Structural insight into insect viruses

Michael G. Rossmann and Yizhi Tao

The first structure of an insect picorna-(small RNA-containing) virus is now available. Although there is considerable similarity in the structures of mammalian and insect picornaviruses, there are also remarkable differences, the most noteworthy being associated with the small, internal, functionally essential, VP4 protein.

Traditionally, biologists have analyzed the evolution of organisms by considering morphological similarities, supported, where possible, by fossil records. For instance, six-legged animals have probably evolved from a common primordial precursor insect that is distinct from tetrapods. The techniques and accuracy of determining phylogenetic trees was revolutionized with the publication by Fitch and Margoliash[1] of a tree linking all major branches of the animal and plant kingdoms, based on a comparison of cytochrome *c* amino acid sequences. However, sequence comparisons are limited to functionally related molecules that retain a 'memory' of a common ancestral molecule. Comparisons of the three-dimensional structures of dehydrogenases and other nucleotide binding proteins[2] demonstrated that polypeptide folding motifs are generally retained over much longer periods of time. Whether or not every two proteins with structural similarity have originated from a common ancestor remains debatable and depends, in part, upon quantifying the measure of similarity and considering other properties, such as common functions. Nevertheless, today one of the first things that is asked about a new structure is whether a similar structure has ever been observed before. This question can certainly be asked about the Cricket paralysis virus structure reported by Tate *et al.*[3] on page 765 in this issue of *Nature Structural Biology*.

Structures common to many viruses

The first spherical virus structure to be determined in three dimensions at near atomic resolution was that of tomato bushy stunt virus (TBSV)[4]. It was a surprise to find, shortly afterwards, that the capsid proteins of TBSV and southern bean mosaic virus (SBMV)[5] not only had a similar tertiary structure (commonly referred to as a 'jelly-roll' fold), but that the 180 copies of the ~30 kDa proteins in each of these viruses were organized in exactly the same way. Both viruses are plant RNA viruses with molecular weights of ~8 million Da, but with no readily detectable sequence similarities. Even more surprising was the discovery that a human common cold (rhino-) virus[6] and a poliovirus[7], members of the animal RNA picornavirus family, have capsid structures that are similar to the RNA plant viruses, such as TBSV and SBMV. While TBSV and SBMV have capsids that consist of 180 copies of the same capsid protein, the animal viruses have 60 copies of each of three different viral proteins (VP1, VP2, VP3) and 60 copies of a small (~70 amino acids) internal viral protein, VP4. The exceptional extent of the structural and functional similarity leaves little doubt that these plant and animal viruses must have evolved from a common primordial virus[6]. More extraordinary still were the further discoveries of the same viral capsid fold in ssRNA insect viruses, ssDNA viruses, and dsDNA viruses[8], and in dsRNA viruses[9], although now there are quite a few viruses (such as bacteriophage MS2[10]) that are known to have capsid structures with completely different folds.

An insect virus structure

The range and variety of insect species are probably greater than in most other classes, yet the study of insect viruses is quite limited[11,12]. Indeed, Pasteur was among the relatively few to study insect pathogens, thereby rescuing the silkworm industry of southern France (see the History section on page 720 of this issue). Now, Tate *et al.*[3] have analyzed the structure of an insect virus that has substantial structural similarities to rhino- and polio-picornaviruses. Picorna-like insect viruses have an ssRNA genome, three major cap-

