Molecular and structural organization of the *Densovirinae*.

P. TIJSSEN[1], M. G. Rossmann[2], A. Simpson[2], G. Fediere[1,3] and M. Bergoin[4].

[1)]Institute Armand-Frappier, Université du Québec, Laval, QC, Canada;
[2)]Department of Biology, Purdue University, West-Lafayette, IN, USA;
[3)]LEC, ORSTOM, Faculty of Agriculture, P.O. Box 26, Cairo, Egypt;
[4)]Unité de Virologie Moléculaire, Laboratoire de Pathologie Comparée, Université Montpellier II, France.

Densoviruses share many physical properties with vertebrate parvoviruses, such as encapsidating ssDNA genomes of about 5 kb in a robust, icosahedral particle with a diameter of about 25 nm (1,2). Their capsids consist of 60 subunits, each of a single protein. Nonetheless, the vertebrate and invertebrate parvoviruses have very little sequence identity and their genome organization is different. Therefore, two subfamilies are recognized within the *Parvoviridae* family, the *Parvovirinae* (parvoviruses of vertebrates) and the *Densovirinae* (parvoviruses of invertebrates).

Densoviruses were among the first parvoviruses recognized, in the mid-sixties (3), but have not been studied as extensively as the vertebrate parvoviruses. It was assumed for a long time that all densoviruses would have similar properties (a homogeneous genus) (4), and only recently have they been subdivided into three genera based on differences in sequences, genome organization, and capsid composition.

The densoviruses of the best-studied genus, the *Densovirus*, contain a relatively large genome (5-6 kb), and encapsidate both positive- and negative-sense genomes, in equal ratios into separate virions. These densoviruses have an ambisense genome (see below). Densoviruses from *Galleria mellonella* and *Junonia coenia* belong to this genus.

The densovirus from *Bombyx mori* (type-1, eg. The Ina isolate) is so different from those of the *Densovirus* genus that it has been classified into a separate genus, the *Iteravirus*. Thus far, it is the only virus in this genus although a densovirus isolated from the cockroach *Periplaneta fuliginosa*, *Pf* DNV-2, may also belong to this genus. Both the positive- and negative-sense strands are separately encapsidated, but *Iteravirus* have a monosense organization (5). In contrast to other densoviruses, the *Iteravirus* genomes do not have Y-like terminal structures although 153 of the 225-nucleotide terminal repeats form imperfect palindromes (6).

A third genus, *Contravirus* or *Brevidensovirus*, contains densoviruses that have a small, 4 kb genome (hence "brevis," Latin for small, short) with a monosense organization and terminal Y-like palindromes. Primarily the negative strand is encapsidated, not unlike many vertebrate parvoviruses. At least two densoviruses, isolated from mosquitoes, belong to this group (7,8).

A fourth group of densoviruses, which have split genomes with a total size more than double of any of the others (see Bando *et al.*, this Symposium), should in our opinion not be considered true densoviruses despite their name. Although some of the physical properties resemble those of true densoviruses, they do not have terminal palindromes and do not replicate by the rolling hairpin mechanisms of other parvoviruses. Their genome organization and expression is also different from parvoviruses.

Many densoviruses have not been studied in sufficient detail to permit their classification. Moreover, molecular biological studies of parvoviruses from other invertebrates such as shrimp and crab are required to permit their definitive classification.

In this study, we present results on the molecular biology of several, as yet unclassified, densoviruses which permitted a better understanding of the diversity of densoviruses and also a re-classification of the densoviruses. Moreover, we present the first atomic structure of a densovirus (*Gm*DNV), obtained through X-ray crystallographic studies, which provides us with tools to study in more detail the structure-function relationships in the *Densovirus* genus and the impact of the structure on the densovirus evolution and tropism.

VIIth International Colloquium on Invertebrate Pathology and Microbial Control
IVth International Conference on *Bacillus thuringiensis*
Sapporo
August 23-28, 1998

GENOME ORGANIZATION OF DNVs.

