Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 7 March 2006
Accepted 11 May 2006

Crystallographic structural organization of human rhinovirus serotype 16, 14, 3, 2 and 1A

A. Janner

Theoretical Physics, Radboud University, Nijmegen, Toernooiveld 1, NL-6525 ED, The Netherlands. Correspondence e-mail: a.janner@science.ru.nl

The architecture of the human rhinovirus is shown to be based on a crystallographic polyhedron (the ico-dodecahedron) with 60 triangular facets and 32 vertices at points of a body-centered icosahedral lattice. The ico-dodecahedron is only slightly different from the $T = 3$ icosadeltahedron of Caspar & Klug [*Cold Spring Harbor Symp. Quant. Biol.* (1962), **27**, 1–24]. The capsid of the virion is encapsulated between two ico-dodecahedra in scaling relation by a factor τ, the golden number. Clusters with axial symmetry of the coat proteins VP1, VP2, VP3 and VP4 are considered (decamers, pentamers, hexamers, trimers and tetramers). Their crystallographic enclosing forms obey the same laws as a number of axial-symmetric proteins, involving planar and linear crystallographic scaling relations and having vertices at points of lattices with an integral metric tensor. These properties also occur for the icosahedral cluster of each coat protein viewed along the symmetry axes (fivefold, threefold and twofold, respectively). The structural organization of the rhinovirus in terms of all these enclosing forms is independent of the serotype (16, 14, 3, 2, 1A) and is typical for a strongly correlated system, as it depends on one single free parameter, taken to be the icosahedral lattice parameter a_0, which relates the geometry with the real structure (up to small variations).

1. Introduction and background

Rotational symmetry plays a fundamental role in biomacromolecules. The icosahedral viruses are among the most impressive examples. Less known is the crystallographic character of the molecular symmetries and the interplay between rotations and scaling transformations in these structures.

In particular, axial-symmetric proteins have been shown to have molecular enclosing forms with non-trivial crystallographic rotation and scaling properties (Janner, 2005*a*,*b*,*c*). Axial symmetry implies that these transformations are essentially planar. Essentially means that one disregards possible additional perpendicular twofold rotations. Crystallographic means a transformation expressible with respect to a basis of n vectors in space as an n-dimensional invertible matrix with integral entries. Accordingly, the vertices of the polyhedral molecular forms have vertices labeled by a set of n integers (the *indices*), which are the integral components of the corresponding position vectors with respect to the n basis vectors. When this is the case, the forms (and the polyhedra) are said to be crystallographic. The most important scalings observed in axial-symmetric proteins are the polygrammal ones (as in regular star polygons). They relate, in particular, the external boundary of the protein to the boundary of a central hole.

The aim of the present work is to investigate the validity of analogous scale-rotational relations in the cubic case of icosahedral viruses. In a strict sense, the icosahedral group is only crystallographic in six dimensions where it leaves a lattice invariant. In three dimensions, it is, however, possible to define six basis vectors which are transformed by the icosahedral point group 532 into integral-linear combinations of the basis vectors. Therefore, according to the definition given above, the icosahedral group is crystallographic. This point of view is supported by the symmetry properties of the icosahedral quasicrystal phases occurring in nature. These quasicrystals also have radial (isotropic) scaling properties with scaling factor τ or τ^3, depending on the icosahedral lattice left invariant (primitive, face centered or body centered), where τ denotes the golden mean $(1 + \sqrt{5})/2$.

Encasing forms for icosahedral viruses have been postulated by Caspar & Klug in a classical paper of 1962 (Caspar & Klug, 1962). Generalizing the construction of the Fuller geodesic dome, Caspar & Klug consider polyhedra whose faces are all equilateral triangles (*deltahedra*) and derive the deltahedra with icosahedral symmetry (*icosadeltahedra*). They show that these polyhedra have $20T$ facets, where T is the triangulation number. The possible values of T are restricted by the condition

$$T = h^2 + hk + k^2, \quad h, k \quad \text{arbitrary integers.} \qquad (1)$$

This restriction follows from the construction of the icosadeltahedron from a planar hexagonal net.

Not all icosahedral viruses obey the Caspar–Klug rule, but most do and remarkably well, in particular for the $T = 3$ case of the picorna viruses, where T corresponds to the number of the major coat proteins VP1, VP2, VP3 of the capsid. A fourth protein VP4 of the 60 protomers lies mostly inside the capsid. The largest genus among the picorna viruses is the rhinovirus. It occurs in about 100 different serotypes. Considered here are the five serotypes 16, 14, 3, 2 and 1A, whose structures have been determined by X-ray diffraction (Kim *et al.*, 1989; Arnold & Rossmann, 1990; Zhao *et al.*, 1996; Hadfield *et al.*, 1997; Verdaguer *et al.*, 2000) with data available from the Brookhaven Protein Data Bank (PDB).

Before being able to compare the rhinovirus with axial-symmetric proteins, one has to answer two basic questions.

1. Is there a crystallographic icosahedral encasing form for the rhinovirus?

2. Does one also have an icosahedral scaling relation between the indexed external envelope and the boundary of the central hole of the capsid?

The first question is due to the fact that the $T = 3$ icosadeltahedron is not crystallographic. It can be viewed as a regular dodecahedron with pentagonal pyramids on each face. The 32 vertices of this polyhedron do not have integral indices with respect to a single set of six basis vectors. The second question arises because of the planar characterization of the capsid. It only involves the three major coat proteins VP1, VP2 and VP3 and does not include the fourth protein VP4 of the protomer, which also plays a role in delimiting the cavity filled by the RNA.

In the present paper, both questions get a positive answer. The crystallographic form enclosing the capsid appears to be a combination of an icosahedron with a regular dodecahedron obtained from the icosahedron by crystallographic scaling. The result is a polyhedron with 32 vertices denoted *icododecahedron*, deviating only slightly from the icosadeltahedron. The whole capsid is then delimited by two such icododecahedra in a radial scaling related by a factor τ.

But there is more. Again the paper of Caspar & Klug shows the conceptual way to follow. One finds on page 3 of Caspar & Klug (1962): *The essential point about grades of organization is that large structures are built out of smaller structures*. Moreover, on page 4: *The structure of simple viruses is principally determined by an ordered packing of protein subunits*. The subunits are obtained by clustering of monomers, say in pentamers, hexamers, trimers, dimers and so on, forming morphological units invariant with respect to subgroups of the icosahedral group, like 5, 52, 3, 32 and 222. In this way, one is back to axial-symmetric biomacromolecules. Indeed, for these clusters, one finds again indexed molecular forms with the expected properties: polygrammal scaling and lattices with integral metric (*integral lattices*), in particular. This applies for each of the four coat proteins and the five serotypes considered. One gets in such a way an overwhelming number of crystallographic forms, impossible to present adequately in a single paper. In addition to a numerical characterization of these forms given in various tables, only some illustrative examples are presented in figures to allow the reader to get a feeling of the very sophisticated structural organization of rhinoviruses revealed by the existence of a whole set of crystallographic enclosing forms. At the level of the polypeptide chains, this ordering, implicitly required by the form confinement, is still hidden. While the serotype depends on the detailed structure of the chains, the present investigation shows that the enclosing forms and their organization is the same for all the five serotypes analyzed. This conserved property is true up to small differences owing to slight changes in the constant scaling factor relating the geometry with the physical structure, and to a more or less ideal fitting of the chains in their crystallographic forms. These facts explain the great structural similarity observed among different serotypes.

2. Structural data, geometry and notation

2.1. Indexed icosahedral forms

In addition to the orthonormal basis $e = \{e_1, e_2, e_3\}$, a symmetry-adapted icosahedral basis $a = \{a_1, \ldots, a_6\}$ is defined with vectors pointing to the six non-aligned vertices of an icosahedron. The components of a vector in the bases e and a are indicated in round and in square brackets, respectively:

$$r = (x, y, z) = [n_1 n_2 n_3 n_4 n_5 n_6]. \tag{2}$$

With this convention, the *icosahedral basis* is given by

$$\begin{aligned} a_1 &= (1, 0, \tau) = [100000], & a_2 &= (\tau, 1, 0) = [010000], \\ a_3 &= (0, \tau, 1) = [001000], & a_4 &= (-1, 0, \tau) = [000100], \\ a_5 &= (0, -\tau, 1) = [000010], & a_6 &= (\tau, -1, 0) = [000001]. \end{aligned} \tag{3}$$

A visualization is easy when the components are given in the orthonormal basis, the crystallographic character becomes explicit only in the icosahedral basis. The *indices* of a point P are the components $n_1, n_2, \ldots, n_6$ of the position vector r_P. The six vectors a_i are linearly dependent on the reals, but they are linearly independent of the rationals. It implies that only rational indices are uniquely determined. Only then does a six-dimensional description in the three-dimensional space make sense. Note that the properties of rational indices are the same as for integral indices as the two sets only differ by a constant integer. In the basis a, the vertices I_k of the icosahedron have indices which are permutations of $[\pm 100000]$. These sets of indices are obtained by applying to one of them the icosahedral group.

The icosahedral group 532 is defined by

$$532 = \{\alpha, \beta \,|\, \alpha^5 = \beta^3 = (\beta\alpha)^2 = 1\}. \tag{4}$$

In the orthonormal basis e, the two generators α and β are represented by the rotation matrices $R_5(e)$ and $R_3(e)$, respectively:

$$R_5(e) = \tfrac{1}{2}\begin{pmatrix} 1 & -\tau & \tau-1 \\ \tau & \tau-1 & -1 \\ \tau-1 & 1 & \tau \end{pmatrix},$$
$$R_3(e) = \tfrac{1}{2}\begin{pmatrix} \tau & 1-\tau & 1 \\ \tau-1 & -1 & -\tau \\ 1 & \tau & 1-\tau \end{pmatrix}. \quad (5)$$

In the basis a, the same rotations are given by integral matrices:

$$R_5(a) = \begin{pmatrix} 1&0&0&0&0&0 \\ 0&0&0&0&0&1 \\ 0&1&0&0&0&0 \\ 0&0&1&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&0&0&1&0 \end{pmatrix}, \quad (6)$$

$$R_3(a) = \begin{pmatrix} 0&0&1&0&0&0 \\ 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&\bar{1}&0 \\ 0&0&0&0&0&\bar{1} \\ 0&0&0&1&0&0 \end{pmatrix}. \quad (7)$$

The 20 threefold axes of the icosahedron allow one to define the vertices of a regular dodecahedron. In particular, by applying the icosahedral group to the vector $a_1 + a_2 + a_3$, one finds the indices of dodecahedral vertices:

$$\tau^2(1,1,1) = [111000],\ \ \tau(0,1,\tau^2) = [101100], \ldots,$$
$$\tau(-\tau^2,0,1) = [0\bar{1}010\bar{1}]. \quad (8)$$

Scaling this dodecahedron by a factor $1/\tau^2$, one gets a new dodecahedron with vertices D_k:

$$(1,1,1) = \tfrac{1}{2}[111\bar{1}\bar{1}\bar{1}],\ \ (0,\tfrac{1}{\tau},\tau) = \tfrac{1}{2}[1\bar{1}11\bar{1}1], \ldots,$$
$$\left(-\tau,0,\frac{1}{\tau}\right) = \tfrac{1}{2}[1\bar{1}\bar{1}1\bar{1}\bar{1}]. \quad (9)$$

The corresponding scaling transformation S_{1/τ^2} is given by the matrices:

$$S_{1/\tau^2}(e) = \begin{pmatrix} 1/\tau^2 & 0 & 0 \\ 0 & 1/\tau^2 & 0 \\ 0 & 0 & 1/\tau^2 \end{pmatrix},$$
$$S_{1/\tau^2}(a) = \tfrac{1}{2}\begin{pmatrix} 3&-1&-1&-1&-1&-1 \\ -1&3&-1&1&1&-1 \\ -1&-1&3&-1&1&1 \\ -1&1&-1&3&-1&1 \\ -1&1&1&-1&3&-1 \\ -1&-1&1&1&-1&3 \end{pmatrix}. \quad (10)$$

By applying $S_{1/\tau^2}(a)$ to the indices of the initial dodecahedron, one indeed gets the indices of the rescaled dodecahedron indicated in (9).

