Thèse

Pour obtenir le grade de

Docteur de l'Université Montpellier II

Discipline : Biologie Moléculaire

Formation Doctorale : Parasitologie

Ecole Doctorale : Sciences Chimiques et Biologiques pour la Santé

Présentée et soutenue publiquement

Par

Rosa Milagros CORRALES

le 8 Juillet 2009

Titre de la thèse

Approche bioinformatique et fonctionnelle pour la caractérisation de protéines sécrétées chez les Trypanosomatidés: implication dans la biologie et la virulence du parasite

JURY

Pr C. BRAUN-BRETON, Université Montpellier II	Président
Pr A. VALENTIN, Université Paul Sabatier, Toulouse	Rapporteur
Pr A. MONROY-OSTRIA, Instituto Politécnico Nacional, Mexico	Rapporteur
Dr J-L LEMESRE DR1, IRD Montpellier.	Examinateur
Dr F. MATHIEU DAUDE, CR1, IRD Montpellier	Examinateur
Dr D. SERENO, CR1, IRD, Montpellier	Directeur de Thèse

ACKNOWLEDGEMENTS

First of all, I would like to thank my family. Without them, it would not have been possible to have everything ‘ready’ to start the big adventure of making a thesis in a big new world. Thanks for giving me the opportunity to have a long and open road of education in my native Salta-Argentina. Muchas gracias!!

To Denis Sereno for accepting to supervise this work. Thanks for the complete freedom to work in the laboratory.

To Françoise Mathieu-Daudé, thanks for teaching me everything in the laboratory. For your presence and support through all three years, not only on the scientific side of things but also on the personal side of life. For your kindness and your friendly sincere smile that always brought me hope.

To Frederique Breniere, thanks for your valuable help when I first arrived in Montpellier. Thanks for your help to obtain the financial support for this work.

To the members of the jury, for accepting to evaluate this thesis research work.

To Dr Miguel Angel Basombrio, thanks for the opportunity to work in the ‘Instituto de Patologia Experimental’ where I did my first scientific work with trypanosomatid parasites.

To Adriano Monte-Alegre, for your valuable help at the beginning of this work. Thanks for helping me when the French language was a perfectly strange thing in my life.

To all the members of the research unit IRD016. Thanks for all the things that I have learned through these years.

To all the ‘latino’ people that I had the pleasure to meet “on the other side” of the world. Thanks to everybody for showing me a little part of different countries: Diego (of course Dieguito!, you are latino), Harling, Andrea, Marta, Ana Maria, Mauricio, Ana and Juan (always together!), Edwin, Joana and Alvaro, Marcelina, Vicky, Julie, Leandro, Edgardo.

Thanks for all the good moments that we shared and for the pleasure of being able to speak Spanish with you all.

Finally, I would like to thank the person that gave me all the personal support in the most difficult moments of this big adventure. Thanks for your enormous and precious patience with this Latina girl. Thanks for showing me that sometimes (just sometimes) it is good to make the things in the ‘English way’. Thanks a lot Phil!!

Financial support for this study was provided by the European Union through a grant by the ‘Programme Alßan’ European Union Programme of High Level Scholarships for Latin America (No. E05D057391AR) and by the IRD DSF.

***“If we knew what it was we were doing,
it would not be called research, would it?”***

Albert Einstein

Table of contents

 3. Results

E. Chapter 3

F. General discussion and perspectives

Résumé en français

Les Trypanosomatidés constituent une famille de protozoaires flagellés, dont certains sont responsables de graves maladies humaines. Les Trypanosomatidés pathogènes pour l'homme sont: *Trypanosoma cruzi*, l'agent de la maladie de Chagas, *Trypanosoma brucei*, responsable de la Trypanosomiase Humaine Africaine (THA), communément appelée maladie du sommeil, et différentes espèces du genre *Leishmania* responsables des diverses formes de leishmanioses. Actuellement, 500 millions de personnes, principalement dans les régions tropicales et sub-tropicales du monde, sont exposées à ces maladies, et l'on estime à 20 millions le nombre de personnes infectées dans le monde (Stuart et al., 2008).

En l’absence de traitement, ces parasitoses sont souvent mortelles. Les quelques médicaments utilisés actuellement dans le traitement de ces maladies sont peu efficaces et toxiques. L’absence de vaccin et l’augmentation des cas de résistance aux rares médicaments disponibles nécessitent le développement de nouveaux moyens thérapeutiques dirigés contre ces parasitoses.

Les Trypanosomatidés font partie des eucaryotes les plus anciens qui possèdent de nombreuses particularités métaboliques et structurales. Parmi les plus remarquables, on peut noter: la présence d’une seule mitochondrie contenant le kinétoplaste, l’expression polycistronique des gènes nucléaires, et la maturation des ARN pré-messagers nucléaires par *trans*-épissage (Lukes et al., 2005).

Le cycle de vie des Trypanosomatidés alterne entre un insecte vecteur et un hôte vertébré (Figure A). Les différentes espèces de *Leishmania* sont transmises par des insectes diptères: les phlébotomes. Selon l’espèce de *Leishmania* incriminée, trois grandes formes cliniques se distinguent : les leishmanioses cutanées (LC), cutanéo-muqueuses (LCM) et viscérales (LV). Ce polymorphisme dépend, non seulement des caractéristiques biologiques des espèces, mais aussi de la réponse immunitaire de l’hôte (Lipoldova and Demant, 2006). Chez l’hôte mammifère, les macrophages sont les principales cibles auxquelles le parasite s’attache avant d’être phagocyté. Une fois internalisés dans la vacuole parasitophore, le phagolysosome, les formes promastigotes se différencient en formes prolifératives amastigotes. Après plusieurs cycles de division, les macrophages éclatent, libérant ainsi les amastigotes qui vont coloniser d’autres macrophages, ou sont ingérés par des phlébotomes, en complétant ainsi le cycle (Bates, 2007).

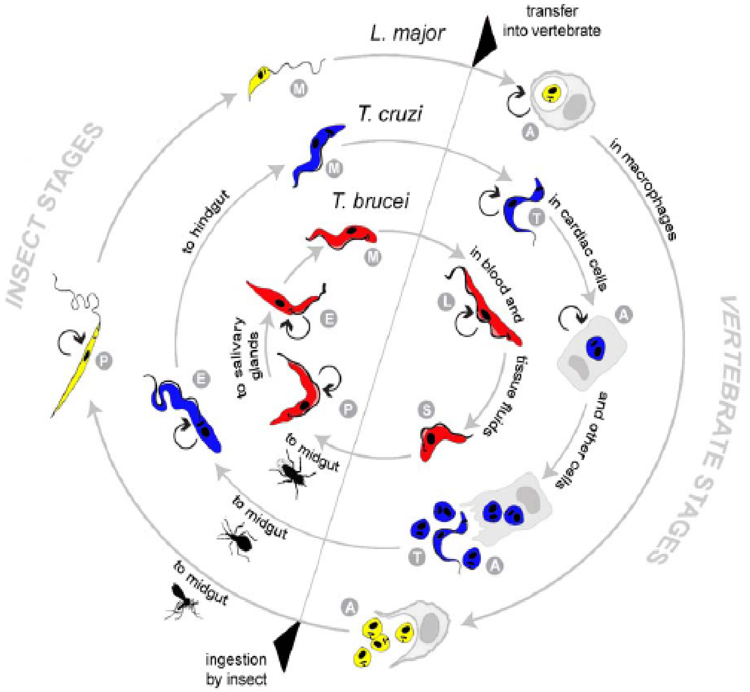


Figure A: Cycles de vie de *Leishmania spp*, *T. cruzi* et *T. brucei*. Les Trypanosomatidés sont des organismes digénétiques, présentant au cours de leur cycle de vie plusieurs formes adaptées à leurs hôtes mammifères et à l’insecte vecteur. *Leishmania* présente des formes promastigotes prolifératives (P) et promastigotes métacycliques (M) chez le phlébotome, et des formes amastigotes (A) dans les macrophages du mammifère infecté. *T. cruzi* présente des formes épimastigotes réplicatives (E) et des formes trypomastigotes métacycliques (M) dans l’insecte *Triatominae*, et des formes amastigotes intracellulaires (A) et trypomastigotes (T) chez l'hôte mammifère. *T brucei* possède exclusivement des formes extracellulaires durant tout le cycle de vie: des formes réplicatives procycliques (P) et épimastigotes (E), et des formes métacycliques (M) chez la mouche tsé-tsé, puis des formes trypomastigotes longues (L) et courtes (S *stumpy*) dans la circulation sanguine de l'hôte mammifère. D'après (El-Sayed et al., 2005).

Le cycle de vie de *T. cruzi* est caractérisé par trois formes de développement: l'épimastigote (forme de multiplication chez l'insecte vecteur), le trypomastigote (forme sanguicole invasive), et l'amastigote (forme de multiplication intracellulaire chez l'hôte vertébré). Les formes épimastigotes se multiplient dans le tube digestif de l'insecte *Triatominae,* puis se différencient en trypomastigotes métacycliques infectants dans l'ampoule rectale du vecteur. Durant l'invasion de la cellule hôte, les parasites se différencient en formes amastigotes, qui, après plusieurs cycles de multiplication, se transforment en formes trypomastigotes. Les trypomastigotes sont libérés dans la circulation sanguine et ainsi disséminés vers divers tissus et organes (cœur, foie, rate,...). Lors de leur passage dans la circulation sanguine, ils peuvent être ingérés par un triatome et réinitier le cycle (Tyler and Engman, 2001).

A la différence de *Leishmania* et *T. cruzi,* qui envahissent les cellules mammifères, *T. brucei* est exclusivement extracellulaire pendant son cycle de vie. Le trypanosome africain est transmis par la mouche tsé-tsé. Le parasite est présent dans l'intestin moyen de l'insecte sous la forme trypomastigote procyclique réplicative, et sous la forme trypomastigote métacyclique dans les glandes salivaires. Les trypanosomes sont inoculés à un mammifère par la piqûre du vecteur. Les parasites se transforment en trypomastigotes circulants dans le sang et dans le système lymphatique, et peuvent ainsi diffuser dans tout l’organisme et atteindre le système nerveux central (Simarro et al., 2008).

Les Trypanosomatidés ont développé différentes stratégies pour assurer leur survie chez leurs hôtes. Au cours de l’infection par les Trypanosomatidés, de nombreux facteurs parasitaires interviennent dans les relations hôte-parasite et dans la modulation de la réponse immune de l’hôte. Parmi ces facteurs, les protéines excrétées/sécrétées par le parasite permettent aux parasites de modifier leur environnement, d'envahir les cellules hôtes et de moduler la réponse immune de l'hôte.

Ainsi, il a été montré que les facteurs extracellulaires libérés dans le milieu de culture de *Leishmania spp*. et *T. cruzi* sont fortement immunogènes. Chez *Leishmania*, ils confèrent une certaine immunité et protection contre l'infection expérimentale (Tonui et al., 2004; Lemesre et al., 2007). Les formes trypomastigotes de *T. cruzi* libèrent plusieurs antigènes dans le surnageant des cultures cellulaires. Ce mélange complexe d'antigènes, nommé TESA (*Trypomastigote Excreted-Secreted Antigens*), est fortement immunogène, et a été utilisé pour le diagnostic de la phase aiguë et chronique de la maladie de Chagas (Nakazawa et al., 2001; Berrizbeitia et al., 2006,). Les membres de la famille des *trans*-sialidases de *T. cruzi* font partie des composants du TESA. Lors de la phase aiguë de la maladie une quantité importante de *trans*-sialidases sont libérées par le parasite, induisant une production d'anticorps spécifiques, qui peuvent être utilisés comme marqueurs de l'infection aiguë (Jazin et al., 1991; Colli, 1993).

A la différence de *T. cruzi* et *Leishmania,* pour lesquels des facteurs extracellulaires sont connus pour être immunogènes, *T. brucei* n'est pas caractérisé par la libération de molécules dans le milieu des cultures. En revanche, *T. brucei* possède un mécanisme de variation antigénique et un mécanisme efficace d'endocytose qui jouent un rôle majeur dans l'infection (Field and Carrington, 2004).

Chez les Trypanosomatidés, seulement quelques protéines sécrétées ont été entièrement caractérisées. La plupart des études portent sur les protéines de surface abondantes, et ont été réalisées chez les formes de multiplication présentes chez l'insecte vecteur, puisqu'elles représentent des formes plus facilement cultivables, *in vitro* en milieu acellulaire (McConville et al., 2002). D'autre part, les mécanismes prévalents pour l'export des protéines dans le milieu extracellulaire par les différents stades du cycle de développement de ces parasites est très mal connu.

Des protéines sécrétées par *Leishmania* jouent un rôle important en modifiant l'environnement dans l'intestin des phlébotomes, afin d'augmenter la transmission des parasites chez l'hôte mammifère (Rogers and Bates, 2007; Rogers et al., 2008). Des membres de la famille de protéines de surface gp63 et protéophosphoglycans sont libérés *in vitro* dans le milieu extracellulaire (Ilg et al., 1999; Jaffe and Dwyer, 2003). Bien que la synthèse et le traffic de ces protéines par l'intermédiaire de la voie sécrétrice classique ont été caractérisés, les mécanismes d'exportation de ces protéines dans l'espace extracellulaire restent méconnus. Chez le phlébotome, la libération extracellulaire de gp63 par des promastigotes pourrait augmenter l'hydrolyse des substrats des protéines et jouer un rôle alimentaire et/ou protecteur (McGwire et al., 2002).

Les protéines sécrétées par *T. cruzi* jouent un rôle essentiel dans l'invasion de la cellule hôte. Ainsi, une hémolysine sécrétée par la forme trypomastigote, Tc-Tox, participe à l'échappement des parasites dans le cytosol de la cellule hôte (Sibley and Andrews, 2000). Des membres de la famille de SAP (protéines riches en Sérine, Alanine, Proline) sont libérés dans le milieu extracellulaire. Certaines de ces protéines sont impliquées dans l'invasion des cellules mammifères en induisant la mobilisation du Ca^{2+}, nécessaire à l'internalisation de *T cruzi* (Baida et al., 2006).

Les protéines de sécrétion, solubles, possèdent une séquence peptidique N-terminale (*peptide signal*) qui permet l'adressage vers la "machinerie" de translocation du réticulum endoplasmique

(RE). Après un transport vésiculaire du RE vers la surface cellulaire via l'appareil de Golgi, les protéines sont libérées dans l'espace extracellulaire par fusion des vésicules sécrétrices dérivées du Golgi avec la membrane plasmique (Schatz and Dobberstein, 1996). Cette voie d'exportation des protéines de la cellule eucaryote représente la voie sécrétrice classique ou voie dépendante du RE/Golgi (Figure B).

La prédiction des protéines sécrétées à partir de la séquence primaire des protéines est un composant majeur de l'annotation automatique des protéines dans les banques de séquence. De nombreux programmes informatiques ont été développés pour prédire ce "peptide signal" dans la séquence d'acides aminés des protéines (Klee and Sosa, 2007).

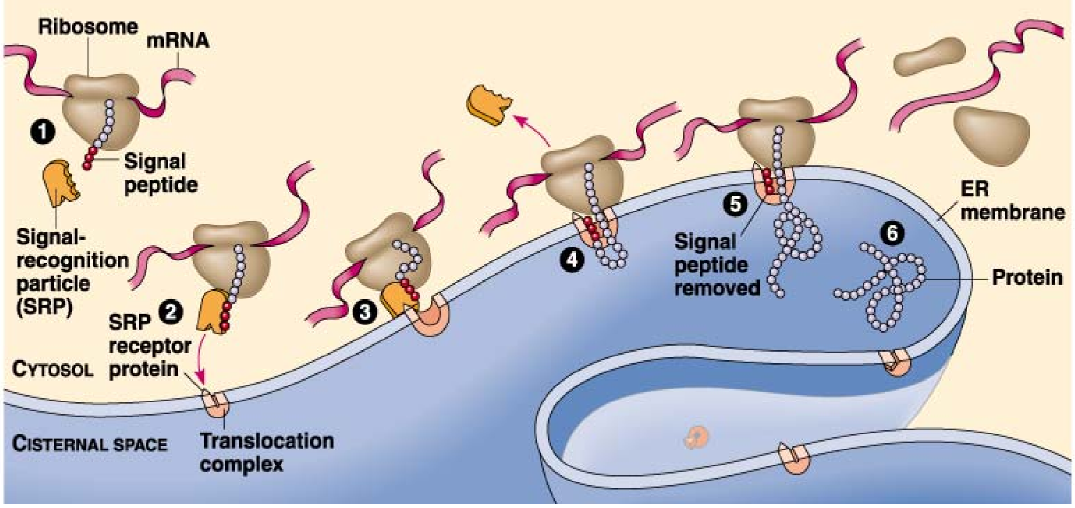


Figure B: Schéma de la voie sécrétrice classique chez les eucaryotes. Source: http://departments.oxy.edu/biology/biology/

Récemment, les séquences des génomes de *Trypanosoma brucei*, *Trypanosoma cruzi* et *Leishmania major* ont été publiés (El-Sayed et al., 2005), permettant de développer de nouvelles approches visant à élucider la biologie de ces parasites. Ainsi, nous avons focalisé notre projet sur la recherche de nouveaux facteurs de virulence des Trypanosomatidés, en particulier chez *Trypanosoma cruzi* et *Leishmania spp*. En utilisant les données issues du séquençage, nous avons recherché les protéines phylogénétiquement conservées chez les Trypanosomatidés, et potentiellement sécrétées par la voie classique. Cette conservation phylogénétique pourrait souligner un rôle important dans la biologie de ces parasites. Bien que les voies de sécrétion utilisées par les Trypanosomatidés soient encore mal connues, différentes protéines essentielles à la virulence parasitaire portent un signal peptide de sécrétion, leur permettant d'être exportées par la voie classique de sécrétion (McConville et al., 2002).

Ainsi, ce travail de thèse s'orientait sur les facteurs de sécrétion et sur la voie de sécrétion du parasite. Mon laboratoire d'origine, en Argentine, mène des collaborations avec l'institut d'accueil (Institut de Recherche pour le Développement) sur différents aspects de deux parasitoses importantes en Argentine: la maladie de Chagas et les Leishmanioses. Le but de cette thèse était d'identifier des nouvelles protéines sécrétés, conservées chez les Trypanosomatidés, représentant des facteurs de virulence ou des facteurs antigéniques, en particulier chez *Leishmania* et *T. cruzi*, endémiques en Argentine.

Résultats

Chapitre 1

Dans le premier chapitre, est décrit le développement et la validation d'une méthodologie pour l'identification des protéines sécrétées, conservées chez les Trypanosomatidés, qui permettra d'effectuer un inventaire de facteurs sécrétés conservés chez les Trypanosomatidés. Nous avons combiné une approche bioinformatique pour sélectionner les protéines possédant un signal peptidique de sécrétion, à une approche expérimentale permettant de valider la sécrétion des protéines candidates.

Plusieurs outils bioinformatiques ont été utilisés pour cribler le génome de *T. cruzi* et sélectionner 13 protéines hypothétiques, conservées chez les Trypanosomatidés, empruntant potentiellement la voie de sécrétion classique. L'expression de ces 13 gènes chez les trois stades parasitaires de *T. cruzi* à été validée par RT-PCR. Pour deux de ces gènes nous avons mis en évidence une expression différentielle au cours du cycle de vie du parasite. Nous avons mis au point une méthodologie permettant de valider expérimentalement les prédictions issues de l'analyse *in silico* du génome de *T. cruzi*. Le processus de transfection et de sélection étant beaucoup moins efficace et beaucoup plus long chez *T. cruzi* que chez *Leishmania*, nous avons utilisé le parasite *Leishmania* pour réaliser les transfections des plasmides contenant les gènes candidats, afin de développer une méthodologie rapide et simple pour l'identification des protéines sécrétées. Pour cela, nous avons fusionné les gènes candidats à une étiquette Histidine, et cloné ces gènes dans le vecteur d'expression pTEX, afin de les surexprimer chez *Leishmania*.

La tubuline a été choisie comme témoin "protéine non-sécrétée". Après sélection des parasites à l'aide de la Néomycine, nous avons recherché la présence des protéines qui possèdent une étiquette Histidine dans le milieu de culture. Pour cela, nous avons mis au point un protocole de concentration de ces protéines à partir du surnageant de culture des parasites, suivie d'une immuno-détection à l'aide d'un anticorps anti-His-tag. Cette approche nous a permis de démontrer que 25% (3/13) des protéines prédites comme potentiellement sécrétées étaient retrouvées dans le milieu extracellulaire. Nous avons également vérifié expérimentalement la sécrétion des gènes orthologues, présents chez *Leishmania*, afin de vérifier la sécrétion dans un système homologue. Nous avons également analysé la capacité de ces protéines à favoriser le parasitisme intracellulaire dans un système d'infection *in vitro*. En utilisant des parasites surexprimant les différents gènes candidats, co-transfectés par le gène codant pour la luciférase, nous avons analysé le développement des parasites lors d'infections expérimentales. Ces résultats montrent que parmi les protéines candidates, l'une d'entre elles est impliquée dans un processus favorisant la survie du parasite *Leishmania* ainsi que sa réplication à l'intérieur de la cellule hôte. En conséquence, cette protéine pourrait représenter un facteur de virulence potentiel conservé chez les Trypanosomatidés. Ces travaux méritent d'être complétés par l'étude de la capacité de cette protéine à favoriser la virulence des parasites *in vivo*, ainsi que l'étude des voies de signalisations ciblées par cette protéine.

Chapitre 2

Le deuxième chapitre porte l'étude de l'immunogénicité des protéines sécrétées identifiées grâce à la méthodologie présentée dans le chapitre 1. Les gènes codant pour les protéines sécrétées par *T. cruzi* ont été produites en système bactérien sous forme de protéines recombinantes, afin d'étudier leurs caractéristiques immunogènes au cours de l'infection par *T. cruz*i. Malgré les difficultés rencontrées dans l'expression de ces protéines cibles chez la bactérie *E. coli* (quantité faible de protéines et protéines tronquées), nous avons pu identifier parmi les trois protéines étudiées une protéine très immunogène chez l'homme. Cette protéine est reconnue par 80% (18/22) des sérums de patients chagasiques testés, provenant de différents pays endémiques d'Amérique Latine, dont le nord de l'Argentine (en collaboration avec mon laboratoire d'origine). De plus cette protéine n'est pas reconnue par des sérums de patients atteints de leishmaniose. Ainsi, cette protéine antigénique chez l'homme représente une cible potentielle pour le développement de nouveaux outils diagnostiques. L’amélioration de la production des protéines recombinantes, en utilisant un système eucaryote par exemple, permettra d’obtenir des protéines de meilleure qualité et en quantité suffisante pour étendre les tests de reconnaissance par les sérums humains, et ainsi confirmer le caractère immunogène de cette protéine par la banque de sérums de patients chagasiques disponible au laboratoire.

Chapitre 3

Chez les Trypanosomatidés, la plupart des connaissances sur la biologie des parasites, y compris les voies de sécrétions et le traffic cellulaire, ont été obtenues à partir d'études chez *T. brucei* (Field et al., 2007). La disponibilité d'outils robustes pour la manipulation génétique chez *T. brucei* ont facilité l'utilisation de ce parasite comme modèle expérimentale pour des analyses fonctionnelles chez les Trypanosomatidés (Field et al., 2007). Plusieurs études ont démontré que les caractéristiques de base de la voie de sécrétion classique sont similaires à ceux des autres eucaryotes (McConville et al., 2002). Par conséquence, comme chez la plupart des eucaryotes, la voie de sécrétion classique pourrait représenter le mécanisme principal pour l’exportation des protéines dans l’espace extracellulaire. Sur la base d'une voie sécrétrice classique active chez les Trypanosomatidés, et l’importance des protéines extracellulaires dans les interactions hôte-parasite, nous avons focalisé ce travail de thèse sur la recherche et l'identification des nouvelles protéines sécrétées par l'intermédiaire de la voie classique. Néanmoins, lors de la réalisation de cette thèse, la première étude protéomique quantitative du sécrétome de *L. donovani* a été publiée (Silverman et al., 2008). Singulièrement, cette étude montre que les promastigotes de *L. donovani* en phase stationnaire de croissance utilisent principalement des mécanismes non-classiques pour l’export des protéines, incluant probablement la libération de microvésicules comme les exososomes (Silverman et al., 2008). Ainsi, le processus de sécrétion chez *Leishmania* et probablement chez *T cruzi* apparait beaucoup plus complexe et se distingue donc des autres eucaryotes. La libération de quantité significative de facteurs antigéniques dans le milieu extracellulaire est bien connu chez *T. cruzi* et *Leishmania* (Colli, 1993; Yokoyama-Yasunaka et al., 1994; Jaffe et Dwyer, 2003). En revanche, *T. brucei* ne libèrerait que peu de molécules *in vitro*. La plupart des études portant sur *T. brucei* sont concentrées sur les protéines de surface VSG qui déterminent l’évasion du parasite à la réponse immune de l'hôte grâce à une variation antigénique remarquable et exclusive à *T. brucei* (Taylor and Rudenko, 2006). L'exocytose étant également impliquée dans l'évasion à la réponse immune, l’étude du processus d’exocytose a été moins caractérisé chez *T. brucei* (Field et al., 2007). Ainsi, l'exoprotéome de *T. brucei* est peu connu, et en conséquence chez les autres Trypanosomatidés également, dû fait de la fréquente extrapolation des résultats obtenus chez *T. bruei* aux deux autres parasites *Leishmania* et *T. cruzi*. Cependant, l'analyse récente de l'exoprotéome de *T. congolense* par une approche protéomique montre que la plupart des protéines identifiées sont apparemment sécrétées par des mécanismes non-classiques, en accord avec les résultats obtenus chez *L. donovani* (Grebaut et al., 2009). L'ensemble de ces résultats souligne le peu de connaissance sur la complexité de l'exoprotéome et des mécanismes de sécrétion utilisés par les Trypanosomatidés pour exporter des protéines vers l’espace extracellulaire.

Dans le troisième chapitre en se basant sur les récents travaux portant sur l’exoprotéome de *Leishmania,* nous présentons une revue décrivant les méthodologies actuelles permettant d'étudier l’ensemble des protéines extracellulaires chez les eucaryotes, ainsi que leurs applications chez les Trypanosomatidés. Les avantages et les limites de telles méthodologies sont discutées. La contribution relative et la régulation des différents mécanismes de sécrétion à l'origine de la constitution de l'exoprotéome de *Leishmania* sont également discutées. Nous soulignons l’importance de l’étude de la complexité du sécrétome et de l'exoprotéome des Trypanosomatidés dans la compréhension des interactions hôte-parasite chez ces pathogènes.

Conclusions et Perspectives

La confirmation expérimentale des gènes prédits par l'analyse *in silico* est une étape importante pour déterminer leurs fonctions dans les génomes des Trypanosomatidés. Ainsi, ce travail de thèse a permis l'exploration d'un grand nombre de protéines hypothétiques dans le but d’identifier de nouvelles protéines sécrétées par l’intermédiaire de la voie classique représentant des facteurs de virulence potentielle et/ou des facteurs antigéniques. En utilisant *T. cruzi* et *Leishmania* comme modèles expérimentaux, nous avons couplé une approche bioinformatique à une validation expérimentale, puis à des études de propriétés biologiques, afin d’évaluer l’importance des protéines sécrétées chez les Trypanosomatidés. Malgré les difficultés rencontrées lors de la production des protéines recombinantes, limitant les expériences visant la caractérisation fonctionnelle des protéines candidates, l'ensemble de ces résultats démontre l'utilité des analyses bioinformatiques couplées à des tests fonctionnels pour l'identification de nouvelles protéines sécrétées, représentant des facteurs de virulence ou des facteurs antigéniques potentiels des Trypanosomatidés.

L’ensemble de ces travaux a permis de développer une méthodologie pour l’identification de nouveaux facteurs de sécrétion communs aux différents membres des Trypanosomatidés. Cette méthodologie représente un nouvel outil qui permettra d’effectuer un inventaire de facteurs sécrétés conservés chez les Trypanosomatidés.

La combinaison de plusieurs outils et de nombreuses méthodologies s'avère nécessaire pour le décryptage des génomes des Trypanosomatidés. Dans ce cadre, les résultats présentés dans cette thèse représentent un point de départ pour des analyses plus exhaustives des protéines sécrétées, représentant des facteurs de virulence et/ou les facteurs antigéniques chez les Trypanosomatidés. Ces études permettront d'élargir nos connaissances sur la biologie et la pathogénie des Trypanosomatidés et donc d'identifier de nouvelles cibles thérapeutiques pour lutter contre ces parasitoses.

Enfin, l'identification d'une protéine immunogène de *T cruzi* représente une étape importante vers le diagnostic spécifique de la maladie de Chagas dans le Nord de l’Argentine, où le diagnostic spécifique de cette maladie constitue actuellement un problème majeur.

A. GENERAL INTRODUCTION

The Trypanosomatidae are a large group of parasitic protozoa, some of which cause important diseases in humans. These include; *Trypanosoma cruzi*, the agent of Chagas disease, *Trypanosoma brucei*, responsible for African trypanosomiasis or sleeping sickness, and species of *Leishmania* that cause the various forms of leishmaniasis. Half a billion people, primarily in tropical and subtropical areas of the world, are at risk of contracting these diseases. It is estimated that more than 20 million individuals are infected with the pathogens that cause them, resulting in extensive suffering and more than 100,000 deaths per year (Stuart et al., 2008).

The name *Trypanosoma* derives from the greek "trypanon", meaning "borer" and "soma", meaning body, and is suggestive of the rotational swimming movements of the parasites. The trypanosomatids are classified as protozoan euglenozoan kinetoplastids with a single flagellum. Kinetoplastids are distinguished by the presence of a DNA-containing region, known as a “kinetoplast,” in their single large mitochondrion (Lukes et al., 2005). For unicellular eukaryotes, they are complex organisms compared to many other pathogens. They undergo morphological changes during their life cycles and are transmitted by different insect vectors. Much of the cellular biology of these parasites is very similar. Nevertheless, these evolutionarily ancient eukaryotic organisms diverged 200 to 500 million years ago (Douzery et al., 2004). Their molecular peculiarities are as interesting as their common features. Understanding the differences and similarities among these disease causing pathogens at the genetic, molecular and cellular levels could provide new approaches to the development of diagnostic techniques, vaccines and drugs to combat these diseases.

The following sections describe some unique features of the cellular biology of these parasites, their life cycles, and different aspects of the disease they cause. Likewise, the background underpinning the development of this thesis is presented.

1. *Trypanosoma cruzi*

1.1 Chagas disease: Pathogenesis and diagnostic

Trypanosoma cruzi is prevalent throughout the Americas. The disease currently affects 8 to 11 million individuals, concentrated in the poorest rural and urban areas of Latin America. More than 100 million people are exposed to infection (Stuart et al., 2008). It mostly affects people whose housing provides a habitat for the *Triatominae* insects, commonly called “kissing bugs” which act as a vector for *T. cruzi*. There are many domestic and wild mammal reservoirs of *T. cruzi* including opossums, cats and dogs. Infection with *T. cruzi* can also result from blood transmission, organ transplantation, vertical transmission and by ingestion of contaminated food or drink (Coura, 2007).

Chagas disease is the result of persistent infection with *T. cruzi* and complex interactions between the pathogen and host immune response, which, in the absence of immune system dysfunctions (e.g., in individuals with AIDS), results in very low parasite numbers but rarely elimination of the infection (Tarleton, 2003). The disease has a very variable clinical presentation. Following infection, there is a short acute phase characterized by an abundant parasitemia, but frequently very mild and unspecific symptoms. An estimated 30% of individuals infected with *T. cruzi* develop clinical Chagas disease. This occurs decades after their initial infection and is often identifiable due to cardiac disease or pathological gut enlargement. Heart pathology is characterized by chronic myocarditis frequently leading to cardiomegaly, congestive heart failure and arrhythmias (Tarleton et al., 2007). In the digestive forms of the disease, dilatation of the esophagus and/or colon (megaesophagus and megacolon, respectively) may be observed in advanced stages. Although an autoimmune etiology was initially hypothesized, the continuous immune attack on persistent parasites is likely to be the primary cause of cumulative tissue damage in chronic Chagas disease (Tarleton, 2003).

The current diagnostic techniques for Chagas disease are inadequate yet crucial for identifying infected individuals, as well as monitoring treatment and other interventions (Tarleton et al., 2007). Direct detection of parasites after the acute stage of infection is very difficult due to their low numbers. Multiple serological tests using different platforms (e.g., ELISA, indirect immunofluorescence, and indirect hemagglutination) are routinely used in the attempt to obtain a consensus result (Pirard et al., 2005). However, the infection status of those positive in only one test is inconclusive (Pirard et al., 2005), and multiple tests on individuals with detectable parasites can give negative results (Salomone et al., 2003). Thus, current tests are inadequate for monitoring treatment efficacy and other interventions. New tests need to be developed to inform policy development for treatment and control strategies.

1.2 Life cycle

The life cycle of *T. cruzi* alternates between an invertebrate host and a vertebrate host. It also involves developmental stages in each host with distinct functional roles (Tyler and Engman, 2001). In the midgut and rectum of the triatomine vector, non-infective epimastigotes undergo a process of differentiation to produce mammalian-infective metacyclic trypomastigotes. Metacyclics are eliminated in the insect feces as it acquires a bloodmeal and parasites gain access to the vertebrate host through breaches in the skin or through the conjunctival mucosa by mechanical introduction. The principal role of the non-dividing trypomastigote form of *T. cruzi* is mammalian cell invasion. These slender, highly motile organisms express a number of cell recognition and signaling molecules that mediate attachment and penetration of a wide variety of cell types (Andrews, 2002). This process occurs either by a lysosome-mediated process involving early recruitment and targeted exocytosis of host cell lysosomes at the parasite attachment site, or by invagination of the host cell plasma membrane (Andrews, 2002). Both routes of entry ultimately deliver trypomastigotes to the lysosomal compartment. Once there, they reside for several hours before disruption of the vacuole occurs and parasites are released into the host cell cytoplasm (Andrews, 2002). Differentiation to amastigotes, the intracellular replicative form, is completed in the host cytoplasm where they multiply by binary fission. The amastigotes differentiate into trypomastigotes. These are released into the bloodstream and infect cells of multiple organs and tissues including the heart, gut, CNS, smooth muscle, and adipose tissue. Once there they become amastigotes again. Triatomine insects become infected when they take a bloodmeal containing the parasite from an infected human or animal, thereby completing the insect–host transmission cycle (Figure 1).

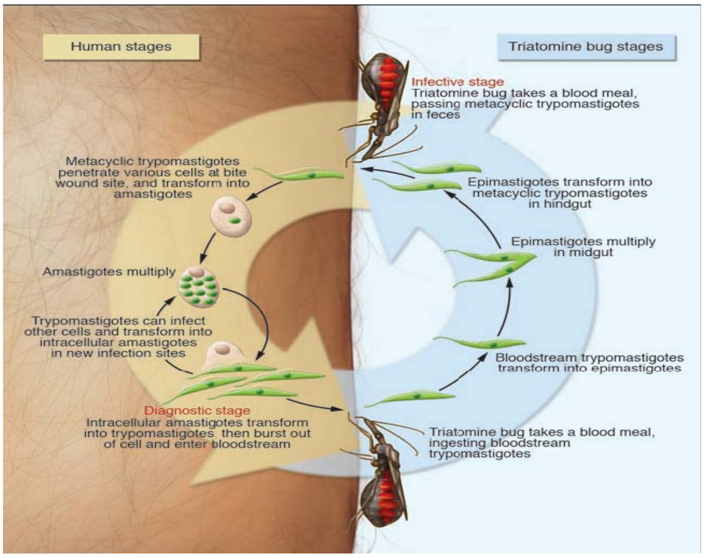


Figure 1: The life cycle of *Trypanosoma cruzi*. Reproduced from (Stuart et al., 2008).

1.3 Treatment, vector control, and the potential for vaccine design

Despite advances in understanding the biology of *T. cruzi*, only the nitrofuran derivative nifurtimox and nitroimidazole benznidazole are currently available for treatment of Chagas disease (Tarleton et al., 2007). Both are administered orally, but have long treatment courses (≥ 60 days) and side effects. In addition, there is variation in the sensitivity of the parasite to the drugs. Benznidazole works effectively against the indeterminate phase of infection in children. Observational studies have demonstrated the efficacy of benznidazole as a treatment for adults with Chagas disease who have been infected with the parasite for a long time (Viotti et al., 2006), but the lower tolerance for side effects in adults and the absence of randomized placebo-controlled studies has adversely affected the frequency of its use in adults.

The reduced transmission of Chagas disease due to intensified vector control in some regions has enhanced the focus on evaluation of drugs against indeterminate and early chronic phase disease. Insecticidal spraying of houses in the Southern Cone of South America dramatically reduced transmission of *T. cruzi* in these areas (Tarleton et al., 2007). Although highly successful at eliminating domestic *Triatoma infestans,* the spraying had limited impact on *T. infestans* in peridomestic sites and on other vector species. Furthermore, insecticide-resistant strains of *T. infestans* have emerged (Picollo et al., 2005). Advances in the understanding of the ecology and transmission of infectious disease and the use of geographic information systems (GISs), along with low-tech transmission control tools (e.g., insecticide-treated dog collars) are promising for vector control (Tarleton et al., 2007).

Substantial progress has been made in understanding the different immune responses that control infection with *T. cruzi* and in identifying the targets of these responses. Such information is crucial if effective vaccines are to be developed (Tarleton, 2007). Testing a prophylactic vaccine for an infection that is difficult to document and is rarely detected until years, or even decades, after the initial infection also presents formidable practical and ethical issues. However, given the cost-effective nature of vaccines and potential innovative applications (e.g., in transmission blocking), vaccines are included in the overall long-term strategy for the control, prevention, and treatment of *T. cruzi* infections.

2. *Leishmania spp.*

2.1 The Leishmaniases: Pathogenesis and diagnostic

Leishmaniasis is an infectious disease prevalent in Europe, Africa, Asia and the Americas, with 2 million new cases reported annually and 350 million people at risk (Stuart et al., 2008). At least 20 *Leishmania* species infect humans, and the spectrum of diseases that they cause can be categorized broadly into three types:

(i) visceral leishmaniasis (VL), the most serious form in which parasites leave the inoculation site and proliferate in liver, spleen and bone marrow, resulting in host immunosuppression and ultimately death in the absence of treatment,

(ii) cutaneous leishmaniasis (CL), in which parasites remain at the site of infection and cause localized long-term ulceration. and

(iii) mucocutaneous leishmaniasis, a chronic destruction of mucosal tissue that develops from the cutaneous disease in less than 5% of affected individuals (McMahon-Pratt and Alexander, 2004).

Host genetic variability and specific immune responses, together with the the sandfly vector and other environmental factors, are known to influence the outcome of infections (Lipoldova and Demant, 2006). However, the main factor determining clinical symptoms is thought to be the species of infecting parasite. For example, the New World parasite *L. braziliensis* is the causative agent of mucocutaneous leishmaniasis, whereas the Old World species *L. major* and *L. infantum*, present in Africa, Europe and Asia, are parasites that cause cutaneous and visceral leishmaniasis, respectively.

Leishmania parasites invade primarily mammalian macrophages by receptor-mediated endocytosis, for example, via complement receptors cleaved by parasite proteases. They multiply in the low-pH, amino acid–rich endolysosomes, to which their metabolism and nutritional demands are adapted (McConville et al., 2007). *L. donovani* and *L. infantum,* which infect cells found in lymphoid tissues, inhibit an effective immune response against themselves. By contrast, patients with CL, in whom the pathogen is largely restricted to the skin and is not found in lymphoid tissues, do not become immunosuppressed and generally develop a curative immune response (McMahon-Pratt and Alexander, 2004).

The broad clinical spectrum of the leishmaniasis makes the diagnosis difficult. Parasitological diagnosis remains the gold standard technique because of its high specificity (Herwaldt, 1999). This is typically undertaken by microscopic examination of Giemsa-stained lesion biopsy smears (CL) or lymph node, bone marrow, and spleen aspirates (VL). Importantly, the sensitivity of microscopy and culture tends to be low and can be highly variable (Herwaldt, 1999). Moreover, a skin test that discriminates between *Leishmania spp.* is not available.

Several serological approaches are commonly used in VL diagnosis, including direct agglutination and commercially available immunochromatographic dipstick tests (Chappuis et al., 2006). Serological tests are rarely used in CL diagnosis because their sensitivity can be variable and because the number of circulating antibodies against CL-causing parasites tends to be low. Their specificity can also be variable, especially in areas where cross-reacting parasites (e.g., *T. cruzi*) are prevalent. Molecular diagnosis of leishmaniasis includes PCR detection of *Leishmania spp*. DNA (Reithinger and Dujardin, 2007). PCR-based protocols have increased the speed and sensitivity of species-specific leishmaniasis diagnosis compared to conventional techniques (Reithinger and Dujardin, 2007). However, molecular approaches remain expensive and require technological expertise in areas where leishmaniasis is endemic (Reithinger and Dujardin, 2007).

2.2 Life cycle

Leishmania parasites are transmitted by the bite of female phlebotomine sand flies. Infection of the mammalian host is initiated by flagellated promastigotes, which develop within the midgut of the sandfly vector, and are deposited into the skin during a sandfly bloodmeal (Bates, 2007). Promastigotes are phagocytosed by macrophages, either directly or after infection of neutrophils initially recruited to the sandfly bite. Promastigotes target vacuolar compartments in the macrophage having the characteristics of mature phagolysosomes. Once there, they differentiate to the smaller aflagellated amastigote stage. Amastigotes proliferate by binary cell division and can spread to other macrophages as well as some other phagocytic (i.e. dendritic cells) or non-professional phagocytic cells (i.e. fibroblasts) (Figure 2) (McConville et al., 2007). The capacity of these pathogens to target and replicate within the mature phagolysosome compartment is remarkable. With the exception of the Gram-negative bacterium *Coxiella burnetti,* no other microbial pathogen resides throughout their replicative cycle in this compartment (McConville et al., 2007).

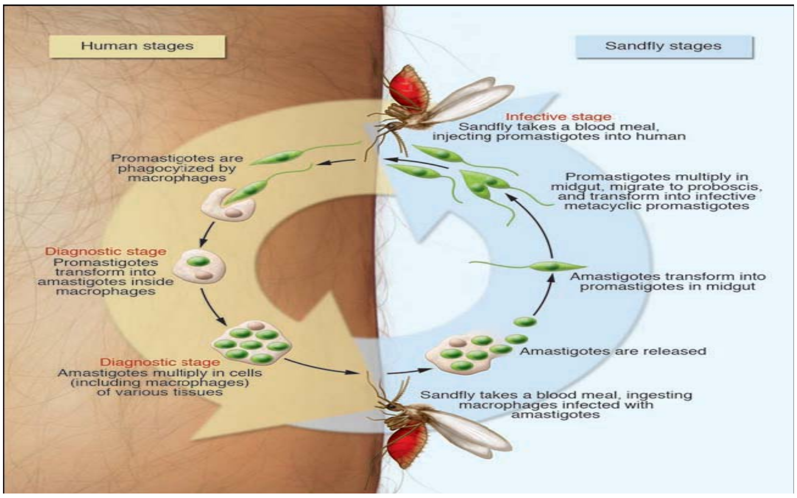


Figure 2: The life cycle of *Leishmania* spp. Reproduced from (Stuart et al., 2008).

2.3 Treatment, vector control, and the potential for vaccine design

Different measures, including residual spraying of dwellings and impregnated dog collars, are used to control phlebotomine sandflies. Given the differences in the epidemiology of VL and CL, strategies for vectorial control are different in countries from the Old World and the Americas (Alexander and Maroli, 2003). Particularly, vectorial control is difficult in the Americas due to the forest environment associated with the vector. Thus, efforts to combat the disease have been concentrated in treatment rather than prevention. The standard drugs for the treatment of both VL and CL are the pentavalent antimonials, sodium stibogluconate and meglumine antimoniate. However, the available drugs are far from satisfactory, since the treatment requires long-term administration, is expensive and highly toxic (Santos et al., 2008). The introduction of generic brands has reduced costs. Nevertheless, the development of *Leishmania* resistance mechanisms towards first-line drugs has accentuated the problem of chemotherapy (Santos et al., 2008).

A vaccine to prevent human leishmaniasis has been a goal for nearly a century and extensive evidence indicates that leishmaniasis can be controlled by vaccination (Palatnik-de-Sousa, 2008). The advances in parasite axenic culture and in the understanding of *Leishmania* pathogenesis coupled with the completion of the genome sequence of several *Leishmania* species and the capacity to perform genetic manipulations have opened new approaches feasible for *Leishmania* vaccine development (Palatnik-de-Sousa, 2008). Vaccination against leishmaniasis seems feasible, in part because the mechanisms underlying protective immunity following infection with *Leishmania spp*. are relatively well understood and because their antigens are highly conserved between species (Sacks and Noben-Trauth, 2002). Understanding of cell-mediated immunological mechanisms for controlling *Leishmania* infection has permitted the identification of several vaccine candidates (Beattie et al., 2008). Ideal antileishmanial vaccines should be safe, affordable and effective against multiple *Leishmania spp*., and effective as either a prophylactic or therapeutic. Their therapeutic potential is particularly attractive, given the problems current therapies have of toxicity, expense, and tendency toward drug resistance (Santos et al., 2008).

