood in terms of "*dispositifs*" [29, p. 299; see also 30] connect-

ing instruments, skills, organisms, notions, reagents, texts, and so on. In turn, "*dispositifs*" function as part of experimental systems, those "machines for the production of difference" that cannot be reduced to either science or technology (since the difference between science and technology is precisely one of their outcomes). A computer-generated molecular model such as *figure 1c* occupies a middle position within the continuum between epistemic and technical things produced by experimental systems [8, p. 110]. Experimental systems, as graphematic articulations (or arrangements of inscription-generating devices), generate their own spaces of representation [8, chapter 7]. They are not, however, pure representational domains, since in an important sense they are a combination of material and graphic traces. Correlatively, experimental systems generate their own recurrent time for "with respect to the movement of material systems, systems of things, or systems of action, time can be viewed as an operator and not simply as a chronological axis of extension in a system of coordinates. ... With 'differential temporality' we are further than ever from the romantic illusion of history as an all-pervading 'totality' dominated by mimetic, metamorphotic, or 'expressive' relations of the parts within an ensemble" [8, pp. 180-2]. We are also obviously far from Braudel, here.

Do the previous considerations also apply to the second set of illustrations? At first, one might think that this is indeed the case. For instance, while the receptors shown in *figure 2a* are purely conventional symbols, the structure of the antibody molecules in *figure 2b* corresponds to one of the accepted renditions of their three-dimensional configuration. However, if we shift our attention from the single elements taken in isolation to the two figures each taken as a whole, the situation is no longer the same. Both illustrations in the second set occupy a sort of middle position between, on the one hand, the "scenes from deep time" analyzed by Martin Rudwick [31] or the Sarah Landry drawings in E.O. Wilson's "Sociobiology" analyzed by Greg Myers [32], and, on the other hand, the schematic renditions of electron microscopic pictures of cells and tissues analyzed by Michael Lynch [33].The former are designed to convey the impression that the situations they portray (prehistorical forests inhabited by now extinct animals, or groups of apes exhibiting various forms of social behavior) could, at least hypothetically (very much so in the case of dinosaurs!), have been witnessed by an observer. Their fullness of detail suggests some contact with reality, yet they are neither mechanical records (photographs) nor records of an actual scene witnessed by an observer. Rather, these drawings "arrange their components in a way a photographer cannot, to make them typical" [32, p. 240]. The schemes of cellular or subcellular entities analyzed by Lynch are often juxtaposed to their corresponding electron-micrograph in order to allow an inexperienced viewer to immediately "see" the relevant features of the polysemic photographic image. They share with the previous category of images the goal of typicality. However, they go further in that direction, or, rather, they move in the opposite direction by stripping away (instead of adding) the gratuitous detail to be found in the photographs, in order to enhance the "eidetic" quality of the diagram [33, p. 162]. The horizon in no longer one of realism but, rather, one of re-specifying the image around a few relevant features, relevant, that is, in relation to a specific argument or fact-producing activity.

The illustrations in the second set, then, are some sort of hybrid: not quite realistic (witness the manipulation of scale and dimensions) but also not schematic (they resort to shading and other graphical tricks in order to recall "real" objects). Both *figures 2a* and *2b*

create a sort of imaginary cellular landscape, mimicking the hidden recesses of the body inhabited by various microscopic and sub-microscopic entities, as in the 1966 Hollywood movie "Fantastic Voyage," that featured human observers reduced to microscopic dimensions and injected into the body with a similarly reduced submarine, from where they could observe cells and tissues on an equivalent scale. It is not so much that these illustrations aim at showing how internal tissues and cells "actually look," since their unrealistic graphic features (cross-cuts through the cells, changes of scale) are obvious, not to speak of the fact that no single image could pretend to capture the heterogeneity of bodily tissues, cells, and organelles. Rather, these cellular and molecular landscapes embody a representational strategy governed by relations of similitude rather than of resemblance [34]. They combine conventional features and symbols with ideal-type reconstructions of the environment within which the symbolized entities operate. They thus provide a utopian connection between the macro-world of laboratory or clinical interventions and the invisible world which is the target of, say, immunology or molecular biology.

Moreover, these utopian landscapes maintain a stable graphic-symbolic operating environment in the face of evolving instrumentalities. As evidenced by the shift from *figure 2a* to *2b*, although we are still within a same or similar representational order, the illustrations feature an increasing number of new and complex entities (both intracellular, such as different kinds of organelles, and intercellular, such as various kind of biochemical messengers, adhesion molecules, etc.). Yet, these obvious differences are secondary to a more fundamental continuity, for while the artist drafting the illustration will, at any given time, borrow the necessary semiotic elements from the repertory of visual forms available at the moment, the overall model has not changed.

If we now apply the analysis in terms of experimental systems to the cellular landscapes of *figures 2a* and *2b*, it appears that the latter are not generated by any particular experimental system. Maybe, then, they should be accounted for by resorting to the notion of an experimental culture that somehow connects the various experimental systems. This solution seems highly unsatisfactory in the view of mounting criticism of all-encompassing cultural or social formations. An entirely different route seems to be called for in tackling this issue, one that does not start from the idea that there must be some consistency, either in kind, in time, or topographical, between systems. Let us restate the question: Do certain kinds of images exist outside experimental systems and do they come with their own temporality, in the same way as experimental systems do? If we grant this, it becomes possible to argue that the spaces of representation corresponding to those images obey different regulatory mechanisms, both in terms of their production and of the temporality they engender. Borrowing Umberto Eco's discussion of invention as code-making, one could go so far as to speak of a semiotic mode of production whereby "something is mapped from something else which was not defined and analyzed before the act of mapping took place" [35, p. 250]. A semiotic mode of production generates specific kinds of events, distinguishable from experimental or instrumental events. They intersect, for sure, but their intersection is defined by heterogeneity and polymorphism. They play different regulatory roles, produce different networks, generate different temporalities. Their analysis calls for different tools that mirror, in their heterogeneity but not in their nature, the heterogeneity of the events.

Consequences

The discussion presented in the previous section is still tentative or, some would say, speculative: rather than provide a full-fledged treatment of the issue of multiple temporalities in historical accounts, its main purpose has been to raise this issue in the hope that it might find some echo in future work. Before taking up the challenge, however, readers will want to know whether the game is worth playing or, in other words, what the possible consequences are of taking this issue seriously. In these concluding remarks, I will limit myself to one obvious domain of application, namely the one concerning the thorny debate between continuity and discontinuity in historical accounts.

Whereas writers in the tradition of French historical epistemology have emphasized discontinuities, scholars reacting against that tradition tend to argue that, in fact, continuities are a more interesting topic for sociohistorical analysis. A few authors have argued that in empirical historical-epistemological research we are confronted with a mix of continuities and discontinuities. This has led, for instance, Joseph Rouse [36] to offer a reinterpretation of Kuhn's notion of paradigm in performative, as opposed to representational terms [37], a move that allows for the replacement of incommensurability (and thus discontinuity) with a certain amount of continuity (as warranted, for instance, by the role played in research by similar instrumental arrangements). But what is at stake here is more than an ecumenical plea for a peaceful coexistence between continuities and discontinuities. The key issue is the attempt to re-specify relations between continuities and discontinuities, for, first, it would appear that new discontinuities introduce new continuities (and vice versa), and, second, that there exist distinctions between the time coordinates according to which continuities and discontinuities display themselves. Continuities and discontinuities, in other words, should not be regarded merely as a matter of historical interpretation but, rather, as full-fledged events in their own right.