The first densoviruses that were sequenced have been *Bm*DNV-1 (5,6) from the Iteravirus genus and *Jc*DNV (9) and *Gm*DNV (10) from the *Densovirus* genus. The sequencing of the *Aedes aegypti* and *A. albopictus* DNVs demonstrated the necessity of a third genus (*Brevidensovirus*) (7, 8). Recently (unpublished results), we have sequenced densoviruses from *Pseudoplusia includens* (*Pi*DNV) (11), *Diatraea saccharalis* (*Ds*DNV) (12), *Acheta domesticus* (*Ad*DNV) (13), *Culex pipiens* (*Cp*DNV) (14), *Mythimna loreyi* (*Ml*DNV) (15), whereas the sequencing of the genome of the densovirus from *Casphalia extranea* (*Ce*DNV) (16) is advanced and those from isolates from *Chilo agamemnon* and *Spodoptera littoralis* are underway.

According to their sequences and genome organization, the *Pi*DNV, *Ds*DNV, *Cp*DNV, *Ad*DNV, and *Ml*DNV all belong to the *Densovirus* genus. The genomic organization of the members of the *Densovirus* genus is displayed in Fig. 1 and is compared to that of a vertebrate parvovirus (porcine parvovirus). A striking feature is the ambisense organization of the *Densovirus*, where the nonstructural proteins (NSs) are coded on the 5'-half of one strand and the structural proteins on the 5'-half of the complementary strand. Although for vertebrate parvoviruses the various nonstructural and structural proteins are expressed through an alternative splicing mechanism, the VPs are expressed through a leaky scanning mechanism of the VP ORF (see Bergoin *et al.*, this Symposium). In the unique region of VP1, a 39-amino acid motif (LGP-motif) is highly conserved between PPV and *Gm*DNV, even though it is interrupted by an intron in the case of PPV. It is noteworthy that several vertebrate parvoviruses (AAV, ADV) as well as some densoviruses in the other genera lack this conserved motif. Conserved regions have also been conserved within the NS between *Gm*DNV and PPV (eg., helicase superfamily III motifs). The *Pi*DNV and *Ds*DNV have a significant sequence homology with *Gm*DNV and *Jc*DNV (all in the order of 70-80%), whereas *Ml*DNV is very closely related to *Gm*DNV (90-95% identity). The close relationship between *Gm*DNV and *Ml*DNV is interesting since their tropism differs greatly, *Gm*DNV being monospecific and infecting only *Galleria mellonella* whereas *Ml*DNV is polyspecific and infects a large number of (pest) Lepidoptera. Interestingly, *Cp*DNV and *Ad*DNV have a shorter genome than the other *Densovirus* members (5-5.5 kb vs 6 kb) and a low sequence homology (for some regions of the genome hardly above background) but their genomic organization is clearly that of the *Densovirus*.

*Ce*DNV, which like the densoviruses from *Sibine fusca* (*Sf*DNV and Periplaneta *fuliginosa* (*Pf*DNV), causes tumour-like cell proliferation in the (mid)gut, was found to have a genome sequence and organization which resembles that of BmDNV-1 (see Fediere and Tijssen, this meeting) and should also be classified into the *Iteravirus* genus. It was observed previously that *Ce*DNV replicates in Bm40 cells (17) and that protein composition of *Ce*DNV resembles that of *Bm*DNV-1 (18).

ATOMIC STRUCTURE OF GmDNV.

The capsid of *Gm*DNV consists of 4 structural proteins that are coded as a nested set of gene products with identical C-termini. The differences among these proteins are thus defined by different N-terminal extensions (see Bergoin *et al.*, this Symposium). The structure of the capsid is thus determined by the common part of the 60 protein subunits. After X-ray crystallography and Fourier back-transformation, the main chain of this protein was unambiguously traced by means of a 3.7 Å, icosahedrally-averaged, electron density map from residue 22 to 436 of VP4. Like most virus capsid proteins, the subunit contains a core β-barrel domain. The program DALI showed that the *Gm*DNV structure had the greatest similarity with that of the canine parvovirus (score of 9.7σ Above random hit). When the rotational symmetry axes of *Gm*DNV and canine parvovirus (CPV) are superimposed, the β-barrel of CPV must be rotated by 7.5 Å And translated outwards by 9.5 Å in order to superimpose it onto the β-barrel of *Gm*DNV. The β-strands in the *Gm*DNV subunit are connected by long loops (up to 100 residues) that are important for the intersubunit contacts as well as the surface of the virion. Both CPV and *Gm*DNV subunits contain loops 1, 2, and 3 but with little or no identity while only CPV has loop 4. In CPV, loop 3 forms the base and loop 4 the spike at the treefold vertices, whereas in *Gm*DNV loop 3 is much shorter and forms a β-annulus type structure around the threefold axis, similar to tomato bushy stunt virus and southern bean mosaic virus, and loop 4 is altogether missing (no spike). The sequence around this 10 Å is particularly variable among DNVs and suggests that this annulus is a flexible part of the structure and might form a portal for DNA or ligand exchange.

