Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 29 April 2011
Accepted 31 August 2011

Form, symmetry and packing of biomacromolecules. IV. Filled capsids of cowpea, tobacco, MS2 and pariacoto RNA viruses

A. Janner

Theoretical Physics, FNWI, Radboud University, Heyendaalseweg 135, NL-6525 AJ Nijmegen, The Netherlands. Correspondence e-mail: a.janner@science.ru.nl

Four icosahedral RNA viruses are considered: the cowpea chlorotic mottle virus, the satellite tobacco mosaic virus, the pariacoto virus and the MS2 bacteriophage. The validity of the phenomenological rules derived in previous publications (crystallographic scaling, indexed forms enclosing axial-symmetric clusters, packing lattices of viral crystals) is confirmed and shown to apply equally well to the coat proteins as to the (ordered) RNA chains.

1. Introduction

In part IV of an investigation of crystallographic relations between enclosing forms, symmetry groups and lattice-periodically packed biomacromolecules, the attention is focused on viruses with partially ordered RNA genomes in the icosahedral capsid. In particular, the following are considered (see Table 1):

(i) The pariacoto virus with a dodecahedral cage of duplex RNA [refcode 1f8v of the Brookhaven Protein Data Bank (PDB) with primary reference Tang *et al.* (2001)].

(ii) The satellite tobacco mosaic virus/RNA complex (PDB refcode 1a34) (Larson *et al.*, 1998).

(iii) The native cowpea chlorotic mottle virus (PDB refcode 1cwp) (Speir *et al.*, 1995).

(iv) The RNA MS2 bacteriophage (PDB refcode 1zdh) (Valegård *et al.*, 1997).

Interest in these viruses has been stimulated by the work of Twarock and Keef of the York Center for Complex System Analysis, York, by that of Stockley, Astbury Center for Structural Molecular Biology, Leeds (Jonoska & Twarock, 2006; Toropova *et al.*, 2008; Keef & Twarock, 2009*a*,*b*; Grayson *et al.*, 2009; Stockley & Twarock, 2009), and by various contributions presented at the Third Mathematical Virology Workshop held in Ambleside, UK, summer 2010, where the importance of the structural connections between the viral capsid and the encapsulated genome clearly emerged, as in particular in a poster by Wardman & Keef (2010) that discussed applications of Keef & Twarock (2009*a*) to pariacoto virus and bacteriophage MS2 based on results by Keef *et al.* (2011).

Several generic structural properties of the capsid, such as crystallographic scaling, indexed enclosing forms of axial-symmetric multimeric clusters and integral crystal lattices based on viral sphere packing, have already been analyzed by the author in a number of publications (Janner, 2004, 2006*a*,*b*, 2008, 2010*a*,*b*).

Table 1
Molecular structure (the length gives the number of backbone positions determined).

PDB key	RNA virus	Chains	Type	Length
1cwp	Cowpea chlorotic	*A*, *B*, *C*	Coat proteins	190
	Mottle virus	*D*, *F*	RNA	4
	(CCMV)	*E*	RNA	2
1a34	Satellite tobacco	*A*	Capsid protein	151
	Mosaic virus	*B*	RNA	10
	(STMV)	*C*	RNA	10
1zdh	Bacteriophage MS2	*A*, *B*, *C*	Coat protein	355
	(MS2)	*R*, *S*	RNA	19
1f8v	Pariacoto virus	*A*, *B*, *C*	Coat protein β	355
	(PaV)	*D*, *E*, *F*	Coat protein γ	40
		R	RNA	25

The aim of this paper is not only to verify, through few illustrative examples, that these properties typically occur in the four viruses indicated above, but also to show that the same phenomenological laws valid for the capsid apply equally well to the genome (more precisely to the ordered RNA chain segments which share the icosahedral symmetry and whose atomic positions could, therefore, be determined). Such structural compatibility between protein and nucleic acid components has already been observed in axial-symmetric proteins complexed with DNA/RNA (Janner, 2005).