The first aspect is clearly captured in Rheinberger's remark that "What we call history is deferred in a constitutive sense: the recent is made into the result of something that did not so happen. And the past is made into a trace of something that had not (yet) occurred." This is so, because "recurrence [...] is at work as part of the time structure of the innermost differential activity of the systems of investigation themselves" [8, p. 178]. Yet, as we have seen, systems of investigation are only part of the story, and differential time structures are generated both from within and from outside these systems. Thus, in describing objects and practices, time should no longer be treated as an external parameter but, rather, as one of the intrinsic parameters defining a particular object or practice. This has the interesting result of disconnecting issues of space and locality from temporal issues. Far from sounding the death knell of history as a scholarly enterprise, as some may fear, the recognition of multiple temporalities appears to open up unexplored possibilities of scholarly inquiry.

Acknowledgements

A preliminary version of this paper was presented at the Boston Colloquium for the Philosophy of Science's conference on "Historiography in the Life Sciences" (April 1998). Many thanks to Fred Tauber for his kind invitation to that conference. Subsequently, several key comments and suggestions were offered by Peter Keating, with whom I entertain a long-standing collaboration, and Daniel Jacobi, a source of many insights concerning immunological imagery. Research for this paper has been made possible by grants from the Hannah Institute for the History of Medicine, the Social Sciences and Humanities Research Council of Canada, and Quebec's Fonds FCAR.

Bibliography

1 Char R. Ce bleu n'est pas le nôtre. In: Char R, Ed. Aromates chasseurs. Paris: Gallimard; 1975. p. 10.
2 Megill A. 'Grand Narrative' and the discipline of history. In: Ankersmit F, Kellner H, Eds. A new philosophy of history. Chicago: The University of Chicago Press; 1995. p. 151-73.
3 Braudel F. Écrits sur l'histoire. Paris: Flammarion; 1969.
4 Klein E. Le temps. Paris: Flammarion; 1995.
5 Knorr-Cetina KD, Mulkay M, Eds. Science observed. Beverly Hills: Sage; 1983.
6 Traweek S. Beamtimes and lifetimes: the world of high energy physics. Cambridge, MA: Harvard University Press; 1988.
7 Young A. The harmony of illusions. Inventing post-traumatic stress disorder. Princeton: Princeton University Press; 1995.
8 Rheinberger HJ. Toward a history of epistemic things. Synthesizing proteins in the test tube. Stanford, CT: Stanford University Press; 1997.
9 Latour B. Did Ramses II die of tuberculosis? On the partial existence of existing and non-existing objects. In: Daston L, Ed. Biographies of scientific objects. Chicago: The University of Chicago Press; 2000; p. 247-69.
10 Kuhn TS, Ed. The essential tension. Chicago: The University of Chicago Press; 1977. p. 293-319.
11 Barnes B. T.S. Kuhn and social science. New York: Columbia University Press; 1982.
12 Bachelard G. La dialectique de la durée. Paris: Presses Universitaires de France; 1963.
13 Elias N. Du temps. Paris: Fayard; 1996.
14 Keating P, Cambrosio A. Helpers and suppressors: on fictional characters in immunology. J Hist Biol 1997; 30: 381-96.
15 Nukaga Y, Cambrosio A. Medical pedigrees and the visual production of family disease in Canadian and Japanese genetic counselling practices. In: Elston MA, Ed. The sociology of medical science and technology. London: Blackwell; 1997. p. 29-55.
16 Castaneda C. The pedigree: time, space, and an 'everyday technology' in the sciences of heredity. In: Weigel S, Kollek R, Eds. Genealogie und Genetik. Berlin: Akademie Verlag; [2000, forthcoming in German translation].
17 Nukaga Y. A genealogy of genealogical practices. The development and use of medical pedigrees in the case Huntingon's disease [dissertation]. McGill University: Department of Sociology; 2000.
18 Foucault M. Les mots et les choses. Paris: Gallimard; 1966.
19 Cambrosio A, Jacobi D, Keating P. Ehrlich's 'beautiful pictures' and the controversial beginnings of immunological imagery. Isis 1993; 84: 662-99.
20 Cambrosio A, Keating P. Of lymphocytes and pixels: the techno-visual production of cell populations. Stud Hist Philos Biol Biomed Sci 2000; 31: 233-70.
21 Ehrlich P, Marshall HT. Ueber die complementophiles Gruppen der Amboceptoren. Berl Klin Wochenschrift 1902; 39: 585-7.
22 Sutherland DR, Keating A. The CD34 antigen: structure, biology, and potential clinical applications. J Hematotherapy. 1992; 1: 115-29.
23 Schatiloff P. Die Ehrlichsche Seitenkettentheorie erlaütert und bildlich dargestellt. Jena: Gustav Fischer; 1908.
24 Sutherland DR. Personal communication, 25 May 2000.
25 Francoeur E. The forgotten tool: the use and development of molecular models. Soc Stud Sci 1997; 27: 7-40.
26 Richardson JS, Richardson DC, Tweedy NB, Gernert KM, Quinn TP, Hecht MH, et al. Looking at proteins: representations, folding, packing and design. Biophys J 1992; 63: 1186-209.
27 Hoffman M. New theory of diabetes etiology riles immunologists. Science 1992; 255: 532-3.
28 Werth B. The billion-dollar molecule: one company's quest for the perfect drug. New York: Simon and Schuster; 1994.
29 Foucault M. Dits et écrits. vol. III. Paris: Gallimard; 1994.
30 Rabinow P. Making PCR. A story of biotechnology. Chicago: The University of Chicago Press; 1996.
31 Rudwick MJS. Scenes from deep time. Early pictorial representations of the prehistoric world. Chicago: The University of Chicago Press; 1992.
32 Myers G. Every picture tells a story: illustrations in E.O. Wilson's sociobiology. In: Lynch M, Woolgar S, Eds. Representation in scientific practice. Cambridge, MA: MIT Press; 1990. p. 231-65.
33 Lynch M. The externalized retina: selection and mathematization in the visual documentation of objects in the life sciences. In: Lynch M, Woolgar S, Eds. Representation in scientific practice. Cambridge, MA: MIT Press; 1990. p. 153-86.

34 Foucault M. This is not a pipe. Berkeley: University of California Press; 1983.
35 Eco U. A theory of semiotics. Bloomington: Indiana University Press; 1979.
36 Rouse J. Knowledge and power. Toward a political philosophy of science. Ithaca: Cornell University Press; 1987.
37 Pickering A. The mangle of practice. Time, agency and science. Chicago: The University of Chicago Press; 1995.

Singular Selves
Historical Issues and Contemporary Debates in Immunology
AM Moulin, A Cambrosio, Eds

Immunology à la Plutarch – biographies of immunologists as an ethical genre

Thomas Söderqvist

What is the history of immunology good for? The participants in the meeting "Immunology: historical issues and contemporary debates" at the Musée Claude Bernard on 3–6 June 1998 represented a wide range of more or less explicit ideas about the purpose of their craft. Some wanted to understand the immunological past for its own sake, others to make a contribution to the larger historical pattern. Some – by taking immunology as a (for the moment) useful, but in principle replaceable, object of inquiry – had more epistemological or sociological axes to grind: for example, Alberto Cambrosio and Peter Keating, who utilized their studies of the recent history of immunology to explore the sociological concept of a(n immunological) "platform" [1]. Others yet were apparently more interested in promoting immunological practice, like Leslie Brent, who wanted to use his conceptual history to inform a younger generation of immunologists about the experimental and theoretical foundations they stand upon, to make them better immunologists [2].

Thus the history of immunology can be written for a variety of purposes. With this paper I wish to add yet another entry to the catalogue of intentions, namely that historical and biographical studies of individual immunologists can also be undertaken for ethical purposes. Drawing on the recent revival of interest in classical historiography and biographical writing, I would argue that biography is not only a historical or literary genre, but has important normative connotations as well. Biographies of immunologists can thus also be written with the intention to provide "moral exemplars" for other immunologists to follow, not in order to make them better immunologists in a restricted sense, i.e., to produce better experiments and theories, but to be exemplars for others how to live a "good life" in science. In other words, biographies of immunologists can have eudaimonistic (from Greek *eudaimonia*, "good fortune, flourishing") purposes.