Both CPV and *Gm*DNV have a channel along the fivefold axis of rotational symmetry of the capsid. For CPV, this channel is about 45 Å long and lined with small amino acids whereas for *Gm*DNV it is only 20 Å long and lined with large hydrophobic amino acids. In both instances the channel has a diameter of about 9 Å. It has been suggested for vertebrate parvoviruses that the N-terminus might extend from the interior of the capsid through the channel and that the stretch of a conserved

glycine-rich sequence would be important for this purpose. It is difficult to envisage that a long and bulky polypeptide (374 amino acids for the extension of *Gm*DNV VP1) could be threaded through this channel after assembly of the capsid. In addition the N-terminal domain of the *Gm*DNV has been swapped with respect to CPV. For CPV, the amino end of βA folded back towards the fivefold axis whereas for *Gm*DNV it is essentially a linear extension and hyrogen-bonds with the neighboring two-fold related subunit.

About 45% of the solvent accessible surface area of the subunit of *Gm*DNV is buried (ie., 15, 20 and 10% at the five-, three-, and twofold interactions, respectively). A putative divalent cation binding site could be coordinated by residues Glu116, Asp156, Gln157 and Asp357.

The ssDNA is icosahedrally well ordered in CPV but not in *Gm*DNV. This may be result of, or require, the presence of polyamines within the capsids. Another distinct feature underneath the subunits is the extension of βA to the neighboring twofold-related subunit ("domain swapping") and thus increases, with respect to CPV, the number of interactions between the subunits.

The threefold spike present in CPV is absent in *Gm*DNV. The intertwining loops of the subunits of *Gm*DNV form an antiparallel β-annulus structure. The function of receptor recognition of this spike in mammalian parvoviruses is apparently carried out elsewhere in densoviruses.

EVOLUTION OF DNVs

Like for vertebrate parvoviruses, the NS proteins are more conserved than the VP proteins but different parts of the VP proteins show considerable variation in the level of conservation. Particularly, the sequences involved in the β-barrel are highly conserved. For VP1, the most divergent stretches are formed by residues 1-134 and 277-375. Particularly variable is the sequence of loop 3 defining the threefold β-annulus. Conservation of particular stretches with one or another densovirus suggests recombination while viruses infected the same host and acceptance of such chimera as advantageous for particular hosts.

The extensive sequence and structural differences outside the β-barrel motif indicate a lengthy, divergent evolutionary development of invertebrate versus vertebrate parvoviruses.

The study of closely-related densoviruses with different host ranges by means of gene domain swapping and site-directed mutagenesis of infectious clones is underway and should lead to a well-defined structure-function relationship and demonstrate how host range can be restricted or how insect viruses may evolve different biological properties. On the other hand, the host may evolve into a resistant phenotype upon a particular co-evolution.