2. Crystallographic scaling

Crystallographic scaling is an important property of icosahedral viruses leading to extensions of the icosahedral point group 235, which allows a classification (Janner, 2006*b*) going beyond the surface-based one of Caspar & Klug (1962). The alternative Keef & Twarock approach *via* affine extended symmetry groups (Keef & Twarock, 2009*a*) has already been applied successfully to the shell structure of the tomato bushy stunt virus (Keef & Twarock, 2009*b*) and to the dodecahedral RNA-cage in the pariacoto virus and the double-shell RNA

density determined by cryo-EM (Wardman & Keef, 2010) based on Keef & Twarock (2009*a*) and Keef *et al.* (2011).

For the pariacoto virus, Fig. 1 shows that the dodecahedron, around which the RNA duplex is wrapped, is scaled by a factor $1/\tau$ with respect to the dodecahedral vertices of the ico-dodecahedron enclosing the capsid, where τ is the golden number $(1+\sqrt{5})/2$. These dodecahedral vertices are generated from $[111\bar{1}1\bar{1}]$ by a six-dimensional integral representation of the icosahedral point group. In a similar way, the $1/\tau$ scaling is represented by the inverse of the scaling matrix $S_\tau(a)$ given in equation (11) of Janner (2006*b*). Accordingly, the vertices of the RNA dodecahedron have indices icosahedrally equivalent to $[000\bar{2}\bar{2}2]$.

The sphere enclosing the pariacoto virus, presented in a projected view in Fig. 1, is inscribed in the ico-dodecahedron mentioned above. We recall that the ico-dodecahedron is obtained from a $T = 3$ Caspar–Klug icosahedral polyhedron (whose icosahedral vertices are generated from [100000] and dodecahedral ones generated from [111000]) by a $1/\tau$ rescaling of the dodecahedral vertices [see Table 1 and Fig. 5 of Janner (2006*b*)]. The fitting between this sphere and the virus has been carried out by eye, not encompassing all external protruding loops which have other properties than the bulk of the major coat proteins. One could adopt a larger sphere including these loops, or make use of a computer-based algorithm as in Keef *et al.* (2011). These procedural changes would only slightly modify the value of the scaling factor between the enclosing forms of the capsid and the RNA, respectively, but not that of the ideal structure, characterized by a $1/\tau$ ratio.

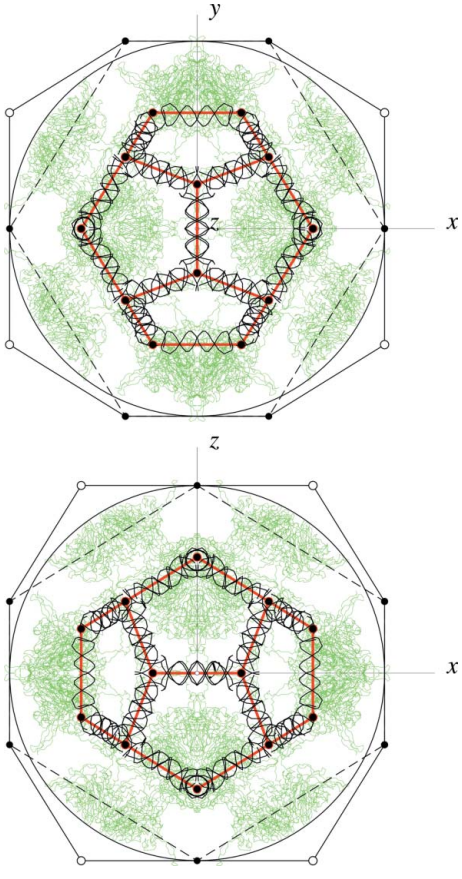

Figure 1
In the pariacoto virus the RNA duplex is wrapped around a dodecahedral cage, scaled by a factor $1/\tau = (1-\sqrt{5})/2$ from the dodecahedral vertices (filled black circles) of the ico-dodecahedron whose inscribed sphere encloses the capsid. Plotted in the background is the coat protein *A* (in green). Shown are projections along the two-fold icosahedral axes *z* and *y*.