The case of Niels K. Jerne

My argument is closely bound up with my experience of writing the biography of the English-Dutch-Danish immunologist Niels K. Jerne [3]. Jerne, who received the Nobel Prize in Physiology or Medicine in 1984, was probably the most renowned and influential immunologist of the 1960s, 1970s, and 1980s. The foundation of his fame was laid when, in 1955, he published a radical alternative to the theories of antibody production that had prevailed until then [4]. Opposing template/instruction theories which postulated that antibody specificity is shaped by the intruding antigen as a "template," Jerne asserted instead that all kinds of specific antibodies already exist, preformed, and that the antigen's only

function is to select the best-fitting kind. This natural selection theory challenged an ingrained immunochemical principle and advocated a more biologically-oriented approach with, as Jerne himself expressed it, "Darwinian overtones" [5].

The selection theory was neglected at first, and Jerne embarked on a career as a medical officer at the World Health Organization (WHO) in Geneva. In 1957, however, his theory was modified by the Australian virologist Macfarlane Burnet into the clonal selection theory [6], and as the selective principle began to be accepted as the foundation of immunology, Jerne's reputation as the new discipline's leading theoretician grew. In the early 1960s, he was asked to organize the WHO's program for immunology, and in 1962 he was called to be chairman of the Department of Microbiology at the University of Pittsburgh Medical School, where he worked out a method for demonstrating individual antibody-producing cells [7]; the "plaque" technique soon became one of the most used methods in burgeoning cellular immunology. A few years later, Jerne returned to Europe as Director of the Paul Ehrlich Institute in Frankfürt-am-Main, now with the expectation of being immunology's "Messiah" [8].

At the annual Cold Spring Harbor symposium on antibodies in 1967 – a conference that marked immunology's full acceptance as a member of the family of life sciences – Burnet declared that the selection theory was the central dogma of immunology and that Jerne was its "onlie begetter" [9]. Jerne confirmed his standing as the leading theoretician in the field when, in 1968, he was asked by the Swiss pharmaceutical company Hoffman-La Roche to create, in Basel, the world's then largest international immunological research institute. Basel Institute for Immunology came to occupy a position central to the next phase of post-war immunology – the fusion of cellular immunology with molecular biology – resulting in two more medical Nobel Prizes: to George Köhler in 1984 and to Susumu Tonegawa in 1987 [10]. Jerne's last major theoretical contribution, the idiotypic network theory, constituted still another radical break with the traditional understanding of the immune system (Anne-Marie Moulin, in my view somewhat hyperbolically, even elevates the network theory to a "Copernican revolution" in immunology) [11]: against the current picture of antibody formation as a response to external influence, Jerne drew the picture of the immune system as a self-generating, cybernetic network. Studies of the immunological network came to occupy a whole generation of immunologists, divided into skeptics and advocates of Jerne's theory [12].

Jerne's position in post-war immunological research is not reflected in any especially comprehensive scientific production. His collected output only amounts to around 85 papers (and no books), a substantial portion of which were published in journals without peer review; in addition, only about a fifth of these are experimental reports. Several of his papers nevertheless had a strong impact on immunological research in the 1960s, 1970s, and 1980s, and some have become modern classics in the immunological literature [13]. Most significantly, Jerne hovered as a critical spirit over the new discipline. Many saw him as its leading intellectual figure, the one who, more than any other immunologist, raised discussions above the level of everyday work. Burnet called him "one of the most intelligent biologists of this century" and "the most intelligent immunologist alive," and even his sharpest critic, Melvin Cohn, saw him as "the dominant figure" in late twentieth century immunology [14].

The purpose of biography

This is not the place to evaluate Jerne's contributions to immunology, nor to discuss the responses to his theories [15]. Instead, I want to summarize some of my experiences in writing Jerne's biography, and particularly to communicate some afterthoughts about what a biography of an immunologist is good for.

Traditionally, biographies of scientists were written by other scientists to commemorate their heroes, a practice that goes back to the funeral speeches ("orationes funebres," "*Leichenpredigten*") for deceased university professors in the sixteenth and seventeenth centuries and to the early stages of the European scientific academies in the late seventeenth century [16]. The "*Éloges*" of the Académie des Sciences in Paris are probably the best known example [17]. This eulogistic practice has survived in the form of innumerable obituaries and encomiastic memoirs in scientific journals and proceedings of scientific societies. A short glance in the Index of Personalities of the consecutive volumes of "Isis Cumulative Bibliography" reveals that brief, commemorative biographical articles remain a huge genre – and brief biographical articles of microbiologists, pathologists, serologists, and immunologists is no exception. But in the course of the nineteenth and twentieth centuries, the purpose of writing lives gradually changed – from eulogy to history – and during the last century most scholarly biographies of scientists, short articles as well as book-length studies, have been written and read as contributions to the history of science. The practice of biographical writing has adopted the prevalent historical standards of the time, and thus the trend in recent decades towards more contextual histories of science has manifested itself in increasingly culturally contextualized biographies of scientists [18].

When I began my research for the Jerne biography, I was occupied with two prevailing ideas among historians of science about the uses of biography: first, that biographies could be used as a lens to gain a broader understanding of science in its social and cultural context, i.e., that the individual scientist is the "single package" from which the "scientific, philosophical, social and political ideas" of the times can be "wrapped up" [19] and second, that biographies could be used for reconstructing the laboratory practices and the cognitive and "investigative pathways" that lead to theories, i.e., what Fredric L. Holmes does in his paradigmatic 900-page, fine-grained reconstruction of the biochemist Hans Krebs's daily work in the laboratory leading to the citric acid cycle model of internal metabolism [20].

Initially, both these views of what a scientific biography is good for were sustained by Jerne's rich collection of papers, donated to the Royal Library in Copenhagen in 1992. This is a unique collection of a late twentieth century biomedical scientist. The approximately 500 not yet registered boxes contain thousands of laboratory and other scientific notes, many thousands of letters to and from almost every immunologist of any importance between the mid-1950s and the 1980s, hundreds of outlines to scientific papers and lecture manuscripts for scientific conferences, and records from a large number of scientific meetings.

A large proportion of the events in the immunologically exciting quarter century from the mid-1950s to the early 1980s are reflected in the Royal Library collection, so that Jerne's life and work could indeed be used as the package from which post-war immunology might be unwrapped. In other words, because Jerne's scientific career coincided with the development of immunology from what one of his contemporary colleagues, Baruj

Benacerraf, called "a comparatively virgin field" at the close of the 1940s to a core biomedical science in the 1970s and 1980s [21], his biography might illuminate the complex social, cultural, and intellectual interplay behind the establishment of the new discipline. The abundance of laboratory notebooks and other notes, including series of successive drafts of manuscripts and conference lectures, could also be utilized for a reconstruction à la Holmes of the interaction between laboratory practice, theoretical arguments, reading experiences, and conversations with other scientists that constituted Jerne's particular "investigative pathway" to the selection and network theories.

Biography as an ethical genre

Elements of these two dominating purposes of biographical writing in the history of science reverberate in the Jerne biography. However, as I worked my way through the tens of thousands of documents in the archive, another – and as I later understood (see below) – much older conception of biography opened itself up to me. Since the age of 16, Jerne had collected almost everything that passed through his hands: not only scientific documents of the kinds mentioned above, but also thousands of private letters, diary records, scraps of papers with passing thoughts, library book-loan receipts, movie ticket stubs, chess records, bills, prescriptions, and so forth. He had kept drafts or copies of almost all his outgoing correspondence, and had even reclaimed his own letters to parents, wives, and close friends. He had, in other words, lived a "biographical life," the life lived in expectation of one's biographers [22].