The ico-dodecahedron mentioned in the *Introduction* is the combined polyhedron with 32 vertices: 12 of the icosahedron and 20 of the rescaled dodecahedron. As the icosadeltahedron of Caspar & Klug, it is a regular dodecahedron with pentagonal pyramids on each face. The resulting 60 triangular facets are not equilateral, but the deviation is small. In Fig. 1, the ico-dodecahedron is shown viewed along a fivefold axis, a threefold axis and a twofold axis, respectively.

2.2. Rhinoviruses

The starting point is the chains A_0, B_0, C_0, D_0 of the polypeptide C^α backbone for the coat proteins VP1, VP2, VP3, VP4, respectively, of human rhinovirus (HRV) taken from the data in the Brookhaven Protein Data Bank for various serotypes.

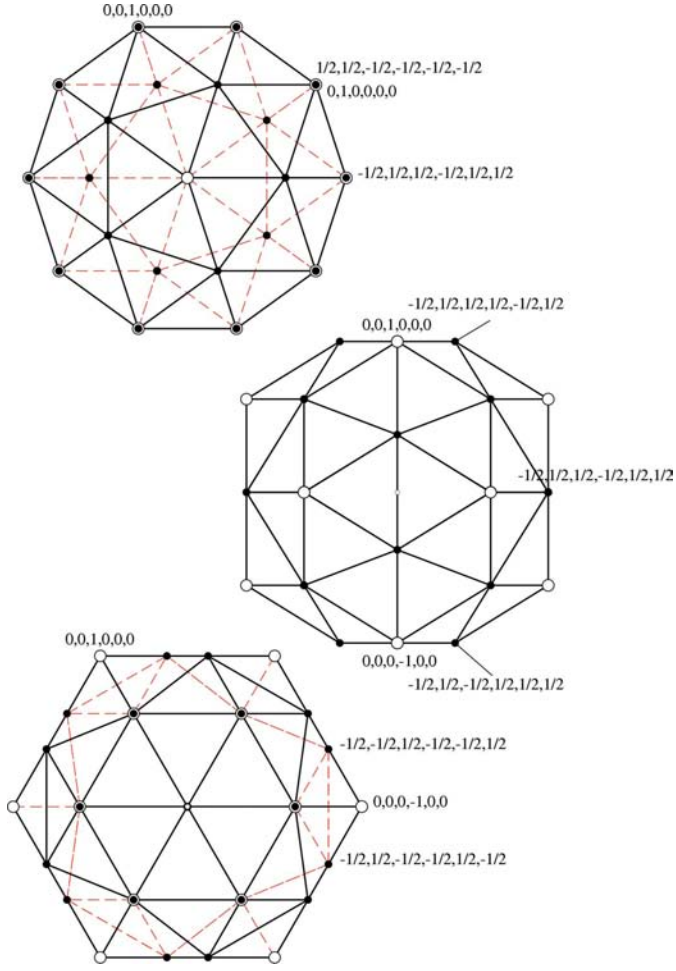

Figure 1
The ico-dodecahedron, a combination of an icosahedron and a dodecahedron, is shown along a fivefold axis (upper part), a twofold axis (central part) and a threefold axis (lower part), respectively. It has 60 triangular facets and 32 vertices. The 12 icosahedral vertices (large circles) have integral indices at points of an icosahedral lattice. The 20 dodecahedral vertices (filled dots) are at body-centered lattice positions and have half-integer indices. The ico-dodecahedron deviates only slightly from the $T = 3$ icosadeltahedron of Caspar & Klug, whose vertices do not have rational indices.

The crystal structure of serotype 1A (HRV1A, PDB 1r1a) was determined by Kim *et al.* (1989); that of serotype 14 (HRV14, PDB 4rhv) by Arnold & Rossmann (1990); that of serotype 3 (HRV3, PDB 1rhi) by Zhao *et al.* (1996); that of serotype 16 (HRV16, PDB 1aym) by Hadfield *et al.* (1997); and finally that of serotype 2 (HRV2, PDB 1fpn) by Verdaguer *et al.* (2000).

The capsid of the virion follows by applying to the Cartesian coordinates (in Å) of these starting chains the 60 elements of the icosahedral group in the matrix representation given in the PDB files. Successive transformations number the chains accordingly. This numbering is the same for the data of the serotypes 1A, 3, 14 and 16. For serotype 2, it requires a relabeling of the axes by the 90° rotation $x \rightarrow y, y \rightarrow -x$ around the z axis and the same set of matrices as for the other serotypes instead of the one indicated in 1fpn.

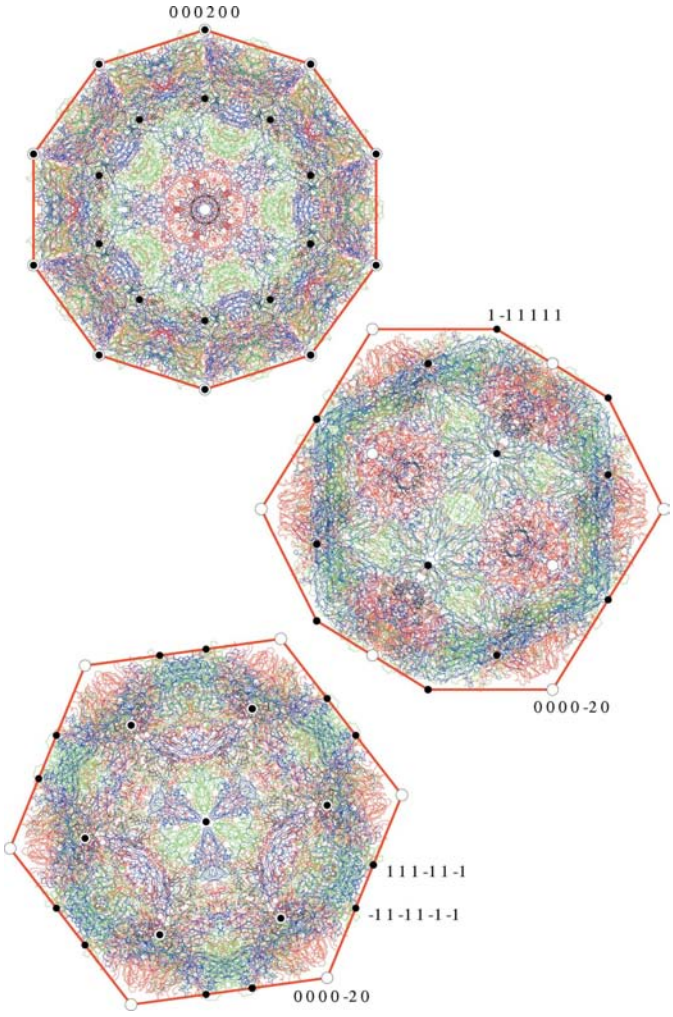

Figure 2
The projected views of the human rhinovirus along the fivefold, the twofold and the threefold axis, respectively, show that the virion is encapsulated into an ico-dodecahedral form with integral indices: even for the icosahedral vertices (large circles) and odd for the dodecahedral vertices (filled dots). Plotted are the backbone chains of the coat proteins: VP1 (red), VP2 (green), VP3 (blue) and VP4 (black) of the serotype 16 (HRV16). This is also the case for the serotypes 1A, 2, 3 and 14. (The orientation and the basis adopted here are not the same as in Fig. 1.)

One gets in this way a unique numbering of the rotational elements R_n of the icosahedral group and of the 240 chains of the capsid for all the serotypes considered. This makes a comparison easier. The crystallographic polyhedral form of the virion is an ico-dodecahedron, as shown in Fig. 2 for the serotype 16 (HRV16). The same is true for the other serotypes. A unit of length a_0 relates geometry with the physical structure and plays the role of an icosahedral lattice parameter. Taking this parameter into account, the icosahedral basis (3) becomes

$$a_1 = a_0(1, 0, \tau) = [100000], \ldots, a_6 = a_0(\tau, -1, 0) = [000001]. \tag{11}$$

For the rhinovirus, the value of a_0 depends on the serotype and also on the four coat proteins of the protomer, but is always about 90 Å, with possible deviations of a few per cent. In Fig. 2, the lattice parameter $a_0 = 90$ Å is used for HRV16.

2.3. Polyhedral form of the capsid

The property found in axial-symmetric proteins (Janner, 2005*a*,*b*,*c*) consisting in a crystallographic scaling relation between central hole and envelope also applies to the two ico-dodecahedra enclosing the capsid of the rhinovirus. It is the simplest scaling relation one possibly expects for an icosahedral virus: a radial scaling by a factor τ, which is indeed crystallographic:

$$S_\tau(e) = \begin{pmatrix} \tau & 0 & 0 \\ 0 & \tau & 0 \\ 0 & 0 & \tau \end{pmatrix},$$
$$S_\tau(a) = \tfrac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \end{pmatrix}. \tag{12}$$

S_τ is integral when expressed in a body-centered icosahedral basis. By applying $S_{1/\tau} = S_\tau^{-1}$ to the indices of the 12 icosahedral vertices I_k, one gets the corresponding ones $J_k = S_\tau^{-1} I_k$ of a $1/\tau$-scaled icosahedron:

$$a_0\left(\frac{1}{\tau}, 0, 1\right) = \tfrac{1}{2}[\bar{1}111\bar{1}1], \quad a_0(1, \tfrac{1}{\tau}, 0) = \tfrac{1}{2}[1\bar{1}1\bar{1}\bar{1}1], \ldots,$$
$$a_0\left(-1, \frac{1}{\tau}, 0\right) = \tfrac{1}{2}[\bar{1}\bar{1}11\bar{1}1]. \tag{13}$$

In a way similar to the 20 dodecahedral vertices D_i of the ico-dodecahedron [exemplified in (9) up to the lattice parameter a_0], one finds the indices of the dodecahedral vertices $F_i = S_\tau^{-1} D_i$ of the $1/\tau$-scaled ico-dodecahedron:

$$a_0\left(\frac{1}{\tau}, \frac{1}{\tau}, \frac{1}{\tau}\right) = [0001\bar{1}1], \quad a_0\left(0, \frac{1}{\tau^2}, 1\right) = [01001\bar{1}], \ldots,$$
$$a_0\left(-1, 0, \frac{1}{\tau^2}\right) = [\bar{1}01010]. \tag{14}$$

The molecular form of the capsid consists of these two icododecahedra: one enclosing the external surface of the capsid and one a factor $1/\tau$ smaller for the internal surface delimiting the RNA cavity. This is shown in Fig. 3 for the capsid of HRV3 viewed along the fivefold, threefold and twofold axes, respectively. In this figure, only the coat proteins are plotted which are situated in the corresponding equatorial regions.