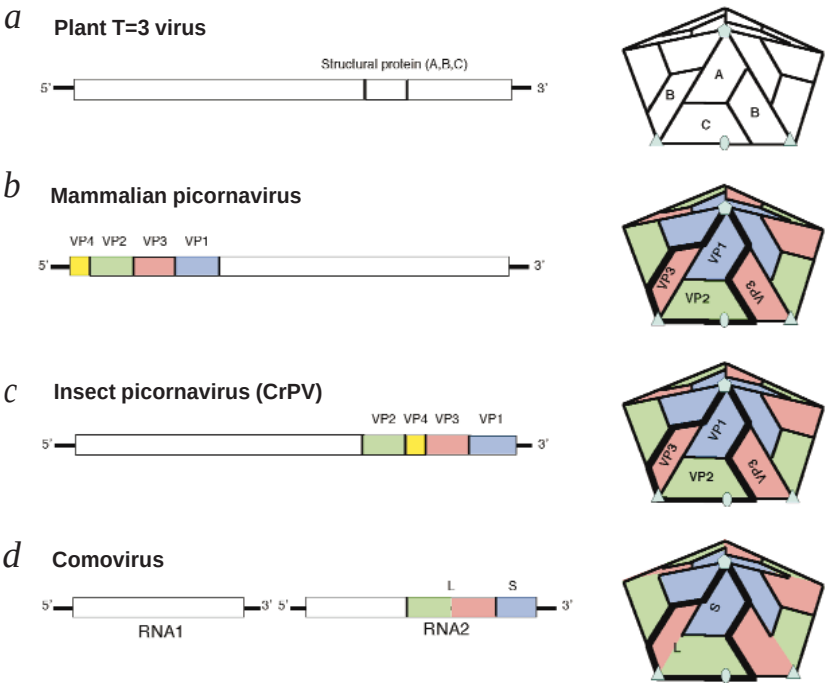

Fig. 1 Comparison of gene order and capsid structures of various +RNA strand viruses whose RNA genome is read from the 5' to 3' end. ***a***, Plant viruses, such as tomato bushy stunt and southern bean mosaic viruses. ***b***, Mammalian picornaviruses, such as rhino-, polio-, and hepatitis A virus. ***c***, Cricket paralysis virus, an insect picornavirus. ***d***, Plant comoviruses, such as cowpea mosaic virus.

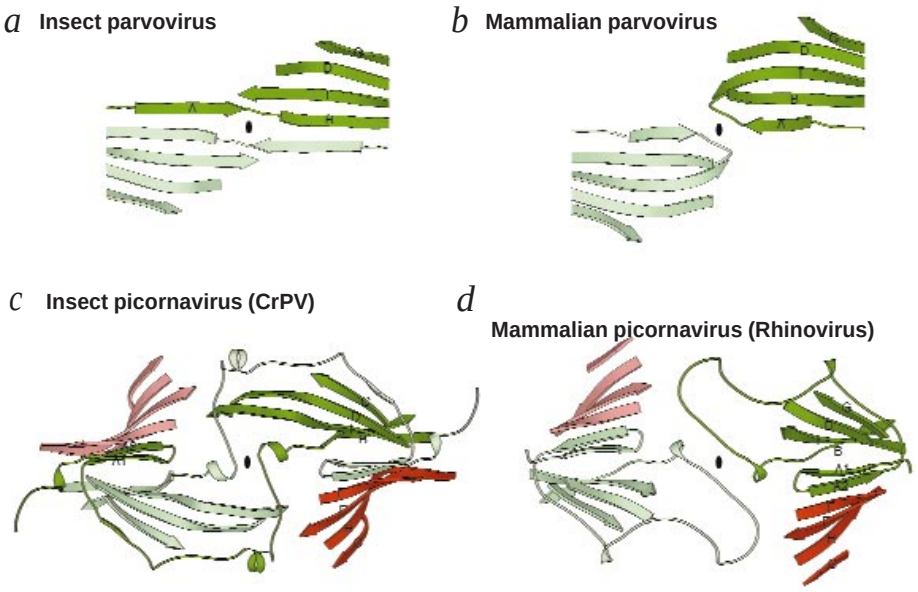

Fig. 2 ***a***,***b***, Domain swapping of the N-terminal arm about the icosahedral two-fold axis in insect and mammalian parvoviruses compared to ***c***,***d***, domain swapping of the VP2 (green) N-terminal arm in insect and mammalian picornaviruses. The β-strands of the common 'jelly roll' β-barrel are named B, I, D, G on one side of the β-barrel and C, H, E, F on the other. VP3 of the picornavirus is red. All the views are from the inside of the viruses looking out.

sid proteins (VP1, VP2, and VP3), an internal small capsid protein (VP4), and a molecular weight that is similar to animal picornaviruses. Tate *et al.*[3] observe that the tertiary and quaternary structures of cricket paralysis virus (CrPV) and animal picornaviruses are remarkably similar, although the organization of each genome[13,14] is considerably different.

The RNA genome of animal picornaviruses has one long open reading frame that is translated into a polyprotein. This protein is post-translationally cleaved into functionally-distinct proteins by proteases that are part of the polyprotein. The N-terminal end of the polyprotein consists of VP0, VP3, and VP1. These proteins are excised out of the polyprotein as a single peptide, which folds into a 6S 'protomer' assembly intermediate. Subsequent cleavage events direct and control the assembly to form, first, 12S pentameric particles and, then, fully-assembled virions[6,15,16]. Finally, VP0 is cleaved into VP4 and VP2, probably due to the suitable juxtaposition of catalytic residues in the assembled particle[17]. This last 'maturation' step causes the virus to be infectious and allows the internal VP4 to be ejected as the first step after recognition of a new host cell. Why VP4 needs to be released on cell entry is not clear, although similar maturation events occur in other RNA viruses[18,19].