3. *Trypanosoma brucei*

3.1 The African trypanosomiasis: Pathogenesis and diagnostic

Human African trypanosomiasis (HAT), also known as sleeping sickness, is caused by *Trypanosoma brucei*, which are transmitted to humans by infected tsetse flies of the genus *Glossina*. Sleeping sickness, coupled with nagana, the animal form of African trypanosomiasis, has been a major obstacle to sub-Saharan African rural development and agricultural production. HAT, are usually found in remote sub-Saharan rural areas where health systems are weak or non-existent (Simarro et al., 2008). The human disease takes two forms, depending on the parasite involved. *T. b. gambiense* causes chronic infection, whereas *T. b. rhodesiense* generally causes a more acute infection. The parasites first develop in the blood, lymph and peripheral organs (stage 1). Rapid parasite growth is countered by host immune responses, but parasite antigenic variation enables immune evasion, resulting in waves of parasitemia (Taylor and Rudenko, 2006). The second stage of the disease starts when parasites cross the blood–brain barrier, invading the central nervous system, where they cause progressive neurological dysfunction, including mental, sensory, and sleep anomalies. Without treatment, the disease is fatal (Simarro et al., 2008).

Current tools for diagnosing HAT are not completely satisfactory, since those for the more dangerous second stage of the disease are invasive and insensitive, whereas those for the first stage of disease do not discriminate between the *T. brucei* subspecies. The card agglutination test for trypanosomiasis is based on a frequent variant surface glycoprotein (VSG) expressed by *T.b. gambiense* and does not require a microscope, but there is no comparable test for *T.b. rhodesiense.* Much progress has been made in the development of molecular tools. Specific genes for both *T. b. gambiense* and *T. b. rhodesiense* have been identified for PCR-based detection of infection. Molecular dipstick tests allow easier reading of the PCR result (Simarro et al., 2008). However, their use is limited due to availability and cost. Diagnosis of the disease is followed by systematic stage determination, requiring a lumbar puncture, an invasive procedure that is not well accepted by patients. As long as there is no safe and effective drug available to treat both stages of the disease, determining disease stage will remain necessary. Stage markers in other body fluids such as serum, urine, or saliva could become ideal tests too, but remain to be identified.

3.2 Life cycle

Over 20 species and subspecies of tsetse flies (*Glossina spp.)* transmit *T. b. gambiense* and *T. b. rhodesiense*, but only approximately 1 in a 1,000 flies exhibit the mature salivary gland infection necessary to transmit the pathogen to humans (Stuart et al., 2008). Trypanosomes enter the fly when it takes a meal of protozoan-containing blood from an infected human or animal. Then over a period of four weeks, they undergo morphological and physiological transformations in the alimentary tract and salivary glands, where they become infective. The reacquisition of infectivity exactly correlates with the expression of a dense coat of variant surface glycoprotein (VSG) on the surface of the parasite. During a blood meal on the mammalian host, an infected tsetse fly injects parasites into skin tissue. The parasites pass via the lymphatic system into the bloodstream, which carries them throughout the body. The parasites produce thousands of antigenic variants by alternate expression and recombination of a repertoire of approximately 1,000 VSG-encoding genes, and this enables them to evade the immune response (Taylor and Rudenko, 2006). The parasites continue to replicate by binary fission. In late-stage disease the parasites invade the CNS and reside in the cerebrospinal fluid and intercellular spaces (Figure 3).

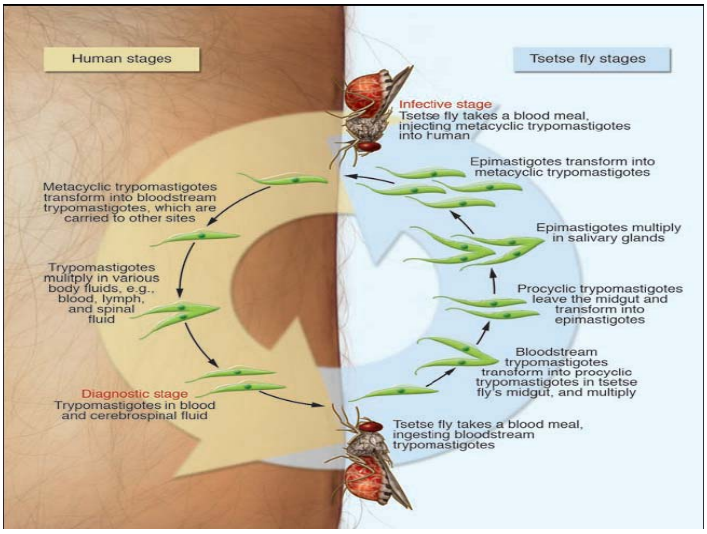


Figure 3: The life cycle of *Trypanosoma brucei*. Reproduced from (Stuart et al., 2008)

3.3 Treatment, vector control, and the potential for vaccine design

Treatment of HAT relies on four licensed compounds, all of which have drawbacks. Toxicity, the need for parenteral administration, lack of a guaranteed supply and increasing incidence of treatment failure with some drugs make the situation difficult (Barrett et al., 2007). Two compounds are used against stage 1 disease: suramin and pentamidine. Against stage 2 disease, melarsoprol (active against *T. b. gambiense* and *T. b. rhodesiense*) and eflornithine (only useful against *T. b. gambiense*) can be used. The only compound in advanced phase III clinical trials is the orally available prodrug, pafuramidine maleate (DB289). However, it is active only against stage 1 disease. New drugs are required, especially for stage 2 patients as it is generally only this cohort who presents themselves at clinics once neurological symptoms are manifest.

Current Tsetse fly control programs involve the use of insecticides through the sequential aerosol spraying technique, insecticide-treated targets, or insecticide-treated animals and the use of traps (Stuart et al., 2008). Despite the considerable progress made in controlling the vector, an ideal methodology easily accessible to the population at risk still does not exist. Development of a vaccine for HAT is highly unlikely because the process of antigenic variation, where parasites repeatedly change the surface coat to evade the immune system. Thus, drug developments are central efforts to control HAT (Barrett et al., 2007).

4. Genomic organization

The trypanosomatids posses a two unit genome: a nuclear genome and a mitochondrial genome. In contrast to many other eukaryotes, the mitochondrial genome is localized in a single mitochondrion: the kinetoplast, specific to the *Kinetoplastida* order.

4.1 Mitochondrial genome

Trypanosomatids belong to the earliest diverging branches of the eukaryotic evolutionary tree which have mitochondria. This is reflected in the organisation of their mitochondrial DNA (called the kinetoplast DNA: kDNA) that consists of a network of two classes of topologically interlocked circular DNA molecules (Lukes et al., 2005). The circular DNA molecules present different sizes, the maxi- and the mini-circles. Approximately 50 copies of the maxicircle DNA (20–40 kb in size depending on the species) are found in each organelle. Minicircles are smaller (0.65–2.5 kb) and found in 5000–10 000 copies per organelle in most trypanosomatid species (Schneider, 2001). The maxicircles are structurally and functionally analogous to the mitochondrial DNA of other organisms. The proteins encoded by the mitochondrial genome are conventional for a mitochondrial genome, their expression, however, involves a complex series of processes. Many genes represent incomplete open reading frames and their primary transcripts have to be remodeled by RNA editing to convert them into translatable mRNAs. The genetic information necessary for RNA editing is specified by short transcripts called guide (g) RNAs, which are mainly encoded by minicircle DNA. Mitochondrial translation is also unconventional. Genes for tRNAs, are completely absent in the mitochondrial genome of trypanosomatids. All tRNAs necessary for translation are imported from the cytosol. The composition of mitochondrial ribosomes is also unusual in that they contain the smallest known rRNAs (Schneider, 2001) (Figure 4).

4.2 Nuclear genome: Comparative genomics

The genome sequences of *L. major*, *T. cruzi* and *T. brucei* were published in 2005 (El-Sayed et al., 2005b). The individual and combined comparative analysis of these three genome sequences has revealed some peculiar aspects of the trypanosomatid biology (El-Sayed et al., 2005b). Despite using different insect vectors, employing diverse mechanisms of immune evasion and causing different pathologies, the genomes share several features. Each kinetoplastid genome has more than 8000 genes with more than 6000 orthologs in common. *L. major* has the fewest genes and *T. cruzi* the most (Table 2). Fewer orthologues are shared between *L. major* and *T. brucei* than for *L. major* and *T. cruzi* (Table 1), perhaps reflecting the common intracellular environment of their mammalian stages (El-Sayed et al., 2005b).

The genes exist in large syntenic blocks that contain 80% of the *T. brucei* and 93% of the *L. major* genes (El-Sayed et al., 2005b). This represents a high degree of conservation for organisms having diverged 200–500 million years ago (Douzery et al., 2004). This organization is consistent with a reliance on post-transcriptional gene regulation (Haile and Papadopoulou, 2007). The genomic organization of these parasites primarily differs in the arrangement of the syntenic clusters, witch results in different diploid chromosome numbers and size (Table 1).

Diversification of gene families is prominent in the *T. cruzi* and *T. brucei* genomes, in which 10-18% of the genomes comprise multigene families of surface proteins (El-Sayed et al., 2005b). These families are represented by the variant surface glycoproteins (VSGs) and Procyclic Acidic Repetitive Proteins (EP/PARP/procyclin) of *T. brucei (*Berriman et al.*, 2005)*, the *trans*-sialidases, mucins, and mucin-associated surface proteins (MASPs) of *T. cruzi* (El-Sayed et al., 2005a). Large gene families in the *L. major* genome are represented by the amastin surface glycoproteins and the promastigote surface antigens (PSA-2). The PSA-2 proteins function in macrophage binding and show structural similarity, but not sequence identity with the *T. cruzi* mucins; there are no closely related orthologs in *T. brucei* (Ivens et al., 2005).

Comparative genomics of the kinetoplastids have also revealed pathways that might be specific adaptations to environments associated with life in their respective insect and vertebrate hosts. For example, only *L. major* seems capable of hydrolyzing disaccharides, possibly because the sandfly vectors of *L. major* feed on nectar and honeydew. Likewise, only *T. cruzi* has the potential to convert histidine to glutamate, reflecting the abundance of histidine in the excreta and hemolymph of its vector. Similarly, only *T. cruzi* lacks *de novo* pathways for polyamine biosynthesis, since polyamines can be obtained from the vector's excreta or from the host cytoplasm (Atwood et al., 2005).

Remarkably, predicted functions have been ascribed to only 40% of the protein-coding genes of the trypanosomatid genomes, but only 5% of the proteins have been confirmed experimentally (El-Sayed et al., 2005b). Most of the remaining genes encode conserved hypothetical proteins, of which slightly more than half are found only in trypanosomatids.

About 2-3% of the Tritryp proteins are related to those found in prokaryotes. Some of these proteins appear to have arisen from horizontal gene transfer, and may represent excellent candidates for drug targets.

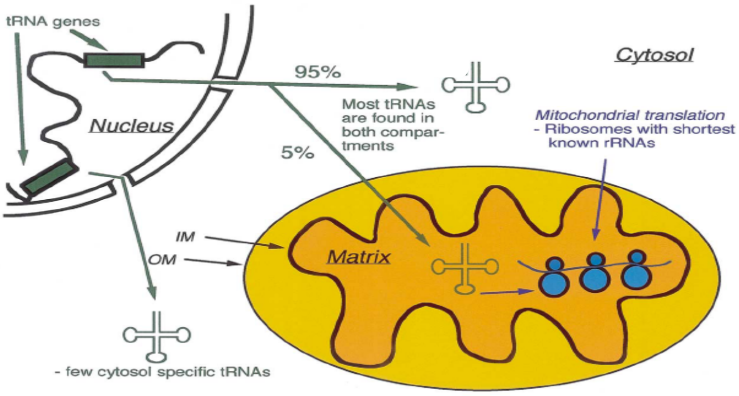


Figure 4: Schematic overview of mitochondrial tRNA import and organellar translation in trypanosomatids. OM: outer membrane; IM: inner membrane. Reproduced from (Schneider, 2001).

	T. brucei	*T. cruzi*	*L. major*
Haploid genome size (Mbp)	25[a]	55	33
N° of chromosomes (per haploide genome)	11[a]	36[b]	28
Protein-coding genes (per haploid genome)	8599[c]	~12,000[c]	8302[c]
Function known	5%	ND[d]	4%
Function inferred	38%	43%	28%
Hypothetical conserved	51%	48%	56%
Orthologues in all Tritryps	73%	54%	80%
Tb+Tc only	5%	4%	-
Tb+Lm only	1%		1%
Tc+Lm only		4%	6%
Species-specific	21%	38%	13%

Table 1: The Tritryp genome statistics: [a] Excludes ~ 100 mini and intermediate-sized chromosomes (totaling ~10Mb). [b] The exact number is not known and homologs can differ substantially in size. [c] Excludes pseudogenes. [d] ND: Not determined

5. Control of gene expression in trypanosomatids

Mechanisms controlling gene expression in trypanosomatids depend on several layers of regulation, with most regulatory pathways acting at a post-transcriptional level. Consequently, these parasites can follow the rapid changes associated with transitions between insect vectors and vertebrate hosts.

The open reading frames of trypanosomatids are organized as large bidirectional polycistronic units. Transcription has been postulated to initiate at strand switch regions on each chromosome in the absence of defined RNA pol II promoters and typical general transcription factors. The maturation of individual mRNAs from polycistronic pre-mRNAs involves two coupled co-transcriptional RNA processing reactions: trans-splicing where a small capped RNA of 39 to 41 nucleotides, the spliced leader RNA (SL-RNA), is added to the 5'-end of the mRNAs, and the 3'-end cleavage/polyadenylation (Figure 4). Polycistronic transcription of up to several hundred genes and a lack of RNA pol-II promoters for protein-coding genes mean that most gene regulation in trypanosomatids occurs post-transcriptionally (Clayton and Shapira, 2007). Indeed, there is currently no evidence for differential regulation of RNA pol II transcription of individual protein-coding genes (Haile and Papadopoulou, 2007). The only genes whose transcription is known to be developmentally regulated are the two major trypanosome surface proteins (EP procyclins and VSGs) from *T. brucei*. These genes are transcribed by RNA polymerase I from dedicated promoters and the transcription rate is developmentally regulated, most likely through chromatin mediated silencing (Glover and Horn, 2006) .

Temperature is a signal shared among trypanosomatids for kinetoplastid differentiation and has been shown to regulate mRNA stability or translation. For example, cold shock increases EP procyclin mRNA stability in bloodstream trypanosomes (Engstler and Boshart, 2004), and heat shock induces translation regulation of amastin mRNAs in *Leishmania* (McNicoll et al., 2005). pH stress silences SL-RNA in *T. brucei* (Lustig et al., 2007) and acidic pH induces the accumulation of the amastigote-specific A2 and amastin transcripts (McNicoll et al., 2005).

Recent studies have revealed the role of extinct retroposons in post-transcriptional regulation of gene expression in *Leishmania* (Bringaud et al., 2007). Examples of retroposon-mediated regulation at post-transcriptional levels are rare in eukaryotes and this type of regulation is unique to *Leishmania* among trypanosomatids.

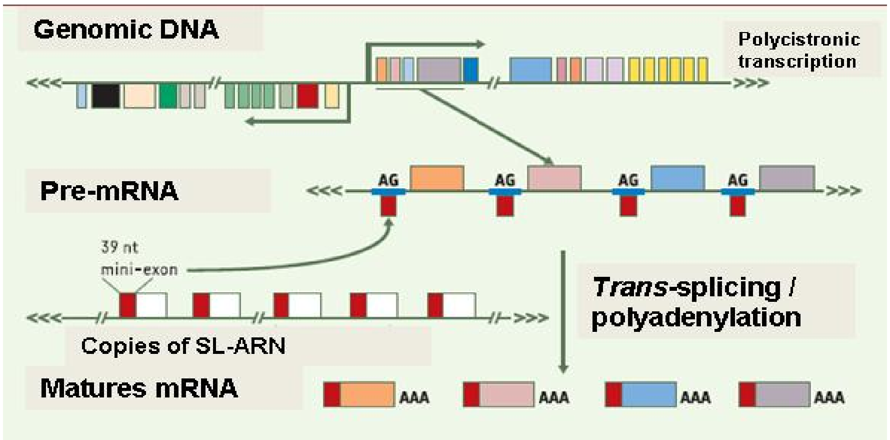


Figure 5: Schematic representation of maturation of individual mRNAs in trypanosomatids.

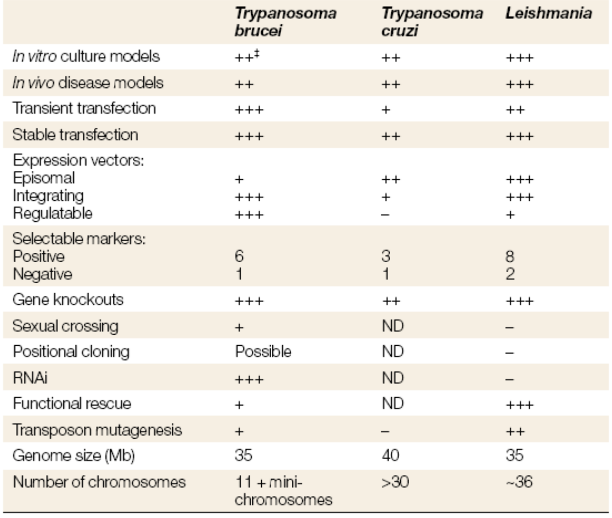
6. Genetic manipulation in trypanosomatids

Several genetic tools have been developed allowing manipulation of trypanosomatid genomes (Table 2). Most of the knowledge about the mechanisms of gene expression in trypanosomatids resulted from the development of DNA transfection methodology (Beverley, 2003). Since then, it has become possible to create and complement mutants, to over-express foreign proteins in the parasites, and to knock out genes. Parasites can be transfected both transiently and stably, using a range of episomal and integrating expression constructs (Table 2) (Beverley, 2003). To date, transgenic parasites have been used to explore mainly parasite gene function and the generation of attenuated *Leishmania* parasites suitable for vaccination studies (Beattie et al., 2008).

A major improvement that facilitated control of genetic manipulation in trypanosomatids was the development of tetracycline-inducible gene expression (Wirtz et al., 1999). This system was originally developed in *T. brucei* and successfully introduced into *T. cruzi* and *L. donovani* (Yan et al., 2002; Taylor and Kelly, 2006). This methodology represents a valuable tool for dissecting the functions of essential genes and for the expression of toxic gene products. Nevertheless, background expression levels and low induction efficiencies are still the main limitation of the tetracycline-inducible systems.

Another powerful tool for genetic manipulation of trypanosomatids is RNA interference (RNAi). The discovery of RNAi as a mechanism of gene silencing, and its successful application in *T. brucei*, has provided a robust tool to assess the functional importance of a particular gene in the development or survival of an organism (Ngo et al., 1998). Briefly, in RNAi, double-stranded RNA (dsRNA) homologous to the coding sequence of a target gene is processed into 20-24 nucleotide long RNAs that work as active guides for mRNA degradation. In *T. brucei* several aspects of parasite biology have been studied using the RNAi approach, including cell cycle and RNA editing process of trypanosomatids (Motyka and Englund, 2004). Despite the immense success of the RNAi technique in *T. brucei*, and *T. congolense* (Inoue et al., 2002), this methodology is not applicable to *T. cruzi* and *L .major*. Genome sequences of *T. cruzi* and *L. major* have revealed the absence of genes involved in RNAi machinery in both genomes. Interestingly, analyses of *L. infantum* and *L braziliensis* genomes identified a putative RNAi pathway only in the *L. braziliensis* genome (Peacock et al., 2007). This finding increases the likelihood of manipulating gene expression by RNAi in *Leishmania*, representing a potential complement to the classical “two-step gene knockout” strategy for disruption of *Leishmania* gene function.

Recently, a high-throughput method for gene function gene discovery using RNAi in *T. brucei* has been developed (Subramaniam et al., 2006). The authors propose that this technique could be used to explore the full genome of *T. brucei*. Moreover, because of the extensive gene conservation among the trypanosomatids, this approach could be extended to advance the annotation of the *T. cruzi* and *L. major* genomes, even though RNAi is not applicable to these species. However, RNAi only down-regulates and does not abolish gene function. Only the full ablation of the diploid copies of determined Open Reading Frames can assure the complete disruption of gene function in trypanosomatids. Further development of genetic tools together with the recent availability of the main pathogenic trypanosomatids will facilitate in depth analysis of the protein-coding genes in these parasites.



	Trypanosoma brucei	***Trypanosoma cruzi***	***Leishmania***
In vitro culture models	++[‡]	++	+++
In vivo disease models	++	++	+++
Transient transfection	+++	+	++
Stable transfection	+++	++	+++
Expression vectors:			
Episomal	+	++	+++
Integrating	+++	+	+++
Regulatable	+++	–	+
Selectable markers:			
Positive	6	3	8
Negative	1	1	2
Gene knockouts	+++	++	+++
Sexual crossing	+	ND	–
Positional cloning	Possible	ND	–
RNAi	+++	ND	–
Functional rescue	+	ND	+++
Transposon mutagenesis	+	–	++
Genome size (Mb)	35	40	35
Number of chromosomes	11 + mini-chromosomes	>30	~36

Table 2: Summary of the genetics tools in trypanosomatids. The "+" symbol indicates that a given method has been established, the number of + symbols reflects an approximate assessment of how well the method works and/or the extent of its use in a given species. ND: Not determined. Reproduced from (Beverley, 2003).

7. Flagellum and flagellar pocket

Two distinctive features in the cell biology of trypanosomatids are the presence of the flagellum and the existence of a prominent invagination of the plasma membrane called the flagellar pocket (Landfear and Ignatushchenko, 2001). The flagellum is a multifunctional organelle that plays critical roles in motility, chemotaxis, cell signaling and attachment of the parasites to the endothelium of their insect hosts (Landfear and Ignatushchenko, 2001). The flagellum emerges from the parasite body trough the flagellar pocket (Figure 6). The flagellar pocket is responsible for uptake of larger nutrients via receptor-mediated endocytosis, for secretion of proteins into the extracellular medium, and for integration of membrane proteins into the cell surface. The flagellar pocket corresponds to a specialized region where most of the endocytic and exocytic activities take place in the trypanosomatids (Field et al., 2007). It is thought that this feature is due to the lack of microtubules attached to this membrane, while the rest of the cell body is supported by a corset of subpellicular microtubules (Field et al., 2007).

Molecular markers have revealed that exocytic and endocytic systems are highly polarized in trypanosomatids and are contained within the posterior region of the cell (Field et al., 2007). The polarized organelles include the Golgi complex, endosomes and lysosomal apparatus, but not the endoplasmic reticulum (ER) which is distributed throughout the cytoplasm (Field and Carrington, 2004) (Figure 3). This polarization allows the remainder of the cell surface to be shielded from the environment in an appropriate manner. Since these parasites exist in multiple environments, they must rapidly and completely change the bulk of their cell surface components in response to a change of host.

Nevertheless, the molecular mechanisms that restrict endocytosis and exocytosis to the flagellar pocket are not known.

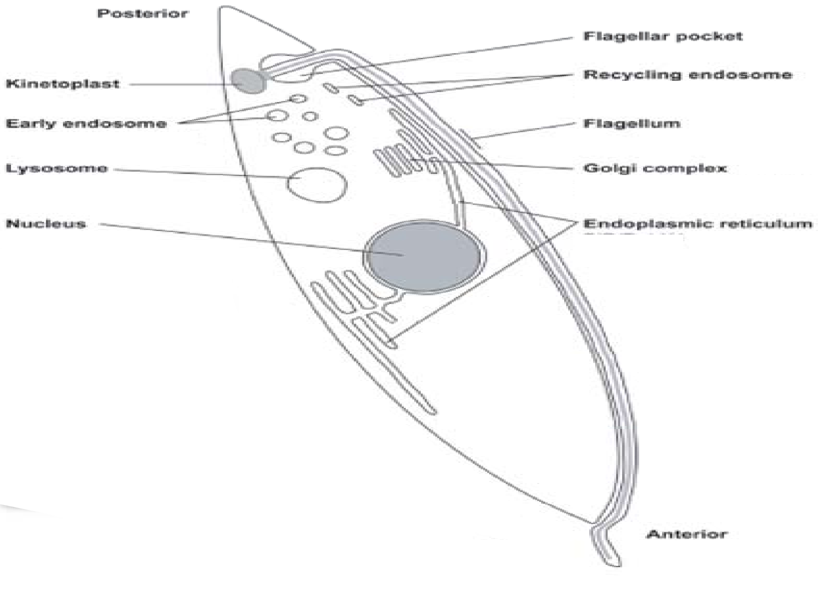


Figure 6: Schematic of a trypanosome cell highlighting the positions of various intracellular organelles. Reproduced from (Field and Carrington, 2004).

8. Surface coats of trypanosomatids

Trypanosomatids have surfaces dominated by glycosylphosphatidylinositol (GPI)-anchored antigens (Figure 7). The surface of the *T. brucei* bloodstream stage is dominated by the variant surface glycoprotein (VSG). By a process of antigenic variation, the VSG coat allows the parasite to persist for long periods in the host bloodstream (Barry et al., 2005). VSG is replaced by procyclin in the major insect (procyclic) stages of *T. brucei* (Haenni et al., 2006). The procyclin coat is thought to protect this stage from hydrolases in the tsetse fly midgut and to be required for maturation to the metacyclic stage (Vassella et al., 2000).

In *T. cruzi,* the major surface antigens are multi-gene families: the mucin-like glycoproteins and the *trans*-sialidases (TS). A primary function of some *trans*-sialidases is to transfer sialic acid from host glycoproteins to the mucin O-linked glycans (Buschiazzo et al., 2002). All developmental stages of *T. cruzi* synthesize a highly abundant class of GPI-based lipid and lipophosphopeptidoglycan (LPPG) that form a densely packed glycocalyx beneath the mucin coat (Lederkremer and Bertello, 2001).

Leishmania present a surface dominated by a family of Lipophosphoglycan (LPG) in the promastigote form and a surface protease (gp63) in both promastigote and amastigote forms. These proteins have been involved in several functions including evasion of immune response and parasite virulence in the insect and the mammalian host (Santarem et al., 2007). Other components are GPI lipids and glucoinositol phospholipid (GIPLs) (Ferguson, 1999).

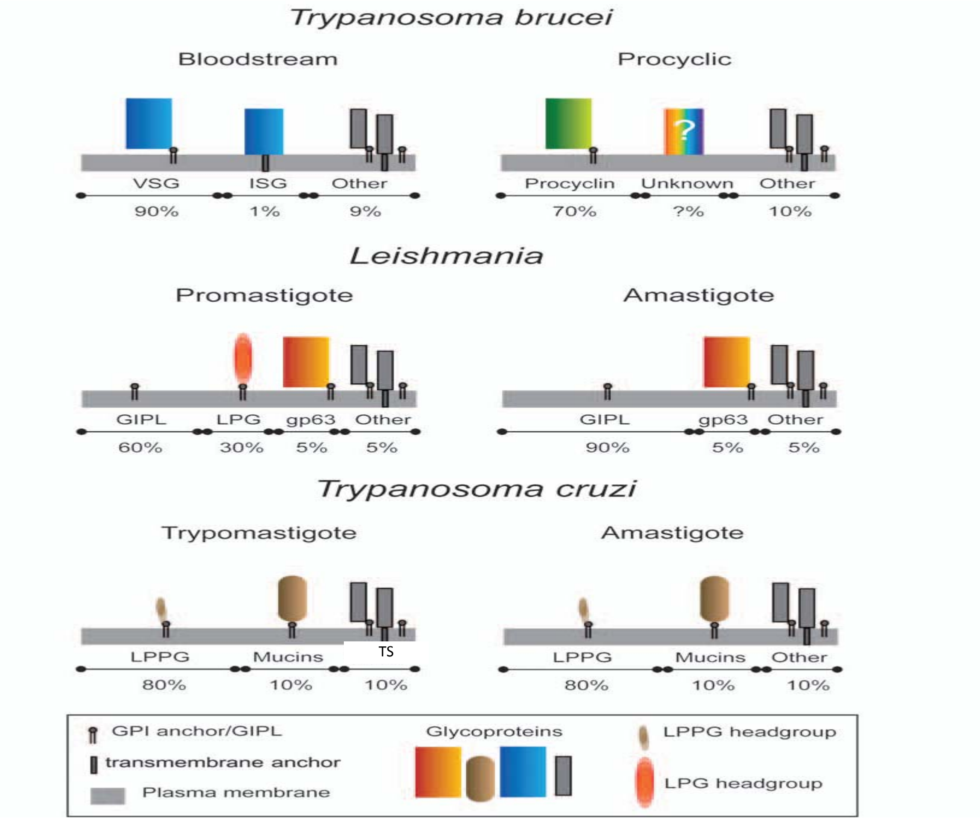


Figure 7: Schematic representation of surface coats of trypanosomatids. Reproduced from (Field et al., 2007).

9. Secretory Pathways

9.1 Classical Secretory Pathway

In Eukaryotes, soluble secretory proteins typically contain N-terminal signal peptides that direct them to the translocation apparatus of the endoplasmic reticulum (ER) (Schatz and Dobberstein, 1996). Following translocation to the ER, the signal peptide is usually cleaved from the mature protein by a signal peptidase. This pathway of protein export from eukaryotic cells is known as the classical or ER/Golgi dependent secretory pathway (Rothman and Orci, 1992). The secretory pathway is primarily responsible for distribution of the newly synthesized products among the endomembrane compartments, as well as delivery to the exterior of the cells (van Vliet et al., 2003). Proteins that are processed in this pathway are commonly referred to as classically secreted proteins. These proteins are by default transported through the Golgi apparatus and exported by secretory vesicles. The presence of secondary signals directs their final localization inside the cell. Some proteins have specific retention signals that hold them back in the ER or the Golgi or divert them to the lysosomes. In general, these retention signals are poorly characterized, one exception being the C-terminal KDEL and HNEL motifs for ER lumen retention (Nilsson and Warren, 1994). Extracellular proteins are released into the extracellular space by fusion of Golgi-derived secretory vesicles with the plasma membrane.

Several studies in trypanosomatids demonstrated that the basic features of the secretory pathway are very similar to other eukaryotes, despite the fact that these organisms represent one of the most divergent eukaryotic lineages (Clayton et al., 1995; McConville et al., 2002). However, the secretory pathway has unusual features in trypanosomatids, such as, the polarized delivery of secretory material to the flagellar pocket (Field et al., 2007). The main features of the molecular machinery of the secretory pathway in trypanosomatids have been studied in *T. brucei*. Trypanosomatids show a well developed ER and Golgi complex system with the formation of coated and uncoated vesicles which subsequently migrate toward the flagellar pocket (Field et al., 2007). The basic feature of the presence of canonical N-terminal signal peptide has been demonstrated for several proteins. In addition, mutational studies of a mammalian signal sequence in *L. major* suggest that the signal peptide sequence can function in a similar manner in *Leishmania* as it does in yeasts and higher eukaryotes (Tobin and Wirth, 1993).

Although some proteins in trypanosomatids have been implicated in the classical secretory pathway, is still unknown the identity of all the proteins involved in this pathway.

Likewise, little is known about the precise nature of vesicles that mediate transport of secreted and membrane proteins between the ER, the Golgi complex and the flagellar pocket.

9.2 Non-classical Secretion

Although prevalent, the classical ER-Golgi dependent pathway of protein export is not exclusive. An increasing number of secreted proteins devoid of a signal peptide have been reported to be exported without the help of ER-Golgi (Nickel and Seedorf, 2008). The main features of non-classical secretion are; (a) the lack of conventional signal peptides, (b) the exclusion of these proteins from classical secretory organelles, such as, the ER and the Golgi, combined with the lack of ER/Golgi-dependent post-translational modifications, (c) resistance of these export processes to brefeldin A, a classical inhibitor of ER/Golgidependent protein secretion (Nickel and Seedorf, 2008).

Non-classical secretion is not limited to a specific family of proteins. Indeed, proteins belonging to diverse functional groups were demonstrated to be released independently of ER-Golgi. Many of the known unconventional secretory proteins are cytokines, growth factors, or other molecules with important signalling roles in physiological processes, such as, inflammation, angiogenesis, cell differentiation, or proliferation (Prudovsky et al., 2008).

Diverse mechanisms have been proposed to explain unconventional secretory processes, including lysosomal secretion, plasma membrane shedding, release in exosomes, as well as secretion through transporters that reside in the plasma membrane (Nickel and Seedorf, 2008). However, their extracellular functions and/or molecular details of their export mechanisms are poorly understood.

In trypanosomatids, an example of an unconventional secretory protein is the hydrophilic acylated surface protein B (HASPB) from *L. major* (Denny et al., 2000). This protein is associated with the outer leaflet of the plasma membrane only in the infectious stages of the parasite lifecycle (Flinn et al., 1994). The protein is synthesized on free ribosomes in the cytoplasm and becomes both myristoylated and palmitoylated at its N-terminus, which is the molecular basis of how HASPB is anchored in the membrane (Denny et al., 2000). A model has been proposed in which HASPB is transferred from the cytoplasm to the outer leaflet of the Golgi membrane, from where it is transported to the plasma membrane via conventional vesicular transport (Denny et al., 2000). This process would insert HASPB into the inner leaflet of the plasma membrane. At present it is completely unclear how HASPB is then translocated across the membrane. Interestingly, HASPB has been shown to be secreted from mammalian cells, demonstrating a conserved pathway between lower and higher eukaryotes (Stegmayer et al., 2005).

9.3 *In-silico* prediction of target signals

In 1999, the Nobel Prize in Physiology or Medicine was awarded to Gunther Blobel for, "the discovery that proteins have intrinsic signals that govern their transport and localization in the cell". As the sub-cellular localization of a protein is an important clue to its function, the characterization and prediction of these intrinsic signals ("zip-codes" of proteins) has become a major task in bioinformatics (Emanuelsson et al., 2007).

There are many different classes of targeting signals. One of the best know protein "zip code" is the secretory signal peptide (SP), which is found in all three domains of life, indicating its ancient universal origin. Signal Peptides function like a postal address label on an envelope by targeting the proteins for translocation across the plasma membrane in prokaryotes and across the ER membrane in eukaryotes (Schatz and Dobberstein, 1996). SP is an N-terminal peptide, typically 15-30 amino acids along, which usually is cleaved off by a signal peptidase during translocation of the protein across the membrane. There is no simple consensus sequence for SPs, but they typically consist of the following three domains: i) a positively charged n-region, ii) a hydrophobic h-region of at least six residues, and iii) an uncharged but polar c-region with some conservation at the -3 and -1 position relative to the cleavage site (Klee and Sosa, 2007). The semi-conserved nature of the N-terminal signal peptide makes it an attractive target for computational algorithms (Klee and Sosa, 2007). Several prediction tools have been developed to predict the presence of signal peptides in amino acid sequences. One of the most accurate predictors is the programme SignalPv3. This programme incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks and hidden Markov models. SignalPv3 produces both classification and cleavage site assignment, while most other methods classify proteins as secretory or non-secretory (Bendtsen et al., 2004b). Although, SP prediction tools can provide a potential list of secreted proteins, these programmes cannot discriminate between extracellular proteins and other proteins that are processed through the ER but not released in the extracellular space (e.g. ER and Golgi resident proteins and lysosomal proteins).

Although several programmes are available to detect secretory SP in protein sequences, the ability to predict non-classical secreted proteins is very limited. SecretomeP is currently the only programme for the prediction of mammalian secretory proteins targeted to the non-classical secretory pathway (Bendtsen et al., 2004a) . Nevertheless, given that mechanisms involved in non-classical secretion are poorly characterized, *in-silico* prediction of unconventional secreted proteins remains a main challenge in bioinformatics.

9.4 Importance of extracellular proteins in Trypanosomatids

The different morphological forms of trypanosomatids represent adaptations to the changing environmental conditions encountered by the parasites within their mammalian and insect hosts. These morphological forms require various effectors that are membrane bound or released extracellularly to ensure their survival and transmission in the different hosts.

It is well known that extracellular components produced by *Leishmania* parasites are highly immunogenic and elicit strong immunity and protection against infection in mice and dogs (Tonui et al., 2004; Lemesre et al., 2007).

In *T. cruzi*, trypomastigote forms release several antigens into the supernatant of infected cell cultures. This complex mixture of antigens, termed TESA (trypomastigote excretory-secretory antigens), is highly immunogenic and has been used for the diagnosis of both acute and chronic Chagas disease (Nakazawa et al., 2001; Berrizbeitia et al., 2006). Members of *trans*-sialidase family are components of the TESA mixture. During the early stages of infection by *T. cruzi* an intense shedding of *trans*-sialidases molecules (also called SAPA: shed acute phase antigens), elicits the production of specific antibodies that can be used as markers of acute infection (Jazin et al., 1991; Colli, 1993). The SAPA responsiveness slowly decreases during the progress of the disease (Jazin et al., 1991). The shift of immune response directed to TESA has been largely analyzed during the course of *T. cruzi* infections in humans. Nevertheless, given the complexity of the TESA mixture, all the components are not currently identified, being mainly characterized by its protease and glycosyltransferase activity (Campetella et al., 1990; Yokoyama-Yasunaka et al., 1994).

The complex mixture of extracellular factors released by *T. cruzi* and *Leishmania spp.* into culture supernatants are known to be highly immunogenic. In contrast, this feature is not seen in *T. brucei*, displaying both antigenic variation and an efficient mechanism for the endocytosis of immune effectors (Field and Carrington, 2004). Antigenic variation in *T. brucei* is determined by the variant surface glycoprotein (VSG) present in the bloodstream forms. The VSG forms a dense coat on the surface of parasites, its antigenic specificity continuously changes and allows these parasites to produce sustained infections (Taylor and Rudenko, 2006). Each VSG is attached to the plasma membrane by a GPI anchor that facilitates the dense packing of VSG dimers into coat arrays. The synthesis and secretion of the VSG is well documented in *T. brucei*. Indeed, studies on the VSG family have played an important role in establishing the general structure of GPI membrane anchors and in delineating the pathway of GPI biosynthesis (Ferguson, 1999).

In trypanosomatids, only a few secreted proteins have been fully characterized. Most of the studies are focused on the surface protein families, especially in the insect vector forms and the VSG of *T. brucei* (McConville et al., 2002). Furthermore, little is known about the mechanism(s) trypanosomatids predominantly used to export proteins extracellularly during their life cycles.

Secreted materials by *Leishmania* play pivotal roles in modifying the environment in the sand fly gut to enhance parasite transmission (Rogers and Bates, 2007; Rogers et al., 2008).

Members of the gp63 and PPG (proteophosphoglycan) protein families are released *in vitro* to the extracellular milieu (Ilg et al., 1994; Jaffe and Dwyer, 2003). Although, the synthesis and traffic of the metalloprotease gp63 along the classical secretory pathway is well documented, little is known about the export mechanisms to the extracellular space. Biochemical analysis of the extra-cellular gp63 of *Leishmania* has revealed two forms of the protein, one released from the cell surface and another that is apparently directly secreted (McGwire et al., 2002). Interestingly, GPI-anchored gp63 and the secreted unanchored gp63 are trafficked via different pathways (Ellis et al., 2002). Within the sandfly vector, release of extracellular gp63 from promastigotes could enhance the hydrolysis of protein substrates and play a nutritional and/or protective role (McGwire et al., 2002).

The *Leishmania* PPG are mucin-like glycoproteins that exist as membrane-bound and soluble forms in both promastigotes and amastigotes (Foth et al., 2002). These proteins are thought to be important in the transmission, invasion, and subsequent intracellular survival of parasites (Ilg et al., 1999). Furthermore, a DNA vaccine coding the N-terminal domain of *L. donovani* PPG provides immunoprotection against experimental visceral leishmaniasis (Samant et al., 2009).

In *T. cruzi*, secreted materials are directly involved in the invasion of target cells. After being in lysosomes, trypomastigotes escape into the cytosol by secreting an acid-activated hemolysin (Sibley and Andrews, 2000). Members of the SAP family (serine-, alanine-, and proline-rich protein) are shed into the extracellular medium. Likewise, some of these proteins may be involved in the invasion of mammalian cells by inducing Ca 2+ mobilization, which is a requirement for *T. cruzi* internalization (Baida et al., 2006).

B. CONTEXT AND THESIS RESEARCH AIMS

In order to complete their life cycle, trypanosomatid parasites have to adapt and develop in an insect vector and a vertebrate host. These organisms have developed various strategies to shape their environment, modulate host immune responses, or invade target cells. Materials secreted by the parasite are involved in these processes and may therefore represent virulence factors, vaccine candidates, and/or new drug targets. Despite the importance of trypanosomatid extracellular components in the host-parasite interaction, only a few extracellular proteins have been fully characterized. In addition, the genome sequences of the three main trypanosomatids have highlighted their genetic relatedness and a large number of hypothetical proteins with unknown function.

The main goal of this thesis work was to identify and characterize biological properties of novel secreted proteins that are conserved among trypanosomatids. We hypothesized that a phylogenetic conservation among *Leishmania spp.*, *T. cruzi* and *T. brucei* would indicate evolutionary selection for these families of proteins and suggest an important role for such secreted proteins in the biology of these parasites. By using a reverse genetic strategy, based on the presence of N-signal peptides in protein sequences, we aimed to identify extracellular proteins involved in the classical secretory pathway.

The first chapter of the present work describes a methodology we designed to identify conserved secreted proteins in trypanosomatids. The usefulness of the methodology to identify potential virulence factors is supported with studies on the infectivity of parasites over-expressing the candidate secreted proteins.

In the second chapter we present data on the immunogenicity of recombinant *T. cruzi* proteins and discuss their potential application for the diagnosis of Chagas disease. Misdiagnosis of *T. cruzi* infections is a major problem especially in areas where Chagas disease and Leishmaniasis overlap, including north-western Argentina. Because extracellular proteins of *Trypanosoma cruzi* represent potential targets for antibodies produced by the host, data on the immunogenicity of secreted proteins could be useful for developing new diagnostic tools for Chagas disease.

In the third chapter, an overview of current methodologies to study the whole population of the extracellular proteins in eukaryotes and its potential application in *Leishmania* are presented. We highlight the importance of understanding the complexity of the *Leishmania* exoproteome in order to elucidate the factors involved in its host-parasite interactions. The importance and relative contribution of protein secretion mechanisms to the export of extracellular proteins are also discussed.

C. CHAPTER 1

Screening and identification of novel secreted proteins in trypanosomatid parasites

1. Article: An experimental approach for the identification of conserved secreted proteins in trypanosomatids. *Molecular Genetics and Genomics,* submitted

In order to identify novel secreted proteins conserved among trypanosomatids, we combined a bioinformatic analysis with functional tests in *Leishmania*. By using the programme SignalP3 we analysed the genome sequences of *L. major*, *T cruzi* and *T. brucei* to generate a list[1]of potentially secreted proteins involved in the classical pathway. Subsequently, *Leishmania* parasites were used to experimentally identify extracellular proteins in recombinant parasites expressing potentially secreted proteins. The methodology we designed allowed us to experimentally identify extracellular proteins in 25% of the selected genes.

2. This methodology led us to apply for a European patent application:

- **Corrales, R.M**, Mathieu-Daude, Sereno, D., 2008. A method for the screening of conserved secreted proteins. European patent application N°08290657.9 pending since 04/07/08.

---

[1] The list of 45 potentially secreted proteins identified by the bioinformatic analysis is presented in the Annexe A.

An experimental approach for the identification of conserved secreted proteins in trypanosomatids

Journal:	*Molecular Genetics and Genomics*
Manuscript ID:	draft
Article Type:	Original Paper
Date Submitted by the Author:	
Complete List of Authors:	Corrales, Rosa; Institut de Recherche pour le Développement, UR016
Keywords:	classical secretion, extracellular proteins

An experimental approach for the identification of conserved secreted proteins in trypanosomatids.

Rosa M Corrales [a], Françoise Mathieu-Daudé [a], Déborah Garcia [a], Simone F Brenière [b], Denis Sereno [a*]

[a] Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Institut de Recherche pour le Développement, 911 Av. Agropolis, 34394 Montpellier, France

[b] Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Représentation en Bolivie, Av Hernando Siles N°5290 Esq Calle 7 Obrajes La Paz, Bolivia.

[*]Corresponding author: Rosa M Corrales, Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Institut de Recherche pour le Développement, 911 Av. Agropolis, 34394 Montpellier, France.

Tel.: +33 4 67 41 61 47; fax: +33 4 67 41 61 47.