3. Orthorhombic, tetragonal and cubic forms for clusters with 222 symmetry

A capsid with the icosahedral point-group symmetry 532 is also invariant with respect to one of the subgroups. By applying the subgroup to one of the 60 monomers of the four coat proteins, one gets a cluster that has the symmetry of the subgroup. For the axial subgroups, one expects cluster forms with crystallographic properties (Janner, 2005*a*,*b*,*c*). The easiest start in the verification of this expectation is to consider the subgroup 222, generated by the twofold rotations 2_x, 2_y, 2_z around the axes x, y, z, respectively. The point group 222 is of order 4, so that one gets 15 different clusters for each coat protein, with chains indicated by their number, fixed by the conventions introduced in the previous section. For example, the cluster $\{0, 32, 37, 46\}$ of the VP1 protein consists of the chains:

$$\begin{aligned} \text{VP1}[0] &= A_0, & \text{VP1}[32] &= 2_y A_0, \\ \text{VP1}[37] &= 2_x A_0, & \text{VP1}[46] &= 2_z A_0. \end{aligned} \tag{15}$$

For each cluster, the vertices of the various encasing polyhedra are at points of a lattice left invariant by the 222 point group. The parameters a, b, c of this lattice are in an integral relation with the half-edge $a_c = \tau a_0$ of the cube in which the ico-

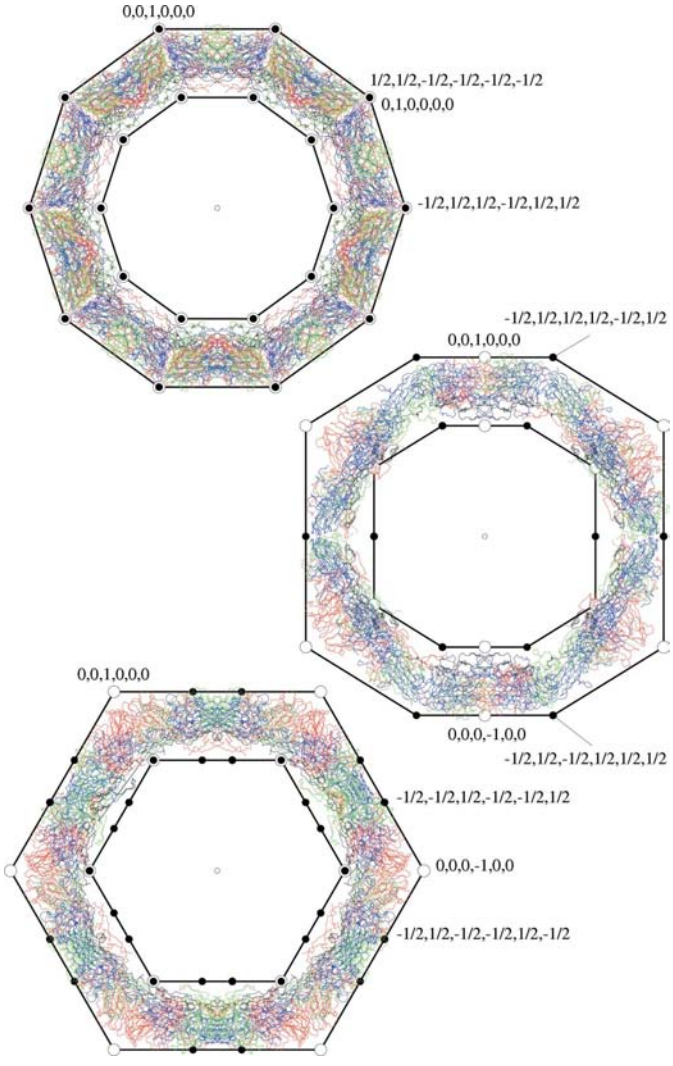

Figure 3
The capsid of the human rhinovirus is encapsulated between two icododecahedra (in the same orientation as in Fig. 1): an external one and an internal one isotropically scaled from the envelope by a factor $1/\tau$, with $\tau = (1 + \sqrt{5})/2$, the golden number. This scaling transformation is crystallographic as it leaves the body-centered icosahedral lattice invariant. Only the vertices and the coat proteins situated in the various equatorial regions are plotted in a way similar to that in Fig. 2.

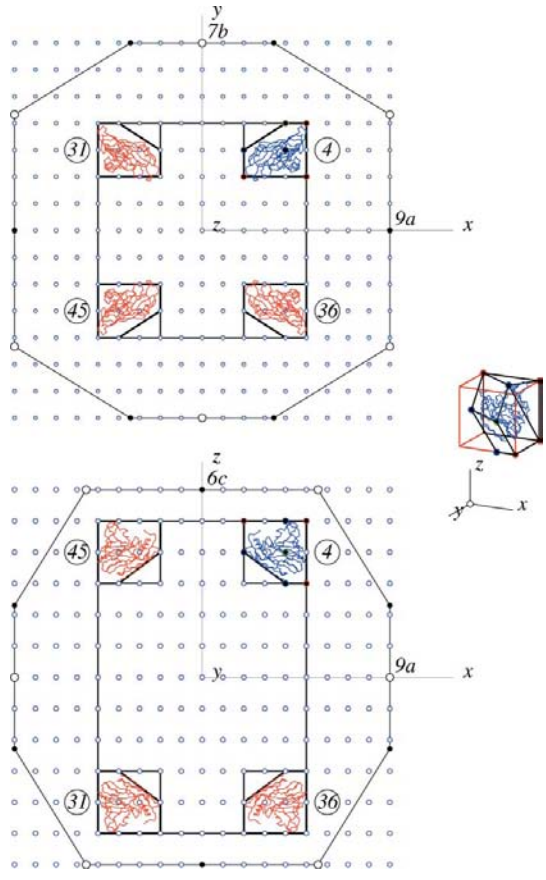

Figure 4
The tetrameric cluster $\{4,31,36,45\}$ of the coat protein VP2 has symmetry 222. It is enclosed in a form with vertices at points of the orthorhombic lattice with parameters a, b, c satisfying the relations $9a = 7b = 6c = a_c$, where $2a_c$ is the edge of the cube encasing the capsid. The boundary of the capsid is shown in projection along the z axis (upper part) and the y axis (lower part), respectively. The prismatic envelope of the tetramer has vertices that also delimit the four monomers. The refined molecular form of each monomer is a polyhedron with eight vertices, shown for VP2[4] in a perspective view at the right-hand side. The indices of these vertices are given in the text. Plotted are the chains for the serotype 3 (HRV3).

dodecahedron is inscribed, with a_0 the icosahedral lattice parameter:

$$a_c = \tau a_0 = z_1 a = z_2 b = z_3 c, \quad \text{integers } z_1, z_2, z_3. \qquad (16)$$

The triple of integers z_1, z_2, z_3 can be used for indicating the corresponding lattice, which is integral because the metric tensor of the basis vectors is proportional to one with integral entries. The form lattices for the orthorhombic clusters of all the coat proteins for the five rhinovirus serotypes have been determined. The result, valid for all the serotypes, is summarized in Table 1.

In the case of VP4, the forms indicated are those of HRV3 and HRV14, for which 43 and 40 residues, respectively, of the total of 68 are visible. In the other cases, these forms have an indicative character only, because too many residues are missing: visible are 19 for HRV1A, 25 for HRV2 and 29 for HRV16. As an illustrative example for the orthorhombic lattice (976), the enclosing forms of the VP2 cluster {4,31,36,45} is shown in Fig. 4 with views along the 2_z and 2_y rotational axes. For the tetragonal lattice (988), the forms of the VP3 cluster {10,15,29,43} are plotted in Fig. 5 in a similar way. Finally, the three combined tetrameric clusters {1,33,38,47}, {7,21,53,58}, {13,18,27,41} of VP1, which share the same cubic lattice (777), are shown in Fig. 6. The chains plotted in these figures belong to different serotypes. As already mentioned, the forms are correspondingly the same for the other serotypes.

From these figures, by counting the number of lattice points from the origin to the various vertices, one deduces the indices of the prismatic form enclosing the tetramer. For example, looking at Fig. 4, one finds that the eight vertices of the prismatic envelope have indices $[\pm 5\ \pm 4\ \pm 5]$, whereas the

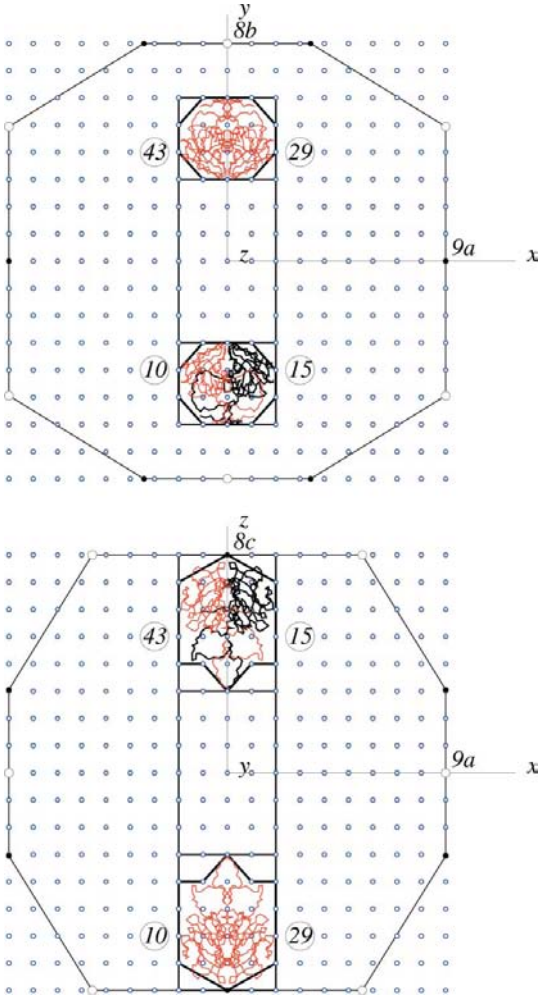

Figure 5
The tetramer {10,15,19,43} of the VP3 coat protein with 222 symmetry has in fact a tetragonal form lattice with parameters satisfying the relation $9a = 8b = 8c = a_c$. In the projection along the z axis (upper part) and the y axis (lower part), the enclosing form of the tetramer splits into two dimeric ones, with vertices at the tetragonal lattice points (one monomer is enhanced). This is shown in a prismatic approximation and for a refined more complex polyhedral form. The result, plotted for HRV16, does not depend on the serotype.

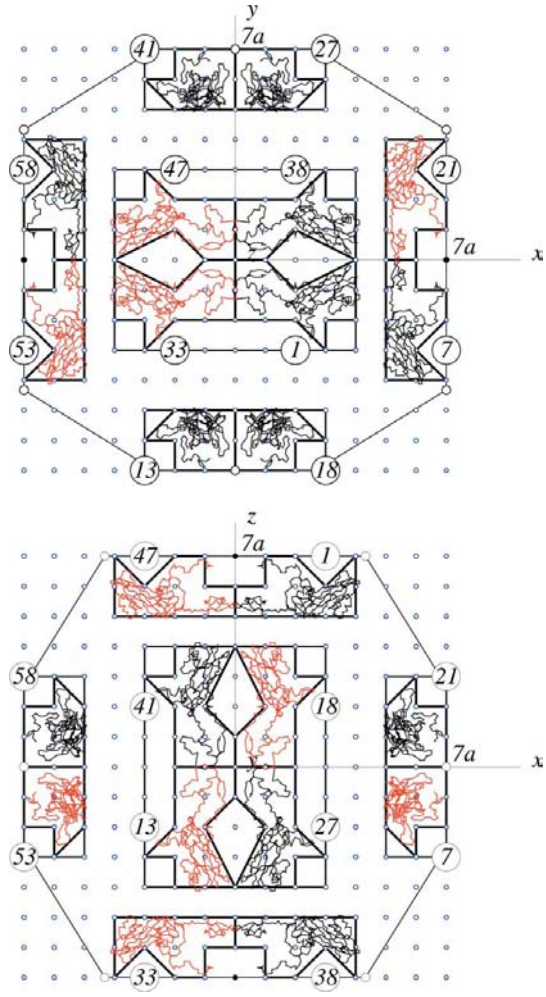

Figure 6
The three tetrameric clusters {1,33,38,47}, {7,21,53,58}, {13,18,27,41} of the VP1 coat protein share the same cubic form lattice $a = a_c/7$, with a_c the half-edge of the cube encasing the capsid. In the projected views along the z and the y axis, respectively, are shown the boundaries of the capsid, those of the various prismatic forms and the more complex refined polyhedral forms enclosing the chains involved. From the enhanced monomers, one sees that these forms allow a monomeric characterization. The result is independent of the serotype 14 plotted in this figure.

Table 1
Orthorhombic, tetragonal and cubic form lattices for clusters with 222 symmetry of the rhinovirus coat proteins ($\tau a_0 = z_1 a = z_2 b = z_3 c$ for a_0 icosahedral and a, b, c orthorhombic lattice parameters, respectively).