I therefore began to entertain the idea that a biography of Jerne might serve other masters than the history, sociology, or practice of immunology. I was particularly struck by the large number of letters and diary entries in which Jerne reflected upon his life, not only at the desk or at the bench, but life in general. In the diaries of his adolescent years in the Netherlands, a Niels Jerne was revealed who felt like an alien in the culture, who read Nietzsche and Dutch romantic poets, and who longed for the sublime. Out of the documents from his Copenhagen years in the late 1930s, 1940s, and early 1950s emerged the picture of a man who long hesitated before choosing science as his life's work. For example, in July 1943, when Jerne had just been employed as a secretary in the Danish State Serum Institute he wrote to his first wife Tjek, wondering "whether it is sensible, humanly speaking, to invest such great energy and powers of concentration in something so specialized, something that doesn't even slightly impinge on your personal sense of life," and expressed concerns about employing his time "in the demanding assignment of familiarizing yourself with a thought-structure that others have already built up to completion; to develop this part of your life as a dabbler in peripheral abstractions, while the pulsing purple-red blood in your veins and the feelings in your heart have to accept 'later' " [23].

Other similar letters and diary entries gradually convinced me that this rich collection of documents allowed me to tell a different story of Jerne's life. In our conversations, too, we entered more and more deeply into his view of himself and his evaluations of his life trajectory, both inside and outside of science. I grew increasingly fascinated with the ways he had come to develop his identity and his understanding of himself, his self-image and his social persona. I became more interested in the way he had orchestrated his beliefs and talents into the shape of an immunologist than by the subsequent success story, culmin-

ating in the Nobel Prize. I began to put questions to myself such as: What choices did Jerne make during his life, and what consequences did these choices have, for himself and others? Which life situations attracted him, and which did he try to avoid or flee from? What brought him to pursue a scientific life instead of a career in business, a life as a doctor, writer or philosopher, or a life of taking care of family and children? How did he bring together, or separate, his life inside and his life outside science? What intellectual and moral virtues and vices did he develop? How did he live his life in such a way as to gain a sense of worth, meaning, and connectedness, and how did he make use of autobiographical narratives to convey this sense of meaning to me and to others? Was there a "narrative unity" in his life, in the sense that there was a pattern in the transaction between his deep, emotional experiences, his social character, and the multitudinous settings of his life [24]?

All these questions, of course, address Jerne's life as a whole, not only his scientific practices and his work. As Keynes' biographer Robert Skidelsky puts it: "With the life, rather than the deeds, the achievement, we have entered a new biographical territory, still largely unexplored" [25]. I call this approach existential biography, as opposed to scientific biography. Inquiries into the way a scientist has tried to make sense out of his or her private and professional life are not primarily aimed at understanding the process and practices of science in the restricted sense; they rather resuscitate a much older conception of the purpose of biographical writing, i.e., that life stories can provide answers to the classical ethical question of what it means to live a good and flourishing life? In other words, biographies which focus on the full lives of scientists are – potentially, implicitly or explicitly – edifying. Biographical narratives can furnish scientists with models with whom, for better or worse, they can compare themselves, somewhat like the way Richard Rorty claims that novels can "take us out of our old selves by the power of strangeness, to aid us in becoming new beings" [26].

The notion of biography as an ethical genre is not a new one. The lives of others have been offered as "moral exemplars" during most of Western history [27]. During the Hellenistic Age, history in general was supposed to lead to moral improvement [28], and biographers, too, clearly had ethical purposes in mind when they wrote their texts. The best known ethical biographer in classical antiquity, L. Mestrius Plutarch, wrote his "Parallel Lives" ("*bioi paralleloi*") to provide model patterns of conduct, both for the author himself and for his readers. A young Greek man aiming for a career in the provincial administration was supposed to emulate by imitation the virtues and noble deeds of famous Greek and Roman soldiers, orators, and statesmen, and correspondingly avoid their vices and failures [29].

Plutarch's biographical project, widely imitated in late antiquity and in Byzantium, was rediscovered in the Latin West in the fifteenth and sixteenth centuries. Rabelais knew his Plutarch well, as did Montaigne. The best known Renaissance emulator of Plutarch was the painter and art historian Giorgio Vasari, whose "Vite de'piú eccelenti architetti, pittori, et scultori italiani," first published in Florence in 1550, transferred the idea of biography as edification from the political sphere to the world of artists. Vasari's main purpose was to "say something useful and helpful to our own artists" and "inspire some of the more able among us to give them every possible encouragement" (a goal not very different from that of Leslie Brent and Arthur Silverstein), but he also wanted "to show how men have acted

wisely or foolishly, with prudence or with compassion and magnanimity." The true spirit of history, Vasari said, "fulfils its real purpose in making men prudent and showing them how to live" [30].

Biographers continued to write along these explicitly morally edifying lines until the end of the nineteenth century. The late eighteenth and early nineteenth centuries saw a proliferation of "Plutarchs" in many countries, for example, "The British Plutarch" (1762), a "*Deutscher Plutarch*" (1822–1824), an "*Oesterreichischer Plutarch*" (1807–1812), etc. [31]. The purpose of the "*Svensk* [Swedish] *Plutarch*" of 1820, for example, was "to rouse love for the nation, to strengthen manly spirits, to awake the feeling of human value and human ability, to fashion civic ways of thinking" [32]. The Scottish author Samuel Smiles, best known for his biographical sketches of engineers and other heroic individuals of the industrial revolution, concluded that "the chief use of biography consists in the *noble models of character* in which it abounds" [33].

Like the commemorative eulogy, however, the character-building biography was soon superseded by the more "objective" and "scientific" view of history that swept over European historical writing in the nineteenth century, and for this reason too, biography was incorporated into mainstream historiography. The ethical component, so pervasive in the genre from Plutarch to the late nineteenth century, has largely disappeared from the explicit agenda of biographical writing today, although there seem to remain traces of it, at least as a more or less hidden subtext, even in so-called scientific biographies.

In recent years, however, there has been a renewed interest in the ethical aspects of historiography [34]. Some of this interest is spurred by a re-reading of Nietzsche and his plea for a "history in the service of life" [35]; others are stimulated by the recent reappraisal of the classical Greek and Roman historiographical tradition. Even Polybius, who was adopted as the special pet of the scientific historians, has turned out to be more overtly moral than the conventional modernist reading allows. In his recent book "Moral Vision in the Histories of Polybius," Arthur Eckstein not only reinterprets Polybius as a thoroughly moral historian, but raises fundamental historiographical questions of the function of praise and blame in history; as Eckstein shows, Polybius was fully aware of the moral component of historiography and was at the same time able to reconcile this with the purposes of critical historical scholarship [36].

The recent revival of ethical historiography and biography is paralleled by an interest in the Plutarchian biographical tradition [37], and the theoretical foundation for this reassessment of ethical biography is, in turn, closely bound to a renewed interest among moral philosophers in virtue ethics [38]. Virtue ethics can best be understood in contrast with the two dominant forms of ethical traditions in modernity, viz., consequentialism and deontology. Both traditions focus on how one should act in given situations. Consequentialists direct their attention to the acts that the agent is required to perform, and more specifically to the acts that will produce the best consequences; deontologists, on the other hand, consider right actions as those that are performed out of respect for the moral law. Both ethical traditions view morality as a set of principles that guide us to perform (and particularly not to perform) certain actions. The basic question for these modernist ethical inquiries is: "How should I act?"

The consequentialist and deontological traditions both fit well with the view of the scientist as an actor, as a producer of knowledge. According to both traditions, cognitive

acts are judged. The only difference is whether these acts should be judged with reference to a set of moral norms or to their scientific, social, or political consequences, i.e., Is a particular investigation in accordance with a pre-given set of moral principles? But from the vantage point of existential biography, according to which the scientist is described as a "whole" moral being, an alternative ethical theory is needed – specifically one which is not act-oriented or founded legalistically, but centered on the person and his/her life. Virtue ethics, an ethical theory that can be traced back to Aristotle's "Nicomachean Ethics," can provide such a theory.