Email address: corrales@mpl.ird.fr

Abstract

The trypanosomatid parasites *Leishmania spp.*, *Trypanosoma cruzi* and *Trypanosoma brucei* have developed various strategies to modulate host immune responses and/or invade target cells. Materials secreted by the parasite are involved in these processes and may represent virulence factors, vaccine candidates and/or new drug targets. Here, we describe an experimental genome-based approach to identify novel secreted proteins that are conserved among trypanosomatids and involved in the classical secretory pathway. Proteins potentially secreted via the endoplasmic reticulum were identified by bioinformatic analysis of the *T. cruzi* genome. A subset of thirteen genes encoding unknown proteins and presenting orthologous members with a signal peptide sequence in *L. infantum, L. major* and *T. brucei* were transfected into *L. infantum*. Detection of the tagged proteins in the extracellular medium confirmed the computer predictions in about 25% of the hits. Secretion was confirmed for two *L. infantum* orthologous proteins using the same experimental system. *In vitro* infectivity properties of recombinant *Leishmania* parasites suggest that one of the secreted proteins increase replication of the parasite inside its target cell. This methodology allows the identification of conserved secreted proteins involved in the classical secretory pathway, representing potential virulence factors in trypanosomatids.

Keywords: classical secretion, secretory signal peptide, extracellular proteins, trypanosomatid hypothetical conserved proteins, His- tagged proteins

Introduction

The Trypanosomatidae comprise a large group of parasitic protozoa, some of which cause important diseases in humans. The organisms that have been most extensively studied are *Trypanosoma brucei*, the causative agent of African sleeping sickness, *T. cruzi*, responsible for Chagas disease in South America, and *Leishmania spp.*, which causes Leishmaniasis in Asia, South America and Mediterranean countries. Half a billion people, primarily in tropical and subtropical areas of the world, are at risk of contracting these diseases (Stuart et al. 2008). It is estimated that more than 20 million individuals are infected, resulting in extensive suffering and more than 100,000 deaths per year (Stuart et al. 2008). In order to complete their life cycle, these parasites have to adapt and develop in an insect vector (a tsetse fly, a triatomine bug, or a sandfly) and in a vertebrate host. These single-celled organisms have developed several strategies to modify their surrounding environment, modulate host immune responses, or interfere with the host's anti-microbial activity. Materials secreted by the parasite play pivotal roles in these processes (Burleigh and Woolsey 2002; McConville et al. 2002). Secreted proteases belonging to the family of cystein- or metallo- proteases are generally thought to be involved in the manipulation of host immune responses in insect and vertebrate hosts (McGwire et al. 2002; Olivier et al. 2005; Kulkarni et al. 2006; Santarem et al. 2007). Biochemical analysis of the extra-cellular gp63 of *Leishmania* has revealed two forms of the protein, one released from the cell surface and another that is apparently directly secreted. The secreted form provides protection for *Leishmania* against insect-derived antimicrobial peptides (Kulkarni et al. 2006). Parasites can also secrete enzymes involved in nutritional processes (Ortiz et al. 2007). *Leishmania*, like other trypanosomatids, are purine auxotrophs, and therefore are entirely dependent upon salvaging these essential nutrients from their hosts (Ortiz et al. 2007). To satisfy its essential purine requirements, *Leishmania* secretes a nuclease that may function externally of the parasite to hydrolyze and access host-derived nucleic acids (Joshi and Dwyer 2007). Secreted materials can also be directly involved in the invasion of target cells. Tc-TOX, a pore-forming protein of *T. cruzi,* allows the parasite to escape the endosome and reach the cytoplasm of host cells, its natural cellular environment (Andrews et al. 1990). Experimental evidence suggests that *Leishmania* also possess a pore-forming protein that contributes to their release from host macrophages (Almeida-Campos and Horta 2000). Together, these findings demonstrate that secreted materials are involved in processes that help the parasite survive in an environment more favorable for its own development. In addition, previous studies indicate that trypanosomatid secreted factors elicit strong immunity and protection against infection in mice and dogs (Tonui et al. 2004; Lemesre et al. 2007). Thus, secreted factors could be a source of antigens for vaccine development as demonstrated in the pathogen *Mycobacterium tuberculosis* (Horwitz et al. 2005). Nevertheless, all trypanosomatid secreted factors involved in virulence and/or representing putative vaccine targets are not currently known.

In eukaryotes, soluble secreted proteins typically contain N-terminal signal peptides that direct them to the translocation apparatus of the endoplasmic reticulum (ER). Following vesicular transport from the ER via the Golgi apparatus to the cell surface, lumenal proteins are released into the extracellular space by fusion of Golgi-derived secretory vesicles with the plasma membrane (Rothman and Orci 1992; Landfear and Ignatushchenko 2001; McConville et al. 2002). In trypanosomatids, it is presumed that molecules destined for the cell surface and secretion follow a typical eukaryotic pathway traveling from the ER to the Golgi apparatus, then to the cell surface via a flagellar reservoir membrane called the flagellar pocket (McConville et al. 2002; Field et al. 2007). Nevertheless, a recent proteomic approach applied to the *Leishmania* secretome suggested that this parasite predominantly uses non-classical secretion pathways to directly export proteins, including the release of microvesicles (Silverman et al. 2008).

The availability of three draft trypanosomatid genome sequences provides valuable data for protein-mining using bioinformatic tools, especially for the localization or prediction of function for hypothetical proteins. Given that a significant number of trypanosomatid protein-coding genes are annotated as hypothetical, additional studies are needed to ascertain their function.

In the present study, we designed an experimental genome-based approach to identify novel secreted proteins that are conserved among the three main trypanosomatid pathogens and involved in the classical secretory pathway. We hypothesized that a phylogenetic conservation among *Leishmania*, *T. cruzi* and *T. brucei* would indicate evolutionary selection for this family of proteins and suggest an important role for such secreted proteins in the biology of these parasites. Our results demonstrate that the bioinformatic analysis, combined with the functional tests, provides a fast and reliable method for the identification of novel *bona fide* secreted proteins, representing potential virulence factors in trypanosomatids. The results are also discussed in relation to the relative importance of the classical and non-classical secretory pathways for the release of proteins into the extracellular environment.

Materials and Methods

In silico sequence analysis

Release V 5.0 of the *T. cruzi* genome was extracted from the integrated *T. cruzi* genome resource TcruziDB (http://tcruzidb.org/tcruzidb/). Protein sequences that do not bear an initial Methionine amino acid were removed manually. Proteins belonging to large families of surface molecules, which include trans-sialidases, mucins, gp63s and mucin-associated surface proteins, were also discarded. Finally ORFs encoding proteins bearing a molecular weight (MW) above 90 kDa were also eliminated. The software SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the presence of a signal peptide and a cleavage site in amino acid sequences (Emanuelsson et al. 2007). Protein sequences having a signal peptide probability greater than 0.8 associated with a cleavage site probability greater than 0.7 were analyzed for the presence of orthologs in the related *Trypanosoma brucei* and *Leishmania major* parasite databases.

Parasite strains and cultures

The *T. cruzi* TcY7 (or Y cl7) clone derived from the Y strain (Garzon et al. 2003) was used throughout this study. Epimastigotes were grown in liver infusion tryptose (LIT) medium supplemented with 10% FCS at 28°C in standard conditions (Camargo 1964) and harvested during the logarithmic growth phase. Metacyclic trypomastigotes, obtained from the differentiation of late stationary phase epimastigotes, were used to initiate infection of mouse fibroblasts (L929). Trypomastigotes and amastigotes were produced and harvested as previously described (Mathieu-Daude et al. 2007). Pellets for RNA purification were processed immediately in lysis buffer. The wild-type (WT) promastigote clone from *L. infantum* (MHOM/MA/67/ITMAP-263) was maintained at 26°C by weekly sub-passages in SDM 79 medium supplemented with 10% heat-inactivated FCS, 100 U/ml penicillin and 100 µg/ml streptomycin (Brun and Schonenberger 1979).

Reverse transcription and PCR amplifications

Total RNA was extracted from epimastigotes, amastigotes and trypomastigotes with the RNeasy kit (Qiagen, Hilden, Germany) according to the manufacturers' instructions, and treated with DNase I (DNA-*free* kit, Ambion Inc., Austin, Texas). Reverse transcription was performed for 1 µg of total RNA using random hexamers and Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturers' instructions. The cDNA (4 µl of 1/10 dilutions) from each stage was amplified by PCR in a 20 µl reaction volume using 10 µl of Master Mix 1X (Promega, Madison, Wisconsin), 0.5 µM gene-specific forward and internal reverse primers (listed in Table 1) using the following cycling conditions: 94°C for 3 min followed by 30 cycles of 94°C for 30 s, 55°C to 58°C (according to the primer pair) for 30 s, 72°C for 45 s and a final elongation at 72°C for 5 min. Amplicons were electrophoresed on 2% agarose gels stained with ethidium bromide.

Cloning and sequencing

The encoding genes selected by *in silico* analysis were cloned into the pTEX expression vector, carrying the Neomycin resistance gene (NEO) (Kelly et al. 1992). Full length ORFs were amplified from genomic DNA with specific reverse and forward primers, including different restriction sites and a 6-Histidine-Tag in the C-terminal region (listed in Tables 1 and 3). PCR reactions were carried out in 20 µl using 0.5 µM of each primer, 0.2 mM dNTP, 0.4 U of Phusion high-fidelity polymerase (Finnzymes, Espoo, Finland) and the following cycling conditions: 98°C for 30 s followed by 25 cycles of 98°C for 10 s, 64°C to 68°C for 15 s, 72°C for 25 to 60 s (according to gene size), and a 72°C elongation for 5 min. Digested and purified fragments were inserted into the dephosphorylated pTEX vector digested with the corresponding restriction enzymes. Cloned sequences were confirmed by restriction digestion and sequencing. Large scale preparations of the different constructs were performed using the plasmid midi kit (Promega).

Transfection procedures

Promastigotes of the *Leishmania infantum* clone were electroporated as described elsewhere (Sereno et al. 2001). Briefly, promastigotes were washed twice with HEPES-NaCl buffer saline (21 mM HEPES, 5 mM KCl, 0.7 mM NaH2PO4, 137 mM NaCl), resuspended at 10^8 cells/ml in HEPES-NaCl electroporation buffer (pH 7.2) supplemented with 6 mM glucose and cooled on ice for 10 min. Cells (10^8) were combined with 15 µg of vector, left on ice for 10 additional min, and electroporated using an Easyject Plus (Eurogentec, Seraing, Belgium) apparatus set at 450 V and 450 µF, for one pulse. The cells were left on ice for a further 10 min and transferred to 4 ml of growth medium. The antibiotic G-418 (20 µg/ml) was added 24 h later, and parasites were sub-cultured at a dilution of 1/10 in 5 ml SDM in the presence of 20 µg/ml G418. Drug-resistant cells were observed 15-20 days later. Parasites were grown in the presence of gradually increasing concentrations of G418 and were routinely maintained in SDM containing 150 µg/ml of G418.

PCR amplifications in transfected parasites

PCR amplifications were carried out to check for the presence of the NEO gene and the corresponding gene in transfected parasites. A fragment of 800 bp corresponding to the NEO gene was amplified with specific forward and reverse primers (F5'ATGATTGAACAAGATGGATTGCACGCAGG3', R5' TCAGAAGAACTCGTCAAGAA 3'). Full length ORFs of the specific genes were amplified with primers listed in Table 1. PCR reactions were carried out in a 20 µl reaction volume using 10 µl of Master Mix 1X (Promega, Madison, Wisconsin), 0.5 µM NEO and gene-specific forward and reverse primers using the following cycling conditions: 94°C for 3 min followed by 30 cycles of 94°C for 30 s, 55°C to

58°C (according to the primer pair) for 30 s, 72°C for 45 s to 2 min (according to gene size) and a final elongation at 72°C for 5 min.

Production of cell free culture supernatants

To analyze the presence of secreted proteins in the supernatant, $1x10^9$ *L. infantum* promastigotes from log-phase culture were collected by centrifugation, washed twice in HEPES-NaCl buffer, re-suspended in 40 ml of HEPES-NaCl (pH 7.2), 11 mM glucose, 200 µg/ml G-418 and incubated for 6 h at 27°C. Parasite viability was then assessed as previously described (Vergnes et al. 2002) and harvested by centrifugation at 2,100 *g* for 10 min at 4°C. The parasite pellet was stored at -80°C for subsequent SDS-PAGE analysis and the supernatant was filtered through a low retention 0.45 µm PVDF filter membrane (Millipore, Boston, Massachusetts). After addition of protease inhibitor cocktail (Sigma-Aldrich) the filtrate was concentrated up to 80-fold using an Ultra-Centrifugal Filter device, according to manufacturers' instructions (Amicon Bioseparations, MilliporeCorp). The concentrated cell-free culture supernatant was frozen and stored at -80°C.

Production of parasite lysates

Cell pellets of wild-type and episomally-transfected *L. infantum* promastigotes were re-suspended in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% Sodium deoxycholate and 0.1 % SDS), incubated on ice for 30 min and sonicated three times for 20 s. The soluble phase was recovered by centrifugation at 10,000g for 30 min (4°C) and the protein concentration was determined using a Bradford protein assay (Bio-Rad Laboratories, Hercules, California).

Gel electrophoresis and western blot analysis

Proteins from parasite lysates (35 µg) or from 80 x concentrated cell-free supernatants (2 µg) were separated on a NuPAGE Bis-Tris gel (4-12%) in MOPS-SDS running buffer (Invitrogen) under reducing conditions (50 mM DTT), and transferred to a PVDF membrane (Hybond-P, Amersham). The membrane was rinsed twice in TBS and incubated for 1h in the anti-His HRP conjugate blocking buffer (Qiagen). The membrane was then incubated in 1/3000 anti-His HRP conjugate antibody (Qiagen) for 1h and washed seven times for 5 min in TBS-T buffer (TBS-0.5 % Tween 20). Signals were detected by chemiluminescence emission using the ECL Plus Western blotting detection system and ECL Hyperfilms (GE Healthcare, UK).

Generation of bioluminescent *L. infantum* promastigotes and *in vitro* infection of human macrophages

A homologous episomal expression system was devised to further examine the infection *in vitro* of two secreted proteins from *L. infantum*. The vector pSP-αHYGαLUC (El Fadili et al. 2002) carrying the firefly-luciferase gene was used to co-transfect *L. infantum* promastigotes over-expressing the secreted proteins LinJ19.0410 (ortholog of Tc00.1047053505789.10) or LinJ36.5780 (ortholog of. Tc00.1047053506155.99). Recombinant parasites were selected for their growth in increasing concentrations of Hygromycin (up to 300 µg/ml) over a period of several weeks. Promastigotes transfected with the pTEX vector alone and the pSP-αHYGαLUC were used as controls for infection experiments. The survival of transfected parasites was evaluated within human leukemia monocyte cell line (THP-1 cells). THP-1 cells were cultured in RPMI 1640 medium supplemented with 10% FCS, 2 mM glutamine, 100IU of penicillin/ml, and 100 µg of streptomycin/ml. THP-1 cells in the log phase of growth were differentiated into macrophages by incubation for 2 days in a medium containing 20 ng/ml of phorbol myristate acetate (Sigma-Aldrich). THP-1 cells treated with PMA were washed with prewarmed medium and then infected with stationary-phase promastigotes of transfected-*Leishmania* in a 24-well plate at a parasite/macrophage ratio of 10:1 for 4 h at 37°C with 5% CO_2. Non-internalized parasites were removed. After different incubation periods (24 h to 96 h) Luciferase activity was determined using the Steady Glo reagent (Promega, Madison,WI), according to the manufacturers' instructions. After 5 min, the plate was read using a Multilabel Counter VICTOR^{2}model 1420 (Perkin Elmer). Results are expressed as the mean of RLU (Relative Luminescence Units) activity of three independent experiments, each performed in triplicate. Statistical significance was analyzed by the Mann-Whitney U test.

Results

Bioinformatic selection for secreted proteins in trypanosomatids

The preliminary analysis of the 19613 *T. cruzi* putative proteins from the CL-Brener genome was performed to discard potential uncompleted sequences. A total of 1796 sequences were removed manually since they did not bear an initial Methionine amino acid. The remaining 17817 (90.8%) protein coding sequences were conserved for subsequent analysis. Housekeeping genes and sequences belonging to large gene families, like the trans-sialidases, mucins, Mucin Asociated Surface Proteins (MASPs), were discarded given that we aimed to identify novel secreted proteins. Finally, we eliminated sequences encoding proteins with a molecular weight above 90 kDa, in order to facilitate subsequent gene cloning. The remaining coding sequences were screened for the presence of both the signal peptide and the peptidase cleavage site with a probability of 0.8 and 0.7, respectively. A total of 216 sequences were obtained by using these criteria. Among them, 91 (42%) were annotated as "hypothetical proteins, conserved" in the data bank. The final criterion for selected proteins likely to be secreted by the classical eukaryotic pathway was the presence of the signal peptide and the signal peptidase site in orthologous members of the related parasites; *Leishmania major*, *L. infantum* and *T. brucei*. Among the 91 sequences, only 45 showed orthologous members with the signal peptide criteria. The 13 proteins bearing the highest probability for the presence of the signal peptide were selected (Table 2) for functional confirmation of *bona fide* secretion. The beta-tubulin *T. cruzi* gene (GeneID Tc00.1047053506563) was added to our sample as a potential negative control for protein secretion.

Transcription of the selected genes in the different forms of *T. cruzi* life cycle

We performed reverse transcription-PCR (RT-PCR) for the 13 *in-silico* selected genes in the different developmental stages of *T. cruzi* in order to verify transcription. The beta-tubulin *T. cruzi* gene (GeneID Tc00.1047053506563), constitutively expressed in all the three stages of *T. cruzi* was used as a positive control for RNA quality. RT-PCR was positive for all genes in the infective trypomastigote and amastigote forms. Two genes (Gene ID Tc00.1047053511901.30 and Tc00.1047053509999.10) were negative for the amplification of cDNA in the non-infective epimastigote form, (Figure 1), suggesting a possible stage-specific expression of these genes. Nevertheless, since trypanosomatid gene expression is almost exclusively regulated post-transcriptionally, further studies at the protein level have to confirm these observations.

Experimental approach for the detection of secreted proteins

We set up a functional test to confirm the presence of selected proteins in the extracellular environment by detection of target proteins in cell-free supernatants. The

13 selected encoding genes of *T. cruzi* and the gene encoding the beta-tubulin (negative control) were amplified from genomic DNA. A sequence encoding a 6xHis-Tag was added at the C-terminal end of each of the encoded genes. This would allow us to detect the protein in total parasite protein extracts or in concentrated cell-free supernatants (CCFS). Amplified PCR products were cloned into the pTEX shuttle vector that is widely used for expression in trypanosomatids (Kelly et al. 1992). Transformation and selection of *T. cruzi* is not as easy to perform as it is with *Leishmania,* mainly due to longer periods required for selecting drug-resistant parasites. Since we aimed to develop a fast and reliable approach to identify trypanosomatid conserved secreted proteins, we used the closely related *Leishmania* parasite as the recipient organism for the experimental validation of our selected proteins. Thus, *L. infantum* promastigotes were separately transformed with pTEX carrying one of the 14 selected *T. cruzi* genes (including the beta-tubulin gene), and the recombinant parasites were selected for resistance to Geneticin G418. Each parasite population was checked for the presence of both the NEO^R gene and the selected gene whose secretion was to be analyzed. A specific 800 bp fragment, indicative of the presence of the NEO^R gene, was detected in the transfected promastigotes and not in wild-type parasites (Figure 2A). Moreover, the presence of each candidate gene in recombinant parasites was confirmed using specific primers designed from *T. cruzi* gene sequences (Figure 2B). PCR performed on wild type *Leishmania* were negative, demonstrating that the amplification was genus specific although the genes are conserved in both trypanosomatids (data not shown). The expression of these genes was screened using an antibody directed against the His-Tag carried by the recombinant proteins (Figure 3A). Western blot analysis demonstrated that; (i) we were able to easily and specifically detect the 6xHis tag protein in the extract derived from recombinant parasites, (ii) recombinant *Leishmania* expressed a relatively high level of

T. cruzi protein, and (iii) the molecular weight of the detected tagged protein corresponded to the expected MW (see Table 1).

We subsequently set up an approach to detect recombinant proteins in cell-free supernatants. In order to limit potential contamination by proteins derived from dying organisms, we restricted incubation in serum-free mediums to 6 hours, and we checked the viability of parasite populations before and after this incubation period. Parasites and cell-free supernatants were collected if the viability of the cell population was above 98%. Western Blot analysis of the concentrated cell-free supernatants revealed that among the 14 proteins only 3 were actively secreted (Tc00.1047053506155.99, Tc00.1047053505789.10 and Tc00.1047053509999.10) (Figure 3B). These proteins represent genuine secreted material, since; (i) the over-expression of the beta-tubulin gene does not induce translocation of the beta-tubulin protein into the extracellular space (difference between Lys and CCFS in Figure 3B), and (ii) the detection of the tagged protein in the cell-free supernatant is not related to the level of its expression by *Leishmania* (low abundance of Tc00.1047053506155.99 in Figure 3B). As expected, we observed a slight molecular weight difference between the tagged protein detected into the whole soluble extract and that detected in the cell-free supernatant that could be explained by the loss of the Signal Peptide (Figure 3B). As anticipated, no cross reactive band was detected in wild-type parasites (Figure 3B). To confirm that the secretion we observed was not related to the heterologous expression system, two *Leishmania* genes (Gene ID LinJ19.0410 and LinJ36.5780) corresponding to the orthologs of Tc00.1047053505789.10 and Tc00.1047053506155.99 genes, were selected to validate our approach. By using the same protocol as above, *Leishmania* expressing the 6xHis tagged proteins were generated (See Table 3). As expected, we detected the presence of the tagged protein in the extracellular medium only in the episomally transfected parasites (Figure 4). Together, these results demonstrate that the approach we used allowed us to identify new and genuinely secreted proteins involved in the endoplasmic reticulum/Golgi-dependent secretory pathway.

Expression of secreted proteins increases ability of recombinant *Leishmania* parasites to infect and survive inside macrophages in vitro.

We attempted to determine whether the expression of *Leishmania* secreted proteins could interfere with the capacity of recombinant parasites to replicate within human macrophages *in vitro*. Both confirmed secreted proteins (LinJ19.0410 and LinJ36.5780) from *L. infantum* were tested by using the luciferase reporter system in transfected parasites over-expressing these proteins. We used bioluminescence as a quantitative indicator of the viability and multiplication of parasites. The number of promastigotes cells and luciferase activity were linearly correlated for the different recombinant parasites before macrophage infection (data not shown). Results of *in vitro* infection showed that over-expression of secreted protein LinJ19.0410 (ortholog of Tc00.1047053505789.10) increases the capacity of *Leishmania* to survive in THP-1 differentiated macrophages as early as 24 h post-infection (Figure 5). Furthermore, a statistically significant increase in luciferase activity of recombinant parasites expressing LinJ19.0410 was maintained throughout the experiments ($P<0.05$). This effect was not observed in parasites over-expressing LinJ36.5780 (ortholog of Tc00.1047053506155.99) where infectivity levels were similar to the control parasites transfected with the pTEX vector alone and the pSP-αHYGαLUC (Figure 5).

Discussion

Comparative analysis have revealed that genomes of the trypanosomatid parasites causing human disease, *Leishmania major*, *Trypanosoma cruzi* and *Trypanosoma brucei* are highly conserved (El-Sayed et al. 2005). About 50% of predicted proteins in these genomes were annotated as "hypothetical proteins" even though evidence of protein expression exists for many of them (Atwood et al. 2005; El-Sayed et al. 2005; Jones et al. 2006). Progress in controlling infections caused by these pathogens requires an improved understanding of the biology of these parasites so as to design novel treatment strategies. Previous findings indicate that excreted/secreted factors may play crucial roles in the biology or virulence of trypanosomatids (Andrews et al. 1990; McConville et al. 2002; Joshi and Dwyer 2007; Rogers et al. 2008). Furthermore, these factors elicit strong immunity and protection against infection in mice and dogs (Tonui et al. 2004; Lemesre et al. 2007). Thus, secreted factors may represent targets for rational drug design and/or a source of antigens for vaccine development.

In trypanosomatids the secretion process is not fully understood and various pathways may contribute to the formation of the “extracellular proteome”. Thus the individual identification of secreted materials would enhance efforts towards understanding mechanisms of protein secretion in these medically important parasites. Therefore, we developed an experimental approach to identify new secreted proteins involved in the classical pathway that are conserved among these three trypanosomatids. We combined a web-based bioinformatics approach that used the Signal IP 3.0 program, one of the most accurate predictors for the presence of a signal peptide sequence (Klee and Ellis 2005), with a functional test that takes advantage of the facility to genetically transform *Leishmania*. We assumed that a phylogenetic conservation among *Leishmania*, *T. cruzi* and *T. brucei* would point to evolutionary selection for this family of proteins and indicate an important role for these proteins in the biology of these parasites. Using these criteria we selected a pool of 13 trypanosomatid conserved hypothetical proteins from the *T. cruzi* database for which secretion was tested.

Various approaches aimed at characterizing such extracellular material were developed. Even if collecting a culture's supernatant is relatively easy to perform, identification of the different factors can be difficult because of the relatively low abundance of the constituents. Additionally, the *in vitro* growth of trypanosomatid developmental stages that occur in mammals is impossible or laborious. Moreover, even if parasites are grown in cell-free and serum-free media (Merlen et al. 1999), the culture's supernatant can be contaminated with materials that are not primarily "secreted" by the parasite. These materials may be shed from the parasite surface or can originate from dead organisms. Thus to avoid such pitfalls it is best to limit the parasite's incubation time in a serum-free media to a few hours. Further characterization by screening cDNA libraries with sera raised against culture medium supernatants has also been performed (Cibrelus et al. 1999; Chenik et al. 2006). However, using this approach proteins with a low abundance or that are poorly immunogenic are likely to be missed. A more exhaustive approach relying on a highly sensitive methodology, like the quantitative mass spectrometry, was recently used to analyze the protein content of the whole conditioned medium from stationary-phase *L. donovani* promastigotes (Silverman et al. 2008). Nevertheless, according to the authors, proteins produced at low abundance and that are mainly exported to the extracellular compartment are likely to be missed, since the method relies on the comparison of conditioned medium versus cell-associated proteins. The method we designed allowed us to identify three new trypanosomatid conserved proteins, likely to be secreted through the classical secretory pathway, via the endoplasmic reticulum. We were confident that the proteins detected in the cell-free supernatant were genuinely secreted since; (i) we were not able to detect the beta-tubulin tagged protein in the cell-free supernatant after over-expression of the gene, (ii) the 6 hours incubation time avoided excessive cell death and contamination by proteins released from dead parasites, (iii) in the extracellular environment we detected orthologous *Leishmania* proteins (LinJ19.0410, LinJ36.5780) suggesting that secretion in this protein family is likely to be evolutionary conserved, and (iv) we detected no relation between the amount of secreted protein and its intracellular expression in the transfected parasites, demonstrating that the translocation of the his-tagged protein into the cell supernatant is not related to the methodology we used**.**

The recent study involving conditioned medium derived from stationary-phase *Leishmania* promastigotes (Silverman et al. 2008), demonstrated that the extracellular proteome was mainly composed of proteins derived from different microvesicles. The parasite growth phase studied by these authors contains the infectious metacyclic parasites and also a high percentage of dying parasites in apoptosis (Zangger et al. 2002). Consequently, the analysis of extracellular material revealed that vesicles are likely to be "apoptotic vesicles" or "blebs", as described by Théry *et al* (Thery et al. 2001), are the main contributors to the "extracellular proteome". This work lead to the general conclusion that 98% of the proteins of the *Leishmania* secretome lacked a targeting signal, indicating that non-classical secretion pathways are likely to be the dominant way by which *Leishmania* export proteins (Silverman et al. 2008). However, some evidence strongly supports the notion that the classical secretory pathway is operational in trypanosomatids and contributes to the composition of the parasite's secretome. For example, a previous screening of a *L. major* cDNA library with antiserum raised against culture supernatant from stationary phase promastigotes led to the detection of 8 proteins that bear a potential signal peptide among the 33 genes identified (Chenik et al. 2006). Moreover some of the well known *Leishmania* proteins found in the extracellular environment have a signal peptide for secretion, such as, gp63 or chitinase (McGwire et al. 2002; Joshi et al. 2005). In the current work we deliberately decided to experimentally validate the secretion of our candidate proteins during the exponential growth phase of *Leishmania* in order to avoid contamination of the cell-free supernatant by apoptotic vesicles or exosomes. In these conditions we did not detect the secretion of tubulin, even in an over-expression model of *Leishmania* transfectants, while tubulin was identified with a significant score in the *Leishmania* secretome studied by the proteomic analysis performed during the stationary-phase (Silverman et al. 2008). Therefore, our results suggest that tubulin might be associated with exosomes and/or apoptotic vesicles generated by promastigotes in the stationary-phase of growth. We suggest that the composition of the parasite "extracellular proteome" is variable and depends both on the parasite stage under consideration and on the relative contribution of the various pathways operating in protein secretion. Further studies are required to highlight the importance of both classical and non classical secretory pathways in different developmental stages of trypanosomatids.

Regarding proteins not detected in the extracellular environment, it is important to point out that proteins bearing a signal peptide are not only targets for secretion but also for transfer to specific organelles, like lysosomes or the cell surface. Thus, these proteins could be retained in specific organelles within the parasite or attach to the cell surface via a GPI-anchor. In this light, analysis of the 13 proteins with GPI-SOM (http://gpi.unibe.ch/) (Fankhauser and Maser 2005) suggests that 2 out of 13 proteins (Tc00.1047053506417.30 and Tc00.1047053506467.29) have a predicted glycosyl-inositol phosphate (GPI) anchor signature sequences. Hence, the presence of the GPI anchored domain may explain the absence of these proteins in the culture supernatant of recombinant parasites.

An interesting observation in our results was the prediction of transmembrane helices in two of our secreted proteins (Tc00.1047053505789.10 and Tc00.1047053509999.10). According to TMHMM 2.0

(http://www.cbs.dtu.dk/services/TMHMM/) (see Table 2) and the DAS-TMfilter server (http://mendel.imp.ac.at/sat/DAS/DAS.html) (Cserzo et al. 2002) these proteins bear 3 to 5 transmembrane domains (strongly hydrophobic profiles). It has been shown that a considerable number of experimentally identified proteins of the *Bacillus anthracis* secretome bear one to ten potential transmembrane helices, suggesting that the presence of a putative transmembrane helix should not be used as a primary criteria to discard candidates of *in silico* inventories of potentially secreted proteins (Walz et al. 2007).

Our results demonstrate that further empirical studies are required to verify bioinformatic predictions. In this regard, this study helps bridge the gap between pure *in silico* prediction and *in vivo* observation. Though, another speculative explanation for the presence of transmembrane domains is the potential insertion of secreted materials into membranes after secretion. Indeed, in *Toxoplasma gondii* the dense granule protein GRA5 is a transmembrane protein that bears a signal peptide and is secreted as a soluble protein into the vacuolar space, before being inserted into the parasitophorous vacuole membrane (Lecordier et al. 1999; Karsten et al. 2004).

The domain or motif search tools available in InterPro (http://www.ebi.ac.uk/interpro/) (Hunter et al. 2009) allowed us to identify a Lipocalin signature in the secreted protein Tc00.1047053505789.10. The Lipocalin family has been defined on the basis of sequence or structural similarity. Lipocalins are a widely distributed group of mostly extracellular proteins, several of which have been implicated in the regulation of host immune responses, such as, alpha-1-microglobulin and alpha-1-acid glycoprotein (Grzyb et al. 2006). In *Rhodnius prolixus*, the insect vector of *T. cruzi*, lipocalins are implicated in the successful accomplishment of a blood meal, interfering with platelet aggregation, blood coagulation and inflammation (Montfort et al. 2000). In addition the pairwise sequence identity within this family is low, often below 30%, despite a well conserved tertiary structure (Flower et al. 2000). The common structure of the lipocalin protein fold is a highly symmetrical all-b structure dominated by a single eight-strand antiparallel b-sheet folded back on itself to form a continuously hydrogen bonded b-barrel (Flower et al. 2000). Nevertheless the secondary structure analysis of Tc00.1047053505789.10 with Psipred (http://bioinf.cs.ucl.ac.uk/psipred/) (Bryson et al. 2005)does not support the inclusion of this protein into the lipocalin family. In addition none of the orthologous genes in *Leishamnia spp* or *T. brucei* shows the lipocalins signature. A clear assignment to this family would only be possible after NMR or x-ray crystallography structure analysis. Further studies are needed to ascertain if this large and diverse group of proteins is present in trypanosomatids and plays a role in the transmission process. A Glucosidase II beta subunit-like protein domain has also been detected in the protein Tc00.1047053506155.99. Although mostly localized to the ER, Glucosidase II was found in endocytic structures beneath the plasma membrane and has been associated with the protein-tyrosine phosphatase CD45 (Baldwin et al. 2000). There is also evidence that in some cell types Glucosidase II beta is capable of being trafficked to the cell surface (Li et al. 1996). Having demonstrated that our methodology is reliable for the identification of genuinely secreted proteins in trypanosomatids, we tested the hypothesis that recombinant *Leishmania* parasites carrying extra copies of *Leishmania* secreted proteins may interfere with their survival or infectivity towards human monocyte-derived macrophages *in vitro*. Results of these assays showed a significant survival advantage to *Leishmania* parasites over-expressing the gene LinJ19.0410 (Ortholog gene of Tc00.1047053505789.10) suggesting that this protein is involved in a process increasing survival and replication of the parasite inside its target cell. Given that *Leishmania* and *T. cruzi* do not target the same host cells and follow different cell invasion processes, further experiments in the *T. cruzi* homologous system are needed to address whether the identified secreted protein Tc00.1047053505789.10 from *T. cruzi* is also involved in host cell invasion and/or replication. Current *in vitro* and *in vivo* studies are in progress to characterize this protein that represents a potential conserved virulence factor in trypanosomatids.

In summary, a group of 13 hypothetical proteins conserved in trypanosomatids were selected by computer-based analysis as potentially secreted via the endoplasmic reticulum system. Extracellular secreted proteins were experimentally confirmed in about 25% (3/13) of the hits by using transgenic *L. infantum* promastigotes expressing *T. cruzi* and *L. infantum* selected genes. Studies of infectivity properties of *Leishmania* secreted proteins suggest that one of these proteins is involved in a process increasing survival and replication of the parasite inside its target cell. Thus, the bioinformatics method combined with the functional tests, provides a fast and reliable method for the identification of novel *bona fide* secreted proteins that may represent potential virulence factors in trypanosomatids.

Acknowledgements

Financial support for this study was provided by the European Union through a grant to RM Corrales by the ‘Programme Alßan’ European Union Programme of High Level Scholarships for Latin America (No. E05D057391AR) and by IRD DSF. We thank B Vergnes for critical review of the manuscript and P Agnew for help revising the manuscript’s English.

References

Almeida-Campos FR, Horta MF (2000) Proteolytic activation of leishporin: evidence that Leishmania amazonensis and Leishmania guyanensis have distinct inactive forms. Mol Biochem Parasitol 111:363-375

Andrews NW, Abrams CK, Slatin SL, Griffiths G (1990) A T. cruzi-secreted protein immunologically related to the complement component C9: evidence for membrane pore-forming activity at low pH. Cell 61:1277-1287

Atwood JA, 3rd et al. (2005) The Trypanosoma cruzi proteome. Science 309:473-476

Baldwin TA, Gogela-Spehar M, Ostergaard HL (2000) Specific isoforms of the resident endoplasmic reticulum protein glucosidase II associate with the CD45 protein-tyrosine phosphatase via a lectin-like interaction. J Biol Chem 275:32071-32076

Brun R, Schonenberger (1979) Cultivation and in vitro cloning or procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication. Acta Trop 36:289-292

Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT (2005) Protein structure prediction servers at University College London. Nucleic Acids Res 33:W36-38

Burleigh BA, Woolsey AM (2002) Cell signalling and Trypanosoma cruzi invasion. Cell Microbiol 4:701-711

Camargo EP (1964) Growth And Differentiation In Trypanosoma Cruzi. I. Origin Of Metacyclic Trypanosomes In Liquid Media. Rev Inst Med Trop Sao Paulo 12:93-100

Chenik M, Lakhal S, Ben Khalef N, Zribi L, Louzir H, Dellagi K (2006) Approaches for the identification of potential excreted/secreted proteins of Leishmania major parasites. Parasitology 132:493-509

Cibrelus P, Precigout E, Sereno D, Carcy B, Lemesre JL, Gorenflot A (1999) Secreted antigens of the amastigote and promastigote forms of Leishmania infantum inducing a humoral response in humans and dogs. Parasite 6:121-129

Cserzo M, Eisenhaber F, Eisenhaber B, Simon I (2002) On filtering false positive transmembrane protein predictions. Protein Eng 15:745-752

El-Sayed NM et al. (2005) Comparative genomics of trypanosomatid parasitic protozoa. Science 309:404-409

El Fadili A, Kundig C, Ouellette M (2002) Characterization of the folylpolyglutamate synthetase gene and polyglutamylation of folates in the protozoan parasite Leishmania. Mol Biochem Parasitol 124:63-71

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2:953-971

Fankhauser N, Maser P (2005) Identification of GPI anchor attachment signals by a Kohonen self-organizing map. Bioinformatics 21:1846-1852

Field MC, Natesan SK, Gabernet-Castello C, Koumandou VL (2007) Intracellular trafficking in the trypanosomatids. Traffic 8:629-639

Flower DR, North AC, Sansom CE (2000) The lipocalin protein family: structural and sequence overview. Biochim Biophys Acta 1482:9-24

Garzon E et al. (2003) Trypanosoma cruzi carrying a targeted deletion of a Tc52 protein-encoding allele elicits attenuated Chagas' disease in mice. Immunol Lett 89:67-80

Grzyb J, Latowski D, Strzalka K (2006) Lipocalins - a family portrait. J Plant Physiol 163:895-915

Horwitz MA, Harth G, Dillon BJ, Maslesa-Galic S (2005) Enhancing the protective efficacy of Mycobacterium bovis BCG vaccination against tuberculosis by boosting with the Mycobacterium tuberculosis major secretory protein. Infect Immun 73:4676-4683

Hunter S et al. (2009) InterPro: the integrative protein signature database. Nucleic Acids Res 37:D211-215

Jones A et al. (2006) Visualisation and analysis of proteomic data from the procyclic form of Trypanosoma brucei. Proteomics 6:259-267

Joshi MB, Dwyer DM (2007) Molecular and functional analyses of a novel class I secretory nuclease from the human pathogen, Leishmania donovani. J Biol Chem 282:10079-10095

Joshi MB et al. (2005) Molecular characterization, expression, and in vivo analysis of LmexCht1: the chitinase of the human pathogen, Leishmania mexicana. J Biol Chem 280:3847-3861

Karsten V, Hegde RS, Sinai AP, Yang M, Joiner KA (2004) Transmembrane domain modulates sorting of membrane proteins in Toxoplasma gondii. J Biol Chem 279:26052-26057

Kelly JM, Ward HM, Miles MA, Kendall G (1992) A shuttle vector which facilitates the expression of transfected genes in Trypanosoma cruzi and Leishmania. Nucleic Acids Res 20:3963-3969

Klee EW, Ellis LB (2005) Evaluating eukaryotic secreted protein prediction. BMC Bioinformatics 6:256

Kulkarni MM, McMaster WR, Kamysz E, Kamysz W, Engman DM, McGwire BS (2006) The major surface-metalloprotease of the parasitic protozoan, Leishmania, protects against antimicrobial peptide-induced apoptotic killing. Mol Microbiol 62:1484-1497

Landfear SM, Ignatushchenko M (2001) The flagellum and flagellar pocket of trypanosomatids. Mol Biochem Parasitol 115:1-17

Lecordier L, Mercier C, Sibley LD, Cesbron-Delauw MF (1999) Transmembrane insertion of the Toxoplasma gondii GRA5 protein occurs after soluble secretion into the host cell. Mol Biol Cell 10:1277-1287

Lemesre JL et al. (2007) Long-lasting protection against canine visceral leishmaniasis using the LiESAp-MDP vaccine in endemic areas of France: double-blind randomised efficacy field trial. Vaccine 25:4223-4234

Li YM et al. (1996) Molecular identity and cellular distribution of advanced glycation endproduct receptors: relationship of p60 to OST-48 and p90 to 80K-H membrane proteins. Proc Natl Acad Sci U S A 93:11047-11052

Mathieu-Daude F et al. (2007) Sequence diversity and differential expression of Tc52 immuno-regulatory protein in Trypanosoma cruzi: potential implications in the biological variability of strains. Parasitol Res 101:1355-1363

McConville MJ, Mullin KA, Ilgoutz SC, Teasdale RD (2002) Secretory pathway of trypanosomatid parasites. Microbiol Mol Biol Rev 66:122-154; table of contents

McGwire BS, O'Connell WA, Chang KP, Engman DM (2002) Extracellular release of the glycosylphosphatidylinositol (GPI)-linked Leishmania surface metalloprotease, gp63, is independent of GPI phospholipolysis: implications for parasite virulence. J Biol Chem 277:8802-8809

Merlen T, Sereno D, Brajon N, Rostand F, Lemesre JL (1999) Leishmania spp: completely defined medium without serum and macromolecules (CDM/LP) for the continuous in vitro cultivation of infective promastigote forms. Am J Trop Med Hyg 60:41-50

Montfort WR, Weichsel A, Andersen JF (2000) Nitrophorins and related antihemostatic lipocalins from Rhodnius prolixus and other blood-sucking arthropods. Biochim Biophys Acta 1482:110-118

Olivier M, Gregory DJ, Forget G (2005) Subversion mechanisms by which Leishmania parasites can escape the host immune response: a signaling point of view. Clin Microbiol Rev 18:293-305

Ortiz D et al. (2007) Molecular genetic analysis of purine nucleobase transport in Leishmania major. Mol Microbiol 64:1228-1243

Rogers ME et al. (2008) Leishmania chitinase facilitates colonization of sand fly vectors and enhances transmission to mice. Cell Microbiol 10:1363-1372

Rothman JE, Orci L (1992) Molecular dissection of the secretory pathway. Nature 355:409-415

Santarem N et al. (2007) Immune response regulation by leishmania secreted and nonsecreted antigens. J Biomed Biotechnol 2007:85154

Sereno D, Roy G, Lemesre JL, Papadopoulou B, Ouellette M (2001) DNA transformation of Leishmania infantum axenic amastigotes and their use in drug screening. Antimicrob Agents Chemother 45:1168-1173

Silverman JM et al. (2008) Proteomic analysis of the secretome of Leishmania donovani. Genome Biol 9:R35

Stuart K et al. (2008) Kinetoplastids: related protozoan pathogens, different diseases. J Clin Invest 118:1301-1310

Thery C et al. (2001) Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles. J Immunol 166:7309-7318

Tonui WK et al. (2004) Immunization with Leishmania major exogenous antigens protects susceptible BALB/c mice against challenge infection with L. major. Infect Immun 72:5654-5661

Vergnes B, Sereno D, Madjidian-Sereno N, Lemesre JL, Ouaissi A (2002) Cytoplasmic SIR2 homologue overexpression promotes survival of Leishmania parasites by preventing programmed cell death. Gene 296:139-150

Walz A et al. (2007) Bacillus anthracis secretome time course under host-simulated conditions and identification of immunogenic proteins. Proteome Sci 5:11

Zangger H, Mottram JC, Fasel N (2002) Cell death in Leishmania induced by stress and differentiation: programmed cell death or necrosis? Cell Death Differ 9:1126-1139

Legends to Figures

Figure 1 - Amplification of the cDNAs encoding the potentially secreted proteins in the different stages of *T. cruzi*

RT-PCR analysis of total RNA from *T. cruzi* (clone derived from the Y strain) epimastigotes (E), trypomastigotes (T) and amastigotes (A). cDNA was amplified using gene-specific PCR primers (Listed in Table 1). Gene ID and expected lengths of cDNA are listed in order in Table 1. M: Molecular marker: Smart Ladder SF.

Figure 2 - PCR analyses in episomally-transfected *L. infantum* promastigotes

(A) Amplification of NEO gene fragment in *L. infantum* episomally transfected promastigotes. (**B)** Amplification of full length transfected genes in *L. infantum* promastigotes. Specific forward and Reverse PCR primers and gene lengths are listed in order in Table 1. WT: Wild Type Parasites. M: Molecular marker: Smart Ladder SL.

Figure 3 - Protein expression in *L. infantum* episomally transfected promastigotes during the exponential phase of development

(**A)** Western blot analysis of His-tagged proteins detected in whole cell lysate. Equal amounts of total protein (35µg) were resolved by electrophoresis in 4-12% gradient gels (Invitrogen), blotted, and developed with anti-HisTag antibody followed by ECL (Amersham). Gene ID and the theoretical molecular weight of detected proteins are listed in order in Table 1. (**B)** Identification of secreted proteins in whole cell lysate (Lys) and concentrated cell-free culture supernatant (CCFS) obtained from promastigotes incubated for 6h in serum-free medium. Note the absence of β Tubulin in the concentrated supernatant of Line 8. Non-transfected *L. infantum* promastigotes

(Wild Type) were used as a negative control. Protein molecular mass standards in kDa are shown on the left of each panel.