3. Indexed forms of axial-symmetric clusters

The existence of enclosing forms with vertices at (projected) points of a form lattice, indexed accordingly, is a fundamental crystallographic property of a great variety of axial-symmetric biomacromolecules, giving rise, in general, to scaling relations expressible in terms of star polygons with Schläfli symbol $\{m/n\}$ (Coxeter, 1963). In the rhinovirus, in particular, these forms are serotype invariant (Janner, 2006*a*).

In the pariacoto virus, indexed enclosing forms have been systematically derived for the decameric, hexameric and tetrameric clusters of the capsid proteins β (with chains A, B, C) and γ (chains D, E, F), and of the RNA chain R, in the C_α and P backbone approximation, respectively, making use of the PDB structural data with refcode 1f8v.

In Fig. 2 the enclosing forms of the coat protein *A* (on the left-hand side) and of the RNA chain *R* (on the right-hand

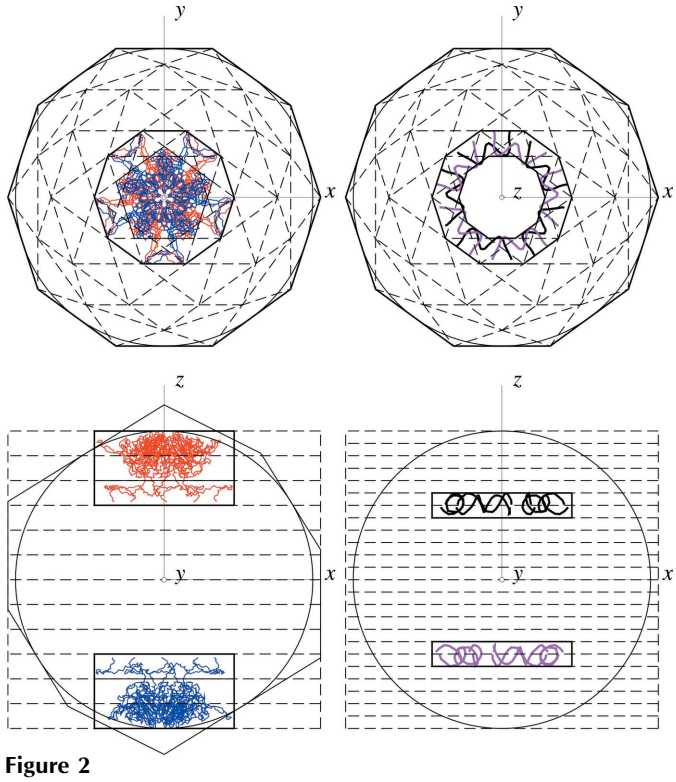

Figure 2
The enclosing form of the decameric cluster {5,6,7,8,9; 10,11,12,13,14} of the coat protein *A* (left-hand side, with pentamers in red and blue, respectively) and of the RNA chain *R* (right-hand side, in black and magenta) of the pariacoto virus are shown along and perpendicular to the icosahedral five-fold axis. The rational indices of the vertices of the enclosing forms in the rotation plane (x, y) follow from the star-polygon scaling relations indicated by dashed lines and in the perpendicular projection from the equidistant intervals between the planes of rotation.

side) of the pariacoto virus are presented together for allowing a comparison which shows the validity of the same law for the capsid and the genome. In the axial projection along the icosahedral five-fold axis (here along the z-coordinate) the external decagonal boundary is related to that of the A and R decamers [which are a pair of pentamers obtained from the icosahedral transformations labeled as {5, 6, 7, 8, 9; 10, 11, 12, 13, 14} (Janner, 2006*a*) and plotted in Fig. 2 in different colors] by successive crystallographic scalings defined by star-polygon transformations. In terms of Schläfli symbols these are given by {10/2}, {10/2}, {10/3}, as indicated by dashed lines. An additional {10/3} scaling relates the external RNA boundary to the decagonal internal one, whereas for the coat protein the same scaling relation leads to a pentagon which delimits the globular region of A from the protruding chain segments. In the perpendicular projection (along the y-coordinate) shown at the bottom of Fig. 2 the boundaries occur at equidistant intervals, $R/6$ for the A clusters and $R/12$ for the RNA chains, respectively, where R is the radius of the spherical form of the capsid inscribed in the viral ico-dodecahedron.