In contrast, the insect picornavirus genome is translated as two separate polyproteins, with the downstream segment coding for the structural proteins. Such a separation of functional and structural genes should not come as a surprise to the senior author of the CrPV paper, Jack Johnson, who had previously studied plant comoviruses, which have a genome that is divided into two RNA molecules[20]. The gene order of the three structural motifs corresponding to VP2, VP3, and VP1 is the same in picornaviruses, comoviruses, and the picorna-like insect viruses, as is the position of these structures in the various capsids (Fig. 1). However, in the case of the insect picornaviruses, VP4 is not at the beginning of the polyprotein, but is inserted between VP2 and VP3. The proximity of the C-terminus of VP4 and the N-terminus of VP3, and the location of a

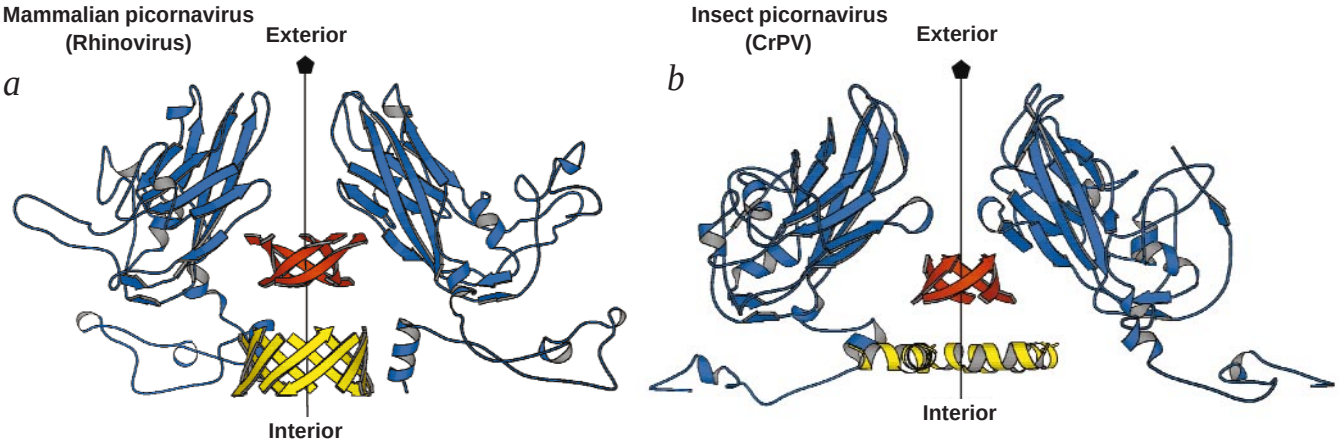

Fig. 3 Comparison of the channel along the five-fold axes of ***a***, mammalian and ***b***, insect picornaviruses. VP1 is blue; the N-termini of VP3, which form a β-cylinder, are red; and VP4 is yellow. VP4 makes a 10-stranded antiparallel N-terminal β-cylinder in mammalian picornaviruses, whereas in the insect CrPV, the C-termini of VP4 form an α-helical disc-like structure. During the initial stages of infection, VP4 of mammalian picornaviruses is externalized, and VP4 in CrPV may play a similar role.

conserved aspartic acid residue near the anticipated cleavage site in CrPV, suggests that the maturation event is the separation of VP4 from the N-terminus of VP3.

In both the mammalian and the insect picornaviruses, the N-terminal end of VP3 forms a β-cylinder by intertwining the five N-terminal ends of symmetry-related VP3s. It has been suggested[16] that liberation of the N-terminal end of VP3 from the polyprotein results in the formation of the β-cylinder and, hence, drives the assembly of five protomers into a pentamer. It is difficult to visualize how this event might take place when the N-terminus of VP3 is encumbered by VP4. Nor does it seem likely that the maturation cleavage takes place before the assembly is completed, as the catalytic residues would not yet be positioned correctly around the active site. However, the occurrence of 'domain swapping' (alternate polypeptide connection between otherwise identical domains) (Fig. 2) might suggest that the different domain organization in CrPV as compared to mammalian picornaviruses requires a different assembly pathway. Whereas most picornaviruses are assembled from pentameric intermediates, the structure of CrPV implies that it is assembled from dimers. If this were so, then the assembly pathway of CrPV would be closer to that of many plant viruses[21], which also depend upon dimeric assembly intermediates. It might be more than a coincidence that domain swapping occurs between mammalian and insect parvovirus structures[22] just as it does between mammalian and insect picornavirus structures (Fig. 2).

The surface of CrPV is substantially different from that of rhino- and polioviruses in that the 'canyon' is completely absent. This feature was first recognized in the structure of human rhinovirus 14 and was predicted to be the site of cellular receptor attachment[6], as the canyon was thought to be a sterically protected location that would be inaccessible to neutralizing antibodies. The predicted receptor attachment site in the canyon was later confirmed[23], but the rationale of the prediction was questioned by Smith *et al.*[24], who observed that the neutralizing antibody penetrated well into the canyon. Indeed, it had also been observed[23] that the footprint of the receptor on the virus surface extended beyond the edge of the canyon, possibly suggesting that residues in the canyon floor contribute more to the energy of receptor binding than do the residues in the overlap region between the receptor and neutralizing antibody. It is, therefore, of interest that the canyon is missing from CrPV, especially noting that the cricket, the host organism of this virus, does not have antibodies.