Figure 4 - Homologous expression of secreted proteins in *L. infantum* episomally transfected promastigotes

L. infantum promastigotes were transfected with genes: LinJ19.0410 and LinJ36.5780 corresponding to secreted proteins Tc00.1047053505789.10 and Tc00.1047053506155.99, respectively. Cell whole lysate (Lys) and concentrated cell-free culture supernatant (CCFS) and electrophoresis procedure were as in Figure 3. Tagged proteins were detected only in recombinant parasites transfected with LinJ19.0410 (58 kDa) (Line 1 and 2) and LinJ36.5780 (28 kDa) (Line 3 and 4). Non-transfected *L. infantum* promastigotes (Wild Type) were used as negative controls (Line 5 and 6). Protein molecular mass standards in kDa are shown on the left.

Figure 5 - Bioluminescence activity of intracellular *Leishmania* expressing episomal luciferase in infected macrophages *in vitro*.

Recombinant *L. infantum promastigotes* over-expressing the secreted proteins pTEX-LinJ19.0410 (▼) or pTEX-LinJ36.5780 (●) were co-transfected with the pSP-YαHYGROαLUC carrying the firefly-luciferase gene. Survival of luciferase-expressing parasites was monitored in infected human monocyte cell line THP-1 differentiated into macrophages as indicated in the Methods section. Promastigotes transfected with the pTEX vector alone and the pSP-αHYGαLUC (□) were used as control for infection experiments. RLU (Relative Luminescence Units) were measured at various time points post-infection using the Steady Glo reagent. Results are expressed as the mean of three independent experiments, each carried out in triplicate.

Tables

Table 1. Primer design of the 13 *T. cruzi* putative secreted proteins conserved in trypanosomatids and *T. cruzi* Tubulin (negative control).

T. cruzi GeneDB Gene ID N°	Primer sequences[a]	F/Rint and F/R product sizes (bp)	MW (kDa)[b]
Tc00.1047053506417.30	F CATGAGCTT*ACTAGT*ATGTTGTCTCTGGCAGAAGTGTGT Rint ACGGTGCCCAAAGGCGTGTA R CACACGG***AAGCTT***TCA**ATGATGATGATGATGATG**GCGACCAAACCTAGCCATAAG	 311 705	 25.7
Tc00.1047053506155.99	F CTGGGG*GAATTC*ATGCGGTGGATTTTTTTGTTACTTGCC Rint CCGATACGTCCACCACCCCTC R CGTCGG*AAGCTT*CTA**GTGATGGTGGTGGTGATG**GACAAGTTCGTGGCATGTAATTG	 336 735	 28.1
Tc00.1047053506467.29	F ACACGG*ACTAGT*ATGATTGTATTGAATGGAATTTCTGAG Rint CTAATAGTCCGAAGTCGTTGCG R CTACAC*AAGCTT*TTA**GTGATGGTGATGGTGATG**GCTGCGCCTCCACACCGTGC	 309 1065	 39.7
Tc00.1047053511901.30	F CTGATAGGC*ACTAGT*ATGTTTCCGGCGCAGGAATTCCT Rint CCCCTTTCAGGTGACCATTACAAGAG R GCCGTC*AAGCTT*TTA**GTGGTGGTGGTGGTGATG**CTCCGCTCCCAACTTCAAACGA	 316 1041	 39.8
Tc00.1047053511871.30	F CTGATAGGC*ACTAGT*ATGCGTCGCACTTTATTTTGTCTG Rint CTCTCCAACTCGTACGGCGA R CTGCAGGC*AAGCTT*CTA**ATGGTGGTGATGGTGATG**TTATGATACCGGCATCAAGTCCCC	 305 1269	 47.0
Tc00.1047053505789.10	F CGCACTC*ACTAGT*ATGCCCTCTGGCAAAGCAACTG Rint TCACTGCTCCGCCCTGGTTTC R CGCTCC*CTCGAG*TTA**GTGGTGATGGTGGTGATG**GGCAGCATTTACCGACCCTGA	 308 1488	 53.6
Tc00.1047053509669.70	F GCTCAGCC*AAGCTT*ATGCGCACTTCTTCTGCCGTGT Rint ATCGGGGAGTTTTGTGCAGGTTGAG R GTGGTCTT*CTCGAG*TTA**GTGATGGTGGTGGTGATG**GTCGACTTTAATGCTCGCGTATA	 322 1944	 73.1
Tc00.1047053507765.20	F CTGCCCAGT*ACTAGT*ATGTCTGCTAAAGCCTCACGGC Rint TCCAGGTAGTCACCCATTCCGTG R CTCAGCC*AAGCTT*TTA**ATGATGATGATGATGGTG**TCGTCTCGCCTCACAGTGCT	 318 1521	 57.2
Tc00.1047053510101.470	F CTGCGCTGG*ACTAGT*ATGTCTGTTAAAGCCTCACGGCG Rint CCATTCCGTGACCGCCGTAGAC R CTCGGT*AAGCTT*TTA**ATGATGATGATGATGGTG**TCGTCTCGCCTCACAGTGCT	 302 1518	 57.3
Tc00.1047053510443.30	F CTCGCTG*GAATTC*ATGCGGTGGGTGATAGTTGTATTTGC Rint CGCCAACAACGTAGTTGCCAAG R ACGGAC*CTCGAG*TTA**GTGATGGTGGTGGTGATG**CTTGTTGAGTTTGGAGCGGCG	 313 612	 23.0
Tc00.1047053509799.50	F CGCGGG*ACTAGT*ATGAAACAAAAAATGCGACGCAAATTG Rint GTGAGGATGGGGAACCAAAAGAGTC R CAGCC*AAGCTT*CTA**GTGATGGTGATGATGATG**GACATTCTTCTTCTTTGTAAAGTAG	 297 687	 26.5
Tc00.1047053509835.30	F CGCGGC*ACTAGT*ATGTATTCATGTTTGTCGCTGAGGC Rint GCAGCAACGGCAACAAAGAGC R CATGGC*AAGCTT*TTA**GTGATGGTGGTGGTGATG**CTCCTCTCTGGGTTTCCTTCG	 324 2031	 71.6
Tc00.1047053509999.10	F CGCGCC*ACTAGT*ATGTACGTCGTGCTTTTTTTCGTTT Rint CGCATATTTCCGCTCCGTTCC R AGCAGTCC*AAGCTT*TTA**GTGATGGTGATGATGATG**GCCGCACCAGCGCTCCAGAA	 305 1227	 46.6
Tc00.1047053506563.40[c]	F GGGTGCC*ACTAGT*ATGCGTGAGATTGTGTGCGTTCAG Rint GGGCGGAAGATCTGCCCGTATG R AGCGCTC*AAGCTT*TTA**GTGATGGTGGTGGTGATG**GTACTGCTCCTCCTCGTCGAACT	 259 1329	 49.6

F, Forward primer including the start codon; Rint, Internal reverse primer for RT-PCR; R, Reverse primer used for the amplification of the full-length ORF.

[a] Restriction sites used for cloning in the pTEX vector in italics and the His-Tag sequence in bold.

[b] Expected molecular weight of the proteins.

[c] Beta-tubulin gene.

Table 2. *T. cruzi* genes selected by *in silico* analysis

T. cruzi GeneDB Gene ID N°.	Orthologous *L. major* and *T. brucei* Gene ID N°		P value		TM^{c}	GPI^{d}	Conserved domains^{e}
			SPP^{a}	CSP^{b}			
Tc00.1047053506417.30	LmjF22.0225	Tb927.8.2180	0.937	0.917	1	Yes	None
Tc00.1047053506155.99	LmjF36.5220	Tb11.01.2470	0.984	0.962	0	No	Glucosidase II beta subunit-like
Tc00.1047053506467.29	LmjF26.2000	Tb09.160.1070	0.811	0.771	0	Yes	Methyltransferase domain
Tc00.1047053511901.30	LmjF24.2160	Tb927.8.6080	0.989	0.898	0	No	Glycerophosphoryl diester phosphodiesterase
Tc00.1047053511871.30	LmjF25.1010	Tb927.3.950	0.979	0.958	0	No	2OG-Fe(II) oxygenase
Tc00.1047053505789.10	LmjF19.0540 LmjF19.0570	Tb927.8.6700 Tb11.39.0005	1.000	0.768	5	No	Lipocalin signature
Tc00.1047053509669.70	LmjF29.1600	Tb927.3.4190	0.999	0.980	9	No	Endomembrane protein 70
Tc00.1047053507765.20	LmjF11.0720	Tb11.02.4400	0.993	0.986	0	No	None
Tc00.1047053510101.470	LmjF11.0720	Tb11.02.4400	0.931	0.919	0	No	None
Tc00.1047053510443.30	LmjF30.3150	Tb927.6.4500	0.903	0.838	1	No	Translocon-associated protein beta (TRAPB)
Tc00.1047053509799.50	LmjF36.5570	Tb10.6k15.1130	0.981	0.931	1	No	None
Tc00.1047053509835.30	LmjF19.0540 LmjF19.0570	Tb927.8.6700 Tb11.39.0005	0.866	0.803	5	No	Heavy-metal-associated domain, Heavy metal transport/detoxification
Tc00.1047053509999.10	LmjF29.1200	Tb927.3.3820	1.000	0.952	3	No	None

SPP^{a} Signal peptide probability predicted by SignalP 3.0.

CSP^{b} Maximal cleavage site probability predicted by SignalP 3.0

TM^{c} Number of transmembrane domains (other than peptide signal sequence) predicted by TMHMM 2.0

GPI^{d} Identification of GPI-anchor signal by GPI-SOM

Conserved domains^{e} from InterPro, PROSITE or Pfam server

Table 3. Gene ID of *L. infantum* orthologous genes and primers used for cloning

Gene ID	Primer sequences[a]	F/R product sizes (bp)	MW (kDa) [b]
LinJ19.0410[c]	F CATGACC*ACTAGT*ATGGCCAAAACAGCGCTTCTC R GCAGTCC*AAGCTT*TTA**GTGATGGTGATGATGATG**AGGTGTTCTCAGGGGTGACGA	1590	58,4
LinJ36.5780[d]	F CATGCTCG*ACTAGT*ATGGGGTGCC GCAGTAGCTG R GCAGTCC*AAGCTT*TTA**ATGATGATGGTGGTGATG**ATCATCCAACATCTGGCACCGC	738	28

F, Forward primer including the start codon; R, Reverse primer used for the amplification of the full-length ORF.

[a] Restriction sites used for cloning in the pTEX vector in italics and the His-Tag sequence in bold.

[b] Expected molecular weight of the proteins.

[c] Ortholog of Tc00.1047053505789.10

[d] Ortholog of Tc00.1047053506155.99

Figure 01

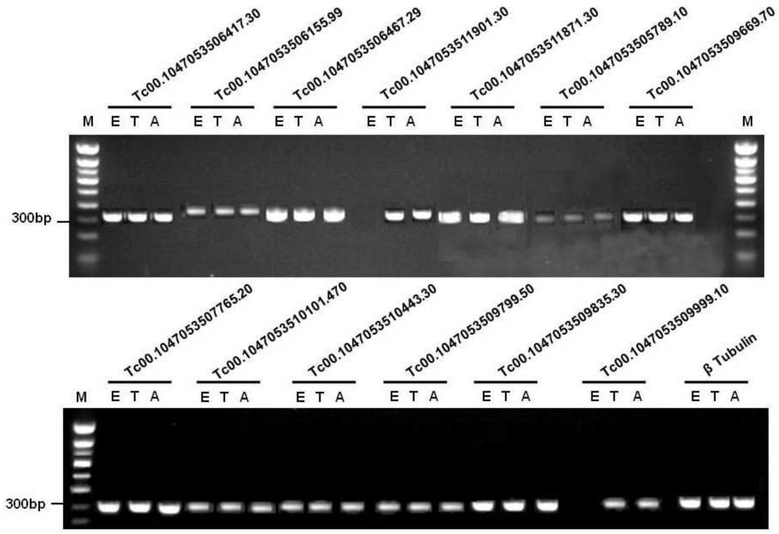

Tc00.1047053506417.30
Tc00.1047053506155.99
Tc00.1047053506467.29
Tc00.1047053511901.30
Tc00.1047053511871.30
Tc00.1047053505789.10
Tc00.1047053509669.70
M
E T A
E T A
E T A
E T A
E T A
E T A
E T A
M
300bp
Tc00.1047053507765.20
Tc00.1047053510101.470
Tc00.1047053510443.30
Tc00.1047053509799.50
Tc00.1047053509835.30
Tc00.1047053509999.10
β Tubulin
M
E T A
E T A
E T A
E T A
E T A
E T A
E T A
300bp

Figure 02

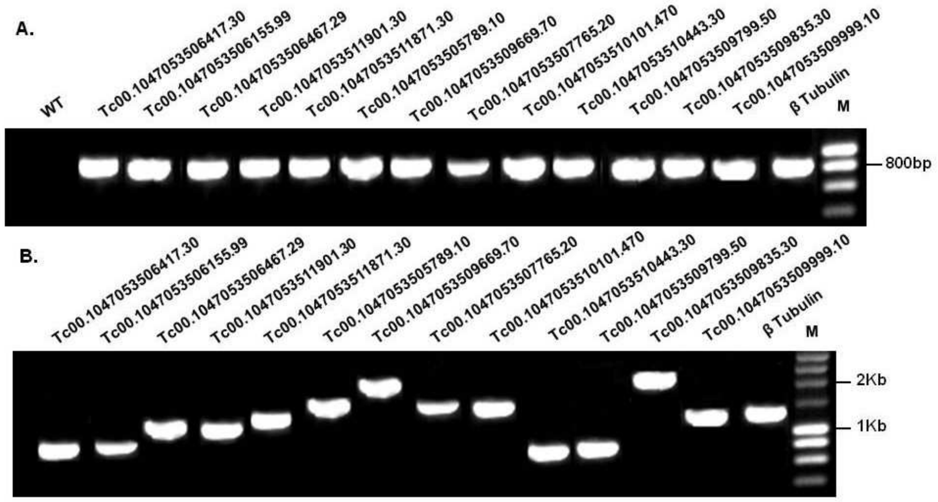

A.
WT
Tc00.1047053506417.30
Tc00.1047053506155.99
Tc00.1047053506467.29
Tc00.1047053511901.30
Tc00.1047053511871.30
Tc00.1047053505789.10
Tc00.1047053509669.70
Tc00.1047053507765.20
Tc00.1047053510101.470
Tc00.1047053510443.30
Tc00.1047053509799.50
Tc00.1047053509835.30
Tc00.1047053509999.10
β Tubulin
M
800bp
B.
Tc00.1047053506417.30
Tc00.1047053506155.99
Tc00.1047053506467.29
Tc00.1047053511901.30
Tc00.1047053511871.30
Tc00.1047053505789.10
Tc00.1047053509669.70
Tc00.1047053507765.20
Tc00.1047053510101.470
Tc00.1047053510443.30
Tc00.1047053509799.50
Tc00.1047053509835.30
Tc00.1047053509999.10
β Tubulin
M
2Kb
1Kb

Figure 03

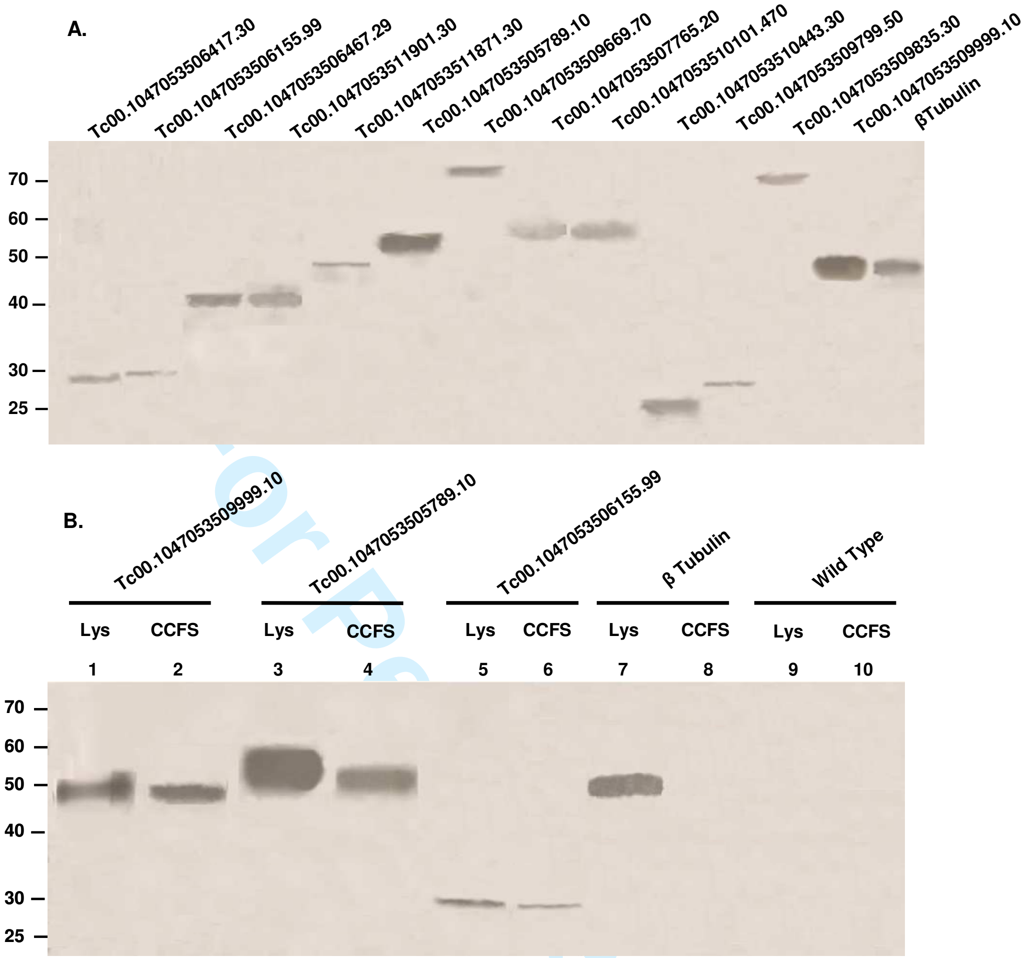

A.
Tc00.1047053506417.30
Tc00.1047053506155.99
Tc00.1047053506467.29
Tc00.1047053511901.30
Tc00.1047053511871.30
Tc00.1047053505789.10
Tc00.1047053509669.70
Tc00.1047053507765.20
Tc00.1047053510101.470
Tc00.1047053510443.30
Tc00.1047053509799.50
Tc00.1047053509835.30
Tc00.1047053509999.10
βTubulin
70
60
50
40
30
25
B.
Tc00.1047053509999.10
Tc00.1047053505789.10
Tc00.1047053506155.99
β Tubulin
Wild Type
Lys
CCFS
Lys
CCFS
Lys
CCFS
Lys
CCFS
Lys
CCFS
1
2
3
4
5
6
7
8
9
10
70
60
50
40
30
25

Figure 04

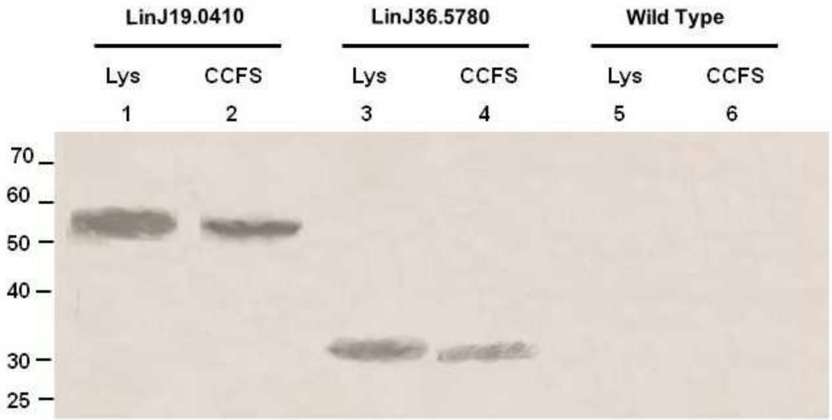
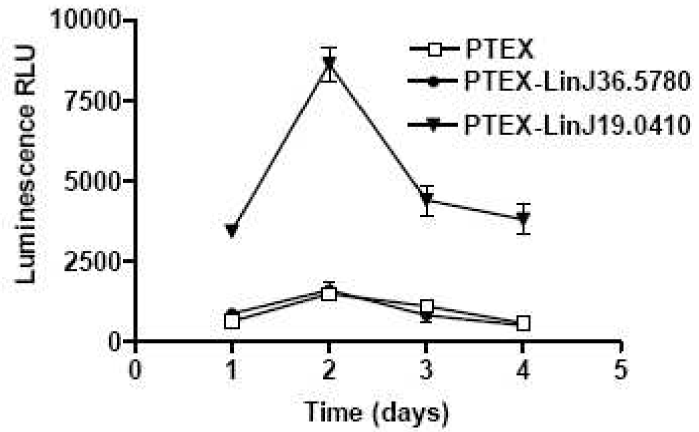

LinJ19.0410
LinJ36.5780
Wild Type
Lys
CCFS
Lys
CCFS
Lys
CCFS
1
2
3
4
5
6
70
60
50
40
30
25

Figure 05


10000
7500
5000
2500
0
Luminescence RLU
0
1
2
3
4
5
Time (days)
PTEX
PTEX-LinJ36.5780
PTEX-LinJ19.0410

3. Additional Results: Infectivity studies of transgenic parasites in different cell cultures

In this section, additional results regarding the infectivity properties of transgenic parasites over-expressing the secreted proteins identified in the chapter 1 are presented.

In order to determine whether the over-expression of secreted proteins could increase the infectious properties of transgenic *Leishmania* parasites, we used a luciferase reporter system to evaluate *in-vitro* infection.

As shown in the Figure 5 of the chapter 1, results of *in vitro* infection showed that over-expression of secreted protein LinJ19.0410 increased the capacity of *Leishmania* to survive within human macrophages *in vitro*. This effect was not observed in parasites over-expressing LinJ36.5780 where infectivity levels were similar to those of control parasites transfected with the pTEX vector alone and the pSP-αHYGαLUC.

To further evaluate these secreted proteins (LinJ19.0410 and LinJ36.5780), we assessed the infectivity of the transgenic parasites in different cell cultures. By using the same infection protocol with slight modifications, we evaluated the survival of transfected parasites within murine macrophage cell line J774 and the canine monocyte-macrophage cell line DH82. Briefly, both lines were cultured in RPMI 1640 complete medium (RPMI 1640 supplemented with 10% heat-inactivated FCS, 2 mM L-glutamine, 2mM sodium pyruvate, 100 units/ml gentamycin). Cell suspensions in the log phase of growth (10^7 cells/ml RPMI 1640) were treated with 25ug/ml mitomycin for 20 min at 37°C, washed with PBS, and distributed into 24-well microplates at $2x10^5$ cells/well in 500 µl of culture medium. The cells were infected with stationary-phase promastigotes (10 promastigotes/cell) and infection was determined by measuring luciferase activity as described in the Material and methods section in Chapter 1.

As shown in Figure 8, parasites expressing the protein LinJ19.0410 displayed higher luciferase activity in canine macrophages at 48 h post-infection, indicating enhanced virulence of transgenic parasites expressing this protein. Nevertheless, this effect was not observed in infected murine macrophages where luciferase activity was similar to control parasites and the transgenic parasites expressing LinJ36.5780 (Figure 8). Taken together, these results suggest that the secreted protein LinJ19.0410 from *L. infantum* increases its infectivity in cell lines derived from both human and dogs, the main natural mammalian hosts of *Leishmania* parasites.

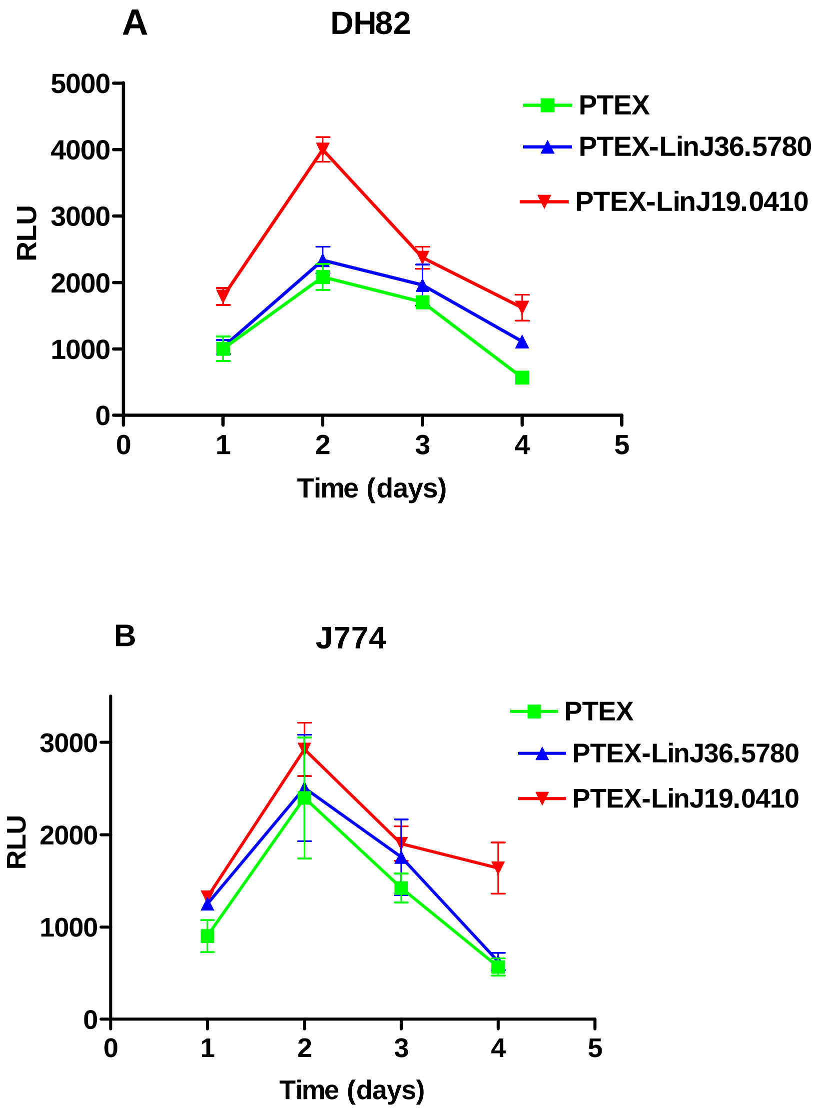


Figure 8: Bioluminescence activity of intracellular *Leishmania* expressing episomal luciferase in infected macrophages *in vitro*.

Transgenic *L. infantum promastigotes* over-expressing the secreted proteins pTEX-LinJ19.0410 (▼) or pTEX-LinJ36.5780 (▲) were co-transfected with the pSP-YαHYGROαLUC carrying the firefly-luciferase gene. Survival of luciferase-expressing parasites was monitored in infected canine monocyte-macrophage cell line DH-82 (A) and murine macrophage cell line J774 (B). Promastigotes transfected with the pTEX vector alone and pSP-αHYGαLUC (■) were used as controls for infection experiments. RLU (Relative Luminescence Units) were measured at various time points post-infection using the Steady Glo reagent. Results are expressed as the mean of three independent experiments, each carried out in triplicate.

4. Heterologous expression of *T. cruzi* secreted proteins in *L. infantum*

The luciferase reporter system was used to asses whether the secreted proteins from *T. cruzi* could modify the infectivity of transgenic *L. infantum* parasites. Although several experiments were performed in order to obtain episomally transfected *T. cruzi* parasites, we were unable to obtain epimastigotes over-expressing the tagged secreted proteins. Transfection procedures in *T. cruzi* are not as efficient and fast as for *Leishmania* parasites. Therefore, we decided to analyze infectivity using a heterologous system with transgenic *L. infantum* promastigotes expressing the *T. cruzi* proteins whose secretion was experimentally confirmed with our methodology.

L. infantum promastigotes expressing *T. cruzi* proteins: Tc00.1047053506155.99, Tc00.1047053505789.10 and Tc00.1047053509999.10 were co-transfected with the firefly luciferase coding vector pSP-αHYGαLUC as described previously (see Material and methods section Chapter 1). Transgenic parasites were selected for their growth in increasing concentrations of Hygromycin (up to 300 μg/ml) over a period of several weeks. Luciferase activity was measured before infection in the transgenic parasites to verify the linear relationship between number of promastigotes and luciferase activity. Likewise, we verified that no significant differences existed in luciferase activity among the transgenic promastigotes before starting infection experiments.

Human monocyte cell line THP-1, canine monocyte-macrophage cell line DH-82 and murine macrophage cell line J774 were infected with transgenic *Leishmania* parasites expressing one of the *T. cruzi* secreted proteins as previously described.

Results of *in-vitro* infections in the different cell lines are depicted in Figure 9. Results are expressed as the mean of three independent experiments, each carried out in triplicate. No significant differences were detected in any of the transgenic parasites when compared to infections with control parasites transfected with the pTEX vector alone and the pSP-αHYGαLUC (Figure 9). Luciferase activity was similar in the different cell cultures infected with any of the transgenic *Leishmania* parasites. A striking observation was detected for parasites expressing the gene Tc00.1047053505789.10 (ortholog of LinJ19.0410). This showed a relatively lower infection success in the three cell lines used for the *in-vitro* infection experiments. This was an unexpected finding given that transgenic *Leishmania* parasites expressing the ortholog of this protein (LinJ19.0410) increase the ability of parasites to survive inside human and canine macrophages. Nevertheless, this observation might be due to the heterologous system used to evaluate *in vitro* infection levels. Thus, the homologous expression of this protein in *T. cruzi* will be necessary to evaluate whether this protein increases the survival of *T. cruzi* parasites inside its target cells, as we detected in *Leishmania*.

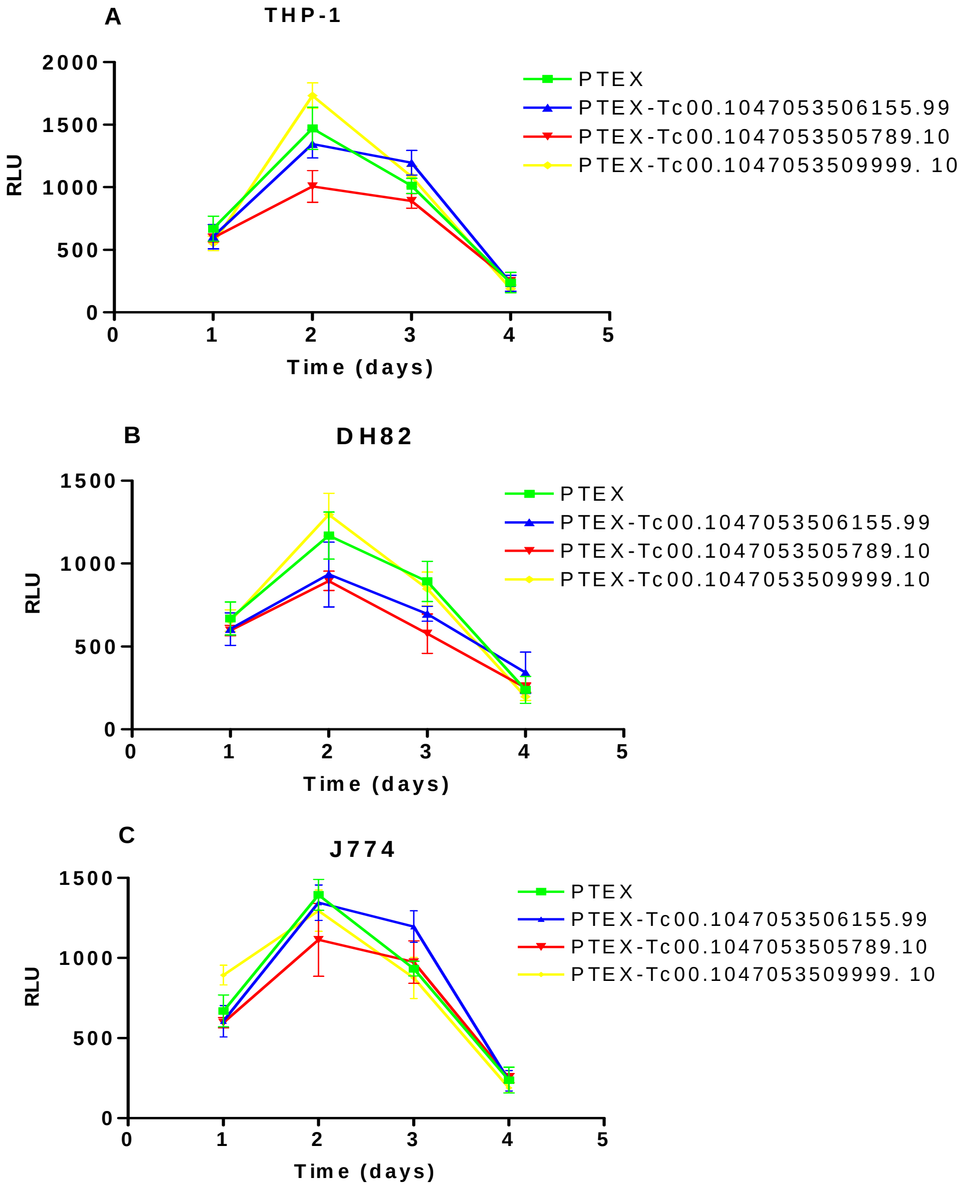


Figure 9: Bioluminescence activity of intracellular *Leishmania* expressing episomal luciferase in infected macrophages *in vitro*.

Transgenic *L. infantum promastigotes* over-expressing the secreted proteins pTEX- Tc00.1047053506155.99 (▲), pTEX- Tc00.1047053505789.10 (▼) or pTEX-Tc00.1047053509999.10 (♦) were co-transfected with the pSP-YαHYGROαLUC carrying the firefly-luciferase gene. Survival of luciferase-expressing parasites was monitored in infected human monocyte cell line THP-1 differentiated into macrophages (A), canine monocyte-macrophage cell line DH-82 (B) and murine macrophage cell line J774 (C). Promastigotes transfected with the pTEX vector alone and pSP-αHYGαLUC (■) were used as controls for infection experiments. RLU (Relative Luminescence Units) were measured at various time points post-infection using the Steady Glo reagent.

D. CHAPTER 2

Study of immunogenic properties of conserved secreted proteins from *Trypanosoma cruzi* identified by a genomic based approach

1. Introduction

Chagas disease is one of the most important parasitic diseases that affect 15 million people in Central and South America. The annual worldwide incidence of new cases is estimated at around 50,000–200,000 (Tarleton et al., 2007). Over the past 15 years, the burden of Chagas disease has been significantly reduced as a consequence of the direct actions promoted by several regional initiatives (Schofield et al., 2006). In the Southern Cone, disease transmission rates by the main vector, *Triatoma infestans*, was controlled in Uruguay, Chile, Brazil and in southern Argentina (Schofield et al., 2006). However limited success was achieved in the Gran Chaco region, which includes northern Argentina, Bolivia and Paraguay. This region is at the core of the *T. infestans* distribution and an area where Chagas disease is still highly prevalent (Tarleton et al., 2007; Vazquez-Prokopec et al., 2009).

The diagnosis of *T. cruzi* infection is difficult because the symptoms of infection are often absent or non-specific, and because the parasites themselves are usually below the level of detection in the infected subjects (Tarleton et al., 2007). Therefore, diagnosis generally depends on the measurement of *T. cruzi*–specific antibodies produced in response to infection.

Conventional serological tests, such as the indirect hemagglutination assay, indirect immunofluorescence assay, and enzyme-linked immunosorbent assay (ELISA), are used widely in countries where the infection is endemic. Most are based on the use of whole or semi-purified antigenic fractions from *T. cruzi* epimastigotes grown in axenic culture. A persistent problem with the conventional assays has been the occurrence of inconclusive and false-positive results (Pirard et al., 2005; Kirchhoff et al., 2006; Caballero et al., 2007). Thus, there is no consensus on which parasite antigen preparation is best for detecting antibodies to *T. cruzi*. The World Health Organization and other expert groups recommend using at least two tests in parallel to confirm *T. cruzi* infection (da Silveira et al., 2001). Due to the lack of a true gold standard for the serologic diagnosis of *T. cruzi* infection, development of new diagnostic tools remains a challenge in trying to control Chagas disease.

High-sensitivity and high-specificity ELISA methods using recombinant or synthetic peptides as antigens have been reported (da Silveira et al., 2001; Umezawa et al., 2004; Cheng et al., 2007; Cooley et al., 2008). The inclusion of recombinant antigens and synthetic peptides for serological diagnosis of *T. cruzi* infections has been advantageous in terms of increasing their specificity (da Silveira et al., 2001). Nevertheless, single recombinant antigens are less sensitive than conventional tests using whole parasite extracts (da Silveira et al., 2001). The use of cocktails of recombinant antigens, mixtures of synthetic peptides or multi-epitope antigens have been shown to increase sensitivity (da Silveira et al., 2001; Umezawa et al., 2004; Cheng et al., 2007).

It is widely assumed that secreted/excreted factors in *T. cruzi* are highly immunogenic. Indeed, trypomastigote forms release several antigens into the supernatant of infected cell cultures. This complex mixture of antigens, termed TESA (trypomastigote excretory-secretory antigens), is highly immunogenic and has been used for the diagnosis of both acute and chronic Chagas disease (Nakazawa et al., 2001; Berrizbeitia et al., 2006). Remarkably, the components of the TESA mixture are currently unknown.

We designed an experimental approach based on bioinformatic analyses of the *T. cruzi* genome sequence to identify secreted proteins that are conserved among trypanosomatids and involved in the classical secretory pathway. Our methodology allowed us to identify three new secreted proteins of unknown function in trypanosomatids. Given that *T. cruzi* secreted proteins are highly immunogenic, we investigated whether these secreted proteins are immunogenic. In the present chapter we present data on the antigenic properties of *T. cruzi* recombinant proteins by analyzing human antibody responses from chagasic patients. Our results show that one of these secreted proteins is immunogenic and may represent a new diagnostic tool for Chagas disease.

2. Material and methods

2.1 Bacterial strains and plasmids

Escherichia coli TOP10 (Invitrogen) was used for cloning and *E. coli* strains used for expression were primarily BL21 (DE3), Rosetta-gami (DE3) pLysS, AD494 (DE3) pLysS, OrgamiB (DE3) pLysS from Novagen, Madison, WI, USA and BL 21 (DE3) pLysS, BL21Star (DE3) pLysS from Invitrogen. Coding sequences for the three secreted proteins from *T. cruzi* were cloned under control of the T7lac promoter. The plasmid vector used was pET-21b and pET-28a (Novagen, Madison, Wis.) which confer resistance to ampicillin and kanamycin, respectively. PET-21b allows expression of the target protein fused to a C-terminal tail of six histidines. The pET-28a vector carries an N-terminal and C-terminal His Tag sequence.

2.2 Cloning of *T. cruzi* genes into the expression vector

T. cruzi genes encoding the secreted proteins were amplified from genomic DNA *(*TcY7 clone derived from the Y strain) by PCR using specific forward y reverse primers (listed in table I). Genes Tc00.1047053506155.99 and Tc00.1047053509999.10 were cloned without the amino-acid sequence corresponding to the predicted secretion N-signal peptide. *T. cruzi* gene Tc00.1047053505789.10 was cloned both with and without the predicted secretion N-signal peptide. PCR reactions were carried out in 20 µl using 0.5 µM of each primer, 0.2 mM dNTP, 0.4 U of Phusion high-fidelity polymerase (Finnzymes, Espoo, Finland) and the following cycling conditions: 98°C for 30 s followed by 25 cycles of 98°C for 10 s, 64°C to 68°C for 15 s, 72°C for 25 to 60 s (according to gene size), and a 72°C elongation for 5 min. Digested and purified fragments were inserted into the dephosphorylated pET-21b vector digested with the corresponding restriction enzymes. Cloned sequences were confirmed by restriction digestion and sequencing. Large scale preparations of the different constructs were performed using the plasmid midi kit (Promega).

Table 1 Gene ID of *T. cruzi* genes and primers used for cloning into the expression vector pET21b

Gene ID	Primer sequences[a]	F/R product sizes (bp)	MW (kDa)[b]
Tc00.1047053506155.99	F AGTCGTCA*CATATG*CCCGATCCTCTGGAGCATGC	834	28.1
	R ATCAACTGACT*GAATTC*GCGACAAGTTCGTGGCATGTAATTGTTC		
Tc00.1047053509999.10	F CATCGCAG*CATATG*GCAGAAGAGGAGGACGTGAGG		
	R GCACTGACT*CTCGAG*GCCGCACCAGCGCTCCAGAA	1182	45.8
Tc00.1047053505789.10	F CAGCTCCTCCA*CATATG*CCCTCTGGCAAAGCAACTGC	1512	55.5
	F WSP CAGCTCCTCCA*CATATG*CAGCGCTCACTCGACTGCCA	1416	52.2
	R GCACTGACT*CTCGAG*GGCAGCATTTACCGACCCTGAC		

F, Forward primer and R, Reverse primer used for the amplification of sequence gene without the predicted secretion N-Signal peptide. FSP Forward primer used for the amplification of the gene Tc00.1047053505789.10 with the predicted N-Signal peptide sequence.

[a] Restriction sites used for cloning in the pET21b vector in italics.

[b] Expected molecular weight of the proteins.

2.3 Expression and purification of *T. cruzi* recombinant proteins

E. coli expression hosts were transformed with plasmids carrying the target protein. Recombinant *E. coli* were propagated overnight at 37°C with shaking in auto-inducing media ZYP-5052 (Studier FW, 2005) containing 100µg/ml of ampicillin. Cleared *E. coli* lysates were obtained from cell pellets (50-ml cultures) re-suspended in 5 ml of buffer containing 100 mM NaH_2PO_4, 10 mM Tris-Cl and 8 M urea pH 8.0. The clarified supernatant was applied to a Ni-NTA agarose column (Qiagen) and the recombinant proteins were eluted at pH 4.5 (Urea 8M). To assess the purity of the proteins, eluted fractions were analyzed in a SDS-PAGE gel by staining with Coomassie blue and/or by immunoblotting with an Anti-His (C-Term)-HRP antibody as detailed below.

2.4 SDS-PAGE

SDS-PAGE analysis was performed on a NuPAGE Bis-Tris gel (4-12%) in MOPS-SDS running buffer (Invitrogen) under reducing conditions (50 mM DTT). The gels were either stained with Coomassie blue R or electroblotted onto a PVDF membrane (Hybond-P, Amersham) for western blot analyses.

2.5 Western blot analysis

The PVDF membranes were rinsed twice in TBS and used for immuno-detection of the recombinant proteins with an Anti-His (C-Term)-HRP antibody (Invitrogen) according to the manufacturers' instructions. Membranes used for detection of anti-*T. cruzi* antibodies in patient sera were rinsed twice in TBS and blocked overnight in TBS-Tween20 0.1% and 5% non-fat milk. Strips of paper (5 mm) were then cut and incubated separately with a 1:2500 dilution of human serum for 2 h at room temperature. Strips were then washed five times (5 min) in PBS-Tween 20 0.1% and incubated for 1 h at room temperature with HRP-conjugated anti-human IgG (heavy and light chain specific) (Nordic Immunology) diluted 20,000 times in TBS-Tween 20 01% and 5% non-fat milk. The strips were washed as before, and the immune complexes were revealed by chemiluminescence emission using the ECL Plus Western blotting detection system and ECL Hyperfilms (GE Healthcare, UK).

2.6 Patient sera

Positive sera were obtained from infected chagasic individuals from an endemic region located in northeast Argentina (Chaco Province), a vector-free city located in north-western Argentina (Salta-city), and from the Bolivian community of Tupiza (Potosi department), an area where vector transmission is controlled. Negative sera were obtained from healthy blood donors of the same regions, respectively. *T. cruzi* infection status was determined by using two conventional tests: commercial ELISA and IHA based on parasite homogenate antigens. Positive sera for both reactions were considered as true positives. Negative sera for both reactions were considered as true negatives.

3. Results

3.1 Expression and purification of recombinant proteins rTc00.1047053506155.99 and rTc00.1047053509999.10 from *T. cruzi*

In order to determine the antigenic properties of the three secreted proteins of *T. cruzi,* we attempted to produce the recombinant proteins in *E. coli* to subsequently analyse their reactivity with chagasic patient sera. After verifying the open reading frame in the plasmid pET21b, protein expression was assessed in transformed bacteria carrying the recombinant plasmid. Two out of the three recombinant proteins were expressed and purified under denaturing conditions in *E. coli* BL21 (DE3). This bacterial strain showed the greatest expression of the proteins when compared to the other strains tested.