Another example is shown in Fig. 3 for the hexameric cluster around the icosahedral trigonal axis (z) and labeled as {1, 21, 41; 18, 38, 58}, of the coat protein A and of the RNA chains B and C, respectively, in the satellite tobacco mosaic virus (PDB refcode 1a34; REVDAT 2009). Also indicated are the points of the hexagonal form lattice, which fit with the ico-dodecahedron of the virus and have lattice points at the vertices of the various enclosing forms.

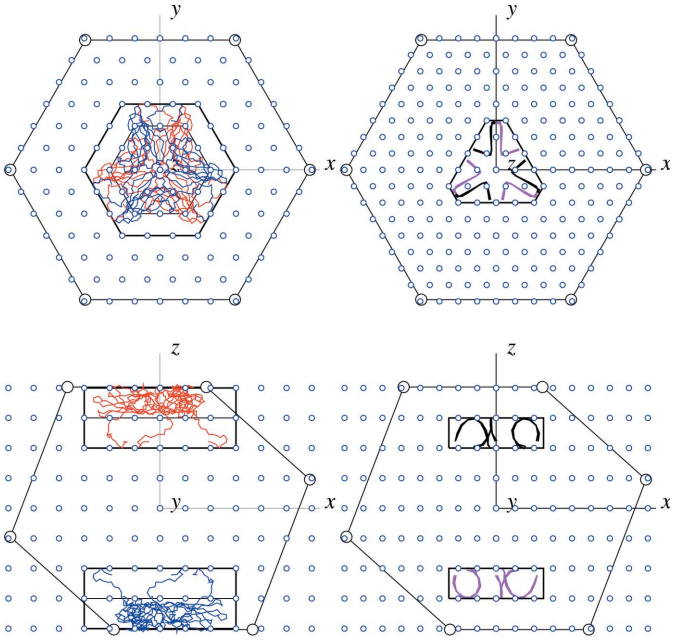

Figure 3
In a similar way as in Fig. 2, the enclosing forms of the hexameric cluster {1, 21, 41; 18, 38, 58} of the coat protein A (left-hand side, with the trimers in red and blue) and of the RNA chains B, C (right-hand side, with trimers in black and magenta, respectively) of the satellite tobacco mosaic virus have vertices at points of a form lattice Λ_F (here, hexagonal). Shown are projections along and perpendicular to the icosahedral three-fold axis. Also indicated is the boundary of the ico-dodecahedron enclosing the capsid, together with icosahedral vertices (large circles).

Again the same law applies to both the nucleic acid and the polypeptide chain. As in the previous figure, the globular region and the protruding chain segments of the coat protein are separated by boundaries with vertices belonging to the same hexagonal form lattice.

4. Crystal packing and packing lattice Λ_P

The compatibility between the icosahedral point group symmetry and the crystal space group of a given virus can be made explicit by considering the lattice periodic packing of the viruses and the corresponding packing lattice Λ_P. As discussed in detail in the articles already published in this same series (Janner, 2010*a*,*b*, 2011), the packing lattice $\Lambda_P(u, v, w)$ has the