The basic question for a virtue ethicist is not "How should I act?," but "How should I live?" and the corresponding answer is: "The good life." So, what does a good life in science look like? One possibility that comes easily to mind is to identify the good life with the Platonic idea of the search for "truth." In "Republic," Plato sets forth the notion of the good man as someone who, through a disciplined purification of intellect and passion, turns his attention to the idea of pure Goodness in his soul [39]. This perception of the Platonic tradition was continued in early Christianity; it is found, for example, in Augustine's view that the highest good is in God: thus, to Augustine to live the good life is to know God and be like him; and to be a philosopher accordingly means to love God. In "City of God," Augustine, inspired by Platonic philosophy, developed the idea of a choice between a political life and a life in the heavenly city, spent in pursuit of truth. Such a contemplative religious life was to be searched for in seclusion, and thus for many centuries a life in the monastery was considered the best life style for those who wanted to search for the highest good.

Now, substitute God with Truth and one obtains a secular, scientific version of the highest good in modernity, namely, to know and to love the "truth," an ideal that shines through almost all biographies of scientists written in the eighteenth, nineteenth, and early twentieth centuries. The modern counterpart to the Platonic-Augustinian search for the good is that of a life in the secular version of the heavenly city, i.e., the academy, the university, the scientific laboratory, and the research library. Accordingly, the immunologist's good life would take place in a well-equipped laboratory in a major immunology department at a leading Western university.

In Jerne's case, institutions like the Department of Standardization at the Danish State Serum Institute in Copenhagen in the late 1940s and early 1950s, the Department of Microbiology at the University of Pittsburgh in the early 1960s, and the Basel Institute for Immunology in the 1970s were secular heavenly cities. During his early years in the Netherlands, he had repeatedly expressed a longing for the sublime, a strongly felt wish to get away from the earthly and the vulgar, in order to reach higher, more abstract levels of existence. The longing for the sublime followed him throughout life, and science gave him an opportunity to realize it – looking back on the experience of working together with members of the phage group in Copenhagen in the early 1950s, he claimed that meeting the molecular biologists and their quest for the gene was like being thrown with "a catapult into higher and deeper regions of ongoing science" [40]. This fervor was contagious: some scientists remember that the experience of going to the Basel Institute under Jerne's leadership was like as pilgrimage to the Mecca of immunology – the institute was a "cathedral of thinking," says one of them [41].

The secularized Platonic-Augustinian conception of a good life devoted to science (and science only) has faded throughout the course of the twentieth century. Today's scientists may still wish to live a monastic life in the laboratory in search of the "truth," but they also search for a more hedonistic good life in science; when graduate students and colleagues are asked why they have chosen to pursue a scientific career instead of going into politics, religion, administration, or business, they give answers such as: "to do some interesting," "to meet interesting people," or "to have fun" [42]. Many of today's scientists see themselves in terms of *Homo ludens* (playing man), and it is probably no coincidence that several autobiographies of contemporary scientists, such as James D. Watson and Richard Feynman, abound with references to the playful aspects of science [43]. These and similar accounts suggest that the notion of a good life in science could be articulated in terms of some variety of hedonism (pleasure ethics). At certain times Jerne surely felt that science was a great playground – many interviewees have given witness to his almost childish passion for playing with ideas in social interaction, and Alain Bussard (otherwise best known for his work on immunological tolerance) remembers that the strongest impression he got of his English-Dutch-Danish roommate at a conference on antibodies in 1958 was a person who primarily "plays with ideas... he plays in his own mind" [44].

The opportunity for living the good life as a scientific *Homo ludens* does not mean that hedonism is restricted to Academia. Today's scientists want to go to the beach, to drink beer together, eat well, enjoy good sex – in other words, they want to combine a good life (Platonian-Augustinian, or hedonistic, or both) in the seminar room or laboratory during day hours with another good life after working hours. Jerne, who loved to play the role of the urbane bohemian, was no exception; contemporary witness reports abound with stories about the many hours spent in cafés, restaurants, and night-clubs in Copenhagen, Pittsburgh, Geneva, and Frankfurt. His favorite story from the 5 years at WHO in Geneva in the late 1950s and early 1960s was that it was "so fantastic to go to three-star restaurants" with his new fiancee:

> I remember... we had a fantastic dinner in a star restaurant, and then we drove back again, though it was 200 kilometers or so.... And we did that many times, in the evening at six we said: 'Shall we go somewhere, Lyon or somewhere', and then we drove along [45].

But pleasure is not all there is in science. It is reportedly also painful. Scientists' autobiographies do not only tell about the pleasure and joy of solving a problem, but also about the intense feelings of pain before the solution comes, the feelings of fear, anxiety, even terror during the process. Pain colors and runs through the life of the scientist, irrespective of his or her scholarly standing. As one scientist says in an interview book:

> There's an aspect of terror [in the] moments of creativity.... Being shaken out from your normal experience enhances your awareness of mortality.... It's like throwing up when you're sick [46].

Pain (or, in this particular case, rather the combination of beauty and pain which Edmund Burke and Immanuel Kant thought of as the "sublime") seems to be an integral part of science. Jerne repeatedly experienced similar periods of pain and pangs of beauty in his career. He recounted a number of depression-like periods; for example, after the

opening of the Basel Institute in the summer of 1971 he was completely at the end of his tether and called off his long-planned participation in the First International Congress of Immunology. The archive is silent about exactly what closed in on him; we only know that his old friend Max Delbrück recommended him to read Beckett's "Molloy" "for the night-side of our lives" [47], and that another old friend, Ernst Sorkin, promised to stand by him in any way at all, if needed, for, as he said, "While things may go outwardly quite splendidly, there are internal commotions in all of us, which can torture or terrorize us" [48].

At other occasions, stress and surges of creative activity went hand-in-hand. For example, Jerne reports that the somatic mutation theory of antibody diversity came to him in "a spell of creativity that lasted until the day before yesterday." As he explained to his friend Gunther Stent:

> Being aware, I followed my own behavior quite carefully; I felt that all the chores (such as farewell speeches in Frankfurt, etc.) were merely nothingness. I had the feeling that I had a good idea somewhere though I did not quite understand what it was. Fact is, that I was very nervous, stopped eating, writing, etc. until 20 July, like a log coming slowly to the surface of a lake, I knew what I wanted to understand [49].

So, a life in science is not sheer pleasure. Remember that in the most influential Hellenistic hedonistic school, i.e., Epicureanism, pleasure is defined negatively as the absence of pain. From an Epicurean point of view, if pain has such a central place in the scientific enterprise, we should rather avoid science; therefore, hedonism is presumably problematic as a candidate for the articulation of the notion of the good life in science.

Internal and external goods in science

An interesting point of departure for a virtue-ethical understanding of the good life in science has been articulated by Alasdair MacIntyre in his seminal revival of Aristotelian virtue ethics [50]. Using MacIntyre's argument about the relation between virtues and the good life, one can distinguish between at least three kinds of "goods" in science. The first kind are those goods that are external to scientific practices, such as honor, reputation, monetary rewards, etc. These goods are obviously important for most scientists. Jerne considered them so important that, in the 1960s, he turned down an offer from Harvard University, then one of the international centers of excellence in immunological research, to take up a position as Director of the Paul Ehrlich Institut in immunologically underdeveloped Frankfurt instead, simply because he found the German offer the most honorable and the best paid.