Figure 1 shows the purification of the recombinant protein rTc00.1047053506155.99 and rTc00.1047053509999.10 with molecular weights of 28 KDa and 45KDa, respectively. The purified recombinant proteins without the signal peptide sequence appeared as bands of expected size.

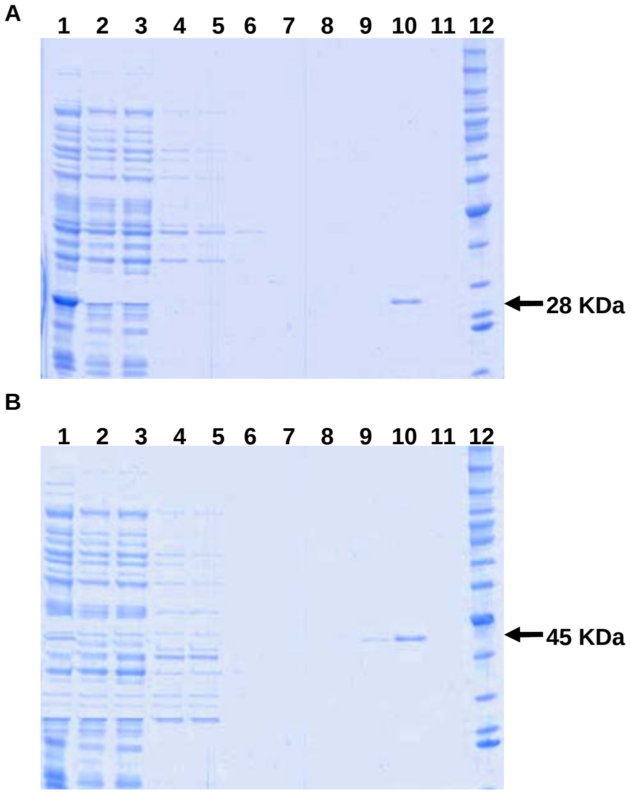


Figure 1: Purification in denaturing conditions of the recombinant proteins: **(A)** rTc00.1047053506155.99 (28 KDa) and **(B)** rTc00.1047053509999.10 (45KDa). The proteins were resolved by electrophoresis in 4-12% gradient gels (Invitrogen) and visualized by Coomassie staining. Line1: cleared lysate. Line 2: flow-through. Lines 3,4,5: Washes. Lines 6,7,8: eluates at pH 5,9. Lines 9,10,11: Eluates at pH 4,5. Line 12: Molecular marker (BenchMark Protein Ladder, Invitrogen).

3.2 Expression of recombinant protein rTc00.1047053505789.10 from *T. cruzi*

The *T. cruzi* gene Tc00.1047053505789.10 codes a hypothetical conserved protein with a predicted molecular weight of 53.5 KDa. Likewise, the protein carries a signal peptide sequence and five putative transmembrane domains. Several strategies were used in order to produce the recombinant protein in *E. coli*. Initially, the full length gene, including the signal peptide sequence was cloned into the plasmid pET21b. The protein sequence was confirmed by sequencing and the recombinant plasmid was transformed into the different expression hosts. Protein expression analyses of the recombinant bacteria revealed limited expression of the target protein that was only detected by western blot with an antibody directed against the His-Tag carried by the recombinant protein (Figure 2). Likewise, western blot analyses revealed the presence of a protein with a molecular weight (43KDa), lower than the expected size (55.5 KDa), suggesting the production of a truncated protein by *E. coli*. As shown in Figure 2, the protein was mainly detected in the insoluble fraction of the bacterial extracts, indicating the aggregation of protein in inclusion bodies.

Although the recombinant protein was present in insoluble aggregates, we attempted to purify the protein under denaturing conditions. However, we were unable to obtain the purified protein. Several factors influence the expression of recombinant proteins in *E. coli*. The bacterial strain, culture media, and temperature are among the factors that can be modified without altering the target protein in the recombinant plasmid (Sorensen H and Al, 2005). The different conditions tested (bacterial strains, induction temperature, culture media) did not improve the expression of the recombinant protein that was detected as insoluble and in small quantities. Therefore, we decided to modify the coding sequence cloned into the plasmid. Because secretory signal peptides may have a toxic effect on the bacteria due to its hydrophobic character, the protein coding gene was cloned into pET21b without the secretory signal peptide. Nevertheless, no expression was detected in recombinant bacteria by western blot analyses. This problem was not associated with the protein coding gene in the recombinant plasmid since the cloned sequence was confirmed by sequencing.

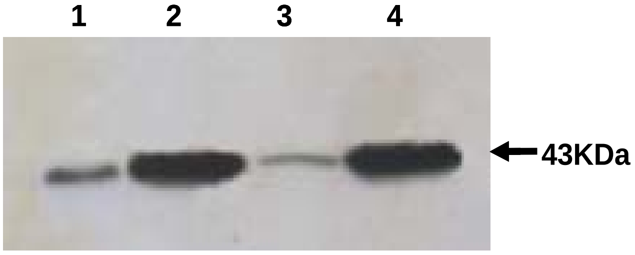


Figure 2. Western blot analysis of rTc00.1047053505789.10. His-tagged proteins were detected in whole soluble extracts (Line 1 and 3) and the insoluble fraction of recombinant *E. coli* BL21(DE3). The Tagged proteins were detected mainly in the insoluble fraction (Line 2 and 4) and with a lower molecular weight (43KDa) than expected (55.5KDa).

3.3 Reactivity of sera from chagasic patients with *Trypanosoma cruzi* recombinant proteins coding conserved secreted proteins

The ability of the recombinant proteins rTc00.1047053506155.99 and rTc00.1047053509999.10 to detect chagasic infections was evaluated by Western Blot. Both proteins were purified in denaturing conditions (Figure 1). The protein rTc00.1047053506155.99 was not recognized by sera from either chagasic or healthy donors. This was an unexpected finding given that this protein is predicted by the TDR Targets Database as one of the most antigenic proteins in the *T. cruzi* genome. Indeed, this protein belongs in the top 0.7% of the most antigenic proteins in the genome of *Trypanosoma cruzi* based on an antigenicity index obtained from putative antigenic peptides for each protein annotated in the genome (http://tdrtargets.org/targets/view?gene_id=49810). In order to test whether denaturing conditions of the recombinant protein could modify its ability to identify antibodies to *T. cruzi*, we purified the protein in native condition. Nevertheless, when testing the antigenic properties of the soluble rTc00.1047053506155.99 protein, no antibodies to *T. cruzi* were detected by western blot analysis. Therefore, this protein was not investigated further.

Although the protein rTc00.1047053506155.99 was not detected as immunogenic, the recombinant protein rTc00.1047053509999.10 was recognized by sera from chagasic patients. As shown in Figure 3, this protein was detected by chagasic sera as a single band at 45 KDa, corresponding to the molecular weight of the recombinant protein. Anti-*T. cruzi* antibodies were detected in 80% (18/22) of the sera from chagasic patients. The sera analysed came from patients in different regions of Latin America. Six sera originated from patients living in North-western Argentina (Salta-city) and diagnosed with heart failure, as detected by electrocardiograms and echocardiograms. Five of these sera were positive by Western blot analyses. Among six sera from children in Northeast Argentina (Chaco-province), only five recognized the recombinant protein. The remaining sera correspond to five Bolivian children and five Mexican patients that recognized the *T. cruzi* protein in four out of the five sera tested.

Sera from healthy donors did not recognize the recombinant protein. Furthermore, we evaluated potential cross-reactivity with antibodies from five patients infected with *L. infantum*. The recombinant protein was not recognized by these sera, suggesting that the antibodies detected by chagasic patients are specific for *T. cruzi* infections. Together these results suggest that the recombinant protein rTc00.1047053509999.10 may be useful to develop a serological test to detect the presence of *T. cruzi*–specific antibodies.

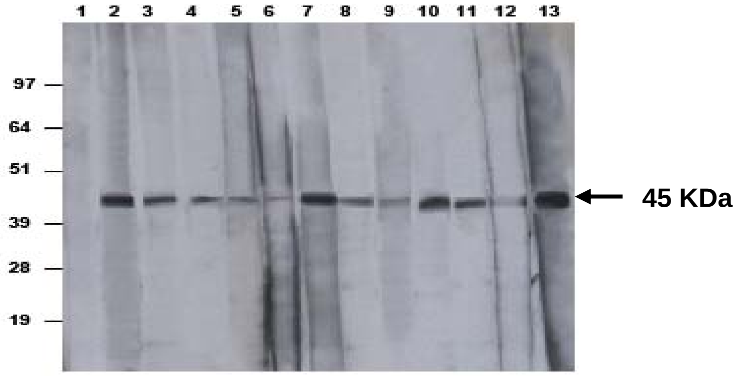


Figure 3. Western blot of the fusion protein rTc00.1047053509999.10 with sera from non chagasic individuals (lanes 1) and sera from chagasic patients (lines 2-13)

4. Discussion

Although Chagas disease was discovered a century ago, diagnosis of infection by *T. cruzi* remains a key issue (Tarleton et al., 2007). Without effective diagnostics, infected individuals cannot be identified and thus treated, and the effectiveness of treatment cannot be efficiently assessed. Moreover, the effectiveness of any control campaign, whether targeted at insect vectors, blocking of transmission, or vaccination of individuals, cannot be measured without competent diagnostics. Transmission via blood transfusion or tissue transplantation has been a point of concern for many years in Latin America, but has only come to the fore in the US and Europe as the number of immigrants infected by *T. cruzi* has increased (Bern et al., 2008).

No assay has been universally accepted as the gold standard for the serologic diagnosis of *T. cruzi* infection, and likewise no assay is viewed as a definitive confirmatory test (Tarleton et al., 2007). Therefore, there is a compelling need for a supplemental assay that would serve for the accurate diagnosis of *T. cruzi* infections. High-sensitivity and high-specificity ELISA methods using recombinant or synthetic peptides as antigens have been shown to improve diagnosis of *T. cruzi* infections (da Silveira et al., 2001; Umezawa et al., 2004; Cheng et al., 2007).

Previously, we have identified three new secreted proteins that are conserved among trypanosomatids and involved in the classical secretory pathway. The function of these proteins cannot be predicted from their sequences, subsequently their annotation in the genome sequences of trypanosomatids are hypothetical conserved proteins. Given that secreted proteins are targets of immune response in *T. cruzi* infections, we tested whether these secreted proteins are immunogenic. We attempted to produce these *T. cruzi* proteins in the classical prokaryote system (*E. coli*) to subsequently test their ability to recognize antibodies in chagasic patients.

Among the various expression systems employed for the over-production of proteins, bacteria still remains the favourite choice. This is because they offer several advantages, including the relative simplicity for protein expression, its relative inexpense, and they can be rapidly cultivated to high densities (Sahdev et al., 2008). Nevertheless, no universal approach has been established for the efficient expression of target proteins in *E. coli*. Several factors affect the expression of target proteins, including the expression host, expression vector and culture conditions (Sahdev et al., 2008). There are two main types of strategy used to improve protein expression; those which avoid modifying the target and those where the target sequence is engineered (Sahdev et al., 2008).

Although there were difficulties encountered in the expression of the *T. cruzi* target proteins in *E. coli* (small quantity and truncated proteins), the elimination of the sequence coding for the secretory signal peptide allowed us to produce two out the three secreted proteins (rTc00.1047053506155.99 and rTc00.1047053509999.10).

Unfortunately, we were unable to obtain the recombinant protein Tc00.1047053505789.10. We were particularly interested in producing this protein because previous studies suggested that this protein is involved in a process increasing *in-vitro* infection of *Leishmania* parasites.

We tested several strategies to improve protein production, including different expression hosts, different culture conditions, and through the elimination of the sequence coding for the signal peptide. In spite of these attempts, the low amount and insoluble form of the protein produced by *E. coli* prevented the purification of protein Tc00.1047053505789.10. A plausible explanation for the difficulties encountered using *E. coli* is its lack of lack of post-translational machinery that may prevent the correct processing and thus expression of the protein. Currently, this protein is being produced using a eukaryote baculoviral system (Invitrogen) that may improve the expression and purification of the *T. cruzi* protein, so its immunogenicity can be determined.

Regarding the recombinant proteins produced, an unexpected result was that the protein rTc00.1047053506155.99 was not found to be immunogenic in either denaturing or native conditions. This protein is predicted by the TDR Targets Database as one of the most immunogenic in the genome of *T. cruzi*. In spite of the fact that this protein possesses 14 predicted antigenic epitopes, the recombinant form was not detected by chagasic sera. These results point out that *in silico* predictions will not always correlate with experimental data. Therefore further experimental validation is needed to confirm genome-based predictions.

Antigenic studies of the recombinant protein rTc00.1047053509999.10 revealed its ability to recognize anti-*T. cruzi* antibodies in chagasic patients from different endemic regions. Antibodies were detected in 80% (18/22) of the sera analysed by western blot. Interestingly the *T. cruzi* protein was not recognized by patients infected with *L. infantum*. Because *T. cruzi* and *Leishmania* parasites are antigenically related, cross-reactivity poses a serious problem for specific serodiagnosis. Moreover, cross-reactivity with *Leishmania* is a major problem for diagnosis in areas where the geographical distribution of Chagas disease and leishmaniasis overlaps, including north-western Argentina (Chiaramonte et al., 1996; Mendes et al., 2007). Therefore, the use of purified and specific antigens is the only means for obtaining an unambiguous and reliable diagnosis of these parasitic diseases (Frank et al., 2003; Marcipar et al., 2005). In this regard the recombinant protein rTc00.1047053509999.10 may represent an interesting antigen to identify anti-*T. cruzi* antibodies. The use of synthetic peptides and recombinant proteins as target antigens in diagnostic assays, such as ELISA, can be more specific than those based on antigens derived from parasite lysates (da Silveira et al., 2001; Meira et al., 2002). Furthermore ELISA assays based on a mixture of recombinant proteins can have high levels of sensitivity (Umezawa et al., 2003; Umezawa et al., 2004). Likewise, immunoblotting has been used successfully as a confirmatory tool to detect antibodies to *T. cruzi* and other pathogens, such as, *Borrelia burgdorferi* the agent of Lyme disease (Aguero-Rosenfeld et al., 2005; Cheng et al., 2007).

Although we analysed the recombinant protein rTc00.1047053509999.10 with a limited number of sera, our results provide evidence to support the further development of an ELISA test to analyse antibodies to *T. cruzi* in a quantitative manner. Analyzing its sensitivity and specificity for the diagnosis of Chagas disease may reveal its utility to develop either an immunoblot assay or an ELISA test based on mixtures of recombinant proteins.

5. References

Aguero-Rosenfeld, M.E., Wang, G., Schwartz, I., Wormser, G.P., 2005. Diagnosis of lyme borreliosis. Clin Microbiol Rev 18, 484-509.

Bern, C., Montgomery, S.P., Katz, L., Caglioti, S., Stramer, S.L., 2008. Chagas disease and the US blood supply. Curr Opin Infect Dis 21, 476-482.

Berrizbeitia, M., Ndao, M., Bubis, J., Gottschalk, M., Ache, A., Lacouture, S., Medina, M., Ward, B.J., 2006. Purified excreted-secreted antigens from Trypanosoma cruzi trypomastigotes as tools for diagnosis of Chagas' disease. J Clin Microbiol 44, 291-296.

Caballero, Z.C., Sousa, O.E., Marques, W.P., Saez-Alquezar, A., Umezawa, E.S., 2007. Evaluation of serological tests to identify Trypanosoma cruzi infection in humans and determine cross-reactivity with Trypanosoma rangeli and Leishmania spp. Clin Vaccine Immunol 14, 1045-1049.

Cheng, K.Y., Chang, C.D., Salbilla, V.A., Kirchhoff, L.V., Leiby, D.A., Schochetman, G., Shah, D.O., 2007. Immunoblot assay using recombinant antigens as a supplemental test to confirm the presence of antibodies to Trypanosoma cruzi. Clin Vaccine Immunol 14, 355-361.

Chiaramonte, M.G., Zwirner, N.W., Caropresi, S.L., Taranto, N.J., Malchiodi, E.L., 1996. Trypanosoma cruzi and Leishmania spp. human mixed infection. Am J Trop Med Hyg 54, 271-273.

Cooley, G., Etheridge, R.D., Boehlke, C., Bundy, B., Weatherly, D.B., Minning, T., Haney, M., Postan, M., Laucella, S., Tarleton, R.L., 2008. High Throughput Selection of Effective Serodiagnostics for Trypanosoma cruzi infection. PLoS Negl Trop Dis 2, e316.

da Silveira, J.F., Umezawa, E.S., Luquetti, A.O., 2001. Chagas disease: recombinant Trypanosoma cruzi antigens for serological diagnosis. Trends Parasitol 17, 286-291.

Frank, F.M., Fernandez, M.M., Taranto, N.J., Cajal, S.P., Margni, R.A., Castro, E., Thomaz-Soccol, V., Malchiodi, E.L., 2003. Characterization of human infection by Leishmania spp. in the Northwest of Argentina: immune response, double infection with Trypanosoma cruzi and species of Leishmania involved. Parasitology 126, 31-39.

Kirchhoff, L.V., Paredes, P., Lomeli-Guerrero, A., Paredes-Espinoza, M., Ron-Guerrero, C.S., Delgado-Mejia, M., Pena-Munoz, J.G., 2006. Transfusion-associated Chagas disease (American trypanosomiasis) in Mexico: implications for transfusion medicine in the United States. Transfusion 46, 298-304.

Marcipar, I.S., Roodveldt, C., Corradi, G., Cabeza, M.L., Brito, M.E., Winter, L.M., Marcipar, A.J., Silber, A.M., 2005. Use of full-length recombinant calflagin and its c fragment for improvement of diagnosis of Trypanosoma cruzi infection. J Clin Microbiol 43, 5498-5503.

Meira, W.S., Galvao, L.M., Gontijo, E.D., Machado-Coelho, G.L., Norris, K.A., Chiari, E., 2002. Trypanosoma cruzi recombinant complement regulatory protein: a novel antigen for use in an enzyme-linked immunosorbent assay for diagnosis of Chagas' disease. J Clin Microbiol 40, 3735-3740.

Mendes, D.G., Lauria-Pires, L., Nitz, N., Lozzi, S.P., Nascimento, R.J., Monteiro, P.S., Rebelo, M.M., Rosa Ade, C., Santana, J.M., Teixeira, A.R., 2007. Exposure to mixed asymptomatic infections with Trypanosoma cruzi, Leishmania braziliensis and Leishmania chagasi in the human population of the greater Amazon. Trop Med Int Health 12, 629-636.

Nakazawa, M., Rosa, D.S., Pereira, V.R., Moura, M.O., Furtado, V.C., Souza, W.V., Barros, M.N., Abath, F.G., Gomes, Y.M., 2001. Excretory-secretory antigens of Trypanosoma cruzi are potentially useful for serodiagnosis of chronic Chagas' disease. Clin Diagn Lab Immunol 8, 1024-1027.

Pirard, M., Iihoshi, N., Boelaert, M., Basanta, P., Lopez, F., Van der Stuyft, P., 2005. The validity of serologic tests for Trypanosoma cruzi and the effectiveness of transfusional screening strategies in a hyperendemic region. Transfusion 45, 554-561.

Sahdev, S., Khattar, S.K., Saini, K.S., 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307, 249-264.

Schofield, C.J., Jannin, J., Salvatella, R., 2006. The future of Chagas disease control. Trends Parasitol 22, 583-588.

Tarleton, R.L., Reithinger, R., Urbina, J.A., Kitron, U., Gurtler, R.E., 2007. The challenges of Chagas Disease-- grim outlook or glimmer of hope. PLoS Med 4, e332.

Umezawa, E.S., Bastos, S.F., Coura, J.R., Levin, M.J., Gonzalez, A., Rangel-Aldao, R., Zingales, B., Luquetti, A.O., da Silveira, J.F., 2003. An improved serodiagnostic test for Chagas' disease employing a mixture of Trypanosoma cruzi recombinant antigens. Transfusion 43, 91-97.

Umezawa, E.S., Luquetti, A.O., Levitus, G., Ponce, C., Ponce, E., Henriquez, D., Revollo, S., Espinoza, B., Sousa, O., Khan, B., da Silveira, J.F., 2004. Serodiagnosis of chronic and acute Chagas' disease with Trypanosoma cruzi recombinant proteins: results of a collaborative study in six Latin American countries. J Clin Microbiol 42, 449-452.

Vazquez-Prokopec, G.M., Spillmann, C., Zaidenberg, M., Kitron, U., Gurtler, R.E., 2009. Cost-effectiveness of chagas disease vector control strategies in northwestern Argentina. PLoS Negl Trop Dis 3, e363.

E. CHAPTER 3

Review:

Deciphering the Leishmania exoproteome: What we know and what we can learn. **Rosa M Corrales**, Denis Sereno, Françoise Mathieu-Daudé, FEMS *Immunology & Medical Microbiology*, submitted

http://mc.manuscriptcentral.com/fems

Deciphering the Leishmania exoproteome: What we know and what we can learn

Journal:	*FEMS Immunology & Medical Microbiology*
Manuscript ID:	draft
Manuscript Type:	Mini Review
Date Submitted by the Author:	
Complete List of Authors:	Corrales, Rosa; Institut de Recherche pour le Développement, Département Sociétés et Santé Sereno, Denis Mathieu-Daudé, Françoise
Keywords:	secretion mechanisms, Leishmania secretome, secreted virulence factors

Deciphering the *Leishmania* exoproteome: What we know and what we can learn

Rosa Milagros Corrales [1*], Denis Sereno [1] and Françoise Mathieu-Daudé [1]

[1] Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Institut de Recherche pour le Développement, 911 Avenue Agropolis, 34394 Montpellier cedex 05, France.

Running title: Deciphering the *Leishmania* exoproteome

Keywords: secretion mechanisms, *Leishmania* secretome, secreted virulence factors

[*] Corresponding author: Rosa M Corrales, Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Institut de Recherche pour le Développement, 911 Av. Agropolis, 34394 Montpellier, France.

Tel.: +33 4 67 41 61 47; fax: +33 4 67 41 61 47.

Email address: corrales@mpl.ird.fr

Abstract

Parasitic protozoa of the genus *Leishmania* are the causative agents of leishmaniasis. Survival and transmission of these parasites in their different hosts requires membrane bound or extracellular factors to interact with and modify their host environments. Over the last decade, several approaches have been applied to the study of all extracellular proteins exported by an organism at a particular time and under defined conditions, collectively termed the secretome or exoproteome. In this review, we focus on emerging data shedding light on the secretion mechanisms involved in the production of the *Leishmania* exoproteome. We also describe other methodologies currently available that could be used to analyze the *Leishmania* exoproteome. Understanding the complexity of the *Leishmania* exoproteome is a key component to elucidate mechanisms used by these parasites to export proteins to the extracellular space during its life cycle. Given the importance of extracellular factors, a detailed knowledge of the *Leishmania* exoproteome may provide novel targets for rational drug design and/or a source of antigens for vaccine development.

1. Introduction

Parasitic protozoa of the genus *Leishmania* are the causative agents of leishmaniasis and are transmitted between mammalian hosts by female phlebotomine sand flies. *Leishmania* infections can range from mild self-healing skin lesions to fatal visceral infections depending on the *Leishmania* species involved (McMahon-Pratt & Alexander, 2004). Currently, it is estimated that 12 million people worldwide are affected and 2 million new cases are believed to occur each year in large areas of the tropics, subtropics and the Mediterranean basin (Stuart*, et al.*, 2008). Control of leishmaniasis has been hampered by the absence of a vaccine, limited efficacy of frontline drugs, and increased transmission as a result of co-infections with HIV (Croft*, et al.*, 2006). Therefore, a detailed understanding of all aspects of *Leishmania* biology is desirable to help formulate new antiparasitic strategies.

To complete their life cycle, *Leishmania* spp. have to adapt and develop in an insect vector and a vertebrate host. In the sandfly, *Leishmania* replicate as motile flagellated extracellular cells known as promastigotes, which reside primarily in the insect alimentary tract. Two main forms can be distinguished, (i) multiplicative, but not mammalian-infective, procyclic promastigotes that are present in the insect midgut, and (ii) non-dividing but mammalian-infective, metacyclic promastigotes found in the thoracic midgut and proboscis of the sandfly (Bates, 2007). When inoculated into a mammalian host through a sandfly bite, the metacyclic promastigotes are phagocytosed by macrophages and differentiate into intracellular aflagellate amastigotes. This amastigote form resides within a vacuole with lysosomal features, that is termed the parasitophorous vacuole (Kima, 2007). Amastigotes proliferate by binary cell division and can spread to other macrophages as well as some other phagocytic and non-professional phagocytic cells (e.g., dendritic cells and fibroblasts) reviewed in (Kima, 2007). The different morphological forms of *Leishmania* represent an adaptation to the changing environmental conditions encountered by the parasites within their two hosts. These parasites require various effectors that are membrane bound or released extracellularly to ensure their survival and transmission in the different hosts. Promastigote surface coat constituents have been the focus of numerous studies, especially the lypophosphoglycans (LPGs) and the promastigote surface protease named glycoprotein 63 (gp63), both of which are involved in the parasite's virulence in the insect and the mammalian host (Descoteaux*, et al.*, 2002, Spath*, et al.*, 2003, Kulkarni*, et al.*, 2006, Santos*, et al.*, 2006). Extracellular components produced by *Leishmania* parasites have been mainly studied for their antigenic properties. Indeed, "exogenous antigens" have proved to be highly immunogenic, elicit strong immunity and protection against infection in mice and dogs (Tonui*, et al.*, 2004, Lemesre*, et al.*, 2007). Therefore, extracellular factors could be a source of antigens for vaccine development as demonstrated for the pathogen *Mycobacterium tuberculosis* (Horwitz*, et al.*, 2005). Despite the potential importance of *Leishmania* extracellular components in these host-parasite interactions, including regulation of host immune response, only a few extracellular proteins have been fully characterized. Moreover, little is known about the mechanism(s) predominantly used by *Leishmania* to export proteins to the extracellular environment during its life cycle.

Over the last decade several approaches have been used to study proteins exported by an organism at a given time or under certain conditions. Collectively these extracellular proteins are termed the secretome or exoproteome (Hathout, 2007). To fully understand functionality of the *Leishmania* exoproteome, it is important to characterize the individual proteins and to address their expression and their role in the different morphological forms. Indeed, the composition of the parasite exoproteome may be variable, depending on the parasite stage under consideration. It could also differ according to the relative contribution of different pathways operating in protein secretion. The study of the exoproteome might provide an overview of the complex relations between *Leishmania* parasites and their different hosts.

Thus, *Leishmania* exoproteome studies may represent a promising strategy to identify novel targets for rational drug design and/or a source of antigens for vaccine development. This review will focus on the mechanisms involved in producing the *Leishmania* exoproteome, as well the methodologies currently available and that could be used for its study. The importance of extracellular components in host-parasite interactions is also discussed.

2. Origins of the secretome

The term "secretome" was first introduced in a study dealing with a genome-based global survey of secreted proteins of *Bacillus subtilis* (Tjalsma*, et al.*, 2000). The authors defined the secretome as a subset of the proteome including secreted proteins and the components of machineries for protein secretion. By using computational methods the authors predicted all the proteins *B. subtilis* exports by searching for signal peptides and cell retention signals in the protein sequences. Further characterization of the *B. subtilis* secretome by using a proteomic approach showed that the original prediction correctly identified about 50% of the proteins (Antelmann*, et al.*, 2001). These studies together with the availability of the complete genome sequences of several organisms opened the door for the identification and analysis of the secretome both in prokaryotes and eukaryotes.

Different approaches and definitions have been used to identify the secretome in different organisms. In eukaryotes, the term secretome has been used to describe different subsets of the proteome, including (i) all proteins processed through the secretory pathway (Klee, 2008), (ii) the proteins processed through the secretory pathway that lack transmembrane domains and/or a GPI anchor signal (Grimmond*, et al.*, 2003, Lee*, et al.*, 2003), or (iii) the subset of proteins identified in the extracellular proteome (Zwickl*, et al.*, 2005, Chevallet*, et al.*, 2007, Paper*, et al.*, 2007). From a proteomic perspective, the mammalian secretome was defined as the quantitative map of distribution of all proteins and lipids of the classical secretory pathway (Simpson*, et al.*, 2007). These studies reveal the term "secretome" has been used (or misused) in a variety of ways.

There is a conspicuous lack of studies involving the secretome of protozoan parasites, despite the importance of its role in parasite virulence and modulation of the host immune response. So far, the most particular definition of secretome refers to the malaria agent *Plasmodium falciparum*. In this parasite, the secretome describes all the proteins exported into the host erythrocyte mediated by an ER-type signal sequence and a downstream Host Targeting (HT)-motif or Plasmodium Export Element (PEXEL) (Hiller*, et al.*, 2004, van Ooij*, et al.*, 2008). For trypanosomatids the term “secretome” was introduced recently by Silverman *et al*. They used a sensitive proteomic approach to identify a large number of extracellular proteins in culture medium conditioned by *L. donovani* (Silverman*, et al.*, 2008). Likewise, a classical proteomic strategy was used to define the secretome as the naturally “excreted/secreted” factors of *T. congolense* and *T. evansi* (Holzmuller*, et al.*, 2008, Grebaut*, et al.*, 2009). In the last two decades, several studies involving trypanosomatids have attempted to identify and characterize “excreted/secreted factors” owing to their potential for vaccine development and/or drug targets. It is well known that *Trypanosoma cruzi* and *Leishmania spp.* not only secrete proteins to the extracellular space but also shed several factors *in vitro* (Jazin*, et al.*, 1991, Yokoyama-Yasunaka*, et al.*, 1994, Jaffe & Dwyer, 2003). For example, members of the LPG family are proteins which are shed *in vitro* into the extracellular medium of *Leishmania* cells (Ilg*, et al.*, 1994). This protein family is involved in the classical secretory pathway to reach their final destination at the cell surface. However the mechanism(s) by which these and other factors are shed by *Leishmania spp.* are largely unknown. Thus, the term “excreted/secreted factors” was used to include all the molecules found outside the cell, including *bona-fide* secreted proteins. For the purpose of this review, we will describe the *Leishmania* secretome based on the definition of “secretion”, meaning the transport of a protein from the inside to the outside of a cell through any of the secretion mechanisms known to date for eukaryotes. Therefore, the secretome will refer to proteins that are actively secreted from the cell using a classical or non-classical mechanism or to secretions mediated by exosomes. The term "exoproteome" will be defined as the set of *Leishmania* proteins present in the extracellular space and thus include non-secreted proteins (e.g. extracellular proteins originating from cell lysis) and those actively secreted (the secretome). Given the complexity of the composition of the extracellular *Leishmania* proteome, we suggest that the use of "exoproteome" is a more suitable term to avoid the misnomer of secretion and secretome.

3. Importance of components of the *Leishmania* exoproteome in the host-parasite interaction

During the *Leishmania* life cycle, parasite survival within the different environments provided by their hosts requires several strategies to block host defence mechanisms. In the mammalian host, amastigote survival within infected cells depends on the outcome of the parasite's interaction with the host cell (Reviewed in (Kima, 2007)). Parasites must promote survival of infected macrophages in the phagolysosome, where the amastigote form resists acidic pH conditions and attack by lysosomal enzymes. Secreted proteases belonging to the family of cystein- or metallo- proteases are generally thought to be involved in the manipulation of host immune responses in both the invertebrate and vertebrate hosts (Mottram*, et al.*, 2004, Olivier*, et al.*, 2005, Santarem*, et al.*, 2007). Likewise, *Leishmania* amastigotes have complex nutritional requirements which must be scavenged from the host cell (Reviewed in (Naderer & McConville, 2008). *Leishmania*, like other trypanosomatids, are purine auxotrophs (Ortiz*, et al.*, 2007). To satisfy its essential purine requirements, *Leishmania* secretes a nuclease that may function externally of the parasite to hydrolyze and access host-derived nucleic acids (Joshi & Dwyer, 2007). Furthermore, there is increasing evidence that some amastigote proteins can be exported to the host cytoplasm and directly modulate host signalling pathways (Mottram*, et al.*, 2004, Naderer & McConville, 2008). Nevertheless, the mechanisms underlying this process are poorly defined.

In the insect vector, *Leishmania* express several virulence factors which may facilitate transmission to and infection of the mammalian host. However, the identity of these factors is still limited. Materials secreted by the parasite play pivotal roles modifying the environment in their sandfly hosts so as to enhance their transmission success (Rogers*, et al.*, 2004). In mature *Leishmania* infections the stomodeal valve is forced open and becomes blocked with parasites embedded in promastigote secretory gel (PSG) (Rogers*, et al.*, 2004). The permanent opening of the stomodeal valve is essential for colonization of the foregut and transmission by regurgitation (Rogers*, et al.*, 2004). The PSG is a viscous mixture of phosphoglycans secreted by the parasites (Ilg*, et al.*, 1996). The major component of the PSG is a secreted glycoprotein called filamentous proteophosphoglycan (fPPG) (Rogers & Bates, 2007). Proteophosphoglycans secreted by *Leishmania* parasites also includes the promastigote secreted acid phosphatase (SAP) and a non-filamentous proteophosphoglycan from amastigotes (aPPG) and promastigotes (pPPG) (Reviewed in (Ilg, 2000). Despite extensive studies on aspects on their biochemistry, structure and genetics, the precise function of these molecules remain elusive. However, it has been shown that aPPG induces vacuole formation in mammalian macrophages (Peters*, et al.*, 1997) and is an activator of complement via the mannose-binding lectin pathway (Peters*, et al.*, 1997). Likewise, aPPG induces neither a B-cell nor a T-cell response in mice, thereby avoiding immune system recognition at this level (Aebischer*, et al.*, 1999).

Secreted chitinase enzyme is involved in the parasite-vector interaction. Recently, chitinase was shown to act as a multifunctional virulence factor benefiting *L. mexicana* throughout its entire life cycle. Joshi *et al* have shown that secreted chitinase is a virulence factor enhancing parasite intramacrophage survival and cutaneous pathology in mice (Joshi, *et al.*, 2005). Furthermore, by using a molecular approach, it was shown that this enzyme not only enhances infection of the mammalian host, but also parasite transmission by modifying the environment in the sandfly vector (Rogers, *et al.*, 2008). Specifically, the overexpression of chitinase enables transgenic parasites to colonize the anterior midgut of the sandfly more quickly, damage the stomodeal valve and affect its blood feeding (Rogers, *et al.*, 2008). A myriad of functions have been described for the gp63 from *Leishmania* spp (Reviewed in (Santos, *et al.*, 2006)). Despite gp63 being one of the best characterized surface molecules of *Leishmania*, its precise role in the life cycle of the parasite during colonization of the insect host remains speculative. Biochemical analysis of the extra-cellular gp63 of *Leishmania* has revealed two forms of the protein, one is a GPI-anchored form that can be released from the cell surface and another that is apparently directly secreted (Ellis, *et al.*, 2002, McGwire, *et al.*, 2002). Within the sandfly vector, release of extracellular gp63 from promastigotes could enhance the hydrolysis of protein substrates and play a nutritional and/or protective role (Kulkarni, *et al.*, 2006, Santos, *et al.*, 2006). Moreover, gp63 may function differently for distinct *Leishmania* species in their interactions with different invertebrate vector species (Joshi, *et al.*, 2002, Hajmova, *et al.*, 2004). Taken together, these studies highlight the importance of studying the interaction between the parasite, its vector and the mammalian host to fully understand the transmission of *Leishmania* parasites. In this scenario, secreted proteins have been shown to play multifunctional roles in the transmission of the parasites and in the infection of mammalian hosts.

4. Secretion mechanisms in eukaryotes and their contribution to the *Leishmania* exoproteome

4.1. *The classical secretory pathway*

In eukaryotes, soluble secretory proteins typically contain N-terminal signal peptides that direct them to the translocation apparatus of the endoplasmic reticulum (ER). Structural and functional studies indicate that eukaryotic signal sequences are composed of three distinct domains, a positively charged amino-terminal domain or residue, a central hydrophobic domain (6 to 12 residues), and a polar carboxy-terminal domain. Despite these characteristics primary sequences show very low homologies and variable lengths (15 to 30 residues). The translocation is mediated by the recognition of the N-terminal signal peptide by a signal recognition particle (SRP), which carries the nascent peptide chain to the Sec61 translocon complex within the ER membrane, across which transport takes place (Schatz & Dobberstein, 1996). Following translocation to the ER, the signal peptide is usually cleaved from the mature protein by a signal peptidase. This pathway of protein export from eukaryotic cells is known as the classical or ER/Golgi dependent secretory pathway (Rothman & Orci, 1992). The secretory pathway is primarily responsible for distribution of the newly synthesized products among the endomembrane compartments, as well as delivery to the exterior of the cells. Proteins processed in this pathway are commonly referred to as classically secreted proteins. These proteins are by default transported through the Golgi apparatus and exported by secretory vesicles. The presence of secondary signals directs their final localization inside the cell. Some proteins have specific retention signals that retain them to the ER or the Golgi, or divert them to lysosomes (Nilsson & Warren, 1994). Extracellular proteins are released into the extracellular space by fusion of Golgi-derived secretory vesicles with the plasma membrane.

Several studies in trypanosomatids demonstrated that the basic features of the secretory pathway are very similar to other eukaryotes, despite the fact that these organisms represent one of the most divergent eukaryotic lineages (Clayton*, et al.*, 1995, McConville*, et al.*, 2002). Mutational analyses of a mammalian signal sequence transfected into *L. major* and used to control an exogenous protein suggest the signal peptide sequence can function in a similar manner in *Leishmania* as it does in yeasts and higher eukaryotes (Tobin & Wirth, 1993). However, the polarized delivery of secretory material to the flagellar pocket of trypanosomatid parasites represents an unusual characteristic of the secretory pathway (Field, *et al.*, 2007).

Although several proteins in *Leishmania* have been identified as components of the secretion machinery, little is know about the number of proteins exported to the extracellular space through the classical pathway. The bulk of knowledge on *Leishmania* N-signal peptide containing proteins is centred on surface proteins, such as lyphophosglycan (LPG) and the surface metalloprotease gp63, which are involved in several process including complement resistance and attachment to the host cell (Descoteaux & Turco, 2002, Spath*, et al.*, 2003, Kulkarni*, et al.*, 2006). Members of these protein families can be constitutively shed and thus be constituents of the exoproteome.

A recent screen of a *L. major* cDNA library with antiserum raised against promastigote culture supernatant led to the detection of 8 proteins that bear a potential signal peptide of which 5 correspond to unknown proteins (Chenik*, et al.*, 2006). Furthermore, by using a genomic based approach, secretion of three proteins was experimentally confirmed among thirteen hypothetical conserved proteins in trypanosomatids bearing an N-terminal signal peptide (Corrales *et al*, submitted). In addition, analysis of the infectivity properties of *Leishmania* secreted proteins suggest that one of these proteins is involved in a process increasing survival and replication of the parasite inside its target cell (Corrales *et al*, submitted). Taken together, these studies support the notion that the classical secretory pathway is operational in *Leishmania* parasites for the export of proteins to the extracellular space. Nevertheless, little is known about the contribution of the classical pathway to the *Leishmania* exoproteome. An exhaustive analysis is required to identify all the proteins involved in the classical pathway and to characterize their function not only in the different parasite stages but also across different *Leishmania* species.

4.2. *Unconventional secretory pathway*

Although prevalent, the classical ER-Golgi dependent pathway is not the only means of protein export in eukaryotes. An increasing number of secreted proteins devoid of a signal peptide have been reported to be exported from eukaryote cells without the help of ER-Golgi (reviewed in (Nickel & Seedorf, 2008). The main features of non-classical secretion are; (a) the lack of conventional signal peptides, (b) the exclusion of these proteins from classical secretory organelles, such as the ER and the Golgi, combined with the lack of ER/Golgi-dependent post-translational modifications, and (c) resistance of these export processes to brefeldin A, a classical inhibitor of ER/Golgi dependent protein secretion (Nickel & Seedorf, 2008). Non-classical secretion is not limited to a specific family of proteins. Indeed, proteins belonging to diverse functional groups were demonstrated to be released independently of ER-Golgi: many of the known unconventional secretory proteins includes cytokines, growth factors, or other molecules with important signalling roles in physiological processes, such as inflammation, angiogenesis, cell differentiation, or proliferation (Prudovsky*, et al.*, 2008). Diverse mechanisms have been proposed to explain unconventional secretory processes, including lysosomal secretion, plasma membrane shedding, release in exosomes as well as secretion through transporters that reside in the plasma membrane (reviewed in (Nickel & Seedorf, 2008).

To date, in trypanosomatids the only characterized protein using an unconventional secretory pathway is the hydrophilic acylated surface protein B (HASPB) of *L. major* (Denny*, et al.*, 2000). This protein is associated with the outer leaflet of the plasma membrane in the infectious metacyclic promastigote and amastigote stages (Flinn*, et al.*, 1994). The protein is synthesized on free ribosomes in the cytoplasm and becomes both myristoylated and palmitoylated at its N-terminus, which is the molecular basis of how HASPB is anchored in the membrane (Denny, *et al.*, 2000). A model has been proposed in which HASPB is transferred from the cytoplasm to the outer leaflet of the Golgi membrane, from where it is transported to the plasma membrane via conventional vesicular transport (Stegmayer, *et al.*, 2005). This process would insert HASPB into the inner leaflet of the plasma membrane. At present it is completely unclear how HASPB is then translocated across the membrane. HASPB has been shown to be secreted from mammalian cells, demonstrating a conserved pathway among lower and higher eukaryotes (Stegmayer, *et al.*, 2005). However, the sequence encoding the HASPB protein is not conserved among all *Leishmania* species and is absent from the *L. braziliensis* genome. A recent study showed that heterologous expression of SH4 domains of *Leishmania* HASPB induced the reorganization of the plasma membrane to produce highly dynamic non-apoptotic membrane blebs (Tournaviti, *et al.*, 2007). Several distinct types of membrane blebs have been described. They can be formed by different mechanisms and have different functions, including cell movement, cytokinesis, cell spreading and apoptosis (reviewed in (Charras, 2008). Interestingly, membrane blebbing is involved in the mechanism of non-classical secretion of cytokine interleukin-1β (Qu, *et al.*, 2007).

Although the mechanisms and molecular components of unconventional secretion are largely unknown, the roles of regulatory components are beginning to emerge (Nickel & Rabouille, 2009). Recent studies have revealed general roles for caspase 1 and the Golgi reassembly stacking protein (GRASP) in the regulation of unconventional protein secretion. Caspase 1 was shown to function as a general regulator of stress-induced unconventional secretion for several cytokines (Keller, *et al.*, 2008). Caspases play a central role in programmed cell death (PCD) but are also involved in cell proliferation and differentiation (Chowdhury, *et al.*, 2008). In *Leishmania* there is evidence that PCD occurs, but little is known about the pathways involved (Deponte, 2008). Comparative genomics of trypanosomatids revealed that the vast majority of the proteins involved in mammalian PCD, including caspases, are not encoded in the genome of *Leishmania spp.* or related protozoa (Deponte, 2008). However, these organisms possess metacaspases that are evolutionary distant orthologues of metazoan caspases and involved in PCD (Lee*, et al.*, 2007, Deponte, 2008). Two putative metacaspase genes can be identified in the genomes of *L. infantum* and *L. donovani,* responsible for visceral leishmaniasis. In contrast, *L. major*, responsible for cutaneous leishmaniasis, possesses only a single metacaspase gene and this has been shown to be essential for cell cycle progression (Ambit*, et al.*, 2008). Apparently metacaspases of *T. brucei* are involved in vesicle trafficking but apparently possess PCD-independent functions (Helms*, et al.*, 2006). Further analyses of trypanosomatid metacaspases are necessary to elucidate their roles and to provide insights into their involvement in regulatory networks, which may include non-conventional secretion.

The GRASP protein is essential for an unconventional secretion pathway during *Dictyostelium discoideum* cellular development (Kinseth, *et al.*, 2007) and is also required for the unconventional secretion of the integral plasma membrane α integrin in *Drosophila melanogaster* during epithelium remodelling (Schotman*, et al.*, 2008). Given that the GRASP family is widely conserved in eukaryotes, these unexpected findings raised several questions. Is GRASP involved in unconventional protein secretion in all eukaryotes, or does its function vary depending upon the cell type? Can the classical secretion pathway regulate the Golgi-independent-secretion? In order to resolve this important issue the role of GRASP proteins in other cell types and organisms should be addressed. In this regard, *Leishmania* may represent an interesting unicellular model to further elucidate function of GRASP proteins in lower eukaryotes. Recently, a quantitative analysis of the *L. donovani* secretome indicated it predominantly uses non-classical secretion pathways to directly export protein (Silverman*, et* *al.*, 2008). It is tempting to suggest that some *Leishmania* proteins involved in the classical pathway, like GRASP, may regulate the non-classical secretion of proteins as demonstrated in *D. discoideum*. Additional studies might reveal conserved functions for GRASP proteins in lower eukaryotes, including regulation of non conventional secretion. The availability of *Leishmania spp.* genome sequences together with genetic tools may shed light on the mystery of the relationship between classical and non-classical secretion pathways.