222 clusters	HRV1A, HRV2, HRV3, HRV14, HRV16									HRV3, HRV14		
	VP1			VP2			VP3			VP4		
Tetramers	z_1	z_2	z_3	z_1	z_2	z_3	z_1	z_2	z_3	z_1	z_2	z_3
{0,32,37,46}	6	6	6	8	6	7	9	9	9	8	7	9
{1,33,38,47}	7	7	7	6	9	11	6	6	6	7	7	9
{2,34,39,48}	9	9	9	8	7	7	9	9	9	8	9	8
{3,30,35,49}	6	9	9	11	7	9	8	9	8	9	9	9
{4,31,36,45}	6	6	6	9	7	6	6	6	6	6	9	10
{6,20,52,57}	6	6	6	7	8	6	9	9	9	9	8	7
{7,21,53,58}	7	7	7	11	6	9	6	6	6	9	7	7
{8,22,54,59}	9	9	9	7	8	7	9	9	9	8	8	9
{9,23,50,55}	9	6	9	9	11	7	8	8	9	9	9	9
{5,24,51,56}	6	6	6	6	9	7	6	6	6	10	6	9
{12,17,26,40}	6	6	6	6	7	8	9	9	9	7	9	8
{13,18,27,41}	7	7	7	9	11	6	6	6	6	7	9	7
{14,19,28,42}	9	9	9	7	7	8	9	9	9	9	8	8
{10,15,29,43}	9	9	6	7	9	11	9	8	8	9	9	9
{11,16,25,44}	6	6	6	7	6	9	6	6	6	9	10	6

vertices of the cross-like hole have indices [$\pm5 \pm2 \pm3$], [$\pm2 \pm2 \pm3$], [$\pm2 \pm4 \pm5$]. The form of this tetramer splits into four forms for the monomers. In particular, the chain VP2[4] is encased in a rectangular prism with vertices at [223], [623], [543], [423], [225], [625], [545], [425]. It is less easy to find the indices of the polyhedron of the refined enclosing form of the same chain. One actually needs to consider all the three projections along the x, y and z axes. For the eight vertices of the refined polyhedron, shown in a perspective view in Fig. 4, one finds the indices: [225], [525], [435], [234], [444], [443], [543], [523].

The situation presented in Fig. 5 is similar for the VP3 tetramer {10,15,19,43}, which has a tetragonal form lattice with $z_1 = 9$, $z_2 = z_3 = 8$. In the case shown in Fig. 6, the form lattice is cubic with $z_1 = z_2 = z_3 = 7$. This allows us to group 12 chains of the VP1 coat protein in the same plot. The prismatic boundaries split into those of the three tetramers {1,33,38,47}, {7,21,53,58} and {13,18,27,41}. The refined enclosing forms of each of the monomers involved are complex polyhedra with vertices at points of the cubic lattice (777).

4. Hexagonal forms for clusters with 32 symmetry

Considered first is subgroup 3 of the icosahedral group. The corresponding clusters are trimers of a coat protein lying around the threefold axes. Two are the crystallographic properties expected from Janner (2005*a*,*b*,*c*) for trimers with axial symmetry 3:

1. an enclosing form with vertices at points of a hexagonal lattice Λ_h;

2. a rational value for the square of the axial ratio $c/a = \gamma$ of the hexagonal lattice parameters a and c: $\gamma^2 = p/q$ for integers p and q.

If this is true for the trimers, the corresponding hexamers with symmetry 32 (obtained by applying to the trimers a twofold rotation perpendicular to the trigonal axis) share the same hexagonal form lattice. It is, therefore, convenient to consider the two expected properties together for a hexamer and for its trimeric components.

For each coat protein, the 10 hexamers and the 20 trimers are specified by the chains numbered in the same way as in the previous two sections. For example, for VP2, the cluster indicated by {0,9,22; 32,50,59} consists of the two trimers VP2[0], VP2[9], VP2[22] and VP2[32], VP2[50], VP2[59], respectively, obtained by applying the appropriate subgroup 32 to the starting chain $B_0 =$ VP2[0].

The basis vectors b_1, b_2, b_3 of the hexagonal lattice Λ_h are chosen in the standard way: $|b_1| = |b_2| = a$, $|b_3| = c$ and scalar products $b_1 \cdot b_2 = -\frac{1}{2}a^2$, $b_1 \cdot b_3 = b_2 \cdot b_3 = 0$, with a and c the hexagonal lattice parameters. The corresponding metric tensor $g_h(a, \gamma)$ is given by

$$g_h(a, \gamma) = a^2 \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & \gamma^2 \end{pmatrix}, \qquad (17)$$

with $\gamma = c/a$ the axial ratio. It is convenient to denote the two indices n_1, n_2 as *planar indices* and the third one n_3 as the *axial index*.

In order to verify the validity of the two crystallographic properties 1 and 2 for the hexamers (and for the trimers), one has to fix the lattice vectors b_1, b_2, b_3 with respect to the ico-dodecahedron of the capsid indicated above.

The axial direction (along b_3) is fixed by choosing the threefold axis associated with one of the 20 dodecahedral vertices. The projection of the ico-dodecahedron along the threefold axis is a regular hexagon with radius r_h delimited by the axial projection of six icosahedral vertices (see Fig. 1). There is, therefore, a relation between r_h and the icosahedral lattice parameter a_0. One finds $r_h = (2\tau/\sqrt{3})a_0$.

The projected hexagon and the lattice Λ_h are both invariant with respect to the same subgroup 3 of 532. Therefore, the

Table 2
Hexagonal form lattices for hexameric clusters with symmetry 32 (and trimeric clusters with symmetry 3) of the rhinovirus coat proteins VP1 and VP2.

$r_h = (2\tau/\sqrt{3})a_0$, with r_h the hexagonal radius of the capsid projected along the threefold axis and a_0 the icosahedral lattice parameter.

32 clusters hexamers	Hexagonal radius r_h	Hexamer height H	Trimer height h	Axial ratio c/a	Lattice $\mathbb{Q}$ class
Coat protein VP1 of HRV1A, HRV2, HRV3, HRV14 and HRV16					
{0,9,22; 32,50,59}	$6a$	$10c = \sqrt{2}r_h$	$3c$	$\frac{3\sqrt{2}}{5}$	$\sqrt{2}$-$\Lambda_{\rm hex}$
{1,5,23; 33,51,55}	$6a$	$6c = \frac{3}{2}r_h$	$2c$	$\frac{3}{2}$	1-$\Lambda_{\rm hex}$
{2,6,24; 34,52,56}	$3a$	$12c = \sqrt{3}r_h$	$3c$	$\sqrt{3}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{3,7,20; 30,53,57}	$10a$	$6c = \sqrt{3}r_h$	c	$\frac{5\sqrt{3}}{3}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{4,8,21; 31,54,58}	$9a$	$6c = \frac{3}{2}r_h$	c	$\frac{9}{4}$	1-$\Lambda_{\rm hex}$
{10,26,45; 15,36,40}	$3a$	$2c = \frac{1}{\sqrt{2}}r_h$	c	$\frac{3\sqrt{2}}{4}$	$\sqrt{2}$-$\Lambda_{\rm hex}$
{11,27,46; 16,37,41}	$10a$	$8c = \frac{2}{5}r_h$	$7c$	$\frac{1}{2}$	1-$\Lambda_{\rm hex}$
{12,28,47; 17,38,42}	$6a$	$10c = \frac{1}{2}r_h$	$9c$	$\frac{3}{10}$	1-$\Lambda_{\rm hex}$
{13,29,48; 18,39,43}	$7a$	$12c = \frac{6}{7}r_h$	$5c$	$\frac{1}{2}$	1-$\Lambda_{\rm hex}$
{14,25,49; 19,35,44}	$4a$	$6c = r_h$	$2c$	$\frac{2}{3}$	1-$\Lambda_{\rm hex}$
Coat protein VP2 of HRV1A, HRV2, HRV3, HRV14 and HRV16					
{0,9,22; 32,50,59}	$6a$	$10c = r_h$	$3c$	$\frac{3}{5}$	1-$\Lambda_{\rm hex}$
{1,5,23; 33,51,55}	$5a$	$10c = \frac{2}{\sqrt{3}}r_h$	$3c$	$\frac{\sqrt{3}}{3}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{2,6,24; 34,52,56}	$8a$	$16c = \frac{8}{5}r_h$	$3c$	$\frac{4}{5}$	1-$\Lambda_{\rm hex}$
{3,7,20; 30,53,57}	$3a$	$6c = \frac{9}{5}r_h$	c	$\frac{9}{10}$	1-$\Lambda_{\rm hex}$
{4,8,21; 31,54,58}	$4a$	$10c = \frac{4}{3}r_h$	$2c$	$\frac{8}{15}$	1-$\Lambda_{\rm hex}$
{10,26,45; 15,36,40}	$7a$	$10c = \frac{1}{2}r_h$	$6c$	$\frac{7}{20}$	1-$\Lambda_{\rm hex}$
{11,27,46; 16,37,41}	$10a$	$6c = \frac{3}{5}r_h$	$2c$	1	1-$\Lambda_{\rm hex}$
{12,28,47; 17,38,42}	$8a$	$6c = \frac{1}{2}r_h$	$4c$	$\frac{2}{3}$	1-$\Lambda_{\rm hex}$
{13,29,48; 18,39,43}	$8a$	$4c = \frac{2\sqrt{2}}{3}r_h$	c	$\frac{4\sqrt{2}}{3}$	$\sqrt{2}$-$\Lambda_{\rm hex}$
{14,25,49; 19,35,44}	$6a$	$4c = \frac{5}{4}r_h$	c	$\frac{15}{8}$	1-$\Lambda_{\rm hex}$

Table 3
Hexagonal form lattices for hexameric clusters with symmetry 32 (and trimeric clusters with symmetry 3) of the rhinovirus coat proteins VP3 and VP4.

$r_h = (2\tau/\sqrt{3})a_0$, with r_h the hexagonal radius of the capsid projected along the threefold axis and a_0 the icosahedral lattice parameter.

32 clusters hexamers	Hexagonal radius r_h	Hexamer height H	Trimer height h	Axial ratio c/a	Lattice $\mathbb{Q}$ class
Coat protein VP3 of HRV1A, HRV2, HRV3, HRV14 and HRV16					
{0,9,22; 32,50,59}	$6a$	$4c = \frac{2}{\sqrt{3}}r_h$	c	$\sqrt{3}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{1,5,23; 33,51,55}	$6a$	$4c = \frac{2}{\sqrt{3}}r_h$	c	$\sqrt{3}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{2,6,24; 34,52,56}	$10a$	$6c = \frac{3}{2}r_h$	c	$\frac{5}{2}$	1-$\Lambda_{\rm hex}$
{3,7,20; 30,53,57}	$6a$	$10c = \sqrt{3}r_h$	$2c$	$\frac{3\sqrt{3}}{5}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{4,8,21; 31,54,58}	$3a$	$10c = \frac{3}{2}r_h$	$2c$	$\frac{9}{20}$	1-$\Lambda_{\rm hex}$
{10,26,45; 15,36,40}	$3a$	$8c = \frac{4}{5}r_h$	$4c$	$\frac{3}{10}$	1-$\Lambda_{\rm hex}$
{11,27,46; 16,37,41}	$6a$	$6c = \frac{1}{2}r_h$	$5c$	$\frac{1}{2}$	1-$\Lambda_{\rm hex}$
{12,28,47; 17,38,42}	$6a$	$6c = \frac{1}{\sqrt{3}}r_h$	$5c$	$\frac{\sqrt{3}}{3}$	$\sqrt{3}$-$\Lambda_{\rm hex}$
{13,29,48; 18,39,43}	$3a$	$12c = \frac{3}{5}r_h$	$5c$	$\frac{3}{20}$	1-$\Lambda_{\rm hex}$
{14,25,49; 19,35,44}	$10a$	$16c = \frac{32}{25}r_h$	$7c$	$\frac{4}{5}$	1-$\Lambda_{\rm hex}$
Coat protein VP4 of HRV3 and HRV14					
{0,9,22; 32,50,59}	$10a$	$4c = r_h$	c	$\frac{5}{2}$	1-$\Lambda_{\rm hex}$
{1,5,23; 33,51,55}	$7a$	$4c = r_h$	c	$\frac{7}{4}$	1-$\Lambda_{\rm hex}$
{2,6,24; 34,52,56}	$6a$	$12c = \frac{6\sqrt{2}}{7}r_h$	c	$\frac{3\sqrt{2}}{7}$	$\sqrt{2}$-$\Lambda_{\rm hex}$
{3,7,20; 30,53,57}	$5a$	$8c = \sqrt{2}r_h$	c	$\frac{5\sqrt{2}}{8}$	$\sqrt{2}$-$\Lambda_{\rm hex}$
{4,8,21; 31,54,58}	$8a$	$14c = \frac{7\sqrt{2}}{9}r_h$	c	$\frac{4\sqrt{2}}{9}$	$\sqrt{2}$-$\Lambda_{\rm hex}$
{10,26,45; 15,36,40}	$9a$	$8c = \frac{8}{15}r_h$	$3c$	$\frac{3}{5}$	1-$\Lambda_{\rm hex}$
{11,27,46; 16,37,41}	$10a$	$6c = \frac{2}{5}r_h$	$4c$	$\frac{2}{3}$	1-$\Lambda_{\rm hex}$
{12,28,47; 17,38,42}	$10a$	$6c = \frac{2}{5}r_h$	$5c$	$\frac{2}{3}$	1-$\Lambda_{\rm hex}$
{13,29,48; 18,39,43}	$9a$	$6c = \frac{6}{11}r_h$	$2c$	$\frac{9}{11}$	1-$\Lambda_{\rm hex}$
{14,25,49; 19,35,44}	$8a$	$6c = r_h$	$2c$	$\frac{4}{3}$	1-$\Lambda_{\rm hex}$

basis vectors b_1, b_2 have the same orientation as the vectors B_1, B_2 which point to two vertices of the hexagon. It appears that for the form lattices of the hexamers the length of b_1 (equal to the length of b_2) is in an integral relation with the hexagonal radius r_h of the capsid. This is a new remarkable property valid for all the serotypes considered of the rhinovirus. Note that an analogous property has already been found for the tetramers with respect to the projections along the twofold rotational axes 2_x, 2_y, 2_z. In the threefold case, one has

$$na = r_h = \frac{2\tau}{\sqrt{3}}a_0 = 1.868\ldots a_0, \quad \text{integer } n, \tag{18}$$

so that for given a_0 the parameters n and γ can be used for characterizing the lattice Λ_h.