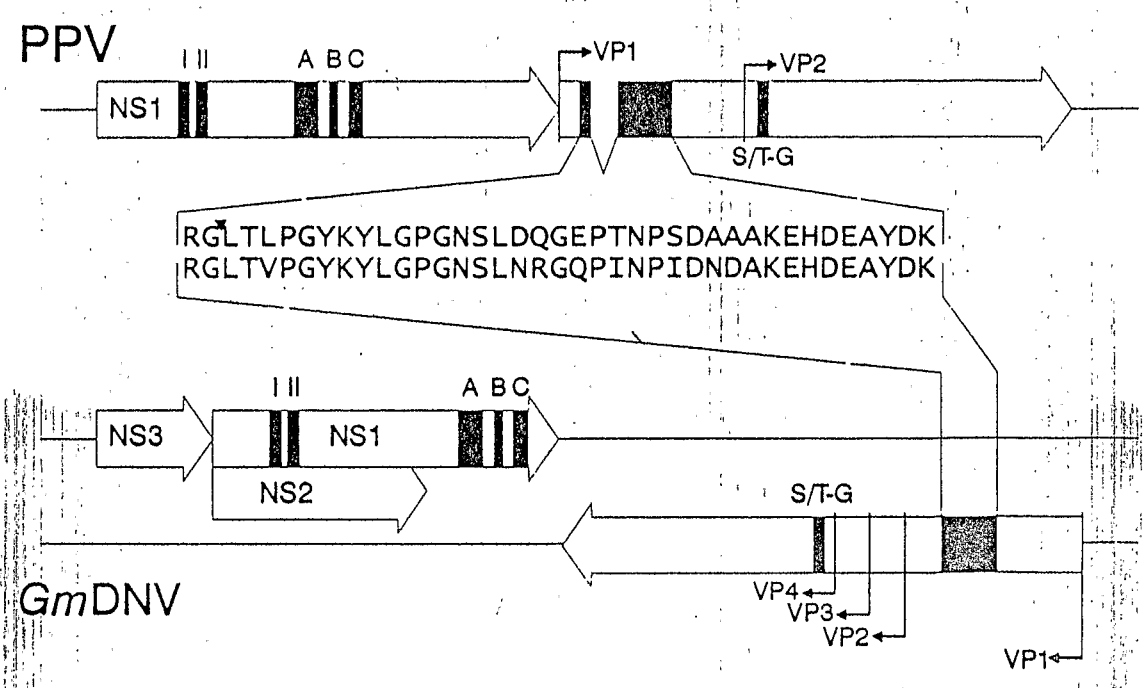

Fig.1 Comparison of genome organization of *Gm*DNV and porcine parvovirus (2, 20). Note that significant motifs may be conserved between vertebrate and invertebrate parvoviruses.Motifs I and II represent initiator (replicator) protein motifs and A, B, and C Helicase superfamily III motifs. The unique portion of both the vertebrate and invertebrate parvoviruses contain generally a conserved LPG region.

BIBLIOGRAPHY

1. Bergoin, M., and Tijssen, P. (in press), *In*: The Insect Viruses (Ball and Miller eds.) Plenum Press, New York
2. Tijssen, P. and Bergoin, M., 1995, *Seminars in Virology*, 6: 347.
3. Meynadier, G., Vago, C., Plantevin, G. and Atger, P., 1964, *Rev. Zool. Agric. Appl.*, 63: 207.
4. Bachmann, P. A., Hoggan, M. D., Melnick, J. L., Pereira, H. G. and Vago, C., 1975, *Intervirology*, 5: 83.
5. Bando, H., Kusuda, J. Gojobori, T., Maruyama, T. and Kawase, S., 1985, *J. Virol.*, 61: 553.
6. Bando, H., Choi, H., Ito, Y. and Kawase, S., 1990, *Virology*, 179: 57.
7. Afanasiev, B. N., Cayloy, E. E., Buchatsky, L. P. and Kozlov, V. V., 1991, *Virology*, 185: 323.
8. Boublik, Y., Jousset, F. X. and Bergoin, M., 1994, *Virology*, 200: 752.
9. Dumas, B., Jourdan, M., Pascaud, A. M. and Bergoin, M., 1992, *Virology*, 191: 202.
10. Tijssen, P. GenBank accession number L32896
11. Tijssen, P. Unpublished results.
12. Boublik, Y., Kouassi, K. N., Cavallaro, C., and Bergoin, M. Unpublished results.
13. Tijssen, P., Jousset, F.X., Bergoin, M. Unpublished results.
14. Baquerizo, E., Jousset, F. X., Boublik, Y., Cousserans, F., and Bergoin, M. Unpublished results.
15. Fediere, G., Veyrunes, J.-C., Bergoin, M., and Tijssen, P. Unpublished results.
16. Fediere, G., and Tijssen, P. Unpublished results.
17. Fediere, G., Lery, X., Quiot J.-M., and Dauthuille, D., 1990, *J. Invertebr. Pathol*, 56: 132.
18. Fediere, G., 1983, Ph.D. thesis, Universite Montpellier II, Montpellier.
19. Simpson, A., Baker, T. S., Chipman, P., Tijssen, P., and Rossmann, M.G., (1998) Submitted to Structure.
20. Bergeron, J., Hebert, B. and Tijssen, P., 1996, *J. Virol.*, 70: 2508.