4.3. *Vesicle mediated protein export*

Another pathway for protein export in eukaryotes is the release of exosomes to the extracellular environment. Exosomes are small microvesicles, released from cells by the fusion of either multivesicular bodies (MVB) or secretory lysosomes with the plasma membrane (Simpson, *et al.*, 2008). Unlike the constitutive trans-Golgi secretory pathway, exosome secretion arises directly from the endocytic pathway. Exosomes have been the subject of intensive research in recent years. Originally described as a removal mechanism of cell surface molecules in reticulocytes, several laboratories showed their importance in a range of biological processes, including development, immunology and cancer. In addition to a common set of membrane and cytosolic molecules, exosomes harbor unique subsets of proteins linked to cell type associated functions (Reviewed in (Simpson, *et al.*, 2008). In the last few years, with the development of proteomic technologies, molecular components of exosomes from a variety of cell types and body fluids have been analysed. Whilst functional roles of exosomes are only recently becoming clear, future investigations are likely to indicate the importance of them as mediators in biological processes. In addition to exosomes, cells can also release other forms of membrane vesicles into the extracellular space; apoptotic blebs, microparticles (MP) and microvesicles (MV) (Simpson, *et al.*, 2008). These vesicles represent a heterogeneous population of vesicles that bud directly from the plasma membrane. However, given that exosomes, apoptotic blebs, MV and MP exhibit the same membrane topology is still unclear if these vesicles represent discreet entities with specialized functions or merely represent a size continuum of the same entity (Simpson, *et al.*, 2008). Recently, the study of *L. donovani* exoproteome supported the notion that stationary phase promastigotes export materials through an exosome-like mechanism (Silverman, *et al.*, 2008). By using a SILAC (stable isotope labelling with amino acids in cells) approach, the authors detected several proteins in the exoproteome that were previously reported as being components of exosomes isolated from B-lymphocytes and adipocytes (Silverman, *et al.*, 2008). In addition, the authors detected similar contents in exosomes from apoptotic dendritic cells that had previously been defined as apoptotic blebs (Thery, *et al.*, 2001). Given that stationary phase promastigotes contain up to 43% apoptotic cells that are involved in the establishment and maintaining an infection, Silverman *et al* proposed that proteins released by *Leishmania* promastigotes through apoptotic blebs were those important for the parasite's virulence (Silverman, *et al.*, 2008). Further studies of *Leishmania* exosomes may shed light on novel functions for these vesicles, such as the modulation of immune response demonstrated in dendritic cells (Chaput, *et al.*, 2006). Finally, exosome studies may explain the presence of proteins in the extracellular space that lack a N-Signal peptide, such as histones, ribosomal proteins and elongation factors (Chenik, *et al.*, 2006, Silverman, *et al.*, 2008).

5. Approaches for the study of the exoproteome and its application in *Leishmania*

Emerging technologies in proteomic research and genome sequencing have greatly accelerated studies of the exoproteome in eukaryotes. Generally, these methods can be categorized into four groups: (i) direct proteomic-based approaches, (ii) genome-based computational predictions, (iii) genetically-based approaches, and (iv) immunologically-based approaches. In the next section we describe the different approaches, their limitations and applications in the study of the *Leishmania* exoproteome.

5.1. Direct proteomic-based approaches

Recently developed methods in proteomics have greatly facilitated the broad scale analysis of proteins in biological samples, including whole tissues and culture media (Reviewed in (Lane, 2005). The massive progress in mass spectrometry (MS), bioinformatics and analytical techniques greatly promoted analyses of the exoproteome, especially the cancer secretome (Latterich*, et al.*, 2008). Currently, three major proteomic technologies are used for secretome analyses; gel-based methods, gel-free MS-based methods, and surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDITOF-MS). Two-dimensional gel electrophoresis (2-DE) coupling MS is the most classic and well-established proteomic approach. This allows identification of hundreds to thousands of proteins from a single gel. Although this method remains the most efficient for separation of complex protein mixtures, the detection of secreted proteins in cell supernatants represents a challenge for current proteomics techniques. This is for several reasons. Firstly, culture media are rich in salts and other compounds, e.g. serum supplement, that interfer with the selective precipitation of proteins. Secondly, the non-secreted proteins liberated in the culture medium upon lysis of a few dead cells contaminate the so-called secretome. To overcome these technical limitations, a strategy involving the elimination of serum constituents through washing and subsequent incubation of cells in a serum-free medium can be used. This has proved useful in studies on the secretome of different cancer cell lines (Xue*, et al.*, 2008) and for the recovery of secreted proteins from *Leishmania* promastigote culture supernatants after 4 to 6 hours of incubation in serum-free media (Merlen*, et al.*, 1999, Chenik*, et al.*, 2006, Silverman*, et al.*, 2008). Nevertheless, the balance between the incubation time required for the cells to secrete detectable amounts of protein and cell survival in the serum-free medium, to avoid proteins from cell lysis, is a key factor.

Another challenge is the detection of low abundance secreted proteins. Given the low concentration of secreted proteins in culture media, supernatant concentration is necessary before subsequent proteomic analyses. Although incubation in serum-free conditioned media combined with concentration of the supernatant by ultra-filtration can significantly reduce contamination by non-secreted proteins, a minimal presence of contaminant proteins may easily mask the proteins of interest, especially when they are in low quantities. Consequently the discrimination of genuine secreted proteins from non-secreted proteins remains the main challenge in proteomic-based technologies.

Stable isotope labelling with amino acids in cell culture (SILAC) has emerged as an alternative methodology for accurate identification of genuinely secreted proteins in quantitative proteomics. SILAC involves culturing cells in media containing either the normal amino acid or its isotopically labelled analogue until essentially all proteins of the cell are labelled. The two populations or samples to be compared are then mixed in equal ratios and analysed by nanoflow liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Reviewed in (Gevaert, *et al.*, 2008)). SILAC-based strategies not only help to profile secreted proteins in mammalian cell cultures, but also facilitate the distinction between secreted proteins and contaminant proteins from the culture media. SILAC methodology was successfully used to analyse the secretome from cultures of stationary-phase promastigotes of *L. donovani*, by comparing the amount of any given protein secreted into the conditioned medium with the amount of the same protein that remained cell associated. A total of 358 proteins were identified by using this approach and based on quantitative analyses the authors concluded that 151 were actively secreted by the parasites (Silverman, *et al.*, 2008). This methodology may represent a useful tool to characterize the secretome at different life cycle stages of *Leishmania* and for studying the related medically important parasites *Trypanosoma brucei* and *T. cruzi*. However, the basis for quantitative analysis using SILAC prevents identification of secreted proteins that display very low intracellular concentrations, independently of their concentration in the conditioned medium. This drawback resulted in the non-identification of some *Leishmania* proteins previously described as secreted when analysing the *L. donovani* secretome (Silverman*, et al.*, 2008).

5.2. Genome-based computational prediction

Computational analyses rely on the prediction of signal peptides, viewed as a hallmark of classically secreted proteins in eukaryotes. The N-terminal signal sequence or signal peptide (SP) holds a conserved set of secondary characteristics that are identifiable by computational algorithms. Over the last decade, the development of new signal peptide prediction programs has been prolific and allowed the identification of potentially secreted proteins from eukaryotes (Reviewed in (Klee & Sosa, 2007). Nevertheless, these programmes use a variety of algorithms and architectures leading to different predictions. To overcome this limitation, several prediction tools, such as trans-membrane domain, GPI anchoring signal and mitochondrial targeting signal predictions can be combined to improve the *in silico* prediction of classically secreted proteins. A computational strategy was used to identify the mouse secretome, by looking for transcripts encoding proteins with signal sequences but without trans-membrane domains (Grimmond*, et al.*, 2003). Lee *et al*. used a set of prediction algorithms to define the potential *Candida albicans* secretome through computational identification of soluble proteins that possessed N-terminal signal sequences and lacked trans-membrane domains, GPI anchor sites and mitochondrial targeting sequences, from the 6165 open reading frames (ORFs) of the yeast genome (Lee*, et al.*, 2003). A similar multistep computational analysis was used to simultaneously predict the secretomes of zebrafish and humans, and to identify orthologs among the potentially secreted proteins (Klee, 2008). Although, genome-based methods can provide a comprehensive list of potentially secreted proteins, the data obtained may contain many false-positive and false-negative candidates. In addition, these approaches make it difficult to discriminate between extracellular proteins and other proteins processed by the ER, but not released into the extracellular space, e.g. ER and

Golgi resident proteins and lysosomal proteins. Another limitation for computational prediction is the lack of programs to predict proteins secreted through non-classical pathways. To date, only one programme is available for the prediction of mammalian proteins targeted by the non-classical secretory pathway (Bendtsen*, et al.*, 2004). Due to the lack of a conserved targeting signal for unconventional secretion, the development of algorithms for the prediction of non-classical secreted proteins remains a challenge.

Although, sequencing of the genome of different *Leishmania* species is now completed, only two genome-based studies have performed *in silico* prediction of potentially secreted proteins in these pathogens. Silverman *et al* analysed the secretome of *L. major* genome by searching protein sequences bearing a signal peptide, lacking a trans-membrane domain and GPI anchor signal (Silverman*, et al.*, 2008). By using these parameters the authors found 217 potentially secreted proteins, of which 141 were annotated as hypothetical proteins. Remarkably, only 14 of the predicted secreted proteins were experimentally identified with a proteomics-based approach. In our group we used a computational approach and confirmed only 25% of trypanosomatid proteins predicted to be secreted were detected in the extracellular environment (Corrales *et al*, submitted). Thus, even if computational predictions can provide a list of potentially secreted proteins, experimental validations is needed to confirm if they are actually secreted.

5.3 Genetically-based approaches: Forward and Reverse Genetics

In eukaryotes, several forward genetic strategies based on the presence of N-signal peptides in protein sequences have been used to identify proteins involved in the classical secretory pathway. A widely used method to identify secreted and transmembrane proteins is the Yeast Signal Sequence Trap (YSST). This method relies on the ability of putative signal sequences to rescue secretion of the yeast invertase enzyme lacking its endogenous signal peptide (Klein*, et al.*, 1996, Jacobs*, et al.*, 1997). Several YSST approaches have been developed and applied to lower and higher eukaryotes to identify secreted and membrane associated proteins (Nene & Bishop, 2001, Smyth*, et al.*, 2003, Yamane*, et al.*, 2005). Nevertheless, the main drawback of YSST is the presence of false negative proteins bearing a predictive signal sequence, which may be explained by heterologous gene expression that does not reflect the natural environment. However, this technology has led to the discovery of genes encoding secretory proteins that were not predicted to bear signal sequences in *Hydra* polyps (Bottger*, et al.*, 2006). Thus genes discovered by YSST methods offer a starting point for more detailed characterization. A morpholino-based gene 'knockdown' strategy was used in zebrafish to assess the role of members of the secretome in vertebrate development (Pickart*, et al.*, 2006). This study provides a framework utilizing zebrafish for the systematic assignment of biological function in a vertebrate genome. In protozoans, a remarkably reverse genetic approach was recently used to analyse the secretome in *Plasmodium falciparum* (van Ooij*, et al.*, 2008). By using the *piggy*Bac transposition system, the authors validated the export of 70% of the predicted secreted proteins (van Ooij*, et al.*, 2008). In the *piggy*Bac transposition system, integration and expression of specific DNA sequences is induced by the lepidopteran transposon and a transposase-expressing helper plasmid (Balu*, et al.*, 2005). Given that efficient transposon-mediated random mutagenesis has been reported in *Leishmania* (Beverley*, et al.*, 2002), this methodology might represent an interesting tool to analyse the exoproteome of amastigotes within macrophages.

5.4 Immune-based approaches: An alternative methodology to identify components of the *Leishmania* exoproteome is based on the highly immunogenic properties displayed by *Leishmania* culture supernatants (Chenik*, et al.*, 2006). The authors attempted to identify extracellular proteins from *L. major* by screening a cDNA expression library with sera raised against culture medium supernatants. By using this strategy 33 proteins were identified, out of which 8 had a signal peptide, suggesting their involvement in the classical pathway.

Remarkably, about 40% of the identified genes encode unknown proteins, pointing out the interest in characterizing the components of the *Leishmania* exoproteome. Unfortunately, antibody-based methodologies are biased toward the identification of the most abundant or highly immunogenic proteins, and thus typically identify only a small subset of proteins in the secretome.

6. Challenges and perspectives

The genome sequences of *Leishmania* species have provided a new starting point to understand the biology of these medically important parasites. Understanding the relative contribution of parasite genes to the wide spectrum of disease caused by *Leishmania* species is one of the most important questions to be answered. Given the relatively minor differences among *Leishmania* species in terms of gene content and RNA gene expression (Lynn & McMaster, 2008), studies of the proteome from different *Leishmania* species are crucial. In this regard and because of the relevance of parasite extracellular components to its interaction with its environment, the in-depth study of the exoproteome represents an open door to further elucidate; i) the function of the large number of hypothetical proteins awaiting for detailed analyses; ii) the dynamics of the exoproteome throughout the *Leishmania* life cycle, iii) the main secretion mechanisms used by these parasites to export proteins into the extracellular space, and their significance in the different developmental forms. Compared with that of bacteria and higher eukaryotes, our knowledge about the *Leishmania* exoproteome is at a rudimentary stage. However, the recent study supporting non-classical secretion pathways as the main mechanism used by *L. donovani* promastigotes to export proteins, may represent the tip of an iceberg ready to emerge. Whether this observation is confirmed for the amastigote form and other *Leishmania* species may turn our research towards understanding the mechanisms and regulation of unconventional secretion. Roles for GRASP and caspase1 proteins in the regulation of Golgi-independent secretion have now been revealed. Although caspases are missing in protists, GRASP proteins are highly conserved in eukaryotes. Analysing their roles in *Leishmania* parasites may further shed light on the regulation of non-classical mechanisms and on the possible crosstalk between classical and non-classical pathways (Summarized in Figure 1). This information will lead to a deeper understanding of how secretion mechanisms are used by trypanosomatids to ensure parasite survival within the different environments encountered throughout its life cycle. Together, a detailed knowledge of the *Leishmania* exoproteome may lead to the development of new anti-parasite strategies.

Acknowledgements

Financial support for this study was provided by the European Union through a grant to RMC by the 'Programme Alßan' European Union Programme of High Level Scholarships for Latin America (No. E05D057391AR) and by IRD DSF. We thank P Agnew for help revising the manuscript's English.

References

[1] Aebischer T, Harbecke D & Ilg T (1999) Proteophosphoglycan, a major secreted product of intracellular Leishmania mexicana amastigotes, is a poor B-cell antigen and does not elicit a specific conventional CD4+ T-cell response. *Infect Immun* **67**: 5379-5385.

[2] Ambit A, Fasel N, Coombs GH & Mottram JC (2008) An essential role for the Leishmania major metacaspase in cell cycle progression. *Cell Death Differ* **15**: 113-122.

[3] Antelmann H, Tjalsma H, Voigt B, Ohlmeier S, Bron S, van Dijl JM & Hecker M (2001) A proteomic view on genome-based signal peptide predictions. *Genome Res* **11**: 1484-1502.

[4] Balu B, Shoue DA, Fraser MJ, Jr. & Adams JH (2005) High-efficiency transformation of Plasmodium falciparum by the lepidopteran transposable element piggyBac. *Proc Natl Acad Sci U S A* **102**: 16391-16396.

[5] Bates PA (2007) Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies. *Int J Parasitol* **37**: 1097-1106.

[6] Bendtsen JD, Jensen LJ, Blom N, Von Heijne G & Brunak S (2004) Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel* **17**: 349-356.

[7] Beverley SM, Akopyants NS, Goyard S*, et al.* (2002) Putting the Leishmania genome to work: functional genomics by transposon trapping and expression profiling. *Philos Trans R Soc Lond B Biol Sci* **357**: 47-53.

[8] Bottger A, Strasser D, Alexandrova O*, et al.* (2006) Genetic screen for signal peptides in Hydra reveals novel secreted proteins and evidence for non-classical protein secretion. *Eur J Cell Biol* **85**: 1107-1117.

[9] Campbell K, Popov V & Soong L (2004) Identification and molecular characterization of a gene encoding a protective Leishmania amazonensis Trp-Asp (WD) protein. *Infect Immun* **72**: 2194-2202.

[10] Chaput N, Flament C, Viaud S*, et al.* (2006) Dendritic cell derived-exosomes: biology and clinical implementations. *J Leukoc Biol* **80**: 471-478.

[11] Charras GT (2008) A short history of blebbing. *J Microsc* **231**: 466-478.

[12] Chenik M, Lakhal S, Ben Khalef N, Zribi L, Louzir H & Dellagi K (2006) Approaches for the identification of potential excreted/secreted proteins of Leishmania major parasites. *Parasitology* **132**: 493-509.

[13] Chevallet M, Diemer H, Van Dorssealer A, Villiers C & Rabilloud T (2007) Toward a better analysis of secreted proteins: the example of the myeloid cells secretome. *Proteomics* **7**: 1757-1770.

[14] Chowdhury I, Tharakan B & Bhat GK (2008) Caspases - an update. *Comp Biochem Physiol B Biochem Mol Biol* **151**: 10-27.

[15] Clayton C, Hausler T & Blattner J (1995) Protein trafficking in kinetoplastid protozoa. *Microbiol Rev* **59**: 325-344.

[16] Croft SL, Sundar S & Fairlamb AH (2006) Drug resistance in leishmaniasis. *Clin Microbiol Rev* **19**: 111-126.

[17] Denny PW, Gokool S, Russell DG, Field MC & Smith DF (2000) Acylation-dependent protein export in Leishmania. *J Biol Chem* **275**: 11017-11025.

[18] Deponte M (2008) Programmed cell death in protists. *Biochim Biophys Acta* **1783**: 1396-1405.

[19] Descoteaux A & Turco SJ (2002) Functional aspects of the Leishmania donovani lipophosphoglycan during macrophage infection. *Microbes Infect* **4**: 975-981.

[20] Descoteaux A, Avila HA, Zhang K, Turco SJ & Beverley SM (2002) Leishmania LPG3 encodes a GRP94 homolog required for phosphoglycan synthesis implicated in parasite virulence but not viability. *Embo J* **21**: 4458-4469.

[21] Duboise SM, Vannier-Santos MA, Costa-Pinto D, *et al.* (1994) The biosynthesis, processing, and immunolocalization of Leishmania pifanoi amastigote cysteine proteinases. *Mol Biochem Parasitol* **68**: 119-132.

[22] Ellis M, Sharma DK, Hilley JD, Coombs GH & Mottram JC (2002) Processing and trafficking of Leishmania mexicana GP63. Analysis using GP18 mutants deficient in glycosylphosphatidylinositol protein anchoring. *J Biol Chem* **277**: 27968-27974.

[23] Field MC, Natesan SK, Gabernet-Castello C & Koumandou VL (2007) Intracellular trafficking in the trypanosomatids. *Traffic* **8**: 629-639.

[24] Flinn HM, Rangarajan D & Smith DF (1994) Expression of a hydrophilic surface protein in infective stages of Leishmania major. *Mol Biochem Parasitol* **65**: 259-270.

[25] Gevaert K, Impens F, Ghesquiere B, Van Damme P, Lambrechts A & Vandekerckhove J (2008) Stable isotopic labeling in proteomics. *Proteomics* **8**: 4873-4885.

[26] Grebaut P, Chuchana P, Brizard JP*, et al.* (2009) Identification of total and differentially expressed excreted-secreted proteins from Trypanosoma congolense strains exhibiting different virulence and pathogenicity. *Int J Parasitol*.

[27] Grimmond SM, Miranda KC, Yuan Z*, et al.* (2003) The mouse secretome: functional classification of the proteins secreted into the extracellular environment. *Genome Res* **13**: 1350-1359.

[28] Hajmova M, Chang KP, Kolli B & Volf P (2004) Down-regulation of gp63 in Leishmania amazonensis reduces its early development in Lutzomyia longipalpis. *Microbes Infect* **6**: 646-649.

[29] Hathout Y (2007) Approaches to the study of the cell secretome. *Expert Rev Proteomics* **4**: 239-248.

[30] Helms MJ, Ambit A, Appleton P, Tetley L, Coombs GH & Mottram JC (2006) Bloodstream form Trypanosoma brucei depend upon multiple metacaspases associated with RAB11-positive endosomes. *J Cell Sci* **119**: 1105-1117.

[31] Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estrano C & Haldar K (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**: 1934-1937.

[32] Holzmuller P, Grebaut P, Peltier JB*, et al.* (2008) Secretome of animal trypanosomes. *Ann N Y Acad Sci* **1149**: 337-342.

[33] Horwitz MA, Harth G, Dillon BJ & Maslesa-Galic S (2005) Enhancing the protective efficacy of Mycobacterium bovis BCG vaccination against tuberculosis by boosting with the Mycobacterium tuberculosis major secretory protein. *Infect Immun* **73**: 4676-4683.

[34] Ilg T (2000) Proteophosphoglycans of Leishmania. *Parasitol Today* **16**: 489-497.

[35] Ilg T, Stierhof YD, Wiese M, McConville MJ & Overath P (1994) Characterization of phosphoglycan-containing secretory products of Leishmania. *Parasitology* **108 Suppl**: S63-71.

[36] Ilg T, Stierhof YD, Craik D, Simpson R, Handman E & Bacic A (1996) Purification and structural characterization of a filamentous, mucin-like proteophosphoglycan secreted by Leishmania parasites. *J Biol Chem* **271**: 21583-21596.

[37] Jacobs KA, Collins-Racie LA, Colbert M, *et al.* (1997) A genetic selection for isolating cDNAs encoding secreted proteins. *Gene* **198**: 289-296.

[38] Jaffe CL & Dwyer DM (2003) Extracellular release of the surface metalloprotease, gp63, from Leishmania and insect trypanosomatids. *Parasitol Res* **91**: 229-237.

[39] Jazin EE, Luquetti AO, Rassi A & Frasch AC (1991) Shift of excretory-secretory immunogens of Trypanosoma cruzi during human Chagas' disease. *Infect Immun* **59**: 2189-2191.

[40] Joshi MB & Dwyer DM (2007) Molecular and functional analyses of a novel class I secretory nuclease from the human pathogen, Leishmania donovani. *J Biol Chem* **282**: 10079-10095.

[41] Joshi MB, Rogers ME, Shakarian AM, Yamage M, Al-Harthi SA, Bates PA & Dwyer DM (2005) Molecular characterization, expression, and in vivo analysis of LmexCht1: the chitinase of the human pathogen, Leishmania mexicana. *J Biol Chem* **280**: 3847-3861.

[42] Joshi PB, Kelly BL, Kamhawi S, Sacks DL & McMaster WR (2002) Targeted gene deletion in Leishmania major identifies leishmanolysin (GP63) as a virulence factor. *Mol Biochem Parasitol* **120**: 33-40.

[43] Keller M, Ruegg A, Werner S & Beer HD (2008) Active caspase-1 is a regulator of unconventional protein secretion. *Cell* **132**: 818-831.

[44] Kima PE (2007) The amastigote forms of Leishmania are experts at exploiting host cell processes to establish infection and persist. *Int J Parasitol* **37**: 1087-1096.

[45] Kinseth MA, Anjard C, Fuller D, Guizzunti G, Loomis WF & Malhotra V (2007) The Golgi-associated protein GRASP is required for unconventional protein secretion during development. *Cell* **130**: 524-534.

[46] Klee EW (2008) The zebrafish secretome. *Zebrafish* **5**: 131-138.

[47] Klee EW & Sosa CP (2007) Computational classification of classically secreted proteins. *Drug Discov Today* **12**: 234-240.

[48] Klein RD, Gu Q, Goddard A & Rosenthal A (1996) Selection for genes encoding secreted proteins and receptors. *Proc Natl Acad Sci U S A* **93**: 7108-7113.

[49] Kulkarni MM, McMaster WR, Kamysz E, Kamysz W, Engman DM & McGwire BS (2006) The major surface-metalloprotease of the parasitic protozoan, Leishmania, protects against antimicrobial peptide-induced apoptotic killing. *Mol Microbiol* **62**: 1484-1497.

[50] Lane CS (2005) Mass spectrometry-based proteomics in the life sciences. *Cell Mol Life Sci* **62**: 848-869.

[51] Latterich M, Abramovitz M & Leyland-Jones B (2008) Proteomics: new technologies and clinical applications. *Eur J Cancer* **44**: 2737-2741.

[52] Lee N, Gannavaram S, Selvapandiyan A & Debrabant A (2007) Characterization of metacaspases with trypsin-like activity and their putative role in programmed cell death in the protozoan parasite Leishmania. *Eukaryot Cell* **6**: 1745-1757.

[53] Lee SA, Wormsley S, Kamoun S, Lee AF, Joiner K & Wong B (2003) An analysis of the Candida albicans genome database for soluble secreted proteins using computer-based prediction algorithms. *Yeast* **20**: 595-610.

[54] Lemesre JL, Holzmuller P, Goncalves RB, Bourdoiseau G, Hugnet C, Cavaleyra M & Papierok G (2007) Long-lasting protection against canine visceral leishmaniasis using the

LiESAp-MDP vaccine in endemic areas of France: double-blind randomised efficacy field trial. *Vaccine* **25**: 4223-4234.

[55] Lynn MA & McMaster WR (2008) Leishmania: conserved evolution--diverse diseases. *Trends Parasitol* **24**: 103-105.

[56] McConville MJ, Mullin KA, Ilgoutz SC & Teasdale RD (2002) Secretory pathway of trypanosomatid parasites. *Microbiol Mol Biol Rev* **66**: 122-154; table of contents.

[57] McGwire BS, O'Connell WA, Chang KP & Engman DM (2002) Extracellular release of the glycosylphosphatidylinositol (GPI)-linked Leishmania surface metalloprotease, gp63, is independent of GPI phospholipolysis: implications for parasite virulence. *J Biol Chem* **277**: 8802-8809.

[58] McMahon-Pratt D & Alexander J (2004) Does the Leishmania major paradigm of pathogenesis and protection hold for New World cutaneous leishmaniases or the visceral disease? *Immunol Rev* **201**: 206-224.

[59] Merlen T, Sereno D, Brajon N, Rostand F & Lemesre JL (1999) Leishmania spp: completely defined medium without serum and macromolecules (CDM/LP) for the continuous in vitro cultivation of infective promastigote forms. *The American Journal of Tropical Medicine and Hygiene* **60**: 41-50.

[60] Mottram JC, Coombs GH & Alexander J (2004) Cysteine peptidases as virulence factors of Leishmania. *Curr Opin Microbiol* **7**: 375-381.

[61] Naderer T & McConville MJ (2008) The Leishmania-macrophage interaction: a metabolic perspective. *Cell Microbiol* **10**: 301-308.

[62] Nene V & Bishop R (2001) Trapping parasite secretory proteins in baker's yeast. *Trends Parasitol* **17**: 407-409.

[63] Nickel W & Seedorf M (2008) Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. *Annu Rev Cell Dev Biol* **24**: 287-308.

[64] Nickel W & Rabouille C (2009) Mechanisms of regulated unconventional protein secretion. *Nat Rev Mol Cell Biol* **10**: 148-155.

[65] Nilsson T & Warren G (1994) Retention and retrieval in the endoplasmic reticulum and the Golgi apparatus. *Curr Opin Cell Biol* **6**: 517-521.

[66] Olivier M, Gregory DJ & Forget G (2005) Subversion mechanisms by which Leishmania parasites can escape the host immune response: a signaling point of view. *Clin Microbiol Rev* **18**: 293-305.

[67] Ortiz D, Sanchez MA, Pierce S, Herrmann T, Kimblin N, Archie Bouwer HG & Landfear SM (2007) Molecular genetic analysis of purine nucleobase transport in Leishmania major. *Mol Microbiol* **64**: 1228-1243.

[68] Paper JM, Scott-Craig JS, Adhikari ND, Cuomo CA & Walton JD (2007) Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus Fusarium graminearum. *Proteomics* **7**: 3171-3183.

[69] Peters C, Stierhof YD & Ilg T (1997) Proteophosphoglycan secreted by Leishmania mexicana amastigotes causes vacuole formation in macrophages. *Infect Immun* **65**: 783-786.

[70] Peters C, Kawakami M, Kaul M, Ilg T, Overath P & Aebischer T (1997) Secreted proteophosphoglycan of Leishmania mexicana amastigotes activates complement by triggering the mannan binding lectin pathway. *Eur J Immunol* **27**: 2666-2672.

[71] Pickart MA, Klee EW, Nielsen AL*, et al.* (2006) Genome-wide reverse genetics framework to identify novel functions of the vertebrate secretome. *PLoS ONE* **1**: e104.

[72] Prudovsky I, Tarantini F, Landriscina M*, et al.* (2008) Secretion without Golgi. *J Cell Biochem* **103**: 1327-1343.

[73] Qu Y, Franchi L, Nunez G & Dubyak GR (2007) Nonclassical IL-1 beta secretion stimulated by P2X7 receptors is dependent on inflammasome activation and correlated with exosome release in murine macrophages. *J Immunol* **179**: 1913-1925.

[74] Rogers ME & Bates PA (2007) Leishmania manipulation of sand fly feeding behavior results in enhanced transmission. *PLoS Pathog* **3**: e91.

[75] Rogers ME, Ilg T, Nikolaev AV, Ferguson MA & Bates PA (2004) Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. *Nature* **430**: 463-467.

[76] Rogers ME, Hajmova M, Joshi MB, Sadlova J, Dwyer DM, Volf P & Bates PA (2008) Leishmania chitinase facilitates colonization of sand fly vectors and enhances transmission to mice. *Cell Microbiol* **10**: 1363-1372.

[77] Rothman JE & Orci L (1992) Molecular dissection of the secretory pathway. *Nature* **355**: 409-415.

[78] Santarem N, Silvestre R, Tavares J, Silva M, Cabral S, Maciel J & Cordeiro-da-Silva A (2007) Immune response regulation by leishmania secreted and nonsecreted antigens. *J Biomed Biotechnol* **2007**: 85154.

[79] Santos AL, Branquinha MH & D'Avila-Levy CM (2006) The ubiquitous gp63-like metalloprotease from lower trypanosomatids: in the search for a function. *An Acad Bras Cienc* **78**: 687-714.

[80] Schatz G & Dobberstein B (1996) Common principles of protein translocation across membranes. *Science* **271**: 1519-1526.

[81] Schotman H, Karhinen L & Rabouille C (2008) dGRASP-mediated noncanonical integrin secretion is required for Drosophila epithelial remodeling. *Dev Cell* **14**: 171-182.

[82] Silverman JM, Chan SK, Robinson DP, Dwyer DM, Nandan D, Foster LJ & Reiner NE (2008) Proteomic analysis of the secretome of Leishmania donovani. *Genome Biol* **9**: R35.

[83] Simpson JC, Mateos A & Pepperkok R (2007) Maturation of the mammalian secretome. *Genome Biol* **8**: 211.

[84] Simpson RJ, Jensen SS & Lim JW (2008) Proteomic profiling of exosomes: current perspectives. *Proteomics* **8**: 4083-4099.

[85] Smyth D, McManus DP, Smout MJ, Laha T, Zhang W & Loukas A (2003) Isolation of cDNAs encoding secreted and transmembrane proteins from Schistosoma mansoni by a signal sequence trap method. *Infect Immun* **71**: 2548-2554.

[86] Spath GF, Garraway LA, Turco SJ & Beverley SM (2003) The role(s) of lipophosphoglycan (LPG) in the establishment of Leishmania major infections in mammalian hosts. *Proc Natl Acad Sci U S A* **100**: 9536-9541.

[87] Stegmayer C, Kehlenbach A, Tournaviti S*, et al.* (2005) Direct transport across the plasma membrane of mammalian cells of Leishmania HASPB as revealed by a CHO export mutant. *J Cell Sci* **118**: 517-527.

[88] Stuart K, Brun R, Croft S*, et al.* (2008) Kinetoplastids: related protozoan pathogens, different diseases. *J Clin Invest* **118**: 1301-1310.

[89] Thery C, Boussac M, Veron P, Ricciardi-Castagnoli P, Raposo G, Garin J & Amigorena S (2001) Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles. *J Immunol* **166**: 7309-7318.

[90] Tjalsma H, Bolhuis A, Jongbloed JD, Bron S & van Dijl JM (2000) Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* **64**: 515-547.

[91] Tobin JF & Wirth DF (1993) Mutational analysis of a signal sequence required for protein secretion in Leishmania major. *Mol Biochem Parasitol* **62**: 243-249.

[92] Tonui WK, Mejia JS, Hochberg L*, et al.* (2004) Immunization with Leishmania major exogenous antigens protects susceptible BALB/c mice against challenge infection with L. major. *Infect Immun* **72**: 5654-5661.

[93] Tournaviti S, Hannemann S, Terjung S, *et al.* (2007) SH4-domain-induced plasma membrane dynamization promotes bleb-associated cell motility. *J Cell Sci* **120**: 3820-3829.

[94] van Ooij C, Tamez P, Bhattacharjee S, *et al.* (2008) The malaria secretome: from algorithms to essential function in blood stage infection. *PLoS Pathog* **4**: e1000084.

[95] Xue H, Lu B & Lai M (2008) The cancer secretome: a reservoir of biomarkers. *J Transl Med* **6**: 52.

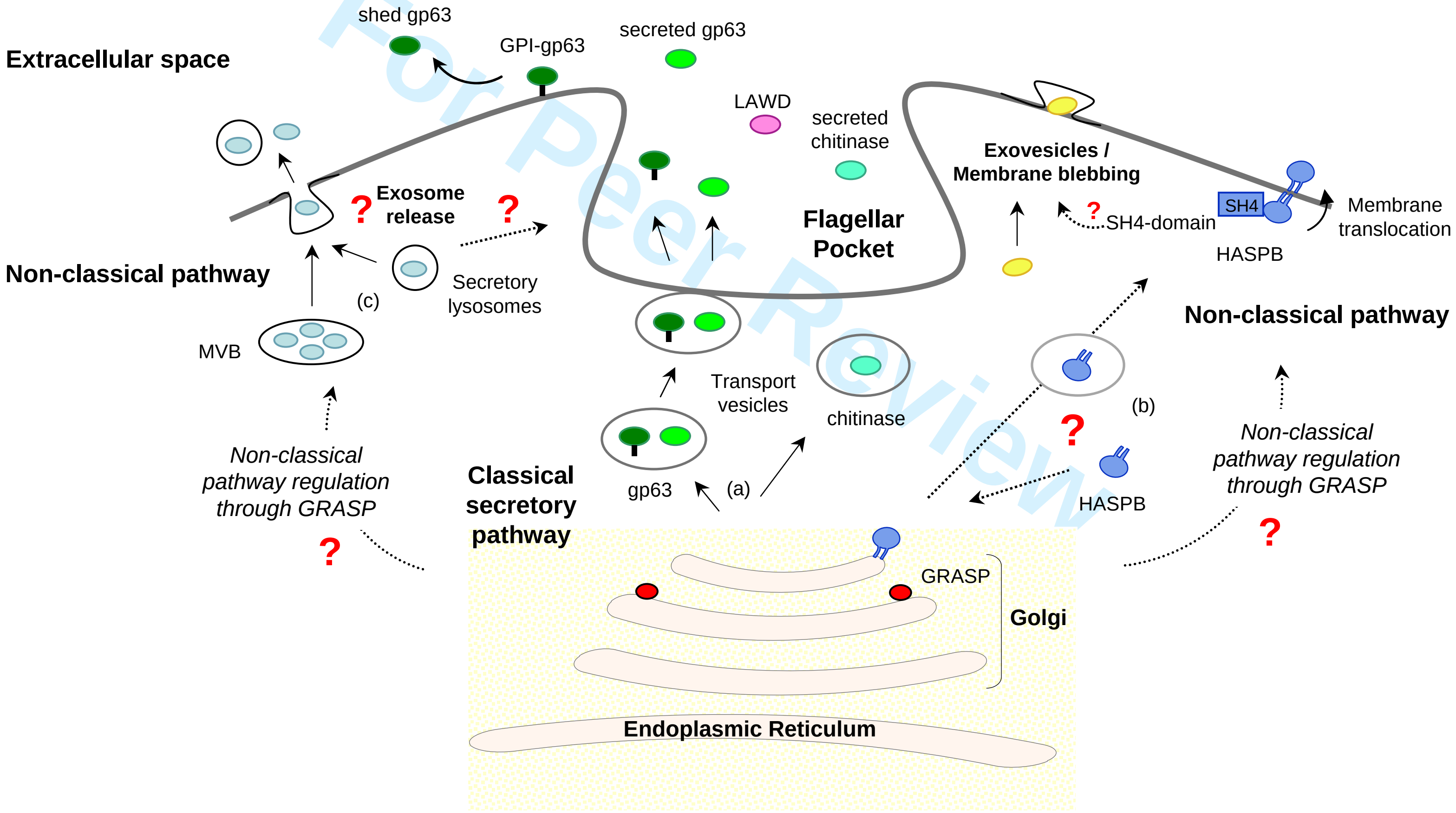
[96] Yamane H, Lee SJ, Kim BD, Tao R & Rose JK (2005) A coupled yeast signal sequence trap and transient plant expression strategy to identify genes encoding secreted proteins from peach pistils. *J Exp Bot* **56**: 2229-2238.

[97] Yokoyama-Yasunaka JK, Pral EM, Oliveira Junior OC, Alfieri SC & Stolf AM (1994) Trypanosoma cruzi: identification of proteinases in shed components of trypomastigote forms. *Acta Trop* **57**: 307-315.

[98] Zwickl H, Traxler E, Staettner S, *et al.* (2005) A novel technique to specifically analyze the secretome of cells and tissues. *Electrophoresis* **26**: 2779-2785.

Legend

Fig. 1. Schematic diagram summarizing the different secretion mechanisms and their putative regulation in *Leishmania*. (a) Proteins secreted through the classical secretory pathway, like gp63 or chitinase, are exported by secretory vesicles and released into the extracellular space of the flagellar pocket by fusion of the vesicles with the plasma membrane. GPI-anchoring proteins are swept out of the flagellar pocket to the cell body and attached to the external surface of the membrane by their GPI moiety. Gp63 is released in a secreted soluble form and a GPI-anchored form. Some proteins devoid of a signal peptide are localized in the flagellar pocket (e.g. LAWD: *Leishmania* antigenic tryptophan-aspartic acid (Campbell*, et al.*, 2004) (b) Hydrophilic acylated surface protein B (HASPB) is synthesized on free ribosomes in the cytoplasm and may be transferred to the outer leaflet of the Golgi membrane and would use conventional vesicular transport to reach the plasma membrane where translocation could occur. The SH4 domain of HASPB induces the production of non-apoptotic membrane blebbing (c) Proteins may be released to the extracellular space through exosomes originating from either lysosomes (e.g. cysteine proteinases (Duboise*, et al.*, 1994) or Multivesicular bodies (MVB) of endosomal origin. Whether GRASP (Golgi reassembly stacking protein), essential to the unconventional secretion pathway of *Dictyostelium* during development, has a role in the non-classical secretion pathway in *Leishmania* remains to be clarified.


shed gp63
GPI-gp63
secreted gp63
Extracellular space
LAWD
secreted
chitinase
Exovesicles /
Membrane blebbing
Exosome
release
?
?
?
SH4-domain
SH4
Membrane
translocation
HASPB
Flagellar
Pocket
Non-classical pathway
Secretory
lysosomes
(c)
MVB
Non-classical pathway
Transport
vesicles
chitinase
(b)
?
HASPB
Non-classical
pathway regulation
through GRASP
?
gp63
(a)
Classical
secretory
pathway
Non-classical
pathway regulation
through GRASP
?
GRASP
Golgi
Endoplasmic Reticulum

F. GENERAL DISCUSSION AND PERSPECTIVES

Neglected diseases are defined as a group of tropical infections which in particular are endemic in low-income populations in developing regions. The lack of satisfactory treatment and the investment required to bring new compounds to market has proven a major disincentive for drug development to treat these neglected diseases.

In Latin America, Chagas disease and Leishmaniasis are neglected diseases representing a significant cause of morbidity and mortality. No vaccines are available and current drug treatments are far from satisfactory. The development of resistance mechanisms to frontline drugs in *Leishmania* and to insecticides by vectors of Chagas disease represents some of the obstacles to combat these diseases.

Clearly, new tools and more effort are needed to combat the trypanosomatid parasites *Trypanosoma cruzi* and *Leishmania spp*., the etiological agents of Chagas disease and Leishmaniasis, respectively. The development of the TDR targets database (supported by the WHO) aiming to integrate genomic data, represents a remarkable effort towards the identification of drug targets in pathogens causing neglected diseases. Nevertheless, given that about 50% of protein-coding genes in trypanosomatids encode hypothetical proteins, the identification of drug targets from genome sequences represents a significant challenge.

The studies conducted in this thesis were mainly devoted to the identification and characterization of novel secreted proteins in trypanosomatids. Because secreted proteins are likely to play pivotal roles in the host-parasite relationship and/or biology of these parasites, we aimed to identify extracellular proteins that are conserved among trypanosomatids. By using a reverse genetic strategy, based on the presence of N-signal peptides in protein sequences, we first designed a genome-based approach for the identification of extracellular proteins. Subsequently, proteins identified with our methodology were analyzed in order to gain further insights regarding their biological importance in *Leishmania* and *T. cruzi*.

1. A new approach for the identification of extracellular proteins in trypanosomatids involved in the classical pathway

The post-genomic era for trypanosomatid parasites has begun with the completed genome sequences of *T. brucei*, *T. cruzi* and *L. major*. The individual and combined comparative analysis of these three genome sequences has revealed some peculiar aspects of trypanosomatid biology. Despite having diverged 200–500 million years ago and causing different diseases, trypanosomatid possess a highly syntenic and conserved genomic organization (El-Sayed et al., 2005b).

Predicted functions have been ascribed by sequence homology to only 40% of the protein-coding genes, but this has been confirmed experimentally for only 5% of the proteins. Most of the remaining genes encode conserved hypothetical proteins, of which slightly more than half are found only in trypanosomatids. Given the significant number of predicted genes in the trypanosomatid genomes annotated as hypothetical or conserved hypothetical, additional molecular, functional and genetic approaches are needed to ascertain their function. Achieving these goals is more feasible now due to a combination of the array of techniques developed over the last decades for the genetic manipulation of trypanosomatids and the arsenal of bioinformatic tools available to exploit genome sequences.

In an attempt to provide a new approach to analyse the large number of hypothetical conserved proteins in trypanosomatids, we designed a genome based approach to identify secreted proteins involved in the classical secretory pathway (presented in Chapter 1).

We focused our research on extracellular secreted proteins because they are likely to play important roles in the host-parasite relationship due to their localization. Furthermore, extracellular factors are known to elicit strong immunity and protection against infection by *Leishmania* in mice and dogs (Tonui et al., 2004; Lemesre et al., 2007). Although secreted proteins may represent putative virulence factors and/or a source of antigens for vaccine development, only a few secreted proteins have been fully characterized in trypanosomatids.

In our search for novel secreted proteins, we hypothesized that a phylogenetic conservation among *Leishmania spp.*, *T. cruzi* and *T. brucei* would indicate evolutionary selection for this family of proteins and suggest an important role for such secreted proteins in the biology of these parasites. With these assumptions in mind, we screened the *T. cruzi* genome and selected a subset of 13 hypothetical proteins conserved among trypanosomatids and potentially secreted via the classical eukaryotic pathway. We confirmed that these proteins were secreted extracellularly in about 25% (3/13) of the selected proteins by using *L. infantum* recombinant parasites expressing *T. cruzi* genes. In order to verify that secretion was not related to the heterologous system, two *L. infantum* ortholog proteins were analyzed with the same experimental system to confirm their secretion in transgenic *L. infantum* promastigotes (homologous system).

Identification of secreted proteins has been hampered in trypanosomatids due to the difficulty in distinguishing genuine secretions from molecules released by lysed, dead, or dying organisms. When developing our methodology we were aware of this potential problem. Therefore, we decided to limit the incubation time of recombinant parasites before detecting the presence of tagged extracellular proteins in culture supernatants. Likewise, the β tubulin protein was used as a negative control for protein secretion. Under these conditions we were able to confirm as extracellular 25% of the selected N-signal peptide containing proteins, suggesting that the classical pathway is operational in trypanosomatids allowing proteins to be exported to the extracellular space.

The plethora of bioinformatic prediction methods developed in the last years for both N-signal peptide and cellular localization exhibit significant variation in sensitivity and specificity due to their reliance on different algorithms (Sprenger et al., 2006; Klee and Sosa, 2007). We based our *in silico* predictions on the Signal peptide server since it has shown to be one of the most accurate for predicting the presence of signal peptides in protein sequences (Klee and Ellis, 2005). However, it is important to remember that the N-signal peptide by itself is not a guarantee for extracellular localization. This is because proteins involved in the classical pathway can be directed to intracellular compartments, such as, lysosomes. The accuracy of predictions for eukaryote proteins possessing an N-terminal signal sequence is significantly improved by using multiple analysis methods (Klee and Ellis, 2005). Thus, for broader analyses of trypanosomatid genomes, the sensitivity of our methodology may be improved by combining several programmes. Likewise, programmes for cellular localization may be included for more in depth *in silico* predictions. An approach that enhances the prediction accuracy of mitochondrial proteins by up to 92% was developed by integrating the available subcellular localization prediction tools (Shen and Burger, 2007).