The description of the hexagonal enclosing forms of the hexameric chains, numbered as explained, requires a specific choice of the direction of the possible basis vectors b_1, b_2, b_3. To begin with, three vectors of the ico-dodecahedron are chosen: v_1, v_2, v_3 pointing to the icosahedral vertices at $a_0(1, 0, -\tau)$, $a_0(0, \tau, 1)$ and the dodecahedral vertex at $a_0(1, 0, 1/\tau^2)$, respectively. The axial direction is clearly given by v_3. The other two vectors v_1, v_2 are split into a parallel and a perpendicular component with respect to v_3. The parallel component is the same for the two vectors:

$$(v_1)_\parallel = (v_2)_\parallel = \tfrac{1}{3}v_3. \tag{13}$$

The perpendicular components define the vectors B_1, B_2 of the projected ico-dodecahedron:

$$\begin{aligned} B_1 &= (v_1)_\perp = \frac{2a_0}{3}(1, 0, -\tau^2), \\ B_2 &= (v_2)_\perp = \frac{a_0}{3}(-1, 3\tau, \tau^2). \end{aligned} \tag{20}$$

One then has indeed

$$\begin{aligned} v_1 &= (v_1)_\parallel + (v_1)_\perp = a_0(1, 0, -\tau), \\ v_2 &= (v_2)_\parallel + (v_2)_\perp = a_0(0, \tau, 1), \end{aligned} \tag{21}$$

and the required metrical relations

$$B_1^2 = B_2^2 = r_h^2, \quad B_1 \cdot B_2 = -\tfrac{1}{2}r_h^2. \tag{22}$$

Finally, the three basis vectors of the form lattice $\Lambda_h(a, c) = \Lambda_h(n, \gamma)$ are

$$\begin{aligned} b_1 &= \frac{B_1}{n} = \frac{a}{\sqrt{3}}\left(\frac{1}{\tau}, 0, -\tau\right), \quad b_2 = \frac{B_2}{n} = \frac{a}{2\sqrt{3}}\left(-\frac{1}{\tau}, 3, \tau\right), \\ b_3 &= \frac{\gamma a}{\sqrt{3}}\left(\tau, 0, \frac{1}{\tau}\right), \end{aligned} \tag{23}$$

with the relations between the parameters:

$$c = \gamma a, \quad a = \frac{2}{n\sqrt{3}}\tau a_0. \tag{24}$$

The results obtained for all the hexameric and trimeric clusters with symmetry 32 and 3, respectively, for all serotypes 1A, 2, 3,

14 and 16 of the rhinovirus are indicated in Table 2 for VP1 and VP2 and in Table 3 for VP3 and VP4, for VP4 only for the serotypes 3 and 14. From the data reported in these tables follows the validity of the two crystallographic properties 1 and 2 formulated above. In these tables, one also finds that the height H of the hexamer in the axial direction is a multiple of the parameter c and in a simple relation with the hexagonal radius of the projected capsid:

$$H = mc = \frac{m}{n}\gamma r_h. \tag{25}$$

Similar considerations apply to the height of the trimer, indicated by h, which is also the height of each monomer in the same direction.

A first example is illustrated in Fig. 7 by the hexameric cluster {10,26,45; 15,36,40} of VP1 and serotype HRV1A. One sees that the parameters a, c of the form lattice satisfy the relations: $3a = r_h, c = h, 2c = H = (1/\sqrt{2})r_h$. So $n = 3, m = 2$ and the axial ratio is $3\sqrt{2}/4$. This lattice is rationally equivalent to the integral lattice denoted $\sqrt{2}$-$\Lambda_{\rm hex}$, characterized by an axial ratio $\sqrt{2}$. [See Janner (2004) for more details.] The hexamer is enclosed between two hexagonal prisms having vertices with the indices $[30 \pm 1]$, $[33 \pm 1]$, ..., $[0 - 3 \pm 1]$ for the external envelope and indices $[20 \pm 1]$, $[22 \pm 1]$, ..., $[0 - 2 \pm 1]$ for the internal one, respectively. This form splits into three dimeric forms with integral indices. In particular, the form enclosing the dimer VP1[26] and VP1[40] has eight vertices indexed by $[33 \pm 1]$, $[03 \pm 1]$, $[02 \pm 1]$, $[22 \pm 1]$.

A second example is given in Fig. 8 for the cluster {3,7,20; 30,53,57} with 32 symmetry of VP4 in HRV3. The hexamer is enclosed in a prism with height $H = 8c = \sqrt{2}r_h$ and hexagonal basis with radius $2a$, where $r_h = 5a$. So $n = 5$, $m = 8$ and the axial ratio is $\gamma = 5\sqrt{2}/8$. The height of the trimer is $h = c$. The form lattice belongs again to the integral lattice class $\sqrt{2}$-$\Lambda_{\rm hex}$.

The third example is a form lattice that belongs to the isometric hexagonal class 1-$\Lambda_{\rm hex}$: one has indeed $H = r_h$. This case occurs for the cluster {0,9,22; 32,50,59} of VP2. The chains plotted in Fig. 9 are those of the rhinovirus HRV16. One has $r_h = 6a$, $H = 10c$ and $h = 3c$. The axial ratio

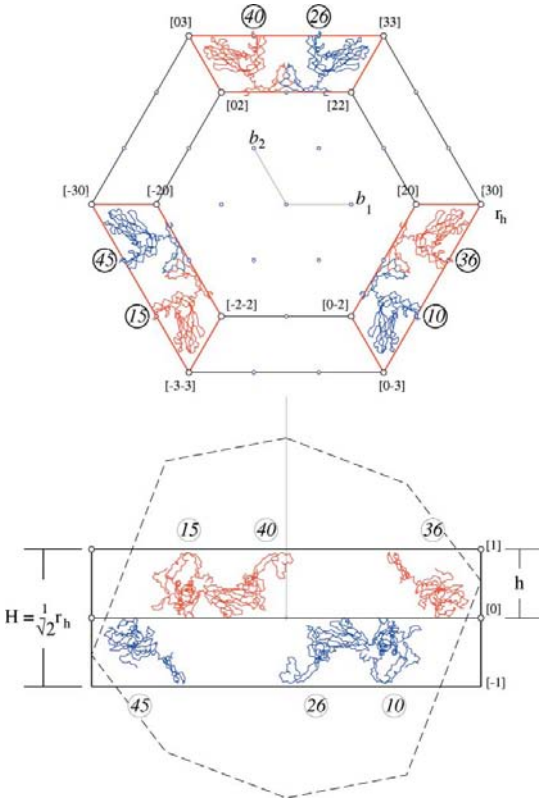

Figure 7
The form lattice for the hexameric cluster {10,26,45; 15,36,40} of VP1 and serotype HRV1A is an integral hexagonal lattice with parameters $a = \frac{1}{3}r_h$ and $c = (1/2\sqrt{2})r_h$, where r_h is the hexagonal radius of the capsid projected along the threefold axis and c is the height h of the monomer (and of the trimer), half the height H of the hexamer. The axial ratio is $\gamma = c/a = 3\sqrt{2}/4$. The planar and axial indices of the enclosing form are indicated in the projection along the threefold axis and perpendicular to it, respectively. Some of the monomers are plotted enhanced. The projected boundary of the capsid is indicated by dashed lines.

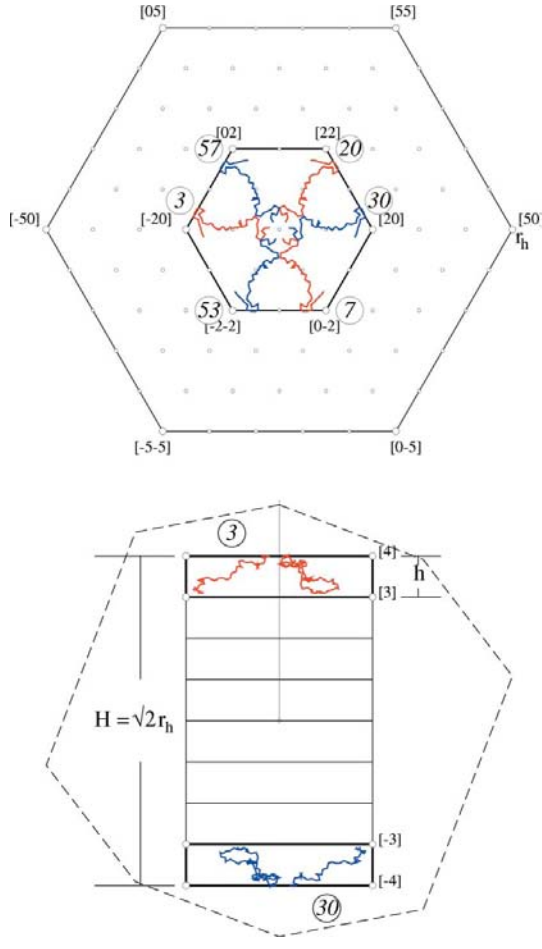

Figure 8
In a way similar to that in Fig. 7, the hexameric cluster {3,7,20; 30,53,57} of the VP4 coat protein, serotype 3, is enclosed in a prismatic form with vertices at points of a hexagonal lattice. One now has $a = b = \frac{1}{5}r_h$ and $c = (\sqrt{2}/8)r_h$. The axial ratio is $c/a = (5/8)\sqrt{2}$, so that the lattice is integral and rationally equivalent to $\sqrt{2}$-$\Lambda_{\rm hex}$. The height h of the monomer and of the trimer is c and that of the hexamer is $H = 8c$.

$c/a = \gamma = 3/5$ is rational, thus rationally equivalent to the ratio 1, as already pointed out. The vertices of the prismatic form enclosing the hexamer have the planar and the axial indices indicated in Fig. 9. One also easily finds the indices of the forms of the single monomers. For example, the chain VP2[22] is a prism with hexagonal indices given by [335], [535], [645], [455], [332], [532], [642], [452]. The last example, shown in Fig. 10, is the cluster {11,27,46; 16,37,41} of VP3, serotype 2. One now has: $r_h = 6a$, $H = 6c = \frac{1}{2}r_h$, $h = 5c$. The form lattice Λ_h has the axial ratio $\gamma = \frac{1}{2}$ and belongs again to the isometric hexagonal class 1-$\Lambda_{\rm hex}$. The planar and axial indices of the various enclosing forms are indicated in Fig. 10. In particular, the prism enclosing the chain VP3[27] has vertices with the hexagonal indices [343], [543], [563], [363], [34$\bar{2}$], [54$\bar{2}$], [56$\bar{2}$], [36$\bar{2}$].

The data reported in Tables 2 and 3 for the four coating proteins and all the hexameric and trimeric clusters with symmetry 32 and 3, respectively, allow one to verify the validity of the properties exemplified so far. These forms are not always the optimal ones for the single monomers.