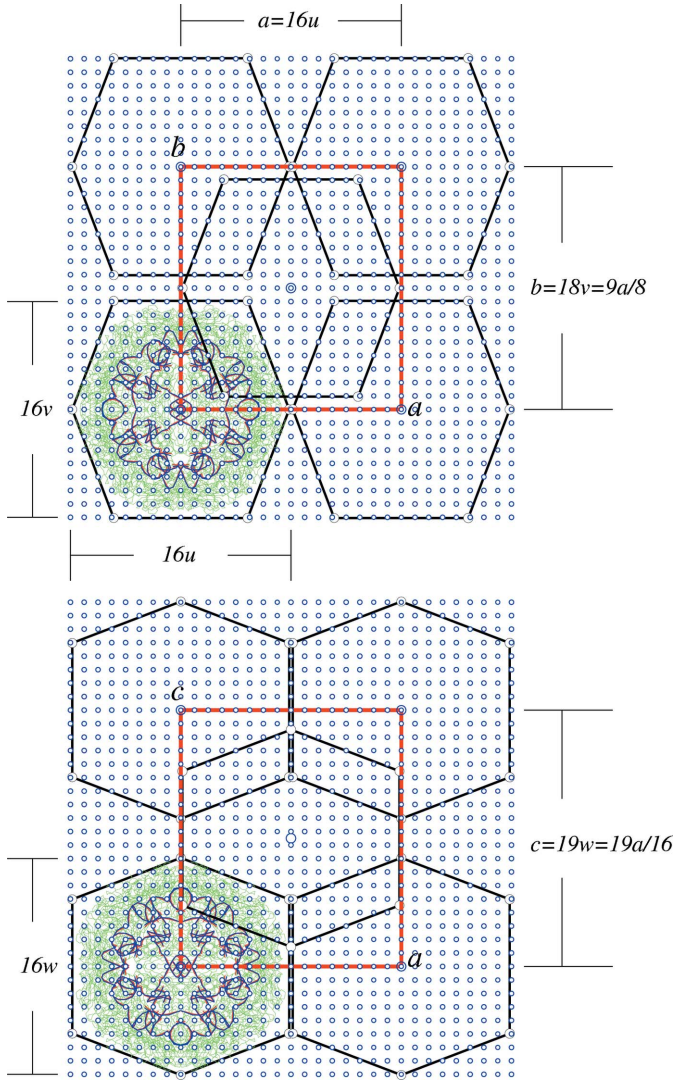

Figure 4
The orthorhombic crystal packing of the satellite tobacco mosaic virus (space group $I222$), enclosed in the ico-dodecahedron of the capsid, is shown together with points of the packing lattice $\Lambda_P(u, v, w)$, the (centered) orthorhombic unit cell (in red) of the crystal lattice $\Lambda(a, b, c)$ and the ico-dodecahedron of the enclosing forms. The relations $a = 16u$, $b = 18v$, $c = 19w$, ensure that Λ is an integral lattice. Also indicated is how the coat protein A (green) and the RNA chains B, C (blue) fill the viral capsid.