Concerning experimental validation, our methodology possesses two main advantages. Firstly, detection of secreted proteins is not limited to a particular parasite stage since the approach relies on the use of transgenic *Leishmania* parasites to detect extracellular proteins. Secondly, it allows the detection of small quantities of extracellular proteins that otherwise may be missed by other methodologies, such as, quantitative mass spectrometry. Recently, the secretome of *L. donovani* promastigotes was analyzed by using a quantitative proteomic approach based on the SILAC methodology (Silverman et al., 2008). Although the authors identified a total of 358 proteins secreted into a promastigote conditioned medium, the SILAC methodology was unable to detect proteins that are well known to be secreted, e.g., chitinase and SAcP (secreted acid phosphatase) (Stierhof et al., 1998; Joshi et al., 2005). This may be explained by the relative ratios (extracellular versus cell-associated protein) used by the SILAC approach to identify extracellular proteins. For example, this methodology is limited in its ability to detect proteins where the majority is secreted extracellularly and nothing detectable remains within the cell. Consequently, no reliable ratio can be calculated. In this regard, our methodology may complement the SILAC approach for the identification of proteins missed when using proteomic-based approaches.

Since no perfect methodology exists, highlighting the advantages of any method is as important as pointing out its limitations. The use of an over-expression system for the detection of extracellular proteins does not reproduce natural conditions of protein production and may lead to the identification of false positive secreted proteins. Although secretion was not detected for the negative control (β tubulin protein), it is important to point out that the definitive confirmation of secretion can only be achieved by analyzing target proteins in wild type parasites using specific antibodies. However, it is likely that secretion could be validated with specific antibodies, as suggested by a recent screening of the *T. cruzi* genome to identify hypothetical secreted or membrane anchored proteins (da Silva et al., 2009). The authors selected one potentially secreted protein based on the presence of an N-Signal peptide and it was characterized as extracellular by using specific polyclonal antibodies (da Silva et al., 2009). Interestingly the recombinant form of the protein adheres to mammalian cells in a dose-dependent manner and up-regulates phagocytosis-like activity, suggesting the involvement of this secreted protein in *T. cruzi* cell invasion (da Silva et al., 2009). This study provides evidence supporting the reliability of analysing the genome sequence of trypanosomatids to identify novel genuinely secreted proteins representing potential virulence factors. Likewise, since only one protein was selected to verify secretion in wild type parasites, this study underlines the limitations of selecting several proteins to experimentally detect genuine secretion. Producing specific antibodies for individual proteins is time-consuming and clearly is not scalable to whole genomes. Our main goal was to set up a fast and reliable methodology to simultaneously identify secreted proteins by selecting several genes in trypanosomatid genomes. Thus, we aimed to provide a potential tool for genome-wide screening for the identification of extracellular proteins that could represent the basis for further detailed studies (e.g. on biological properties). Once additional data provides evidence on the relevance of the target proteins, it would then be worthwhile developing specific antibodies for functional studies of these proteins in wild type parasites.

2. Extracellular proteins as potential virulence factors and/or immunogenic components

Once we set up our methodology, we identified three proteins of unknown function that are conserved among trypanosomatids and likely to be secreted via the classical eukaryotic pathway. Subsequently, we aimed to gain insight into whether these proteins could be important for the biology of these parasites. The lack of bioinformatic data about function for genes coding hypothetical conserved proteins among trypanosomatids represents a significant challenge for subsequent functional analyses. Domain or motif search tools identified a signature and a protein domain in two of the three conserved proteins. The protein-coding gene Tc00.1047053505789.10 from *T. cruzi* presents a lipocalin signature. Lipocalins are a large and diverse group of mostly extracellular proteins, exhibiting great structural and functional diversity (Grzyb et al., 2006). Interestingly, some lipocalin proteins are involved in regulating host immune responses (Grzyb et al., 2006). In spite of a conservative tertiary structure, the identity of amino acid sequences is low among lipocalins, and most often it does not exceed 30%, which makes it difficult to predict lipocalins from protein sequences (Flower et al., 2000). This may explain the absence of a lipocalin signature in Tc00.1047053505789.10 ortholog genes from *Leishmania spp.* and *T. brucei* genomes. The only sure way to determinate whether the protein Tc00.1047053505789.10 from *T. cruzi* is a lipocalin would require analysis of NMR or x-ray crystallography data. Remarkably, the genome of *T. cruzi*, *L. major* and *T. brucei* possess 29, 9 and 13 genes, respectively, containing a lipocalin signature in their sequences. Further analyses are required to ascertain whether lipocalins are present in trypanosomatid parasites and to decipher their functions, which may include regulation of immune responses.

A Glucosidase II (GII) beta subunit-like protein domain is present in the protein-coding gene Tc00.1047053506155.99. GII is a heterodimer; composed of catalytic (GIIα) and ER retention (GIIβ) subunits (D'Alessio et al., 1999). This enzyme plays a pivotal role in the processing of *N*-oligosaccharides in the Endoplasmic reticulum (ER) (D'Alessio et al., 1999). GII is not functionally characterized in *T. cruzi* although GII activity has been detected from intact *T. cruzi* cells and cell-free assays (Bosch et al., 1988). In contrast, GII from *T. brucei* is non-essential for the growth of null mutant GII in the bloodstream form, but it is required for correct maturation of VSG (Jones et al., 2005). The extracellular localization detected for the protein Tc00.1047053506155.99 is unexpected since GII is mostly localized to the ER. Nevertheless, GII was found in endocytic structures beneath the plasma membrane and has been associated with protein-tyrosine phosphatase CD45 (Li et al., 1996). A striking observation is the prediction for only the GIIβ subunit and the absence of the catalytic subunit (GIIα) in the Tc00.1047053506155.99 protein sequence. This suggests prediction of GIIβ is not accurate, which seems unlikely since its orthologs in *L. infantum* and *T. brucei* show GII related predicted domains. Further studies of these proteins are required to ascertain whether they are genuine GII enzymes or have different functions in trypanosomatids.

Due to scarce information in gene sequences coding for secreted proteins, we analysed their biological properties to gain additional information about their possible relevance for parasite survival. On one hand we investigated whether over-expression of the secreted proteins in *Leishmania* parasites increased their ability to replicate inside their host-cell (Chapter 1). On the other hand, we attempted to determinate the immunogenicity of the *T. cruzi* secreted proteins. Because secreted/excreted factors of *T. cruzi* are well known as immunogenic, this information could be useful for developing new diagnostic tools for Chagas disease (Chapter 2).

Infectivity tests showed that the gene LinJ19.0410 from *L. infantum* (Ortholog gene of Tc00.1047053505789.10) is involved in a process increasing replication of the parasite inside human and murine macrophages. Given that the protein LinJ19.0410 does not have a predicted protein domain, is difficult to explain its role for the parasite. Enhanced infectivity was not detected by analysing *L. infantum* parasites over-expressing the *T. cruzi* ortholog gene. We attempted to obtain transgenic *T. cruzi* parasites expressing the protein Tc00.1047053505789.10 to analyse their infectivity. Although we obtained drug-resistant parasites, expression of the tagged protein was not detected. Genetic manipulation in *T. cruzi* is not as easy to perform as in *Leishmania* parasites (Beverley, 2003). Indeed, only a few genes have been experimentally characterized and validated in *T. cruzi*, mainly due to the lack of simple methods for gene manipulation needed for reverse genetic studies. Hence, we analysed the effect of heterologous expression of the protein Tc00.1047053505789.10 and the other secreted genes by using the same system in *L. infantum* promastigotes. Transgenic *Leishmania* parasites over-expressing either of the *T. cruzi* secreted proteins showed no significant increase in infectivity when compared to control parasites (additional results Chapter 1). Since *Leishmania* and *T. cruzi* infect different target cells, further studies in *T. cruzi* are needed to confirm whether the protein Tc00.1047053505789.10 is a virulence factor as suggested in *Leishmania*. Recently, a Multisite Gateway (MS/GW) based method was developed to improve specific gene deletions in epimastigotes of *T. cruzi* (Xu et al., 2009). By using MS/GW knock out constructs obtained in only five days, the authors successfully obtained simple knockouts of *T. cruzi* epimastigotes for the dihydrofolate reductase-thymidylate synthase (*dhfr-ts*) and enoyl-CoA hydratase (*ech*) genes (Xu et al., 2009). This methodology could be used to produce knockout parasites for the gene Tc00.1047053505789.10 and determinate whether this protein is a *T. cruzi* virulence factor. Similarly, further studies are needed to investigate the immunogenicity of this protein. The different strategies we tried to produce and purify the Tc00.1047053505789.10 recombinant form were unsuccessful (described in Chapter 2).

Currently, this protein is being produced in a baculoviral system, which allows the expression of proteins in insect cells. The baculoviral system has been widely used for the production of recombinant proteins in insect cells due to the advantage of using the cellular machinery of insect cells to produce post-translational modifications of proteins, (absent in *E. coli*) (Hu, 2005).

Concerning the two remaining *T. cruzi* recombinant proteins, we were able to test their ability to recognize *T. cruzi* antibodies by analysing sera from chagasic patients. Results showed that the recombinant protein Tc00.1047053506155.99, predicted as one of the most antigenic in the *T. cruzi* genome, was not recognized by chagasic sera (Chapter 2). In contrast, the protein Tc00.1047053509999.10 proved to be immunogenic, since it is able to bind to IgGs present in sera from chagasic patients from different endemic areas. Furthermore, the recombinant antigen was not detected by patients infected by *L. infantum*. Hence, this protein represents a potential antigen for the development of serological diagnostic tools based on mixed recombinant proteins. Conventional serologic tests for *T. cruzi* suffer from low sensitivity and specificity, the latter being due to cross-reactions with other pathogens, including *Leishmania* parasites (Tarleton et al., 2007). Misdiagnosis of *T. cruzi* infections is a major problem in areas where both diseases overlap, including north-western Argentina (Chiaramonte et al., 1996; Frank et al., 2003). This has a great socioeconomic impact, since many employers reject workers with positive serology for Chagas disease. Higher specificities can be obtained with tests based on recombinant antigens (da Silveira et al., 2001). Nevertheless, the use of single recombinant antigens showed geographical variation in sensitivity, possibly because the genetic heterogeneity among *T. cruzi* strains present in endemic regions (Verani et al., 2009). Therefore, the inclusion of several recombinant antigens in ELISA tests has been suggested as a strategy to improve sensitivity for the diagnosis of Chagas disease (Umezawa et al., 2003; Umezawa et al., 2004; Cooley et al., 2008).

Collectively our data on the biological properties of secreted proteins provide evidence that our methodology allowed us: i) to identify a potential virulence factor of *L. infantum* (gene LinJ19.0410), as suggested by *in-vitro* infection tests, (ii) to identify a *T. cruzi* immunogenic protein, as showed by analysis of recombinant proteins (Tc00.1047053509999.10). Although we centred our search on secreted proteins that are conserved among trypanosomatids, our results are based on *T. cruzi* and *Leishmania* as models for experimental validation. The next step would be to analyse these secreted proteins in *T. brucei*. A major feature that experimentally differentiates *T. brucei* from *T. cruzi* and *Leishmania* is the presence of an active RNA interference (RNAi) mechanism (Motyka and Englund, 2004). The availability of an RNAi system in *T. brucei* has provided a robust tool to assess the functional importance of a particular gene in the development or survival of the parasite. Subramanian *et al* carried out the first systematic study of gene function in *T. brucei* using a high-throughput RNAi approach to test whether ORFs potentially coding for a gene were essential for the organism (Subramaniam et al., 2006). A total of 210 genes annotated on chromosome I were targeted in the bloodstream form of the parasite and the effects of knocked-down genes were tested using several assays and protocols. Over 30% of the chromosome I genes generated a phenotype when targeted by RNAi; most commonly affecting cell growth, viability, and/or cell cycle progression. Remarkably, the majority of *T. brucei* genes with associated growth defects are hypothetical conserved proteins among trypanosomatids, some of which may represent potential drug targets (Subramaniam et al., 2006).

To further analyse the subset of hypothetical conserved proteins in trypanosomatids, it would be interesting to combine the experimental system proposed by Subramanian *et al* with our methodology. Combining these approaches would make it possible to analyse proteins whose secretion is conserved in trypanosomatids and whether they are required for cell growth, viability, and/or cell cycle progression *in-vitro*. Subsequently, target proteins identified with these methodologies could be validated as virulence factors in *Leishmania* by using the experimental approach described by Zhang *et al* (Zhang et al., 2008). By using transgenic *L. major* parasites expressing *L. donovani* genes, the authors identified a novel protein with unknown function increasing pathogenesis in mice. Aiming to identify specific genes of *Leishmania* parasites involved in pathogenesis, the authors initially selected seven genes present in *L. infantum* but absent in the *L. major* genome. During the course of the study, one of the selected genes was reclassified as being present in the *L. major* genome, and interestingly only this conserved gene increased visceral infection in transgenic *L. major* parasites (Zhang et al., 2008). These results provide evidence that ubiquitously shared genes among trypanosomatids are virulence factors reinforcing the utility of analysing conserved genes.

3. Secretion in trypanosomatids: Classical or non-classical mechanisms?

Most knowledge of the cell biology of trypanosomatids, including secretory pathway and cellular trafficking, has been obtained from studies involving *T. brucei* (Field et al., 2007). The robust tools available for genetic manipulation of *T. brucei* have facilitated the use of this parasite as a model system for functional analyses in trypanosomatids (Field et al., 2007). Several studies have demonstrated that the basic features of the trypanosomatid secretory-endocytic pathways are very similar to those found in other eukaryotes (McConville et al., 2002). Therefore, as in higher eukaryotes, one would expect the classical secretion pathway is an important mechanism in the export of secreted proteins to the extracellular space. On the basis of an active classical secretory pathway in trypanosomatids and the importance of extracellular proteins in host-parasite interactions, we centred our research on the identification of novel secreted proteins.

During the course of this thesis, the first quantitative proteomic analysis of the *L. donovani* secretome was published (Silverman et al., 2008). Surprisingly, this study indicated that *L. donovani* promastigotes in the stationary phase of growth predominantly use non-classical secretion mechanisms to direct export protein, including the release of exosome-like microvesicles (Silverman et al., 2008). Thus, secretion in *Leishmania*, and probably *T. cruzi*, is likely to be a heterogeneous process. It is well known that *T. cruzi* and *Leishmania* parasites release significant amounts of antigenic factors into the extracellular space (Colli, 1993; Ilg et al., 1994; Yokoyama-Yasunaka et al., 1994; Jaffe and Dwyer, 2003). In contrast to *T. cruzi* and *Leishmania*, apparently *T. brucei* does not release detectable levels of extracellular molecules *in-vitro*. Most studies involving *T. brucei* have focused on VSG proteins determining immune evasion by antigenic variation (Taylor and Rudenko, 2006). Because endocytosis is involved in immune evasion, exocytosis has received substantially less attention in *T. brucei* (Field et al., 2007). Therefore, little is known about the exoproteome of *T. brucei* and consequently for trypanosomatids as a whole. This is due to the common extrapolation of results from *T. brucei* to other trypanosomatids, including *Leishmania* and *T. cruzi*. However, analysis of the *T. congolense* exoproteome showed that most of the proteins identified by mass-spectrometry are apparently secreted through unconventional mechanisms, as suggested in *L. donovani* (Grebaut et al., 2009). These findings underline our scarce knowledge on the complexity of the exoproteome and secretion mechanisms used by trypanosomatids to export proteins to the extracellular environment.

Given the relevance of extracellular factors to the biology of these parasites, understanding the complexity of the trypanosomatid exoproteome is a key component to clarify how secretion mechanisms are used to ensure parasite survival within their hosts. Based on recent findings from the *Leishmania* exoproteome, a review deciphering extracellular components produced by *Leishmania* parasites, including the so called "secretome" is presented in the Chapter 3. The secretion mechanisms described to date in *Leishmania*, as well as recent data shedding light on regulatory components of unconventional secretion and their potential involvement in *Leishmania* secretion, are also discussed.

G. CONCLUDING REMARKS

Experimental confirmation of *in-silico* predicted genes is very important for determining their potential function in trypanosomatid genomes. Hence, the research conducted during this thesis was aimed at exploring the large number of hypothetical proteins to identify secreted proteins representing potential virulence and/or antigenic factors. By using *Leishmania* and *T. cruzi* as models, we set up an experimental system combined with tests to assess the importance of secreted proteins in trypanosomatids. Despite the difficulties encountered to produce recombinant forms of the secreted proteins, which limited the functional characterization of the target proteins, collectively, the data presented provides evidence to support the following conclusions:

(i) The genomic-based-approach designed for the identification of extracellular proteins detected 25% of the predicted secreted proteins in the extracellular space;

(ii) The classical secretory pathway is active in trypanosomatids, allowing proteins containing a predicted N-Signal peptide to be exported to the extracellular space;

(iii) The methodology developed allows the identification of potential virulence factors in *Leishmania,* as indicated by *in-vitro* infectivity tests;

(iv) The methodology is useful for the identification of immunogenic *T. cruzi* proteins, representing potential targets for the development of new diagnostic tools for Chagas disease;

It would be worthwhile to:

(i) determine by gene-knockout whether the gene Tc00.1047053505789.10 from *T. cruzi* and its orthologs are virulence factors in trypanosomatids;

(ii) analyse the immunogenicity of the recombinant form of the protein Tc00.1047053505789.10 produced in a eukaryotic expression system (baculovirus/insect cells). Likewise, the recombinant form would be useful to produce specific antibodies for functional analysis (e.g. cellular trafficking);

(iii) validate the usefulness of the recombinant protein Tc00.1047053509999.10 for the diagnosis of Chagas disease through the development of an ELISA test.

Efforts towards understanding the protein coding genes of trypanosomatids would require combining different tools. In this regard, data provided in this thesis represent a starting point for broader analyses of novel secreted proteins representing virulence and/or antigenic factors in trypanosomatids. Hopefully, further functional screening will greatly enhance both our understanding of the basic biology of trypanosomatids and the identification of new drug targets. Finally, identification of an immunogenic *T. cruzi* protein achieved in this thesis represents a valuable step towards improved diagnosis of infection in my native region in Argentina where the specific detection of Chagas disease and Leishmaniasis remains a challenge.

H. ADDITIONAL WORK

Article:

Congenital Chagas disease involves *Trypanosoma cruzi* sub-lineage IId in the northwestern province of Salta, Argentina. **Rosa M. Corrales**, Maria C. Mora, Olga Sanchez Negrette, Patricio Diosque, Diego Lacunza, Myrna Virreira, Simone F Brenière and Miguel A Basombrio. 2009. *Infect, Genet Evol*., 9(2):278-282

During this thesis I had the opportunity to work in collaboration with my former laboratory in Argentina. One of the main research projects there is the study of congenital Chagas disease, representing the main source of *T. cruzi* infections in Northwestern Argentina. In this article we described the genetic characterization of strains involved in the vertical transmission of *T. cruzi*. Our results showed that all *T. cruzi* isolates obtained from congenitally infected newborns in Northwestern Argentina belong to the *T. cruzi* IId lineage. These results provide further evidence supporting the hypothesis that human infection by *T. cruzi* in the Southern Cone countries of Latin America is due principally to *T. cruzi* II.

Infection, Genetics and Evolution 9 (2009) 278–282

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Congenital Chagas disease involves *Trypanosoma cruzi* sub-lineage IId in the northwestern province of Salta, Argentina

Rosa M. Corrales [a,*], Maria C. Mora [b], Olga Sanchez Negrette [b], Patricio Diosque [c], Diego Lacunza [b], Myrna Virreira [d], Simone F. Brenière [e], Miguel A. Basombrio [b]

[a] *Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Institut de Recherche pour le Développement, 911 Av. Agropolis, 34394 Montpellier, France*
[b] *Instituto de Patología Experimental (INSIBIO-CONICET), Universidad Nacional de Salta, Av. Bolivia 5150, CP 4400 Salta, Argentina*
[c] *Unidad de Epidemiología Molecular, Instituto de Patología Experimental, Universidad Nacional de Salta, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Bolivia 5150, CP 4400 Salta, Argentina*
[d] *Faculté de Médecine, Université Libre de Bruxelles, 808 route de Lennik CP 611, B-1070 Brussels, Belgium*
[e] *Département Sociétés et Santé, UR016 Caractérisation et contrôle des populations de vecteurs, Représentation en Bolivie, Av Hernando Siles N°5290 Esq Calle 7 Obrajes La Paz, Bolivia*

ARTICLE INFO

Article history:
Received 1 November 2008
Received in revised form 15 December 2008
Accepted 17 December 2008
Available online 31 December 2008

Keywords:
Congenital Chagas disease
Lineages
Trypanosoma cruzi IId
Multilocus enzyme electrophoresis (MLEE)
Random amplified polymorphic DNA (RAPD)

ABSTRACT

Trypanosoma cruzi is genetically classified into six discrete phylogenetic lineages on the basis of different genetic markers. Identifying lineages circulating among humans in different areas is essential to understand the molecular epidemiology of Chagas disease. In the present study, 18 *T. cruzi* isolates from congenitally infected newborns in the northwestern province of Salta-Argentina were studied by multilocus enzyme electrophoresis (MLEE) and random amplified polymorphic DNA (RAPD). All isolates were typed by MLEE and RAPD as belonging to *T. cruzi* IId. Analysis of minor variants of TcIId using probes hybridizing with hypervariable domains of kDNA minicircles, detected three variants with a similar distribution among the isolates. Our findings confirm the presence of *T. cruzi* IId among congenitally infected newborns in northwestern Argentina and support the assumption that human infection by *T. cruzi* in the Southern Cone countries of Latin America is due principally to *T. cruzi* II.

1. Introduction

Congenital transmission is the principal mode of infection by *Trypanosoma cruzi*, the agent of Chagas disease, in areas where transmission by vectors and blood transfusions are controlled. In endemic regions, the transmission rates of congenital Chagas disease is extremely variable, ranging from 1% in Brazil to 4–12% in Argentina, Bolivia, Chile, and Paraguay (Gurtler et al., 2003; Carlier, 2005; Brutus et al., 2008). In addition, the morbidity and mortality of congenital infection also varies from being asymptomatic to severe and lethal clinical forms of disease (Torrico et al., 2004; Sanchez Negrette et al., 2005; Virreira et al., 2006). Host factors, such as the level of placental defense, and/or the maternal and fetal capacities to develop a specific immune response for the control of parasitic multiplication, can be involved in these differences (Hermann et al., 2002, 2004). An alternative hypothesis is that these differences are related to genetic polymorphism among parasite strains (Macedo and Pena, 1998).

Clonal populations of *T. cruzi* harbor considerable genetic polymorphism and different parasite genotypes might play an important role in creating variation in the pathology they cause in different regions (Macedo and Pena, 1998). Multilocus enzyme electrophoresis (MLEE) is currently the gold standard for strain typing which has lead to sub-division of the taxon into two major lineages, *T. cruzi* I and *T. cruzi* II (Barnabe et al., 2000; Brisse et al., 2000). Each lineage is genetically heterogeneous and despite a failure to identify sub-divisions within *T. cruzi* I, five sub-groups of *T. cruzi* II have been proposed on the basis of biochemical and molecular markers (Brisse et al., 2000). These six discrete genetic subdivisions, or lineages, have been proposed as a reference framework for genetic variability and biological characterization of *T. cruzi* stocks. Therefore, identifying lineages circulating among humans in different areas is important to evaluate the possible influence of infective genotypes on the broad spectrum of observed clinical disease.

---

* Corresponding author. Tel.: +33 4 67 41 61 47; fax: +33 4 67 41 61 47.
E-mail addresses: corrales@mpl.ird.fr, milagroscorrales@yahoo.com.ar (R.M. Corrales).

1567-1348/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2008.12.008

In previous works we have reported a congenital transmission rate of 9% and a family clustering of vertical transmission of *T. cruzi* in the province of Salta, Argentina (Mora et al., 2005; Sanchez Negrette et al., 2005). Since knowledge about the genetic diversity of *T. cruzi* isolates circulating within a country is essential to understand the molecular epidemiology of Chagas disease, we aimed to characterize *T. cruzi* lineages in 18 isolates from congenitally infected newborns in the province of Salta by using three different molecular markers. MLEE and RAPD analyses were used as standard markers for lineage identification. Additionally, new sensitive probes designed to distinguish variants within the *T. cruzi* IId sub-lineage were used to asses genetic variability of *T. cruzi* isolates.

2. Materials and methods

2.1. Study area

This study was conducted in the city of Salta, in northwestern Argentina, with approximately 500,000 inhabitants. Triatomine vectors were eradicated from the city in 1970s, and no re-infestations have been registered in the area where this study was performed.

2.2. Sample collection

Samples were collected from July 1997 to December 2001 in the framework of a collaborative project to improve diagnostic methods for Congenital Chagas disease in Salta city (Mora et al., 2005). Mothers were identified as bearing *T. cruzi* infection by using standard parasite-specific serological tests (ELISA and indirect hemagglutination IHA) in duplicate samples as previously described (Mora et al., 2005). Samples from newborns were mainly obtained from umbilical cords and occasionally from venous punctures performed during the first 15 days of life. Umbilical cords were collected in the Hospital Materno Infantil of Salta, as previously described (Mora et al., 2005). All congenital infections were subjected to treatment. Informed consent was obtained from mothers of babies participating in this study, under a protocol approved by the Ministry of Public Health of the Province of Salta.

2.3. Parasite isolation and reference strains

Eighteen isolates of *T. cruzi* were obtained from umbilical cords and venous puncture following hemoculture in liver infusion tryptose (LIT) medium. The isolated parasites were maintained by passage in LIT medium at 28 °C. Table 1 summarizes the geographical origin and clinical data of the 18 stocks. Four strains were used for reference: X10cl1 (*T. cruzi* I), CANIIIcl1 (*T. cruzi* IIa), Mncl2 (*T. cruzi* IId) and CL-Brener (*T. cruzi* IIe). X10cl1 and CANIIIcl1 correspond to the formerly described zymodemes I and III, respectively (Miles et al., 1978).

2.4. Sample preparation

Stocks were harvested by centrifugation (2800 × g, 20 min, 4 °C) and washed in PBS (Na_2HPO_4 10 mM, NaCl 150 mM, pH 7.2). Cells were lysed on ice for 20 min in an equal volume of hypotonic enzyme stabilizer (EDTA 2 mM, dithiothreitol 2 mM, 1-aminocaproic acid 2 mM). The soluble fraction was stored at −70 °C until used in MLEE analyses, whereas the pellet of lysed cells was used for DNA extraction, as described elsewhere (Brisse et al., 2000). DNA concentration was estimated by spectrophotometry at 260 nm.

Table 1
Origin of 18 *Trypanosoma cruzi* stocks examined and clinical features of respective congenital chagasic patients.

Stocks	Geographical origin of the mother	Clinical form
CONG 1	Salta, Santa Victoria	Asymptomatic
CONG 2	Salta, Colonia Sta Rosa	Anemia
CONG 3	Salta, Cerrillos	Asymptomatic
CONG 4	Salta, Cerrillos	Asymptomatic
CONG 5	Salta, Metan	Asymptomatic
CONG 6	Salta, Gral San Martin	Anemia
CONG 7	Salta, Campo Quijano	Asymptomatic
CONG 8	Salta, Oran	Anemia
CONG 9	Salta, Gral San Martin	Asymptomatic
CONG 10	Salta	Asymptomatic
CONG 11	Salta	Asymptomatic
CONG 12	Salta	Asymptomatic
CONG 13	Salta	Anemia
CONG 14	Salta	Asymptomatic
CONG 15	Salta, Gral San Martin	Jaundice
CONG 16	Salta, Santa Victoria	Asymptomatic
CONG 17	Salta, Lesser	Asymptomatic
CONG 18	Salta, Santa Victoria	Asymptomatic

2.5. Protocol for isoenzyme analysis

Multilocus enzyme electrophoresis analysis was performed as described elsewhere, with slight modifications (Ben Abderrazak et al., 1993). The following 9 enzyme systems were used: diaphorase (EC 1.6.99.2, DIA), glutamate dehydrogenase, NAD^+ (EC 1.4.1.2, GDH-NAD^+), glutamate dehydrogenase $NADP^+$ (EC 1.4.1.4, GDH-$NADP^+$), glutamate oxaloacetate transaminase (EC 2.6.1.1, GOT), glucose-6 phosphate dehydrogenase (EC 1.1.1.49, G6PD), glucose-6-phosphate isomerase (EC 5.3.1.9, GPI), malate dehydrogenase (EC 1.1.1.37, MDH), malic enzyme (EC 1.1.1.40, ME) and phosphoglucomutase (EC 5.4.2.2, PGM). These 9 enzyme systems correspond to 11 different genetic loci, since diaphorase and malic enzymes both exhibit activity of two distinct loci.

2.6. Random amplified polymorphic DNA analysis

We selected six primers that yielded the most discriminating patterns as previously described (Brisse et al., 2000). The primers correspond to the A, B, F and U kits from Operon Technologies: A10 (GTGATCGCAT), B11 (GTAGACCCGT), B19 (ACCCCCGAAG), F5 (CCGAATTCCC), U11 (AGACCCAGAG) and U14 (TGGGTCCCTC). The amplification reactions were performed in a final volume of 60 μl containing 0.9 units Taq Polymerase (Boehringer), 100 μM each dNTP, 200 nM primer, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris ± HCl, pH 8.3, and 20 ng template DNA. Forty-five cycles (denaturation: 1 min at 94 °C; annealing: 1 min at 36 °C; elongation: 2 min at 72 °C) were followed by a final elongation step of 7 min at 72 °C. Random amplified polymorphic DNA products were analyzed by electrophoresis in 1.6% agarose gels in TAE buffer (Tris-acetate 40 mM, EDTA 1 mM), stained with ethidium bromide and visualized by ultra-violet light.

2.7. Hybridizations with synthetic oligodeoxynucleotide probes for variants of lineage IId

Identification of variants of the sub-lineage TcIId was performed by hybridization of DNA amplicons resulting from the Tc121/Tc122 PCR amplification, with oligonucleotide probes as previously described (Virreira et al., 2006). The TcIId variants were grouped according to their relative hybridization pattern with probes Oli-IId-1 and Oli-IId-2, as described by Virreira et al., 2006.

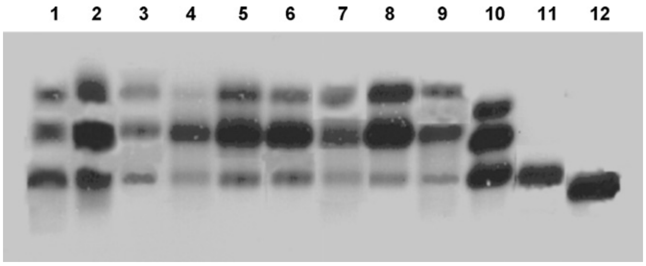


Fig. 1. Patterns obtained for the glucose phosphate isomerase (GPI) locus for 8 of the 18 *T. cruzi* stocks under study. Lines 1–8 correspond to stocks listed in order in Table 1. Line 9: Mncl2 (reference strain for *T. cruzi* IId), Line 10: CL-Brener (reference strain for *T. cruzi* IIe), Line 11: CANIIIcl1 (reference strain for *T. cruzi* IIa), Line 12: X10cl1 (reference strain for *T. cruzi* I).

2.8. Data analysis

MLEE patterns were compared with those described for the major lineages and for additional subdivisions of *T. cruzi* II lineage (Barnabe et al., 2000; Brisse et al., 2000) using the four reference strains. Every band was counted, excepting central bands in heterozygous patterns that showed three bands. These patterns in *T. cruzi* have been attributed to heterozygous genotypes for dimeric enzymes, in which central bands do not correspond to an allele (Tibayrenc et al., 1985).

3. Results

3.1. Multilocus enzyme electrophoresis analysis

Multilocus enzyme electrophoresis analysis was performed at 11 isoenzyme loci on all 18 isolates. According to the patterns observed, an allelic interpretation of MLEE variability was performed based on the hypothesis that *T. cruzi* is a diploid organism (Tibayrenc et al., 1986). Each enzyme extract was analyzed at least twice in parallel with the reference strains. Among all the isolates studied ($n = 18$) only one multilocus genotype was detected from the analysis of zymograms. All stocks presented identical electromorphic patterns at the 11 loci analyzed, corresponding to patterns for the Mncl2 reference. Fig. 1 shows the electrophoresis of glucose-6-phospate isomerase (Gpi). Since all isolates exhibited the specific character Gpi 2/4 genotype of the *T. cruzi* IId lineage all stocks were assigned to this Discrete Typing Unit.

3.2. Random amplified polymorphic DNA polymorphism

As all isolates studied were typed within the same lineage *T. cruzi* IId, we performed RAPD with 6 decameric primers in order to assess genetic polymorphism of *T. cruzi* isolates. All amplifications were performed at least twice confirming the reproducibility of the

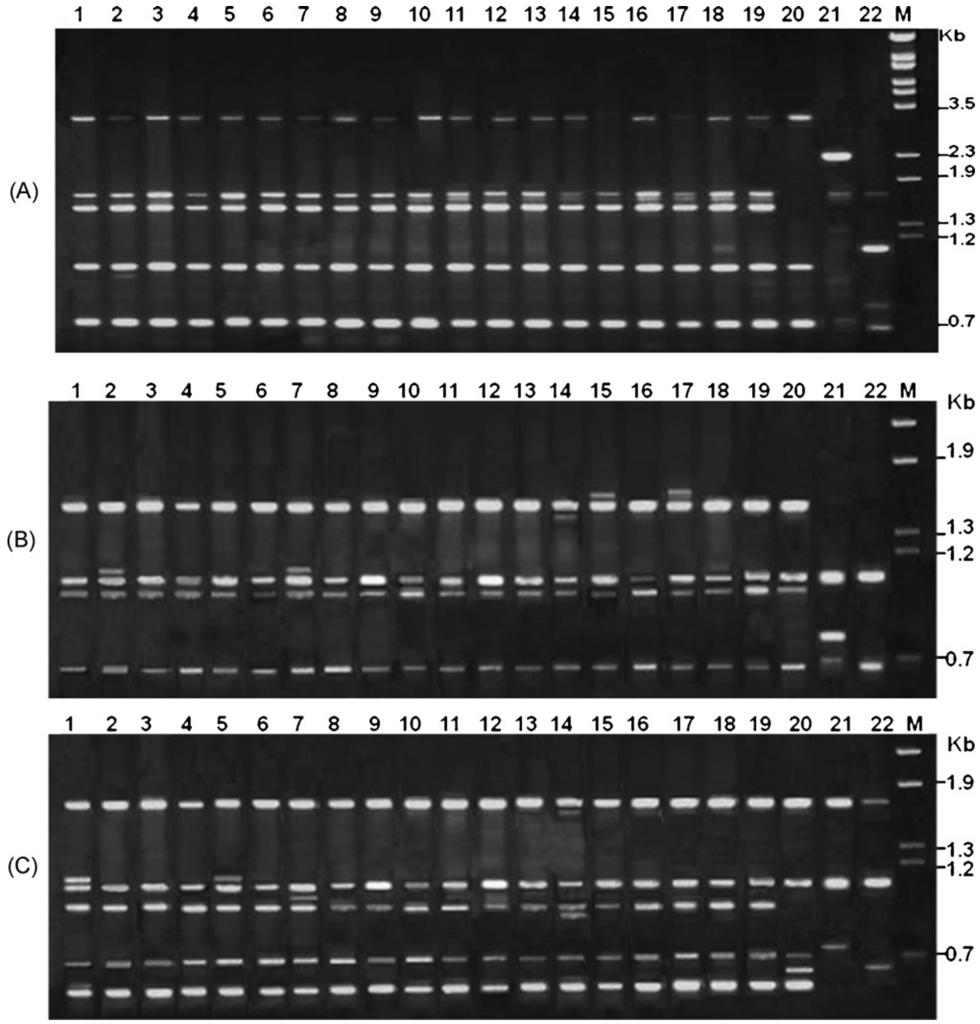


Fig. 2. Agarose gels (1.6%) stained with ethidium bromide showing random amplified polymorphic DNA (RAPD) patterns generated by primers F5 (A), B19 (B) and U14 (C) for 18 *T. cruzi* stocks under study. The numbers above the lanes 1–18 correspond to the stock numbers listed in order in Table 1. Line 19: Mncl2 (reference strain for *T. cruzi* IId), Line 20: CL-Brener (reference strain for *T. cruzi* IIe), Line 21: CANIIIcl1 (reference strain for *T. cruzi* IIa), Line 22: X10cl1 (reference strain for *T. cruzi* I). M: Molecular marker: phage lambda DNA digested with BstEII.

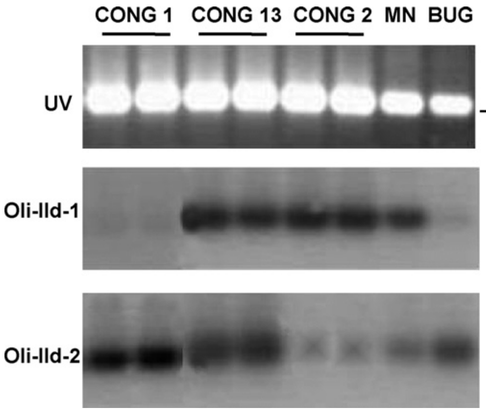


Fig. 3. Relative hybridization of kDNA amplicons with synthetic oligonucleotides probes. The hypervariable regions of *T. cruzi* kDNA were amplified (first line) and the intensity of amplicon fluorescence is compared with autoradiography after hybridization with indicated oligodeoxynucleotides probe Oli-IId-1 and Oli-IId-2 (second and third lines). Amplicons were analyzed in duplicate. MN: Reference strain MNcl2. BUG: Reference strain Bug2148.

patterns. This analysis generally found limited genetic variability among all isolates, since patterns generated with all six primers were very similar. Patterns generated using DNA from Mncl2 reference strains were very similar to those generated for all congenital *T. cruzi* stocks analyzed in the present study. Moreover, the *T. cruzi* IId lineage was specifically identified for all isolates by the presence of an identical pattern obtained with F5 primer that detects specific diagnostic RAPD fragments as previously described (Fig. 2A). Therefore, these results confirm the identity of all stocks as belonging to the *T. cruzi* IId lineage, as previously demonstrated by MLEE analyses. Nevertheless, we detected intraspecific genetic polymorphism among *T. cruzi* isolates using the B19 and U14 primers. Fig. 2B and C shows the patterns obtained with these primers. Pattern differences obtained with these primers were characterized by the absence of some bands in a few isolates, or the presence of an extra band for others.

3.3. T. cruzi typing by hybridization of synthetic oligodeoxynucleotide probes identifying variants of sub-lineage IId

Further characterization with synthetic oligodeoxynucleotide probes was performed for all isolates in order to identify variants of sub-lineage IId, as previously described (Virreira et al., 2006). These probes allow sensitive hybridization of kDNA minicircle derived-amplicons obtained from PCR with primers Tc121/Tc122. The hybridization patterns of the reference strains with Oli-IId-1 and Oli-IId-2 probes allow samples to be classified as the following *T. cruzi* IId variants: MN-like, Bug-like, or TPK-like. Isolates displaying the pattern MN-like and Bug-like hybridized more strongly with Oli-IId-1 and Oli-IId-2, respectively, whereas the TPK-like hybridized equally with both probes. The amplification of the 320-bp was performed in duplicate for all isolates in order to obtain the hybridization pattern. This method allowed *T. cruzi* IId typing of all isolates with the following variant distribution: 6 MN-like, 5 Bug-like, and 7 TPK-like. Hybridization patterns of some stocks are shown in Fig. 3.

The pattern of variation detected in the stocks by RAPD analysis was not correlated with that detected by the hybridization approach (data not shown).

4. Discussion

In the present work we report the genotypic analysis of 18 *T. cruzi* stocks isolated from congenitally infected newborns in the province of Salta, Argentina by using three molecular markers. All isolates were unequivocally typed by MLEE and RAPD as belonging to the *T. cruzi* IId sub-lineage, showing the clear predominance of this sub-lineage among the studied congenitally infected neonates.

Despite the limited number of isolates investigated and the lack of statistical data analysis (both sample size and genetic variability are too small for this purpose) our RAPD analysis showed low genetic polymorphism among isolates with 2 out of 6 decameric primers used in our study. Further analysis by using a new comparative hybridization approach (Virreira et al., 2006) allowed us to detect three variants of the IId with a similar distribution among the isolates studied. This result confirms the previous heterogeneity of this sub-lineage, suggesting the existence of families of lineage-specific sequences (Telleria et al., 2006). Surprisingly, the lack of concordance between the polymorphism revealed by RAPD and the hybridization approach suggests no genetic linkage among these markers, as would be expected for a clonal propagation of this parasite. Further studies are needed to understand the relationship among these molecular markers.

Genetic variability studies in human cultured isolates have been found to belong to a single lineage in other studies (Barnabe et al., 2001; Bosseno et al., 2002; Montilla et al., 2002). Nevertheless identification of *T. cruzi* lineages from cultured stocks may underestimate the parasite diversity originally present in the host. Therefore, we could not discard the possibility that some newborns in our study could be infected by parasites from more than one *T. cruzi* lineage. Indeed, mixed infections with *T. cruzi* I and *T. cruzi* II lineages have been reported in Bolivia and Chile (Breniere et al., 1998; Torres et al., 2004).

Despite the identification of *T. cruzi* lineages from cultured stocks, our results agree with previous works indicating *T. cruzi* IId as the main lineage infecting newborns in Argentina and Bolivia (Virreira et al., 2006; Burgos et al., 2007). These authors found a high prevalence of *T. cruzi* IId (94% and 95%, respectively) in congenital infections by using PCR-based technology that avoids the possible selection of a particular parasite lineage.

In addition, these studies reveal that *T. cruzi* IId was the main lineage in adults of the same area and of mothers who gave birth to uninfected children (Virreira et al., 2006; Burgos et al., 2007). Therefore, the predominance of the IId lineage detected in our study cannot be attributed to a preferential association of this lineage with congenital infection, but rather as indicative of the general distribution of *T. cruzi* lineages in the province of Salta, Argentina.

In this regard, our results agree with previous findings suggesting that, in the Southern Cone countries of Latin America, human infection by *T. cruzi* is due principally to the *T. cruzi* II lineage. In these countries, molecular and immunological markers corroborated that human chagasic infection is due principally to *T. cruzi* II (Breniere et al., 1998; Zingales et al., 1998; Di Noia et al., 2002; Buscaglia and Di Noia, 2003). This contrasts with the genetic characterization of human isolates in Mexico, Colombia and Venezuela where a high predominance of the *T. cruzi* I lineage has been reported (Bosseno et al., 2002; Montilla et al., 2002; Anez et al., 2004).

It is important to point out that *T. cruzi* typing studies are performed mainly from blood samples and consequently the presence of other lineages in infected organs cannot be ruled out. Recent data have shown the presence of different lineages of *T. cruzi* in blood and brain tissue from an Argentinean patient with chagasic reactivation due to AIDS. Interestingly, *T. cruzi* I was only found in the cerebrospinal fluid from this patient and not in peripheral blood, suggesting a tropism of this lineage for target organs (Burgos et al., 2008). These authors suggest that the *T. cruzi* I lineage circulates at higher frequencies in the Southern countries of America and that the low rate of detection of this lineage was due to the low bloodstream parasitic load as a consequence of higher tropism for target organs (Burgos et al., 2008).

Several host as well parasitic factors might be involved in vertical transmission of *T. cruzi*. Regarding the genetic variability of parasite strains, previous studies do not support a direct association between the *T. cruzi* lineage or minicircle signature with the occurrence of congenital infection (Virreira et al., 2006; Burgos et al., 2007). In cases of transmission, both studies reveal that whole maternal bloodstream populations were transmitted to newborns since sub-lineages of *T. cruzi* were found to be identical in mothers and their neonates.

These findings support the hypothesis that *T. cruzi* infection in human fetus/neonates, or adults, depends more on host resistance/susceptibility governed by its genetic background, as well as environmental and social factors (Hermann et al., 2002; Campbell et al., 2004; Cruz-Robles et al., 2004). This assumption is also reinforced by previous data showing that congenital transmission of *T. cruzi* is associated with high parasitic loads and peripheral deficient immunological responses in mothers (Hermann et al., 2004). In addition, data from our group suggest the association of the HLA-DRB1*08 haplotype from the major histocompatibility complex (MHC) with the congenital transmission of *T. cruzi* in the province of Salta (Sanchez Negrette, unpublished data). Taken together, these data indicate that several other variable features from both parasites and humans have an important influence on the clinical outcome of Chagas disease, in addition to the genetic diversity of *T. cruzi* clonal populations.