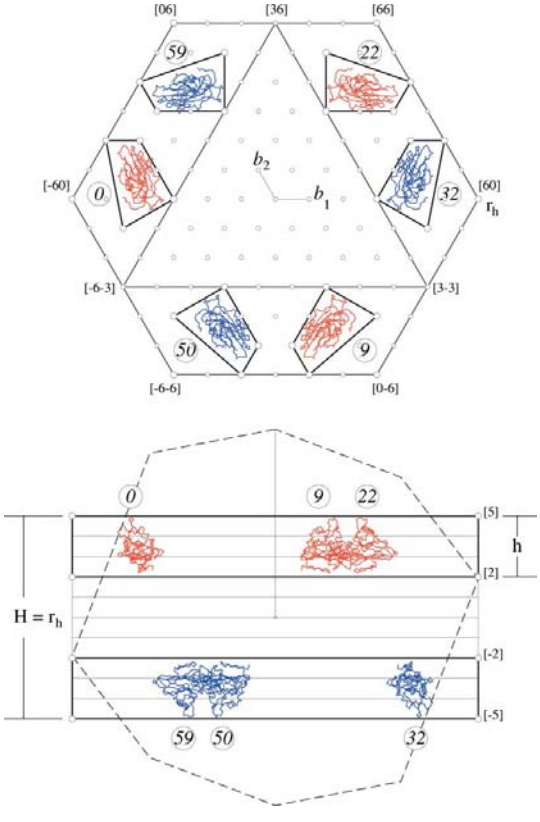

Figure 9
The height H of the VP2 hexamer {0,9,22; 32,50,59}, serotype 16, is equal to the radius r_h of the hexagonal enclosing form, the same as of the projected capsid along the threefold axis and of the central equilateral triangular basis. The planar and axial indices for the six monomers involved are expressed with respect to the form lattice $\Lambda_h(a, c)$ with $a = b = \frac{1}{6}r_h$ and $c = \frac{1}{10}r_h$. The axial ratio is $\gamma = \frac{3}{5}$, so that $\Lambda_h(a, c)$ is rationally equivalent to the isometric hexagonal lattice 1-$\Lambda_{\rm hex}$.

5. Pentagonal forms for clusters with 52 symmetry

For each coat protein, there are 6 decamers with symmetry 52 and 12 pentamers with symmetry 5 arranged around the fivefold axes. In a way similar to that in the previous sections, these clusters are specified by sets of numbers where a given number indicates the rotation yielding the chain from the starting one numbered as 0. In the present case, a decamer consists of two dyadically related pentamers labeled by five numbers, one for each chain, like in {5,17,22,41,49; 12,27,35,51,59}. The clusters are delimited by polyhedra with vertices at a point of a decagonal lattice Λ_d, with lattice parameters a, c and axial ratio $\gamma = c/a$, as in the previous hexagonal case. The difference is that a decagonal lattice requires five basis vectors: one along the fivefold axis and four in a plane perpendicular to it. This set of vectors is linearly independent of the rationals (not of the reals) and, therefore, only meaningful for positions with rational indices (as in the case of the icosahedral lattice). Moreover, the lattice

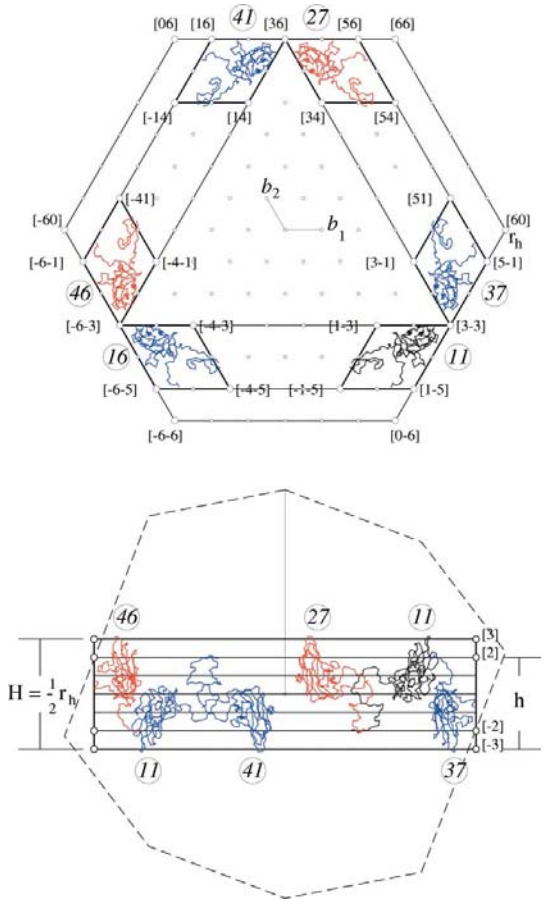

Figure 10
The VP3 hexamer {11,27,46; 16,37,41} of HRV2 is enclosed in a way similar to the hexamer of Fig. 9. The form splits into six prismatic monomeric ones with vertices at points of a hexagonal lattice, rationally equivalent to the isometric 1-$\Lambda_{\rm hex}$. The lattice parameters are $a = b = \frac{1}{6}r_h$ and $c = \frac{1}{12}r_h$, so that the axial ratio is $\frac{1}{2}$. The height of the hexamer is $H = \frac{1}{2}r_h$ and that of the monomer and of the trimer is $h = \frac{5}{12}r_h$. The planar and axial indices of the prismatic forms enclosing the monomers are indicated. The dashed line marks the projected boundary of the capsid.

points are dense in planes perpendicular to the fivefold axis, so that only positions with small indices are structurally relevant. A new strategy is required for getting the indices of the forms.

In the standard choice, the five decagonal basis vectors d_k of $\Lambda_d(a, c)$ have length and angles determined by the metrical conditions

$$\begin{aligned} |d_k| &= a, \quad d_k \cdot d_{k+1} = a^2 \cos\frac{2\pi}{5} \\ |d_5| &= c, \quad d_k \cdot d_5 = 0, \quad k \neq 5. \end{aligned} \tag{26}$$

The orientation of these basis vectors is fixed by the icosahedral vertices of the ico-dodecahedron projected on the plane through the origin perpendicular to a fivefold axis: ten of the projected vertices define a regular decagon with radius r_d and the remaining two its center (see Fig. 1). Decagon and decagonal lattice Λ_d are left invariant by the chosen subgroup 52 of 532. One can, therefore, follow similar logical steps as in the hexagonal case. There is, first of all, a relation between the radius r_d and the icosahedral lattice parameter a_0:

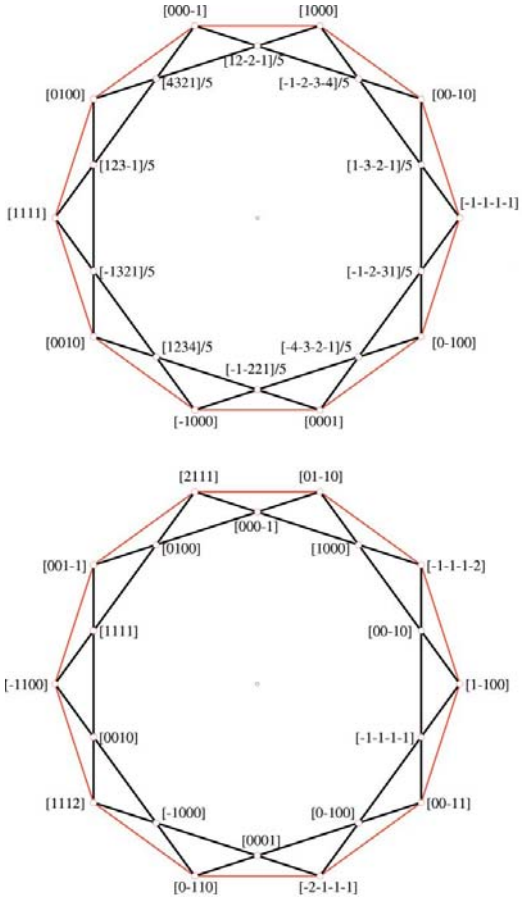

Figure 11
Two ways of indexing a $\{10/2\}$ star decagon are indicated. In the upper part, the decagonal basis vectors point to vertices of the larger decagon. The indices of the central decagon are fractional, with 5 as denominator. In the lower part, the decagonal basis vectors selected are those defined by the smaller decagon and the indices of all vertices are integers. This shows that the polygrammal scaling $S_{\{10/2\}}$, which transforms the external decagon into the internal one, corresponds to a centering transformation for the two decagonal lattices generated by the corresponding bases.

$$r_d = \frac{2\tau}{\sqrt{\tau+2}} a_0 \approx 1.701 a_0. \tag{27}$$

In order to fix the orientation of the vectors d_k, one decomposes the basis vectors a_2, a_3, a_4, a_5 of the icosahedral lattice generated by the basis defined in (3) into a component parallel to the fivefold axis along $a_1 = a_0(1, 0, \tau)$ and one perpendicular to it. All parallel components are given by

$$(a_k)_\parallel = a_1 \cos\delta = \frac{\tau}{\tau+2} a_0(1, 0, \tau), \quad k = 2, 3, 4, 5, \tag{28}$$

with δ the angle of a_k with the fivefold axis. The perpendicular components $(a_k)_\perp$ together with a_1 define a decagonal basis for the lattice $\Lambda_d(r_d, c)$ adapted to the icosahedral symmetry. This basis is not necessarily a basis for the form lattice $\Lambda_d(a, c)$ so that, in general, the vertices of the forms enclosing the decameric and pentameric clusters are indexed by fractions instead of by integers. It is a basis usually denoted as the *conventional basis* for $\Lambda_d(a, c)$ considered as a centering of the primitive lattice $\Lambda_d(r_d, c)$. One finds:

$$\begin{aligned} d_1 = (a_2)_\perp &= a_0\left(\frac{3\tau+1}{5}, 1, -\frac{\tau+2}{5}\right) \\ &= \frac{\sqrt{\tau+2}}{10} r_d\left(\tau+2, \frac{5}{\tau}, -2\tau+1\right) = [1000, 0], \\ d_2 = (a_3)_\perp &= a_0\left(\frac{-2\tau+1}{5}, \tau, \frac{-\tau+3}{5}\right) \\ &= \frac{\sqrt{\tau+2}}{10} r_d(\tau-3, 5, 3\tau-4) = [0100, 0], \\ d_3 = (a_4)_\perp &= a_0\left(-\frac{2\tau+4}{5}, 0, \frac{4\tau-2}{5}\right) \\ &= \frac{\sqrt{\tau+2}}{10} r_d(-4\tau+2, 0, -2\tau+6) = [0010, 0], \\ d_4 = (a_5)_\perp &= a_0\left(\frac{-2\tau+1}{5}, -\tau, \frac{-\tau+3}{5}\right) \\ &= \frac{\sqrt{\tau+2}}{10} r_d(\tau-3, -5, 3\tau-4) = [0001, 0], \\ d_5 &= \frac{1}{\sqrt{\tau+2}} c(1, 0, \tau) = [0000, 1]. \end{aligned} \tag{29}$$

One verifies the metrical relations $(a_k)_\parallel + (a_k)_\perp = a_k$ and (26) for $a = r_d$. For clearness, the planar indices n_1, n_2, n_3, n_4 are separated from the axial index n_5 by a comma.

As already stated, the basis $d = \{d_1, d_2, \ldots, d_5\}$ generates the primitive decagonal lattice $\Lambda(r_d, c)$. By centering, one gets the other decagonal form lattices $\Lambda(a, c)$. The general relation $na = r_h$, integral n of the hexagonal case is replaced by

$$\mu\, a = r_d, \quad \mu \text{ decagrammal scaling factor.}$$

The derivation of these scaling factors is a technical problem, not suitable for treatment here. The particularly important case of the star decagon $\{10/2\}$ is presented as an example.