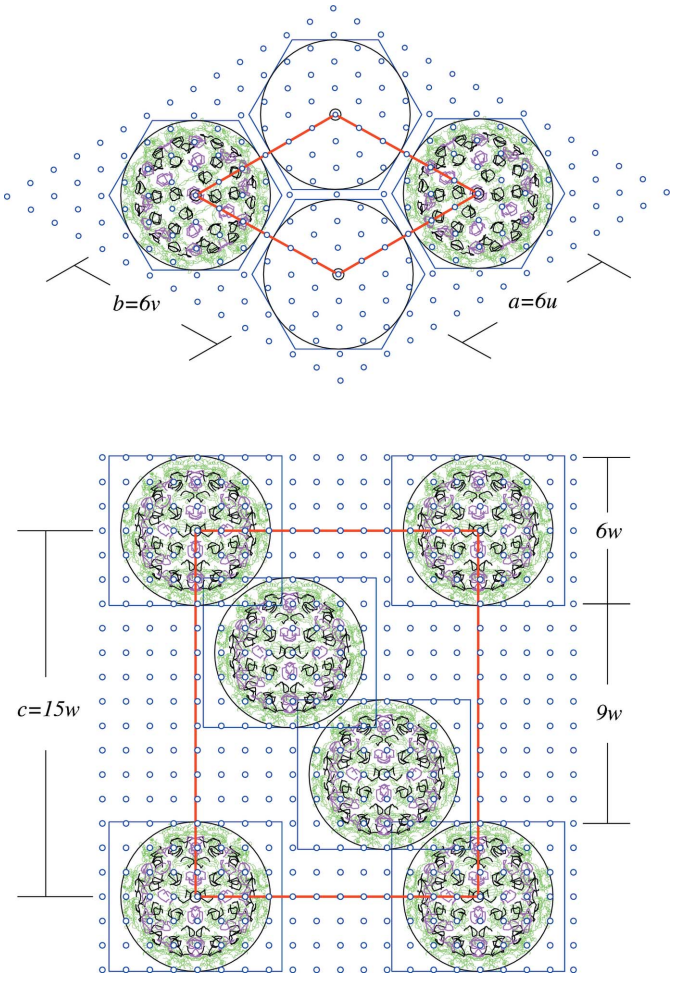

Figure 5
In a similar way as in Fig. 4, the hexagonal crystal packing of the MS2 bacteriophage (space group *H*32) is shown in the spherical and in the ico-dodecahedral enclosing form of the virus. Indicated are the integral relations between the crystal lattice $\Lambda(a, b, c)$ and the packing lattice $\Lambda_P(u, v, w)$. Represented is how the coat protein *A* (green) and the RNA chains *R* (black) and *S* (magenta) fill the capsid.

crystal lattice $\Lambda(a, b, c)$ as a sublattice and fits with enclosing forms of the packed viruses. The parameters of these two lattices are connected by integers N_a, N_b, N_c according to

$$a = N_a u, \qquad b = N_b v, \qquad c = N_c w. \tag{1}$$

The crystal packing structure is shown in Fig. 4 for the satellite tobacco mosaic virus, together with the projected ico-dodecahedral enclosing forms, and in Fig. 5 for the MS2 bacteriophage, with the packing lattice related to spherical viral forms.

5. Conclusion

The structures of the considered four viruses further corroborate the phenomenological laws derived in previous publications and, in particular, the surprising connection between details of the viral structure with the space-group symmetry of its crystal.

References

Caspar, D. L. D. & Klug, A. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
Coxeter, H. S. M. (1963). *Regular Polytopes*, p. 60. New York: Macmillan.
Grayson, N. E., Taormina, A. & Twarock, R. (2009). *Theor. Comput. Sci.* **410**, 1440–1447.
Janner, A. (2004). *Acta Cryst.* A**60**, 198–200.
Janner, A. (2005). *Acta Cryst.* D**61**, 269–277.
Janner, A. (2006*a*). *Acta Cryst.* A**62**, 270–286.
Janner, A. (2006*b*). *Acta Cryst.* A**62**, 319–330.
Janner, A. (2008). *Comput. Math. Methods Med.* **9**, 167–173.
Janner, A. (2010*a*). *Acta Cryst.* A**66**, 301–311.
Janner, A. (2010*b*). *Acta Cryst.* A**66**, 312–326.
Janner, A. (2011). *Acta Cryst.* A**67**, 174–189.
Jonoska, N. & Twarock, R. (2006). arXiv:q-bio/0611021v1.
Keef, T. & Twarock, R. (2009*a*). *J. Math. Biol.* **59**, 287–313.
Keef, T. & Twarock, R. (2009*b*). *Emerging Topics in Physical Virology*, edited by P. G. Stockley & R. Twarock, pp. 59–84. London: Imperial College Press.
Keef, T., Wardman, J., Ranson, N. A., Stockley, P. C. & Twarock, R. (2011). *PLos Biol.* Submitted.
Larson, S. B., Day, J., Greenwood, A. & McPherson, A. (1998). *J. Mol. Biol.* **277**, 37–59.
Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. (1995). *Structure*, **3**, 63–78.
Stockley, P. G. & Twarock, R. (2009). Editors. *Emerging Topics in Physical Virology*, pp. 332. London: Imperial College Press.
Tang, L., Johnson, K. N., Ball, L. A., Lin, T., Yeager, M. & Johnson, J. E. (2001). *Nat. Struct. Biol.* **8**, 77–83.
Toropova, K., Basnak, G., Twarock, R., Stockley, P. G. & Ranson, N. A. (2008). *J. Mol. Biol.* **375**, 824–836.
Valegård, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G. & Liljas, L. (1997). *J. Mol. Biol.* **270**, 724–738.
Wardman, J. & Keef, T. (2010). *3rd Mathematical Virology Workshop*, Ambleside, UK. Poster.