From a virtue-ethical point of view, however, external goods are uninteresting, because they do not involve the expression of virtues; on the contrary, they often collide with the development of a virtuous life in science. More important to my discussion is therefore another kind of goods in MacIntyre's typology, viz., those which are internal to scientific practices (intrinsic goods) and which are achieved by means of the expression of virtues specific to these practices, such as being an honest, courageous, and skillful experimentalist, or being an able, just, and generous professor in relation to students. The category of intrinsic goods also includes the satisfaction that goes with being virtuously absorbed in these practices. Jerne, for example, often described the immense satisfaction he felt when he tried to fit theoretical curves to experimental data; it was not just "fun" in the hedonistic sense, it was also a feeling of pride and total mastery of the world. Maybe it is this aspect

of the good life that some scientists and graduate students refer to when they say that they want "to develop themselves" or "to express themselves." French molecular biologist François Jacob says in his autobiography:

> Science meant for me the most elevating form of revolt against the incoherence of nature... taking part in the new developments that were shaping up in biology... I felt, deeply rooted in myself, the sense of being where something was happening.... The opportunity to prove what I could do [51].

The goods associated with scientific practices are not enough to characterize a good life in science, however. Not only has history taught us that scientific practices can be utilized for altogether evil purposes (Nazi medicine is a tragic example), but scientific excellence also often collides with the moral virtues of ordinary life. One of Jerne's collaborators, Dr. N.N., was known as one of the most blessed experimental molecular geneticists of the time and may thus have experienced the "Zen of genetical engineering" in his work, but he was apparently also a bad guy: thus, he reportedly threw out (and thereby destroyed) other researchers' samples from the refrigerator if he could not find space for his own. Confronted by a furious young researcher who complained about Dr. N.N.'s outrageous behavior, Jerne laconically answered, "But he's the best scientist" [52]. Internal excellence evidently trumped all other concerns.

Evidently something more is needed to characterize a good life in science. Here we might again rely on MacIntyre. To ask "What is the good for me?" is to ask how one can bring about unity in one's life. Not any sort of unity, but a "narrative unity," i.e., a unity which consists of giving accounts of one's actions in terms of one's past and future aims. More specifically, the good life is a narratively unified life spent in seeking for "the" good, i.e., "the" good which will enable us to evaluate and order other goods in relation to each other, for example, the higher-level good that enables us to evaluate the virtuousness in using PCR skillfully to obtain excellent research results in relation to the virtuousness in being a just and generous laboratory leader. "The" good can, of course, not be absolutely defined; on the contrary, the *telos* of life implicit in this definition of the good life is contingent on historically given moral traditions. Nevertheless, it is precisely through this quest for "the" good that we, pace MacIntyre, are able to develop the good life. Thus, instead of a biological *telos*, which Aristotle meant was the aim of a human life (the "finis ultimus" which Hobbes later rejected), MacIntyre introduces a cultural *telos*, a socially contextual *telos*, "a conception of "the" good which will enable us to understand the place of integrity and constancy in life" and "which will enable us to order other goods" [53] – a quest which is never given, but the result of education and self-knowledge (the late Michel Foucault would probably say "self-construction").

The idea of "the unity of virtues" is of particular importance for the articulation of the good life in science, because it opens up for a discussion of virtues in science that may qualify Lorraine Daston's notion of "the moral economy" of science. Daston identifies a set of qualities, such as precision, accuracy, impersonality, impartiality, and communicability having "an almost unbroken history in the sciences as well as in public life" up to the present [54]. But despite calling these qualities virtues ("quantifying virtues"), hers is not a virtue-theoretical understanding of virtues. These and similar qualities in scientific work are rather what MacIntyre would call "professional skills" because, according to a

virtue-ethical understanding, one cannot differentiate between one set of virtues in scientific practices as distinct from another set of virtues in ordinary life. A virtue is not a disposition that guarantees success in one particular type of situation only. As MacIntyre puts it, someone who "genuinely" possesses a virtue "can be expected to manifest it in very different types of situation" [55].

Thus, if an immunologist displays courage or honesty in her work at the bench, but not in her daily life outside the laboratory, she is not – from a virtue-ethical point of view – a courageous or honest person, and consequently, her quest for the good life as the narrative unity of life is severely hampered. As Iris Murdoch, another exponent of the renaissance for virtue ethics puts it, honesty not only seems to be "much the same virtue in a chemist as in a historian," but there is also a close similarity between "the honesty required to tear up one's theory and the honesty required to perceive the real state of one's marriage, though doubtless the latter is much more difficult." And, continues Murdoch,

> A serious scholar has great merits. But a serious scholar who is also a good man knows not only his subject but the proper place of his subject in the whole of his life [56].

This widened sphere of knowledge that Murdoch speaks of is similar to that which the ancients called *phronesis* (prudence, moral wisdom) – a wisdom which at least Aristotle thought superior to scientific knowledge. And a scholar who possesses *phronesis* is, according to Socrates (who, in *Apology* and *Crito*, is said to have equated moral wisdom, virtue, the good life, and happiness), the quintessential happy person living a good life.

Did Niels Jerne live a good life? According to a Platonic-Augustinian conception of the good, he was certainly happy, because he had more experiences of contact with the sublimity of "higher regions" of biomedicine than most scientists of his generation, and rarely experienced the drudgery of everyday, uninspiring and non-consequential experimental work. As an hedonist he was happy too: like a *puer aeternus* (eternal boy) he spent many hours at the blackboard in the laboratory, in the department's seminar room, or at meetings of international science playing with ideas and concepts [57]; and as an inveterate city-dweller he knew how to enjoy himself in the nightlife of Geneva, Pittsburgh, and Frankfurt. He also had his share of the extrinsic goods of science: good pay, beautiful home, heaps of scientific honors. No wonder that so many scientists of a younger generation admired him, not only for his scientific work (which, as I said, I will not try to evaluate here), but also for his personality.

But seen from the vantage point of "the unity of virtues," Jerne's life is more problematic as a moral exemplar for immunologists to follow. In spite of his success and apparent outer happiness, he was also haunted by unhappiness and misfortune (*kakodaimonia*). His first wife's suicide in 1945, and particularly his awareness of having provoked it, continued to haunt him with a feeling of guilt, and it is well known that he drank too much and too freely. The biography offers a full range of examples of a less than good life that makes one think of the poet's words: "The intellect of man is forced to choose / Perfection of the life, or of the work" [58].

Worst of all, he did not even rest assured about the perfection of the work. The Nobel Prize in 1984 apparently did not satisfy him, because towards the end of his life, the question of his posthumous reputation was increasingly gnawing at him [59]. In the beginning

of August 1994, his third wife, Alexandra, suddenly called me up. "Niels doesn't like that title of the book, 'What struggle to escape'," she proclaimed in her usual direct way. I asked to speak with Jerne, and a few minutes later he came to the telephone.

"What do you mean?" I asked. "Why don't you like the title of the book?"

There was a silence. "You want to call it 'What struggle to escape', right?" he replied at last, and went on, "Isn't that from a poem by Keats?" [60]

"Yes."

"And the line before it is 'What mad pursuit'?"

"Yes."

"Which is also the title of Francis Crick's autobiography?" [61]

"That's right," I answered, still not sure just what his drift was. Then it came, plainly and clearly, without the least uncertainty in his voice:

"I don't want to be second to Francis Crick" [62].

This our last conversation – a few weeks later Jerne died at his retirement home in Castillon-du-Gard in Languedoc – illustrates MacIntyre's point that the biographical genre is "neither hagiography nor saga, but tragedy" [63]. As mentioned above, Jerne was recognized as "one of the most intelligent biologists of this century," once even hailed as a "Messiah" for European immunology, and thereby he fulfilled all reasonable requirements for "the perfection of the work." Nevertheless, he felt threatened by the thought that someone else might come in as Number One in the race for fame and glory. Jerne truly wanted to live an unusual and excellent life – he hated mediocrity – but in spite of all his success, in the end he felt as if his life was unfulfilled.