The Schäfli symbol $\{n/m\}$ denotes the star polygon obtained from a regular n-gon by joining each vertex i with the next $i + m$ one (Coxeter, 1961). The scaling factor $\mu = \mu_{\{n/m\}}$

relates the external regular polygon to the central one. For $n = 10$, $m = 2$, the decagrammal scaling factor $\mu_{\{10/2\}}$ is given by

$$\mu_{\{10/2\}} = \frac{\cos(\pi/5)}{\cos(\pi/10)} = \frac{\tau}{\sqrt{\tau+2}} = \frac{r_d}{2a_0} = 0.8507\ldots \qquad (31)$$

The planar scaling transformation $S_{\{10/2\}}$ with scaling factor $\mu_{\{10/2\}}$, when expressed in the basis $d = \{d_1, \ldots, d_5\}$, is given by the matrices

$$S_{\{10/2\}}(d) = \frac{1}{5}\begin{pmatrix} \bar{1} & 4 & \bar{1} & \bar{1} & 0 \\ \bar{2} & 3 & 3 & \bar{2} & 0 \\ \bar{3} & 2 & 2 & 2 & 0 \\ \bar{4} & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$
$$S^{-1}_{\{10/2\}}(d) = \begin{pmatrix} 0 & 0 & 1 & \bar{2} & 0 \\ 1 & 0 & 1 & \bar{1} & 0 \\ \bar{1} & 1 & 1 & \bar{1} & 0 \\ 0 & \bar{1} & 2 & \bar{1} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (32)$$

which are the centering matrices for this case. Accordingly, the star decagon $\{10/2\}$ has vertices with fractional indices in the conventional basis and integral ones in the basis of the centered lattice $\Lambda_d(a, c)$ with

$$a = \mu_{\{10/2\}} r_d = \frac{r_d^2}{2a_0}, \qquad (33)$$

as indicated in Fig. 11 in the planar version. The star decagon $\{10/3\}$ also plays an important role in the crystallographic properties of the pentagonal and decagonal clusters considered. The scaling factor $\mu_{\{10/3\}}$ is the inverse of the golden number τ:

$$\mu_{\{10/3\}} = \frac{\cos(3\pi/10)}{\cos(\pi/10)} = \frac{1}{\tau} = 0.618\ldots \qquad (34)$$

The planar scaling transformation $S_{\{10/3\}}$ leaves the primitive decagonal form lattice $\Lambda_d(r_d, c)$ invariant and in the basis d is represented by an invertible integral matrix.

The height H of the decamer and the height h of the pentamer, which is also the height of the monomers in the same axial direction, are a multiple of the lattice parameter c:

$$H = pc, \quad h = qc, \quad p, q \text{ integers}. \qquad (35)$$

The results obtained for all the decamers and pentamers with symmetry 52 and 5, respectively, are reported in Table 4. The table shows that the external envelopes (with radius r_e) are in decagrammal scaling relation with the corresponding central boundaries (with radius r_0) and that the form lattices are integral, *i.e.* have a scaling ratio squared given by a simple fraction. In many cases, one finds $\gamma = \frac{1}{16}$ and $\gamma = \frac{1}{8}$, both being rationally equivalent to the isometric decagonal lattice 1-$\Lambda_{\rm dec}$. Actually, the refined boundaries of monomers or dimers still possibly have vertices with integral and half-integral indices when expressed with respect to the form lattices $\Lambda_d(a, c)$ indicated in Table 4.

Fig. 12 presents a simple example of the centered decagonal form lattice $\Lambda_d(a, c)$ with the lattice parameter $a = \mu_{\{10/2\}} r_d$ discussed above. The enclosing form of the VP1 decamer {6,18,23,42,45; 13,28,36,52,55} splits into forms for dimers with monomers pairwise related by a twofold rotation perpendicular to the fivefold axis. The lattice parameter c is determined by the distance of the pentamers from the plane of the dyads. It is a situation observed several times in axial-symmetric proteins (Janner, 2005*a*,*b*,*c*). Moreover, the total height $H = 22c$ of the decamer ensures that the form lattice is integral and has the axial ratio $\gamma = \frac{1}{16}$. The next case is presented to show that the parameter a of the form lattice is

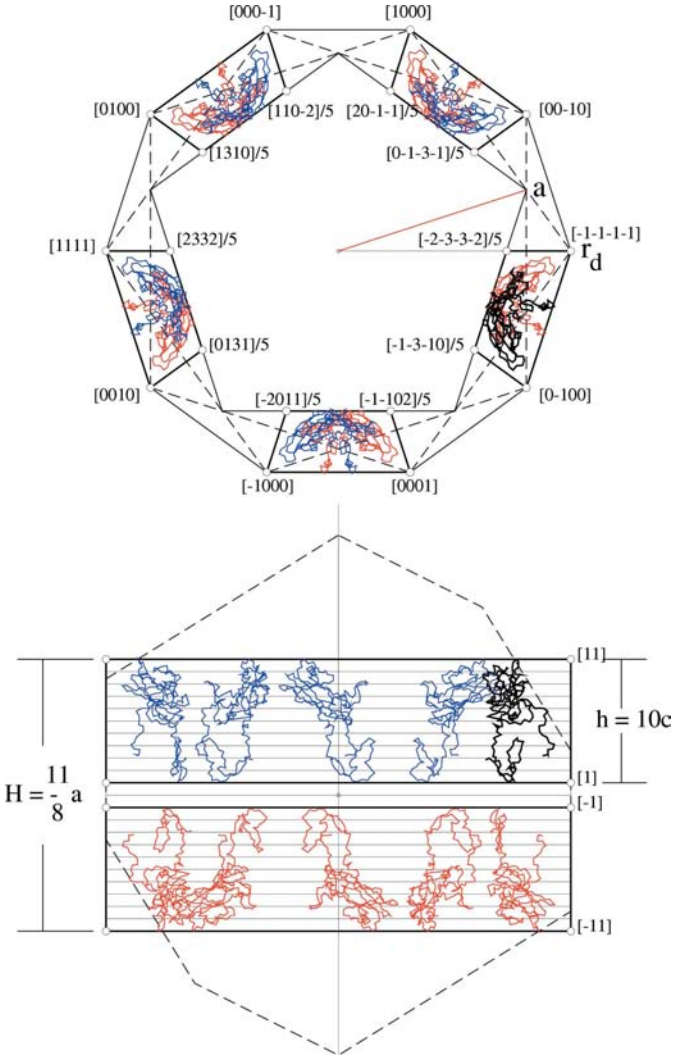

Figure 12
The boundaries of the VP1 decamer {6,18,23,42,45; 13,28,36,52,55} split into five prismatic enclosing forms with vertices at points of the $\{10/2\}$-centered decagonal lattice $\Lambda_d(a, c)$ (compare with Fig. 11), with lattice parameters $a = \mu_{\{10/2\}} r_d$ and $c = \frac{1}{16}a$, where r_d is the radius of the decagon delimiting the capsid projected along the fivefold axis. Accordingly, $\Lambda_d(a, c)$ is rationally equivalent with the isometric decagonal lattice 1-$\Lambda_{\rm dec}$. The height of the decamer is $H = \frac{11}{8}a = 22c$ and that of each monomer is $h = 10c$ (one monomer is enhanced). The dashed lines indicate the star decagon $\{10/2\}$ and the decagonal boundary of the projected capsid, respectively.

not only determined by the planar indices. Indeed, the VP1 decamer {9,16,21,40,48; 11,26,39,50,58} is confined in the axial view by two decagons related by the decagrammal scaling $S_{\{10/3\}}$ with scaling factor $\mu_{\{10/3\}} = 1/\tau$. It is, therefore, possible to assign the integers indicated in Fig. 13 for the planar indices of the vertices, typical for the points of the primitive decagonal lattice $\Lambda_d(r_d, c)$. The height of the decamer, however, $H = \mu_{\{10/2\}} r_d$, indicates that the parameter a is that of the centered lattice. With $H = a = 8c$, one then gets the same axial ratio $\frac{1}{8}$ as in several other cases (see Table 4). Again, the lattice parameter c gives the distance of the pentamers from the plane of the dyads. In the decagonal plane, the dimeric forms appearing in Fig. 12 and Fig. 13 have radially a perfect match.

The third example presented in Fig. 14 has been chosen because it nicely shows the various decagrammal scaling relations between the vertices of the form enclosing the VP3 decamer {0,1,2,3,4; 30,31,32,33,34}, despite (or perhaps

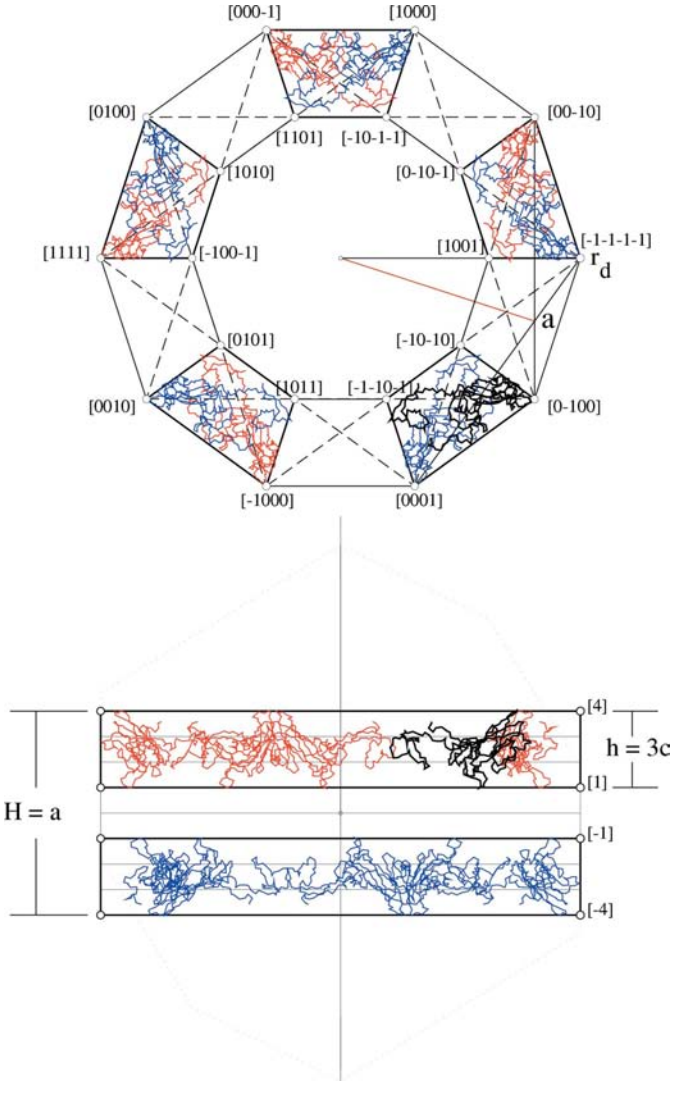

Figure 13
In axial projection, the boundaries of the VP1 decamer {9,16,21,40,48; 11,26,39,50,58} are complementary to those of the decamer plotted in Fig. 12. The height H of the decamer is equal to the parameter $a = \mu_{\{10/2\}} r_d$ of the form lattice which is integral and a 10/2-centered one. The height of the monomer is $h = 3c$ and the axial parameter is $\gamma = \frac{1}{3}$. The external decagonal boundary has the same radius r_d as the projected capsid. The central decagonal boundary follows from the external one by the decagrammal transformation $S_{\{10/3\}}$ with scaling factor $1/\tau$. All vertices are at points of the lattice $\Lambda_d(a, c)$. The star decagon $\{10/3\}$ and the boundaries of the projected capsid are indicated by dashed lines.