The β-cylinder, constructed of the N-termini of VP3, surrounds the five-fold axes in CrPV as well as in mammalian picornaviruses (Fig. 3). In CrPV, the VP3 β-cylinder is surrounded by an α-annulus of greater radius, formed by the α-helical C-termini of VP4 (Fig. 3*b*). In mammalian picornaviruses, this position is also occupied by VP4, but in this case it is the N-terminus of VP4, and it forms a β-cylinder instead of an α-annulus (Fig. 3*a*).

There has been extensive speculation as to which aperture of the viral capsid is used by VP4 molecules when they emerge from the virion in the initial stages of infection[17,25]. One possibility is that the VP4s use the channels along the five-fold axes (Fig. 3). If this is the case, and if CrPV also loses VP4 early on, then it is surprising, but apparently immaterial, that first part of VP4 to emerge from the virion is the N-terminus in one case and the C-terminus in the other (Fig. 3).

The preceding discussion indicates that one of the important lessons learned from the CrPV structure determined by Tate *et al.*[3] is that comparative structural studies can be essential to the analysis of the processes required for viral propagation. Moreover, such analyses can elaborate, in unexpected ways, on viral evolution.

Michael G. Rossmann and Yizhi Tao are in the Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA.

Correspondence should be addressed to M.G.R. email: mgr@indiana.bio.purdue.edu

1. Fitch, W.M. & Margoliash, E. *Science* **155**, 279-284 (1967).
2. Rossmann, M.G., Moras, D. & Olsen, K.W. *Nature* **250**, 194-199 (1974).
3. Tate, J. *et al. Nature Struct. Biol.* **6**, 765–774 (1999).
4. Harrison, S.C., Olson, A.J., Schutt, C.E., Winkler, F.K. & Bricogne, G. *Nature* **276**, 368-373 (1978).
5. Abad-Zapatero, C. *et al. Nature* **286**, 33-39 (1980).
6. Rossmann, M.G. *et al. Nature* **317**, 145-153 (1985).
7. Hogle, J.M., Chow, M. & Filman, D.J. *Science* **229**, 1358-1365 (1985).
8. Rossmann, M.G. & Johnson, J.E. *Annu. Rev. Biochem.* **58**, 533-573 (1989).
9. Grimes, J.M. *et al. Nature* **395**, 470-478 (1998).
10. Valegård, K., Liljas, L., Fridborg, K. & Unge, T. *Nature* **345**, 36-41 (1990).
11. Vago, C. & Bergoin, M. In *Advances in virus research*. (eds Smith, K.M. & Lauffer, M.A.) 247-303 (Academic Press, New York; 1968).
12. Johnson, J.E. & Reddy, V. In *The insect viruses.* (ed. Miller, L.) 171-223 (Plenum, New York; 1998).
13. Koonin, E.V. & Gorbalenya, A.E. *FEBS Lett.* **297**, 81-86 (1992).
14. Sasaki, J., Nakashima, N., Saito, H. & Noda, H. *Virology* **244**, 50-58 (1998).
15. Jacobson, M.F. & Baltimore, D. *Proc. Natl. Acad. Sci. USA* **61**, 77-84 (1968).
16. Arnold, E. *et al. Proc. Natl. Acad. Sci. USA* **84**, 21-25 (1987).
17. Basavappa, R. *et al. Prot. Sci.* **3**, 1651-1669 (1994).
18. Schneemann, A., Zhong, W., Gallagher, T.M. & Rueckert, R.R. *J. Virol.* **66**, 6728-6734 (1992).
19. Zlotnick, A. *et al. J. Biol. Chem.* **269**, 13680-13684 (1994).
20. Chen, Z., Stauffacher, C.V. & Johnson, J.E. *Sem. Virol.* **1**, 453-466 (1990).
21. Rossmann, M.G. *et al. J. Mol. Biol.* **165**, 711-736 (1983).
22. Simpson, A.A., Chipman, P.R., Baker, T.S., Tijssen, P. & Rossmann, M.G. *Structure* **6**, 1355-1367 (1998).
23. Olson, N.H. *et al. Proc. Natl. Acad. Sci. USA* **90**, 507-511 (1993).
24. Smith, T.J., Chase, E.S., Schmidt, T.J., Olson, N.H. & Baker, T.S. *Nature* **383**, 350-354 (1996).
25. Giranda, V.L. *et al. Proc. Natl. Acad. Sci. USA* **89**, 10213-10217 (1992).