I would not be surprised if this type of tragedy is more widespread in today's scientific culture than we normally wish to acknowledge. I am also convinced that this kind of life can function as a moral exemplar, too. Moral models need not be uniformly positive, like Samuel Smiles' edifying Victorian biographies [64], but can portray more complex human fates, portraits that present the reader with moral dilemmas, like John Heilbron's of Max Planck as a study in "heroic tragedy" [65]. In Tim Duff's recent reading of Plutarch's "Parallel Lives," the "moralism" of these ancient portraits is not a simple exposition of advice or injunctions to be put into effect, but a "challenging moralism," a food for reflection and "a kind of gentle exploration of the realities of human life and the moral dilemmas." The "Lives" were "designed to make the reader ask new and rather challenging moral questions" [66]. In a similar fashion, biographies of immunologists can provide food for reflection for readers in the biomedical professions who are increasingly asking questions about what is good, what is bad, and what constitutes a flourishing life in science [67].

Bibliography

1 See this volume, pp. 259-286. Arthur Silverstein expressed a similar purpose in his Preface to Silverstein AM. A history of immunology. San Diego: Academic Press; 1989.
2 See this volume, pp. 144-152.
3 Söderqvist T. Hvilken kamp for at undslippe: en biografi om immunologen og nobelpristageren Niels Kaj Jerne [What struggle to escape: a biography of the immunologist and Nobel prize winner Niels Kaj Jerne]. Copenhagen: Borgens Forlag; 1998. [English translation by David Mel Paul forthcoming].
4 Jerne NK. The natural-selection theory of antibody formation. Proc Natl Acad Sci USA 1955; 41: 849-57.

5 Jerne NK. The natural selection theory of antibody formation: ten years later. In: Cairns J, Stent GS, Watson J, Eds. Phage and the origins of molecular biology. Cold Spring Harbor: Cold Spring Harbor Laboratory of Quantitative Biology; 1966. p. 301-12 (quote from p. 303).
6 Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. Austr J Sci 1959; 20: 67-8.
7 Jerne NK, Nordin AA. Plaque formation in agar by single antibody-producing cells. Science 1963; 140: 405.
8 Richard Prigge to Jerne, 20 January 1965. In: Jerne's collection, Royal Library, Copenhagen.
9 Burnet FM. The impact on ideas of immunology. Cold Spring Harb Symp Quant Biol 1967; 32: 1-8 (quote from p. 2).
10 The young Georges Köhler who (together with César Milstein) shared the prize with Jerne in 1984 was a Basel product. The same applies to Susumu Tonegawa, who received the prize in 1987 for the discovery that the variability in the repertoire of antibodies is the result of a gene lottery involving a very small number of genes. For details about the fusion between immunology and molecular biology, see Podolsky S, Tauber AI. The generation of diversity. Cambridge, MA: Harvard University Press; 1997.
11 Moulin AM. Le dernier langage de la médecine : histoire de l'immunologie de Pasteur au sida. Paris: Presses universitaires de France; 1991 (quote from p. 339).
12 The spectrum ranges from American immunologist Melvin Cohn, who declared that the theory "mocks all and comforts none" (Cohn M. The concept of functional idiotype network for immune regulation mocks all and comforts none. Ann Immunol 1985; 137 : 64-76; quote from p. 64) to British immunologist John Humphrey, who explained that it was the result of an "unusually bold and clear mind" (Humphrey J. Serendipity in immunology. Annu Rev Immunol 1984; 2: 1-21; quote from p. 10).
13 E.g., Jerne NK, The natural-selection theory of antibody formation. Proc Natl Acad Sci USA 1955; 41: 849-57; Jerne NK, Nordin AA. Plaque formation in agar by single antibody-producing cells. Science 1963; 140: 405; and Jerne NK. Towards a network theory of the immune system. Ann Immunol 1974; 125C: 373-89.
14 Burnet FM. Changing patterns: an atypical autobiography. Melbourne: Heineman; 1968 (quote from p. 249 and p. 203); Cohn M. Foreword: clippings from one immunologist's journal. In: Langman RE. The immune system. San Diego: Academic Press; 1989. p. xiii-xlv (quote from p. xxxiv).
15 For some of these responses, See Söderqvist 1998 (note 3, above).
16 See, e.g., Schmidt-Grave H. Leichenreden und Leichenpredigten Tübinger Professoren (1550–1750): Untersuchungen zur biographischen Geschichtsschreibung in der frühen Neuzeit. Tübingen: Mohr; 1974.
17 See Paul CB. Science and immortality: the éloges of the Paris Academy of Sciences (1699-1791). Berkeley: University of California Press; 1980.
18 See further, Söderqvist T. Existential projects and existential choice in science: science biography as an edifying genre. In: Shortland M, Yeo R, Eds. Telling lives in science: essays on scientific biography. Cambridge: Cambridge University Press; 1996. p. 45-84.
19 Hankins TL. In defence of biography: the use of biography in the history of science. Hist Sci 1979; 17: 1-16 (quote from p. 5).
20 Holmes FL, Hans Krebs, vol. 1. New York: Oxford University Press; 1991 (particularly p. xv-xx). For a fine-grained study of the development of Jerne's natural selection theory of antibody formation, see Söderqvist T. Darwinian overtones: Niels K. Jerne and the origin of the selection theory of antibody formation. J Hist Biol 1994; 27: 481-529.
21 Benacerraf B. Son of the angel. Boston: privately printed; 1990 (quote from p. 104).
22 Pletsch C. On the autobiographical life of Nietzsche. In: Moraitis G, Pollock GH, Eds. Psychoanalytic studies of biography. Madison, CT: International Universities Press; 1987 (quote from p. 415).
23 Jerne to Tjek Jerne, 7 July 1943. In: Jerne's collection, Royal Library, Copenhagen.
24 For the notion of "narrative unity" in life, see MacIntyre A. After virtue. 2nd ed. Notre Dame: University of Notre Dame Press; 1984, esp. chapter 15. The idea of a "narrative unity" of the biographical subject is, of course, diametrically opposed to a post-structuralist and post-modern apprehension of the subject (see further, Söderqvist T. Existential projects. 1996 (above, note 18), but not necessarily a "realization story" and thus not in conflict with Shapin's point that personal development takes place as a "complex continual flux of transactions between individual and setting" (Shapin S. Essay review: personal development and intellectual biography: the case Robert Boyle. Br J Hist Sci 1993; 26: 335-45; quote from p. 337).
25 Skidelsky R. Only connect: biography and truth. In: Homberger E, Charmey J, Eds. The troubled face of biography. London: Macmillan; 1988. p. 1-16 (quote from p. 14).
26 Rorty R. Philosophy and the mirror of nature. Princeton: Princeton University Press; 1980 (quote from p. 360).