In conclusion, data reported in this study confirms for the first time the presence of the *T. cruzi IId* lineage among congenitally infected newborns in the province of Salta, Argentina. These data support the assumption that in the Southern Cone countries of Latin America, human infection by *T. cruzi* is due principally to *T. cruzi* II. Nevertheless, further studies are needed to ascertain the role of both parasite and human factors in the development of the clinical outcome of Chagas disease in addition to the genetic diversity of *T. cruzi* as defined by their lineages.

Acknowledgements

Financial support for this study was provided by CONICET (National Research Council) and Project 1238 of CIUNSa (Research Council of the National University of Salta). RMC is supported by CONICET and the Programme Alban, the European Union Programme of High Level Scholarships for Latin America, scholarship No. E05D057391AR. The technical assistance of Alejandro Uncos and Federico Ramos is gratefully acknowledged. We thank P. Agnew for help revising the manuscript's English.

References

Anez, N., Crisante, G., da Silva, F.M., Rojas, A., Carrasco, H., Umezawa, E.S., Stolf, A.M., Ramirez, J.L., Teixeira, M.M., 2004. Predominance of lineage I among *Trypanosoma cruzi* isolates from Venezuelan patients with different clinical profiles of acute Chagas' disease. Trop Med Int Health 9, 1319–1326.

Barnabe, C., Brisse, S., Tibayrenc, M., 2000. Population structure and genetic typing of *Trypanosoma cruzi*, the agent of Chagas disease: a multilocus enzyme electrophoresis approach. Parasitology 120 (Pt 5), 513–526.

Barnabe, C., Neubauer, K., Solari, A., Tibayrenc, M., 2001. *Trypanosoma cruzi*: presence of the two major phylogenetic lineages and of several lesser discrete typing units (DTUs) in Chile and Paraguay. Acta Trop 78, 127–137.

Ben Abderrazak, S., Guerrini, F., Mathieu-Daude, F., Truc, P., Neubauer, K., Lewicka, K., Barnabe, C., Tibayrenc, M., 1993. Isoenzyme electrophoresis for parasite characterization. Methods Mol Biol 21, 361–382.

Bosseno, M.F., Barnabe, C., Magallon Gastelum, E., Lozano Kasten, F., Ramsey, J., Espinoza, B., Breniere, S.F., 2002. Predominance of *Trypanosoma cruzi* lineage I in Mexico. J Clin Microbiol 40, 627–632.

Breniere, S.F., Bosseno, M.F., Telleria, J., Bastrenta, B., Yacsik, N., Noireau, F., Alcazar, J.L., Barnabe, C., Wincker, P., Tibayrenc, M., 1998. Different behavior of two *Trypanosoma cruzi* major clones: transmission and circulation in young Bolivian patients. Exp Parasitol 89, 285–295.

Brisse, S., Barnabe, C., Tibayrenc, M., 2000. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. Int. J. Parasitol. 30, 35–44.

Brutus, L., Schneider, D., Postigo, J., Romero, M., Santalla, J., Chippaux, J.P., 2008. Congenital Chagas disease: diagnostic and clinical aspects in an area without vectorial transmission, Bermejo, Bolivia. Acta Trop. 106, 195–199.

Burgos, J.M., Altcheh, J., Bisio, M., Duffy, T., Valadares, H.M., Seidenstein, M.E., Piccinali, R., Freitas, J.M., Levin, M.J., Macchi, L., Macedo, A.M., Freilij, H., Schijman, A.G., 2007. Direct molecular profiling of minicircle signatures and lineages of *Trypanosoma cruzi* bloodstream populations causing congenital Chagas disease. Int. J. Parasitol. 37, 1319–1327.

Burgos, J.M., Begher, S., Silva, H.M., Bisio, M., Duffy, T., Levin, M.J., Macedo, A.M., Schijman, A.G., 2008. Molecular identification of *Trypanosoma cruzi* I tropism for central nervous system in Chagas reactivation due to AIDS. Am. J. Trop. Med. Hyg. 78, 294–297.

Buscaglia, C.A., Di Noia, J.M., 2003. *Trypanosoma cruzi* clonal diversity and the epidemiology of Chagas' disease. Microbes Infect. 5, 419–427.

Campbell, D.A., Westenberger, S.J., Sturm, N.R., 2004. The determinants of Chagas disease: connecting parasite and host genetics. Curr. Mol. Med. 4, 549–562.

Carlier, Y., 2005. Factors and mechanisms involved in the transmission and development of congenital infection with *Trypanosoma cruzi*. Rev. Soc. Bras. Med. Trop. 38 (Suppl. 2), 105–107.

Cruz-Robles, D., Reyes, P.A., Monteon-Padilla, V.M., Ortiz-Muniz, A.R., Vargas-Alarcon, G., 2004. MHC class I and class II genes in Mexican patients with Chagas disease. Hum. Immunol. 65, 60–65.

Di Noia, J.M., Buscaglia, C.A., De Marchi, C.R., Almeida, I.C., Frasch, A.C., 2002. A *Trypanosoma cruzi* small surface molecule provides the first immunological evidence that Chagas' disease is due to a single parasite lineage. J. Exp. Med. 195, 401–413.

Gurtler, R.E., Segura, E.L., Cohen, J.E., 2003. Congenital transmission of *Trypanosoma cruzi* infection in Argentina. Emerg. Infect. Dis. 9, 29–32.

Hermann, E., Truyens, C., Alonso-Vega, C., Even, J., Rodriguez, P., Berthe, A., Gonzalez-Merino, E., Torrico, F., Carlier, Y., 2002. Human fetuses are able to mount an adult like CD8 T-cell response. Blood 100, 2153–2158.

Hermann, E., Truyens, C., Alonso-Vega, C., Rodriguez, P., Berthe, A., Torrico, F., Carlier, Y., 2004. Congenital transmission of *Trypanosoma cruzi* is associated with maternal enhanced parasitemia and decreased production of interferon-gamma in response to parasite antigens. J. Infect. Dis. 189, 1274–1281.

Macedo, A.M., Pena, S.D., 1998. Genetic variability of *Trypanosoma cruzi*: Implications for the pathogenesis of Chagas disease. Parasitol. Today 14, 119–124.

Miles, M.A., Souza, A., Povoa, M., Shaw, J.J., Lainson, R., Toye, P.J., 1978. Isozymic heterogeneity of *Trypanosoma cruzi* in the first autochthonous patients with Chagas' disease in Amazonian Brazil. Nature 272, 819–821.

Montilla, M.M., Guhl, F., Jaramillo, C., Nicholls, S., Barnabe, C., Bosseno, M.F., Breniere, S.F., 2002. Isoenzyme clustering of Trypanosomatidae Colombian populations. Am. J. Trop. Med. Hyg. 66, 394–400.

Mora, M.C., Sanchez Negrette, O., Marco, D., Barrio, A., Ciaccio, M., Segura, M.A., Basombrio, M.A., 2005. Early diagnosis of congenital *Trypanosoma cruzi* infection using PCR, hemoculture, and capillary concentration, as compared with delayed serology. J. Parasitol. 91, 1468–1473.

Sanchez Negrette, O., Mora, M.C., Basombrio, M.A., 2005. High prevalence of congenital *Trypanosoma cruzi* infection and family clustering in Salta, Argentina. Pediatrics 115, e668–e672.

Telleria, J., Lafay, B., Virreira, M., Barnabe, C., Tibayrenc, M., Svoboda, M., 2006. *Trypanosoma cruzi*: sequence analysis of the variable region of kinetoplast minicircles. Exp. Parasitol. 114, 279–288.

Tibayrenc, M., Breniere, F., Barnabe, C., Lemesre, J.L., Echalar, L., Desjeux, P., 1985. Isozymic variability of *Trypanosoma cruzi*: biological and epidemiological significance. Ann. Soc. Belg. Med. Trop. 65 (Suppl. 1), 59–61.

Tibayrenc, M., Ward, P., Moya, A., Ayala, F.J., 1986. Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure. Proc. Natl. Acad. Sci. U.S.A. 83, 115–119.

Torres, J.P., Ortiz, S., Munoz, S., Solari, A., 2004. *Trypanosoma cruzi* isolates from Chile are heterogeneous and composed of mixed populations when characterized by schizodeme and Southern analyses. Parasitology 128, 161–168.

Torrico, F., Alonso-Vega, C., Suarez, E., Rodriguez, P., Torrico, M.C., Dramaix, M., Truyens, C., Carlier, Y., 2004. Maternal *Trypanosoma cruzi* infection, pregnancy outcome, morbidity, and mortality of congenitally infected and non-infected newborns in Bolivia. Am. J. Trop. Med. Hyg. 70, 201–209.

Virreira, M., Alonso-Vega, C., Solano, M., Jijena, J., Brutus, L., Bustamante, Z., Truyens, C., Schneider, D., Torrico, F., Carlier, Y., Svoboda, M., 2006. Congenital Chagas disease in Bolivia is not associated with DNA polymorphism of *Trypanosoma cruzi*. Am. J. Trop. Med. Hyg. 75, 871–879.

Zingales, B., Souto, R.P., Mangia, R.H., Lisboa, C.V., Campbell, D.A., Coura, J.R., Jansen, A., Fernandes, O., 1998. Molecular epidemiology of American trypanosomiasis in Brazil based on dimorphisms of rRNA and mini-exon gene sequences. Int. J. Parasitol. 28, 105–112.

I. BIBLIOGRAPHY

Alexander, B., Maroli, M., 2003. Control of phlebotomine sandflies. Med Vet Entomol 17, 1-18.

Andrews, N.W., 2002. Lysosomes and the plasma membrane: trypanosomes reveal a secret relationship. J Cell Biol 158, 389-394.

Atwood, J.A., 3rd, Weatherly, D.B., Minning, T.A., Bundy, B., Cavola, C., Opperdoes, F.R., Orlando, R., Tarleton, R.L., 2005. The Trypanosoma cruzi proteome. Science 309, 473-476.

Baida, R.C., Santos, M.R., Carmo, M.S., Yoshida, N., Ferreira, D., Ferreira, A.T., El Sayed, N.M., Andersson, B., da Silveira, J.F., 2006. Molecular characterization of serine-, alanine-, and proline-rich proteins of Trypanosoma cruzi and their possible role in host cell infection. Infect Immun 74, 1537-1546.

Barrett, M.P., Boykin, D.W., Brun, R., Tidwell, R.R., 2007. Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. Br J Pharmacol 152, 1155-1171.

Barry, J.D., Marcello, L., Morrison, L.J., Read, A.F., Lythgoe, K., Jones, N., Carrington, M., Blandin, G., Bohme, U., Caler, E., Hertz-Fowler, C., Renauld, H., El-Sayed, N., Berriman, M., 2005. What the genome sequence is revealing about trypanosome antigenic variation. Biochem Soc Trans 33, 986-989.

Bates, P.A., 2007. Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies. Int J Parasitol 37, 1097-1106.

Beattie, L., Evans, K.J., Kaye, P.M., Smith, D.F., 2008. Transgenic Leishmania and the immune response to infection. Parasite Immunol 30, 255-266.

Bendtsen, J.D., Jensen, L.J., Blom, N., Von Heijne, G., Brunak, S., 2004a. Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel 17, 349-356.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004b. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340, 783-795.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Bohme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E., El-Sayed, N.M., 2005. The genome of the African trypanosome Trypanosoma brucei. Science 309, 416-422.

Berrizbeitia, M., Ndao, M., Bubis, J., Gottschalk, M., Ache, A., Lacouture, S., Medina, M., Ward, B.J., 2006. Purified excreted-secreted antigens from Trypanosoma cruzi trypomastigotes as tools for diagnosis of Chagas' disease. J Clin Microbiol 44, 291-296.

Beverley, S.M., 2003. Protozomics: trypanosomatid parasite genetics comes of age. Nat Rev Genet 4, 11-19.

Bosch, M., Trombetta, S., Engstrom, U., Parodi, A.J., 1988. Characterization of dolichol diphosphate oligosaccharide: protein oligosaccharyltransferase and glycoprotein-processing glucosidases occurring in trypanosomatid protozoa. J Biol Chem 263, 17360-17365.

Bringaud, F., Muller, M., Cerqueira, G.C., Smith, M., Rochette, A., El-Sayed, N.M., Papadopoulou, B., Ghedin, E., 2007. Members of a large retroposon family are determinants of post-transcriptional gene expression in Leishmania. PLoS Pathog 3, 1291-1307.

Buschiazzo, A., Amaya, M.F., Cremona, M.L., Frasch, A.C., Alzari, P.M., 2002. The crystal structure and mode of action of trans-sialidase, a key enzyme in Trypanosoma cruzi pathogenesis. Mol Cell 10, 757-768.

Campetella, O., Martinez, J., Cazzulo, J.J., 1990. A major cysteine proteinase is developmentally regulated in Trypanosoma cruzi. FEMS Microbiol Lett 55, 145-149.

Chappuis, F., Rijal, S., Soto, A., Menten, J., Boelaert, M., 2006. A meta-analysis of the diagnostic performance of the direct agglutination test and rK39 dipstick for visceral leishmaniasis. Bmj 333, 723.

Chiaramonte, M.G., Zwirner, N.W., Caropresi, S.L., Taranto, N.J., Malchiodi, E.L., 1996. Trypanosoma cruzi and Leishmania spp. human mixed infection. Am J Trop Med Hyg 54, 271-273.

Clayton, C., Hausler, T., Blattner, J., 1995. Protein trafficking in kinetoplastid protozoa. Microbiol Rev 59, 325-344.

Clayton, C., Shapira, M., 2007. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Mol Biochem Parasitol 156, 93-101.

Colli, W., 1993. Trans-sialidase: a unique enzyme activity discovered in the protozoan Trypanosoma cruzi. Faseb J 7, 1257-1264.

Cooley, G., Etheridge, R.D., Boehlke, C., Bundy, B., Weatherly, D.B., Minning, T., Haney, M., Postan, M., Laucella, S., Tarleton, R.L., 2008. High Throughput Selection of Effective Serodiagnostics for Trypanosoma cruzi infection. PLoS Negl Trop Dis 2, e316.

Coura, J.R., 2007. Chagas disease: what is known and what is needed--a background article. Mem Inst Oswaldo Cruz 102 Suppl 1, 113-122.

D'Alessio, C., Fernandez, F., Trombetta, E.S., Parodi, A.J., 1999. Genetic evidence for the heterodimeric structure of glucosidase II. The effect of disrupting the subunit-encoding genes on glycoprotein folding. J Biol Chem 274, 25899-25905.

da Silva, C.V., Kawashita, S.Y., Probst, C.M., Dallagiovanna, B., Cruz, M.C., da Silva, E.A., Souto-Padron, T.C., Krieger, M.A., Goldenberg, S., Briones, M.R., Andrews, N.W., Mortara, R.A., 2009. Characterization of a 21kDa protein from Trypanosoma cruzi associated with mammalian cell invasion. Microbes Infect.

da Silveira, J.F., Umezawa, E.S., Luquetti, A.O., 2001. Chagas disease: recombinant Trypanosoma cruzi antigens for serological diagnosis. Trends Parasitol 17, 286-291.

Denny, P.W., Gokool, S., Russell, D.G., Field, M.C., Smith, D.F., 2000. Acylation-dependent protein export in Leishmania. J Biol Chem 275, 11017-11025.

Douzery, E.J., Snell, E.A., Bapteste, E., Delsuc, F., Philippe, H., 2004. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc Natl Acad Sci U S A 101, 15386-15391.

El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L., Blandin, G., Westenberger, S.J., Caler, E., Cerqueira, G.C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D.A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J.F., de Jong, P., Edwards, K., Englund, P.T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M.J., Lorenzi, H., Louie, T., Machado, C.R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J.C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J.L., Rinta, J., Robertson, L., Salzberg, S.L., Sanchez, D.O., Seyler, A., Sharma, R., Shetty, J., Simpson, A.J., Sisk, E., Tammi, M.T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P.N., Wickstead, B., Wortman, J., White, O., Fraser, C.M., Stuart, K.D., Andersson, B., 2005a. The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. Science 309, 409-415.

El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E.A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D.C., Haas, B.J., Tran, A.N., Wortman, J.R., Alsmark, U.C., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J.M., Cerqueira, G.C., Creasy, T., Delcher, A.L., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A.C., Kummerfeld, S.K., Pereira-Leal, J.B., Nilsson, D., Peterson, J., Salzberg, S.L., Shallom, J., Silva, J.C., Sundaram, J., Westenberger, S., White, O., Melville, S.E., Donelson, J.E., Andersson, B., Stuart, K.D., Hall, N., 2005b. Comparative genomics of trypanosomatid parasitic protozoa. Science 309, 404-409.

Ellis, M., Sharma, D.K., Hilley, J.D., Coombs, G.H., Mottram, J.C., 2002. Processing and trafficking of Leishmania mexicana GP63. Analysis using GP18 mutants deficient in glycosylphosphatidylinositol protein anchoring. J Biol Chem 277, 27968-27974.

Emanuelsson, O., Brunak, S., von Heijne, G., Nielsen, H., 2007. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2, 953-971.

Engstler, M., Boshart, M., 2004. Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei. Genes Dev 18, 2798-2811.

Ferguson, M.A., 1999. The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. J Cell Sci 112 (Pt 17), 2799-2809.

Field, M.C., Carrington, M., 2004. Intracellular membrane transport systems in Trypanosoma brucei. Traffic 5, 905-913.

Field, M.C., Natesan, S.K., Gabernet-Castello, C., Koumandou, V.L., 2007. Intracellular trafficking in the trypanosomatids. Traffic 8, 629-639.

Flinn, H.M., Rangarajan, D., Smith, D.F., 1994. Expression of a hydrophilic surface protein in infective stages of Leishmania major. Mol Biochem Parasitol 65, 259-270.

Flower, D.R., North, A.C., Sansom, C.E., 2000. The lipocalin protein family: structural and sequence overview. Biochimica et Biophysica Acta 1482, 9-24.

Foth, B., Piani, A., Curtis, J.M., Ilg, T., McConville, M., Handman, E., 2002. Leishmania major proteophosphoglycans exist as membrane-bound and soluble forms and localise to the cell membrane, the flagellar pocket and the lysosome. Int J Parasitol 32, 1701-1708.

Frank, F.M., Fernandez, M.M., Taranto, N.J., Cajal, S.P., Margni, R.A., Castro, E., Thomaz-Soccol, V., Malchiodi, E.L., 2003. Characterization of human infection by Leishmania spp. in the Northwest of Argentina: immune response, double infection with Trypanosoma cruzi and species of Leishmania involved. Parasitology 126, 31-39.

Glover, L., Horn, D., 2006. Repression of polymerase I-mediated gene expression at Trypanosoma brucei telomeres. EMBO Rep 7, 93-99.

Grebaut, P., Chuchana, P., Brizard, J.P., Demettre, E., Seveno, M., Bossard, G., Jouin, P., Vincendeau, P., Bengaly, Z., Boulange, A., Cuny, G., Holzmuller, P., 2009. Identification of total and differentially expressed excreted-secreted proteins from Trypanosoma congolense strains exhibiting different virulence and pathogenicity. Int J Parasitol.

Grzyb, J., Latowski, D., Strzalka, K., 2006. Lipocalins - a family portrait. Journal of Plant Physiology 163, 895-915.

Haenni, S., Renggli, C.K., Fragoso, C.M., Oberle, M., Roditi, I., 2006. The procyclin-associated genes of Trypanosoma brucei are not essential for cyclical transmission by tsetse. Mol Biochem Parasitol 150, 144-156.

Haile, S., Papadopoulou, B., 2007. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. Curr Opin Microbiol 10, 569-577.

Herwaldt, B.L., 1999. Leishmaniasis. Lancet 354, 1191-1199.

Hu, Y.C., 2005. Baculovirus as a highly efficient expression vector in insect and mammalian cells. Acta Pharmacol Sin 26, 405-416.

Ilg, T., Stierhof, Y.D., Wiese, M., McConville, M.J., Overath, P., 1994. Characterization of phosphoglycan-containing secretory products of Leishmania. Parasitology 108 Suppl, S63-71.

Ilg, T., Handman, E., Stierhof, Y.D., 1999. Proteophosphoglycans from Leishmania promastigotes and amastigotes. Biochem Soc Trans 27, 518-525.

Inoue, N., Otsu, K., Ferraro, D.M., Donelson, J.E., 2002. Tetracycline-regulated RNA interference in Trypanosoma congolense. Mol Biochem Parasitol 120, 309-313.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., Myler, P.J., 2005. The genome of the kinetoplastid parasite, Leishmania major. Science 309, 436-442.

Jaffe, C.L., Dwyer, D.M., 2003. Extracellular release of the surface metalloprotease, gp63, from Leishmania and insect trypanosomatids. Parasitol Res 91, 229-237.

Jazin, E.E., Luquetti, A.O., Rassi, A., Frasch, A.C., 1991. Shift of excretory-secretory immunogens of Trypanosoma cruzi during human Chagas' disease. Infect Immun 59, 2189-2191.

Jones, D.C., Mehlert, A., Guther, M.L., Ferguson, M.A., 2005. Deletion of the glucosidase II gene in Trypanosoma brucei reveals novel N-glycosylation mechanisms in the biosynthesis of variant surface glycoprotein. J Biol Chem 280, 35929-35942.

Joshi, M.B., Rogers, M.E., Shakarian, A.M., Yamage, M., Al-Harthi, S.A., Bates, P.A., Dwyer, D.M., 2005. Molecular characterization, expression, and in vivo analysis of LmexCht1: the chitinase of the human pathogen, Leishmania mexicana. Journal of Biological Chemistry 280, 3847-3861.

Klee, E.W., Ellis, L.B., 2005. Evaluating eukaryotic secreted protein prediction. BMC Bioinformatics 6, 256.

Klee, E.W., Sosa, C.P., 2007. Computational classification of classically secreted proteins. Drug Discov Today 12, 234-240.

Landfear, S.M., Ignatushchenko, M., 2001. The flagellum and flagellar pocket of trypanosomatids. Mol Biochem Parasitol 115, 1-17.

Lederkremer, R.M., Bertello, L.E., 2001. Glycoinositolphospholipids, free and as anchors of proteins, in Trypanosoma cruzi. Curr Pharm Des 7, 1165-1179.

Lemesre, J.L., Holzmuller, P., Goncalves, R.B., Bourdoiseau, G., Hugnet, C., Cavaleyra, M., Papierok, G., 2007. Long-lasting protection against canine visceral leishmaniasis using the LiESAp-MDP vaccine in endemic areas of France: double-blind randomised efficacy field trial. Vaccine 25, 4223-4234.

Li, Y.M., Mitsuhashi, T., Wojciechowicz, D., Shimizu, N., Li, J., Stitt, A., He, C., Banerjee, D., Vlassara, H., 1996. Molecular identity and cellular distribution of advanced glycation endproduct receptors: relationship of p60 to OST-48 and p90 to 80K-H membrane proteins. Proceedings of the National Academy of Sciences U S A 93, 11047-11052.

Lipoldova, M., Demant, P., 2006. Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis. Nat Rev Genet 7, 294-305.

Lukes, J., Hashimi, H., Zikova, A., 2005. Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. Curr Genet 48, 277-299.

Lustig, Y., Sheiner, L., Vagima, Y., Goldshmidt, H., Das, A., Bellofatto, V., Michaeli, S., 2007. Spliced-leader RNA silencing: a novel stress-induced mechanism in Trypanosoma brucei. EMBO Rep 8, 408-413.

McConville, M.J., Mullin, K.A., Ilgoutz, S.C., Teasdale, R.D., 2002. Secretory pathway of trypanosomatid parasites. Microbiol Mol Biol Rev 66, 122-154; table of contents.

McConville, M.J., de Souza, D., Saunders, E., Likic, V.A., Naderer, T., 2007. Living in a phagolysosome; metabolism of Leishmania amastigotes. Trends Parasitol 23, 368-375.

McGwire, B.S., O'Connell, W.A., Chang, K.P., Engman, D.M., 2002. Extracellular release of the glycosylphosphatidylinositol (GPI)-linked Leishmania surface metalloprotease, gp63, is independent of GPI phospholipolysis: implications for parasite virulence. J Biol Chem 277, 8802-8809.

McMahon-Pratt, D., Alexander, J., 2004. Does the Leishmania major paradigm of pathogenesis and protection hold for New World cutaneous leishmaniases or the visceral disease? Immunol Rev 201, 206-224.

McNicoll, F., Muller, M., Cloutier, S., Boilard, N., Rochette, A., Dube, M., Papadopoulou, B., 2005. Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in Leishmania. J Biol Chem 280, 35238-35246.

Motyka, S.A., Englund, P.T., 2004. RNA interference for analysis of gene function in trypanosomatids. Curr Opin Microbiol 7, 362-368.

Nakazawa, M., Rosa, D.S., Pereira, V.R., Moura, M.O., Furtado, V.C., Souza, W.V., Barros, M.N., Abath, F.G., Gomes, Y.M., 2001. Excretory-secretory antigens of Trypanosoma cruzi are potentially useful for serodiagnosis of chronic Chagas' disease. Clin Diagn Lab Immunol 8, 1024-1027.

Ngo, H., Tschudi, C., Gull, K., Ullu, E., 1998. Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proc Natl Acad Sci U S A 95, 14687-14692.

Nickel, W., Seedorf, M., 2008. Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. Annu Rev Cell Dev Biol 24, 287-308.

Nilsson, T., Warren, G., 1994. Retention and retrieval in the endoplasmic reticulum and the Golgi apparatus. Curr Opin Cell Biol 6, 517-521.

Palatnik-de-Sousa, C.B., 2008. Vaccines for leishmaniasis in the fore coming 25 years. Vaccine 26, 1709-1724.

Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M.A., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabbinowitsch, E., Arrowsmith, C., White, B., Thurston, S., Bringaud, F., Baldauf, S.L., Faulconbridge, A., Jeffares, D., Depledge, D.P., Oyola, S.O., Hilley, J.D., Brito, L.O., Tosi, L.R., Barrell, B., Cruz, A.K., Mottram, J.C., Smith, D.F., Berriman, M., 2007. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nat Genet 39, 839-847.

Picollo, M.I., Vassena, C., Santo Orihuela, P., Barrios, S., Zaidemberg, M., Zerba, E., 2005. High resistance to pyrethroid insecticides associated with ineffective field treatments in Triatoma infestans (Hemiptera: Reduviidae) from Northern Argentina. J Med Entomol 42, 637-642.

Pirard, M., Iihoshi, N., Boelaert, M., Basanta, P., Lopez, F., Van der Stuyft, P., 2005. The validity of serologic tests for Trypanosoma cruzi and the effectiveness of transfusional screening strategies in a hyperendemic region. Transfusion 45, 554-561.

Prudovsky, I., Tarantini, F., Landriscina, M., Neivandt, D., Soldi, R., Kirov, A., Small, D., Kathir, K.M., Rajalingam, D., Kumar, T.K., 2008. Secretion without Golgi. J Cell Biochem 103, 1327-1343.

Reithinger, R., Dujardin, J.C., 2007. Molecular diagnosis of leishmaniasis: current status and future applications. J Clin Microbiol 45, 21-25.

Rogers, M.E., Bates, P.A., 2007. Leishmania manipulation of sand fly feeding behavior results in enhanced transmission. PLoS Pathog 3, e91.

Rogers, M.E., Hajmova, M., Joshi, M.B., Sadlova, J., Dwyer, D.M., Volf, P., Bates, P.A., 2008. Leishmania chitinase facilitates colonization of sand fly vectors and enhances transmission to mice. Cell Microbiol 10, 1363-1372.

Rothman, J.E., Orci, L., 1992. Molecular dissection of the secretory pathway. Nature 355, 409-415.

Sacks, D., Noben-Trauth, N., 2002. The immunology of susceptibility and resistance to Leishmania major in mice. Nat Rev Immunol 2, 845-858.

Salomone, O.A., Basquiera, A.L., Sembaj, A., Aguerri, A.M., Reyes, M.E., Omelianuk, M., Fernandez, R.A., Enders, J., Palma, A., Barral, J.M., Madoery, R.J., 2003. Trypanosoma cruzi in persons without serologic evidence of disease, Argentina. Emerg Infect Dis 9, 1558-1562.

Samant, M., Gupta, R., Kumari, S., Misra, P., Khare, P., Kushawaha, P.K., Sahasrabuddhe, A.A., Dube, A., 2009. Immunization with the DNA-encoding N-terminal domain of proteophosphoglycan of Leishmania donovani generates Th1-Type immunoprotective response against experimental visceral leishmaniasis. J Immunol 183, 470-479.

Santarem, N., Silvestre, R., Tavares, J., Silva, M., Cabral, S., Maciel, J., Cordeiro-da-Silva, A., 2007. Immune response regulation by leishmania secreted and nonsecreted antigens. J Biomed Biotechnol 2007, 85154.

Santos, D.O., Coutinho, C.E., Madeira, M.F., Bottino, C.G., Vieira, R.T., Nascimento, S.B., Bernardino, A., Bourguignon, S.C., Corte-Real, S., Pinho, R.T., Rodrigues, C.R., Castro, H.C., 2008. Leishmaniasis treatment--a challenge that remains: a review. Parasitol Res 103, 1-10.

Schatz, G., Dobberstein, B., 1996. Common principles of protein translocation across membranes. Science 271, 1519-1526.

Schneider, A., 2001. Unique aspects of mitochondrial biogenesis in trypanosomatids. Int J Parasitol 31, 1403-1415.

Shen, Y.Q., Burger, G., 2007. 'Unite and conquer': enhanced prediction of protein subcellular localization by integrating multiple specialized tools. BMC Bioinformatics 8, 420.

Sibley, L.D., Andrews, N.W., 2000. Cell invasion by un-palatable parasites. Traffic 1, 100-106.

Silverman, J.M., Chan, S.K., Robinson, D.P., Dwyer, D.M., Nandan, D., Foster, L.J., Reiner, N.E., 2008. Proteomic analysis of the secretome of Leishmania donovani. Genome Biology 9, R35.

Simarro, P.P., Jannin, J., Cattand, P., 2008. Eliminating human African trypanosomiasis: where do we stand and what comes next? PLoS Med 5, e55.

Sprenger, J., Fink, J.L., Teasdale, R.D., 2006. Evaluation and comparison of mammalian subcellular localization prediction methods. BMC Bioinformatics 7 Suppl 5, S3.

Stegmayer, C., Kehlenbach, A., Tournaviti, S., Wegehingel, S., Zehe, C., Denny, P., Smith, D.F., Schwappach, B., Nickel, W., 2005. Direct transport across the plasma membrane of mammalian cells of Leishmania HASPB as revealed by a CHO export mutant. J Cell Sci 118, 517-527.

Stierhof, Y.D., Wiese, M., Ilg, T., Overath, P., Haner, M., Aebi, U., 1998. Structure of a filamentous phosphoglycoprotein polymer: the secreted acid phosphatase of Leishmania mexicana. J Mol Biol 282, 137-148.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gurtler, R.E., McKerrow, J., Reed, S., Tarleton, R., 2008. Kinetoplastids: related protozoan pathogens, different diseases. J Clin Invest 118, 1301-1310.

Subramaniam, C., Veazey, P., Redmond, S., Hayes-Sinclair, J., Chambers, E., Carrington, M., Gull, K., Matthews, K., Horn, D., Field, M.C., 2006. Chromosome-wide analysis of gene function by RNA interference in the african trypanosome. Eukaryot Cell 5, 1539-1549.

Tarleton, R.L., 2003. Chagas disease: a role for autoimmunity? Trends Parasitol 19, 447-451.

Tarleton, R.L., 2007. Immune system recognition of Trypanosoma cruzi. Curr Opin Immunol 19, 430-434.

Tarleton, R.L., Reithinger, R., Urbina, J.A., Kitron, U., Gurtler, R.E., 2007. The challenges of Chagas Disease-- grim outlook or glimmer of hope. PLoS Med 4, e332.

Taylor, J.E., Rudenko, G., 2006. Switching trypanosome coats: what's in the wardrobe? Trends Genet 22, 614-620.

Taylor, M.C., Kelly, J.M., 2006. pTcINDEX: a stable tetracycline-regulated expression vector for Trypanosoma cruzi. BMC Biotechnol 6, 32.

Tobin, J.F., Wirth, D.F., 1993. Mutational analysis of a signal sequence required for protein secretion in Leishmania major. Mol Biochem Parasitol 62, 243-249.

Tonui, W.K., Mejia, J.S., Hochberg, L., Mbow, M.L., Ryan, J.R., Chan, A.S., Martin, S.K., Titus, R.G., 2004. Immunization with Leishmania major exogenous antigens protects susceptible BALB/c mice against challenge infection with L. major. Infect Immun 72, 5654-5661.

Tyler, K.M., Engman, D.M., 2001. The life cycle of Trypanosoma cruzi revisited. Int J Parasitol 31, 472-481.

Umezawa, E.S., Bastos, S.F., Coura, J.R., Levin, M.J., Gonzalez, A., Rangel-Aldao, R., Zingales, B., Luquetti, A.O., da Silveira, J.F., 2003. An improved serodiagnostic test for Chagas' disease employing a mixture of Trypanosoma cruzi recombinant antigens. Transfusion 43, 91-97.

Umezawa, E.S., Luquetti, A.O., Levitus, G., Ponce, C., Ponce, E., Henriquez, D., Revollo, S., Espinoza, B., Sousa, O., Khan, B., da Silveira, J.F., 2004. Serodiagnosis of chronic and acute Chagas' disease with Trypanosoma cruzi recombinant proteins: results of a collaborative study in six Latin American countries. J Clin Microbiol 42, 449-452.

van Vliet, C., Thomas, E.C., Merino-Trigo, A., Teasdale, R.D., Gleeson, P.A., 2003. Intracellular sorting and transport of proteins. Prog Biophys Mol Biol 83, 1-45.

Vassella, E., Den Abbeele, J.V., Butikofer, P., Renggli, C.K., Furger, A., Brun, R., Roditi, I., 2000. A major surface glycoprotein of trypanosoma brucei is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia. Genes Dev 14, 615-626.

Verani, J.R., Seitz, A., Gilman, R.H., LaFuente, C., Galdos-Cardenas, G., Kawai, V., de LaFuente, E., Ferrufino, L., Bowman, N.M., Pinedo-Cancino, V., Levy, M.Z., Steurer, F., Todd, C.W., Kirchhoff, L.V., Cabrera, L., Verastegui, M., Bern, C., 2009. Geographic variation in the sensitivity of recombinant antigen-based rapid tests for chronic Trypanosoma cruzi infection. Am J Trop Med Hyg 80, 410-415.

Viotti, R., Vigliano, C., Lococo, B., Bertocchi, G., Petti, M., Alvarez, M.G., Postan, M., Armenti, A., 2006. Long-term cardiac outcomes of treating chronic Chagas disease with benznidazole versus no treatment: a nonrandomized trial. Ann Intern Med 144, 724-734.

Wirtz, E., Leal, S., Ochatt, C., Cross, G.A., 1999. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol Biochem Parasitol 99, 89-101.

Xu, D., Brandan, C.P., Basombrio, M.A., Tarleton, R.L., 2009. Evaluation of high efficiency gene knockout strategies for Trypanosoma cruzi. BMC Microbiol 9, 90.

Yan, S., Martinez-Calvillo, S., Schnaufer, A., Sunkin, S., Myler, P.J., Stuart, K., 2002. A low-background inducible promoter system in Leishmania donovani. Mol Biochem Parasitol 119, 217-223.

Yokoyama-Yasunaka, J.K., Pral, E.M., Oliveira Junior, O.C., Alfieri, S.C., Stolf, A.M., 1994. Trypanosoma cruzi: identification of proteinases in shed components of trypomastigote forms. Acta Trop 57, 307-315.

Zhang, W.W., Peacock, C.S., Matlashewski, G., 2008. A genomic-based approach combining in vivo selection in mice to identify a novel virulence gene in leishmania. PLoS Negl Trop Dis 2, e248.

J. ANNEXE A

List of the 45 potentially secreted proteins identified by the bioinformatic analysis presented in Chapter 1.

T. cruzi GeneDB Gene ID N°.	Orthologous *L. major* and *T. brucei* Gene ID N°		P value	
			SPP[a]	CSP[b]
Tc00.1047053506417.30	LmjF22.0225	Tb927.8.2180	0.937	0.917
Tc00.1047053506155.99	LmjF36.5220	Tb11.01.2470	0.984	0.962
Tc00.1047053506467.29	LmjF26.2000	Tb09.160.1070	0.811	0.771
Tc00.1047053511901.30	LmjF24.2160	Tb927.8.6080	0.989	0.898
Tc00.1047053511871.30	LmjF25.1010	Tb927.3.950	0.979	0.958
Tc00.1047053505789.10	LmjF19.0540	Tb927.8.6700	1.000	0.768
	LmjF19.0570	Tb11.39.0005		
Tc00.1047053509669.70	LmjF29.1600	Tb927.3.4190	0.999	0.980
Tc00.1047053507765.20	LmjF11.0720	Tb11.02.4400	0.993	0.986
Tc00.1047053510101.470	LmjF11.0720	Tb11.02.4400	0.931	0.919
Tc00.1047053510443.30	LmjF30.3150	Tb927.6.4500	0.903	0.838
Tc00.1047053509799.50	LmjF36.5570	Tb10.6k15.1130	0.981	0.931
Tc00.1047053509835.30	LmjF19.0540	Tb927.8.6700	0.866	0.803
	LmjF19.0570	Tb11.39.0005		
Tc00.1047053509999.10	LmjF29.1200	Tb927.3.3820	1.000	0.952
Tc00.1047053511517.120	LmjF30.2670	Tb927.6.3940	0.990	0.709
Tc00.1047053506129.89	LmjF31.0900	Tb11.02.3760	0.998	0.994
Tc00.1047053506405.140	LmjF34.3660	Tb927.4.1090	1.000	0.740
Tc00.1047053506627.20	LmjF04.0290	Tb09.160.5260	0.980	0.879
Tc00.1047053506679.180	LmjF36.1020	Tb10.70.1640	0.999	0.929
Tc00.1047053506733.30	LmjF13.0970	Tb11.02.1350	0.995	0.949
Tc00.1047053506789.160	LmjF05.1070	Tb927.7.6920	0.838	0.800
Tc00.1047053506855.60	LmjF21.1360	Tb10.70.7500	0.999	0.802
Tc00.1047053506855.100	LmjF21.1320	Tb10.70.7460	0.994	0.968
Tc00.1047053506789.160	LmjF05.1070	Tb927.7.6920	0.838	0.800

T. cruzi GeneDB Gene ID N°.	Orthologous *L. major* and *T. brucei* Gene ID N°		P value	
			SPP[a]	CSP[b]
Tc00.1047053507163.80	LmjF36.3657	Tb11.01.1590	0.936	0.904
Tc00.1047053509751.20	LmjF36.3657	Tb11.01.1590	0.983	0.972
Tc00.1047053507251.10	LmjF35.0640	Tb927.4.1110	1.000	0.964
Tc00.1047053510101.90	LmjF11.1030	Tb927.5.2280	0.996	0.960
Tc00.1047053507993.200	LmjF18.0140	Tb10.05.0100	0.908	0.731
Tc00.1047053503791.30	LmjF36.0480	Tb10.70.2320	0.971	0.791
Tc00.1047053508307.90	LmjF36.4760	Tb10.6k15.0160	0.994	0.857
Tc00.1047053506297.220	LmjF11.1030	Tb11.02.4120	0.996	0.750
Tc00.1047053508547.140	LmjF07.0330	Tb927.8.1530	0.801	0.703
Tc00.1047053509641.79	LmjF33.1140	Tb10.26.0900	0.998	0.943
Tc00.1047053509965.170	LmjF30.1590	Tb927.6.2560	0.878	0.848
Tc00.1047053510065.10	LmjF19.0540	Tb927.8.6700	1.000	0.768
	LmjF19.0570	Tb11.39.0005		
Tc00.1047053510119.60	LmjF21.0280	Tb10.70.5720	0.878	0.872
Tc00.1047053510885.110	LmjF12.1160	Tb927.1.4370	0.925	0.912
Tc00.1047053511303.70	LmjF17.1270	Tb927.5.2550	0.996	0.848
Tc00.1047053511315.20	LmjF17.0910	Tb927.5.2280	0.994	0.878
Tc00.1047053511527.60	LmjF32.3880	Tb11.01.7980	0.960	0.878
Tc00.1047053511755.50	LmjF17.0910	Tb927.5.2280	0.994	0.859
Tc00.1047053511857.70	LmjF13.0970	Tb11.02.1350	0.997	0.945
Tc00.1047053504103.3	LmjF36.4640	Tb10.6k15.0280	0.992	0.874
Tc00.1047053504137.100	LmjF36.5570	Tb10.6k15.1130	0.983	0.982

SPP[a] Signal peptide probability predicted by SignalP 3.0.

CSP[b] Maximal cleavage site probability predicted by SignalP 3.0

RESUME en français

Approche bioinformatique et fonctionnelle pour la caractérisation de protéines sécrétées chez les Trypanosomatidés: implication dans la biologie et la virulence du parasite

Les Trypanosomatidés *Leishmania spp.*, *Trypanosoma cruzi* et *Trypanosoma brucei* sont des protozoaires parasites responsables de pathologies humaines, dont le cycle de vie alterne entre un insecte vecteur et un hôte vertébré. Ces pathogènes ont développé différentes stratégies pour assurer leur survie chez leurs hôtes. Des facteurs de sécrétion permettent aux parasites de modifier leur environnement, de moduler la réponse immune de l'hôte, et d'envahir les cellules hôtes. Les protéines sécrétées représentent donc des cibles potentielles pour la conception de nouvelles thérapies et/ou de nouveaux vaccins. Nous avons développé une approche basée sur des analyses bioinformatiques afin d'identifier des protéines conservées chez les Trypanosomatidés, impliquées dans la voie de sécrétion classique (dépendante du réticulum endoplasmique et de l'appareil de Golgi). Cette méthode nous a permis d'identifier trois nouvelles protéines conservées, sécrétées dans le milieu extracellulaire. L'étude des propriétés biologiques de ces protéines suggère que l'une d'entre elles est impliquée dans un processus favorisant la survie du parasite *Leishmania* et sa réplication à l'intérieur de la cellule hôte. Les études portant sur l'immunogénicité des protéines sécrétées par *T. cruzi* ont montré que l'une des protéines est très antigénique chez l'homme. L'ensemble de ces résultats démontre l'utilité des analyses bioinformatiques couplées à des tests fonctionnels pour l'identification de nouvelles protéines sécrétées, représentant des facteurs de virulence ou des facteurs antigéniques potentiels des Trypanosomatidés. Ces résultats sont discutés à la lumière de récents travaux apportant des éléments de réponse sur les mécanismes de sécrétion à l'origine de l'"exoproteome" de *Leishmania*. Une meilleure connaissance de la complexité du sécrétome et de l'exoprotéome des Trypanosomatidés constitue un élément majeur vers la compréhension des interactions hôte-parasite chez ces pathogènes.

TITRE en anglais

Bioinformatic and functional approach for the characterization of secreted proteins in Trypanosomatids: implications for the biology and virulence of the parasite

RESUME en anglais

The trypanosomatid parasites causing human disease, *Leishmania spp.*, *Trypanosoma cruzi* and *Trypanosoma brucei* are protozoa that complete their life cycle in an insect vector and a vertebrate host. These pathogens have developed various strategies to modify their environment, influence host immune responses, or invade target cells. Materials secreted by these parasites are involved in such processes and may represent targets for vaccines and/or rational drug design. In this thesis, we designed an experimental approach based on bioinformatic analyses to identify conserved trypanosomatid proteins involved in the endoplasmic reticulum/Golgi-dependent secretory pathway. This method allowed us to identify three new trypanosomatid proteins in the extracellular space, demonstrating the utility of this approach for the identification of secreted proteins. Studies of the biological properties of these proteins suggest that one of them is involved in a process increasing survival and replication of *Leishmania* parasites inside its target cell. Immunogenicity studies of *T. cruzi* secreted proteins showed that one of these proteins is antigenic, suggesting its suitability for the development of a new diagnostic tool for Chagas disease. Altogether, these results indicate the utility of bioinformatic analyses combined with functional tests for the identification of novel secreted proteins, representing potential virulence factors or antigens in trypanosomatids. The results are also discussed in relation to emerging data shedding light on the secretion mechanisms involved in the production of the *Leishmania* exoproteome. We highlight the importance of understanding the complexity of the trypanosomatid secretome and exoproteome in order to elucidate the factors involved in its host-parasite interactions.

DISCIPLINE: Biologie moléculaire

MOTS-CLES: Trypanosomatidés, protéines sécrétées, voie de sécrétion classique, exoproteome, facteurs de virulence, immunogénicité

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE: Institut de Recherche pour le Développement (IRD), UR016 «Caractérisation et Contrôle des Populations de Vecteurs»,
911 Avenue Agropolis, BP 64501, 34394 Montpellier cedex 5