27 For the term "moral exemplar," see Alderman H. By virtue of virtue. In: Statman D, Ed. Virtue ethics: a critical reader. Edinburgh: Edinburgh University Press; 1997. p. 145-64.
28 Luce TJ. The Greek historians. London and New York: Routledge; 1997 (quote from p. 116).
29 For a brilliant, recent analysis of the moral purpose of Plutarch's Parallel Lives, see Duff T. Plutarch's lives: exploring virtue and vice. Oxford: Clarendon Press; 1999.
30 Vasari G. Lives of the artists. Baltimore: Penguin Books; 1988 (quotes from Part I, p. 46-7, and Part II, p. 83).
31 Dilly E, Ed. The British Plutarch; or, biographical entertainer. Being a select collection of the lives at large of the most eminent men, natives of Great Britain and Ireland; from the reign of Henry VIII to George II. both inclusive: whether distinguished as statesmen, patriots, warriors, divines, poets, philosophers. 12 vols. London: Edward Dilly; 1762. Niemeyer, C. Deutscher Plutarch enthaltend die Geschichten merkwürdiger Deutschen. 2 vols. Halle; 1822–1824. Freyherrn von Hormayr J. Oesterreichischer Plutarch oder Leben und Bildnisse aller Regenten und der berühmtesten Feldherren, Staatsmänner, Gelehrten und Künstler des österreichischen Kaiserstaates. Wien: Anton Doll; 1807–1812.
32 Thomaeus JJ. Svensk Plutarch: historisk läsning för svenska ynglingar, innehållande lefvernesbeskrifningar öfver fäderneslandets största män. Stockholm: Haeggström; 1820 (quote from p. iv).
33 Smiles S. Self-help. London: John Murray; 1883 (quote from p. 371).
34 See, e.g., Wyschogrod E. An ethics of remembering history: history, heterology, and the nameless others. Chicago: University of Chicago Press; 1998.
35 See, e.g., Palladino P. On writing the histor(ies) of modern medicine. Rethinking History 1999; 3: 271-88, and Palladino P. Icarus "flight": on the dialogue between historians and the historical actor. Rethinking History 2000; 4: 21-36.
36 Eckstein AM. Moral vision in the histories of Polybius. Berkeley: University of California Press; 1995.
37 See Duff, 1999 (note 29, above).
38 For a review of the literature, see Statman 1997 (note 27).
39 For a discussion, see, e.g., Murdoch I. Metaphysics as a guide to morals. London: Chatto and Windus; 1992.
40 Jerne to Thomas Söderqvist, 15 January 1991. In: Söderqvist's private collection.
41 Interview with Tommaso Meo, 19 June 1992.
42 From interviews with post-graduates at my own university.
43 Feynman RP (with Leighton R). Surely you're joking, Mr. Feynman: adventures of a curious character. New York: Norton; 1985. Watson JD. The double helix: a personal account of the discovery of the structure of DNA. London: Weidenfeld and Nicolson; 1968.
44 Interview with Alain Bussard, 23 June 1992.
45 Interview with Jerne, 9 February 1988.
46 Dash J. A life of one's own: three gifted women and the men they married. New York: Harper and Row; 1973 (quote from p. 318).
47 Max Delbrück to Jerne, January 1971. In: Jerne's collection, Royal Library, Copenhagen.
48 Ernst Sorkin to Jerne, 26 August 1971. In: Jerne's collection, Royal Library, Copenhagen.
49 Jerne to Gunther Stent, 8 August 1969. In: Stent's private collection.
50 MacIntyre, 1984 (note 24, above).
51 Jacob F. The statue within: an autobiography. New York: Basic Books; 1988.
52 Interview with Jerne, 11 February 1988.
53 MacIntyre, 1984 (note 24, above; quote from p. 219).
54 Daston L. The moral economy of science. Osiris 1995; 10: 3-24 (quote from p. 8-9).
55 MacIntyre, 1984 (note 24, above; quote from p. 205).
56 Murdoch I. The sovereignty of good. London: Routledge and Kegan Paul; 1970 (quote from p. 96).
57 I am grateful to Gunnar Broberg for directing my attention to the *puer aeternus* aspect of Jerne's character (see further, von Franz ML. Puer aeternus. Santa Monica,CA: Sigo Press; 1981).
58 Yeats WB. The choice [1932]. In: Allt P, Alspach RK, Eds. The variorum edition of the poems of W.B. Keats. New York: Macmillan; 1957 (quote from p. 495).
59 See also Söderqvist T. A Nobel prize is no guarantee against the anxiety of being forgotten. In: Nielsen H, Nielsen K, eds. Neighbouring Nobel: the history of thirteen Danish Nobel prizes. Aarhus: Aarhus University Press; 2001.
60 "What men or gods are these? What maidens loth? / What mad pursuit? What struggle to escape?" Keats J. Ode to a Grecian urn. In: Gittings R, Ed. The odes of Keats and their earliest known manuscripts. London: Heinemann; 1970.
61 Crick F. What mad pursuit: a personal view of scientific discovery. London: Weidenfeld and Nicolson; 1988.
62 Telephone interview with Jerne, 6 August 1994.

63 MacIntyre, 1984 (note 24, above; quote from p. 213).
64 E.g., Smiles S. Life of a Scotch naturalist: Thomas Edward, associate of the Linnean Society. London: John Murray; 1876.
65 Heilbron JL. The dilemmas of an upright man. Berkeley: University of California Press; 1986.
66 Duff, 1999 (note 29, above; quotes from p. 68-71 and p. 243).
67 I am grateful to David M Paul for permission to use passages from his English translation of my biography of Jerne (Söderqvist, 1998; note 3). The research for this paper was supported by grants from the Danish Research Councils and from the Wellcome Trust.

Author Index

Brent, Leslie B.
London, UK
Fax number: 171 609 8313

Cambrosio, Alberto
McGill University, Montreal, QC, Canada
e-mail: alberto.cambrosio@mcgill.ca

Cohen, Ed
Rutgers University, New Brunswick, NJ, USA
e-mail: edcohen@rci.rutgers.edu

Cohn, Melvin
The Salk Institute for Biological Studies, La Jolla, CA, USA
e-mail: mel_cohn@qm.salk.edu

Creager, Angela N.H.
Princeton University, Princeton, NJ, USA
e-mail: creager@princeton.edu

Crist, Eileen
Virginia Tech, Blacksburg, VA, USA
and Williams College, Williamstown, MA, USA
e-mail: ecrist@williams.edu

Frommel, Dominique
International Children's Centre, Paris, France
e-mail: arthouro@gmx.net

Jackson, Mark
University of Exeter, Exeter, UK
e-mail: m.a.jackson@exeter.ac.uk

Jacobi, Daniel
University of Avignon, Avignon, France
e-mail: daniel.jacobi@univ-avignon.fr

Keating, Peter
University of Quebec at Montreal, Montreal, QC, Canada
e-mail: r32260@er.uqam.ca

Menut, Philippe
Alfortville, France
e-mail: psmenut@club-internet.fr

Moulin, Anne-Marie
CNRS and Institut de recherche pour le développement, Paris, France
e-mail: moulin@paris.ird.fr

Potter, Michael
National Cancer Institute, Bethesda, MD, USA
Fax number: 301 402 1031

Schaffner, Kenneth F.
George Washington University, Washington DC, USA
e-mail: kfs12@attglobal.net, medhum@gwu.edu

Silverstein, Arthur M.
Johns Hopkins University, Baltimore, MD, USA
e-mail: arts@jhmi.edu

Söderqvist, Thomas
University of Copenhagen, Copenhagen, Denmark
e-mail: t.soderqvist@pubhealth.ku.dk

Stanton, Jennifer
London School of Hygiene and Tropical Medicine, London, UK
e-mail: jenny.stanton@lshtm.ac.uk

Tauber, Alfred I.
Boston University, Boston, MA, USA
e-mail: ait@bu.edu

Triadou, Patrick
Hôpital Necker, Paris, France
e-mail: patrick.triadou@nck.ap-hop-paris.fr

Key Word Index

COLLECTION MUSÉE CLAUDE-BERNARD

Already published:

La nécessité de Claude Bernard. Sous la direction de J. Michel, Université Lyon 2. Actes parus chez Méridiens-Klincksieck, 1991

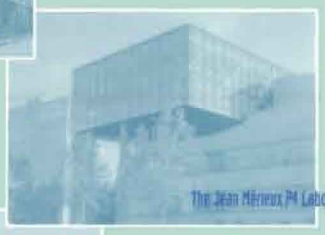
La philosophie du remède. Sous la direction de J.C. Beaune, Université Lyon III. Champ-Vallon, 1993

Les thérapeutiques : savoirs et usages. Sous la direction de O. Faure, Université Lyon 2. Fondation Mérieux, 1999

To be published:

Hygiénistes et hygiénisme : histoire et actualité. Sous la direction de P. Bourdelais, École des hautes études en sciences sociales. Belin, 2001

Headquarters of the Foundation in Lyon, France



The Jean Mérieux P4 Laboratory in Lyon, France

"Les Pensières" on the banks of Lake Annecy, France

Claude Bernard Museum, Saint-Julien-en-Beaujolais, France

ISBN : 2-84-299-210-5 — Prix TTC : 49 € [321,42 FF] — X 91