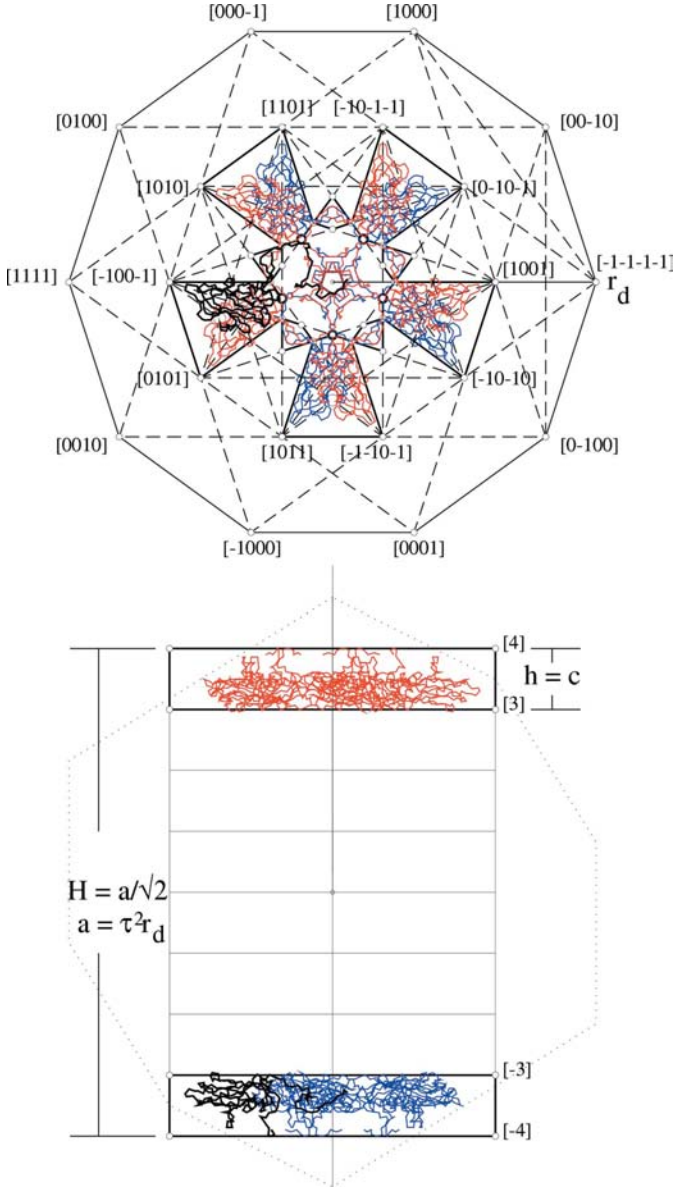

Figure 14
The VP3 decamer {0,1,2,3,4; 30,31,32,33,34} is a nice example of the crystallographic character of the enclosing form. The lattice parameter c is given by the height h of the monomer. The height H of the decamer is $8c$ and equal to $(1/\sqrt{2})a$ with the lattice parameter $a = \tau^2 r_d$, where r_d is the decagonal radius of the capsid. The decagrammal scaling $S_{\{10/3\}}$ leaves the decagonal lattice invariant, so that the lattice parameter $a' = r_d$ is equivalent to a and $\Lambda_d(a, c)$ is an integral lattice, rationally equivalent to $\sqrt{2}$-$\Lambda_{\rm dec}$. All vertices in a given decagonal plane (perpendicular to the fivefold axis) are connected by decagrammal scaling relations. The most important ones are indicated in the upper part by dashed lines. Those of the lower part give the projected boundary of the capsid.

Table 4
Decagonal form lattices for decameric clusters with symmetry 52 (and pentameric clusters with symmetry 5) of the rhinovirus coat proteins.

$r_d = (2\tau/\sqrt{\tau+2})a_0$, with r_d the decagonal radius of the capsid projected along the fivefold axis and a_0 the icosahedral lattice parameter.

Clusters with symmetry 52 decamers and pentamers	a lattice parameter a/r_d	Envelope radius r_e r_e/r_d	Central hole radius r_0 r_0/r_d	Decamer height H	Pentamer height h	Ratio c/a γ
VP1 coat protein (HRV1A, HRV2, HRV3, HRV14, HRV16)						
{0,1,2,3,4; 30,31,32,33,34}	1	$(r_d/2a_0)(1/\tau)$	0	$10c = 2r_d$	$2c$	$\frac{1}{5}$
{5,17,22,41,49; 12,27,35,51,59}	$r_d/2a_0$	1	$r_d/2a_0$	$8c = a$	$6c$	$\frac{1}{8}$
{6,18,23,42,45; 13,28,36,52,55}	$r_d/2a_0$	1	$r_d/2a_0$	$22c = \frac{11}{8}a$	$10c$	$\frac{1}{16}$
{7,19,24,43,46; 14,29,37,53,56}	1	1	$(r_d/2a_0)(1/\tau)$	$24c = \frac{3}{2}r_d$	$6c$	$\frac{1}{16}$
{8,15,20,44,47; 10,25,38,54,57}	$r_d/2a_0$	1	$(r_d/2a_0)(1/\tau)$	$14c = \frac{7}{4}a$	$4c$	$\frac{1}{8}$
{9,16,21,40,48; 11,26,39,50,58}	$r_d/2a_0$	1	$1/\tau$	$8c = a$	$3c$	$\frac{1}{8}$
VP2 coat protein (HRV1A, HRV2, HRV3, HRV14, HRV16)						
{0,1,2,3,4; 30,31,32,33,34}	$r_d/2a_0$	$1/\tau$	$(r_d/2a_0)(1/\tau)^2$	$34c = \frac{17}{8}a$	$6c$	$\frac{1}{16}$
{5,17,22,41,49; 12,27,35,51,59}	$r_d/2a_0$	1	$(r_d/2a_0)^2$	$6c = \frac{3}{8}a$	$6c$	$\frac{1}{16}$
{6,18,23,42,45; 13,28,36,52,55}	$r_d/2a_0$	1	$1/\tau$	$18c = \frac{9}{8}a$	$6c$	$\frac{1}{16}$
{7,19,24,43,46; 14,29,37,53,56}	$r_d/2a_0$	$r_d/2a_0$	$(1/\tau)/(r_d/2a_0)$	$14c = \frac{7}{4}a$	$2c$	$\frac{1}{8}$
{8,15,20,44,47; 10,25,38,54,57}	$r_d/2a_0$	$r_d/2a_0$	$(r_d/2a_0)(1/\tau)$	$14c = \frac{7}{4}a$	$4c$	$\frac{1}{8}$
{9,16,21,40,48; 11,26,39,50,58}	$r_d/2a_0$	1	$(r_d/2a_0)^2$	$10c = \frac{5}{8}a$	$5c$	$\frac{1}{16}$
VP3 coat protein (HRV1A, HRV2, HRV3, HRV14, HRV16)						
{0,1,2,3,4; 30,31,32,33,34}	τ^2	$1/\tau$	$(r_d/2a_0)(1/\tau)^6$	$8c = \frac{1}{\sqrt{2}}a$	c	$\frac{1}{8\sqrt{2}}$
{5,17,22,41,49; 12,27,35,51,59}	$r_d/2a_0$	1	$1/\tau$	$14c = \frac{7}{8}a$	$10c$	$\frac{1}{16}$
{6,18,23,42,45; 13,28,36,52,55}	$r_d/2a_0$	1	$\cos(\pi/5)\cos(\pi/10)$	$16c = a$	$7c$	$\frac{1}{16}$
{7,19,24,43,46; 14,29,37,53,56}	$r_d/2a_0$	$\cos(\pi/10)$	$(1/\tau)\cos(\pi/10)$	$14c = \frac{7}{4}a$	$4c$	$\frac{1}{8}$
{8,15,20,44,47; 10,25,38,54,57}	$r_d/2a_0$	$\cos(\pi/10)$	$(1/\tau)\cos(\pi/10)$	$14c = \frac{7}{4}a$	$4c$	$\frac{1}{8}$
{9,16,21,40,48; 11,26,39,50,58}	$r_d/2a_0$	1	$(r_d/2a_0)^2$	$16c = a$	$5c$	$\frac{1}{16}$
VP4 coat protein (HRV3, HRV14)						
{0,1,2,3,4; 30,31,32,33,34}	1	$(r_d/2a_0)^2(1/\tau)$	$(r_d/2a_0)(1/\tau)^5$	$12c = \frac{3}{2}r_d$	c	$\frac{1}{8}$
{5,17,22,41,49; 12,27,35,51,59}	1	$\cos(\pi/5)$	$(r_d/2a_0)\cos(\pi/5)$	$6c = \frac{1}{\sqrt{3}}r_d$	$4c$	$\frac{1}{6\sqrt{3}}$
{6,18,23,42,45; 13,28,36,52,55}	$1/\tau$	$\cos(\pi/5)$	$1/\tau$	$6c = \frac{5}{4}a$	$2c$	$\frac{5}{24}$
{7,19,24,43,46; 14,29,37,53,56}	1	$1/\tau$	0.47 (?)	$10c = \frac{6}{5}r_d$	$2c$	$\frac{3}{25}$
{8,15,20,44,47; 10,25,38,54,57}	1	$(1/\tau)/(r_d/2a_0)$	$(1/\tau^2)/(r_d/2a_0)$	$8c = \frac{6}{5}r_d$	$2c$	$\frac{3}{20}$
{9,16,21,40,48; 11,26,39,50,58}	$1/\tau$	$\cos(\pi/5)$	$1/\tau$	$10c = a$	$3c$	$\frac{1}{10}$

thanks to) the complexity of the beautiful form. This form has vertices at points of the primitive decagonal lattice $\Lambda_d(r_d, c)$, the same as the lattice $\Lambda_d(a, c)$ for the parameter $a = \tau^2 r_d$ indicated in Fig. 14, owing to the lattice invariance with respect to the decagrammal scaling $S_{\{10/3\}}$.

6. Axial-symmetric forms for clusters with 532 symmetry

As already mentioned, two fundamental properties of proteins with axial symmetry have been observed (Janner, 2005*a*,*b*,*c*):

1. enclosing forms with vertices at points of a lattice $\Lambda(a, c)$ invariant with respect to the axial symmetry;

2. simple fractional value of the axial ratio squared, $\gamma^2 = (c/a)^2$, of the lattice parameters a, c. The lattice $\Lambda(a, c)$ is, therefore, rationally equivalent to an integral lattice.

These properties have been verified so far for tetramers, hexamers and decamers of the coat proteins of rhinovirus. The cluster formed by the full icosahedral set of each coat protein has of course the axial symmetry of the subgroups 222, 32 and 52, respectively, as already pointed out at the beginning. One may wonder whether the properties mentioned above are also true for an icosahedral cluster. This is indeed the case as indicated in Table 5.

The refined enclosing forms of the icosahedral clusters with vertices at points of the lattices indicated in Table 5 are particularly interesting when viewed along the axis of the corresponding rotational symmetry because this reveals the existence of a number of channels. The situation is illustrated in Fig. 15:

(*a*) for VP1 and 222 symmetry by a cubic form lattice with parameter $a = a_c/11 = \tau a_0/11$;

(*b*) for VP2 and 32 symmetry by a hexagonal form lattice with $a = r_h/11$;

(*c*) for VP3 and 52 symmetry by a set of vertices at points of a decagonal lattice and in various decagrammal scaling relations.

In this figure, the planar indices are only indicated for a limited number of vertices.

7. Results and perspectives

Central to this paper are the concepts of *crystallographic scaling* and of *integral lattice*. A scaling transformation of the coordinates is crystallographic if it is faithfully represented by a matrix with integral entries and determinant larger than or equal to 1. A lattice is integral if the metric tensor of the basis vectors is proportional to one expressed by integers. It is

Table 5
Form lattices for the icosahedral clusters of the rhinovirus coat proteins considered with respect to the axial-symmetry subgroups 222, 32 and 52.

Same notation as in the previous tables.

Axial symmetry	Parameters	VP1	VP2	VP3	VP4
222	$a_c/a = a_c/b = a_c/c$	11	8	11	9
32	r_h/a	11	11	11	9
	H/r_h	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$\sqrt{2}$
	$\gamma = c/a$	$11\sqrt{3}$	$11\sqrt{3}$	$11\sqrt{3}$	$9\sqrt{2}$
52	r_d/a	1	$1/\tau^2$	$r_d/(2\tau a_0)$	1
	H/r_d	2	$\tau^2/\sqrt{2}$	$2\tau a_0/r_d$	3/2
	$\gamma = c/a$	2	$1/\sqrt{2}$	1	3/2

nature that reveals the relevance of these concepts in a subtle interplay between integer numbers, geometry and biology. In one dimension, rabbits have been the model for the Fibonacci numbers and their generation, which is in fact a crystallographic scaling transformation. More recently, self-similar patterns in proteins have been recognized, analogous to those already observed in quasicrystals, opening the door for a crystallographic characterization of protein's enclosing forms with vertices at lattice points. Surprisingly enough, these lattices appear to be integral. This observation has been the starting point for an investigation of the abundance of integral lattices in all known crystal structures (de Gelder & Janner, 2005*a*,*b*). In the present paper, it is the icosahedral geometry of viruses that interacts with their biological characterization and shines light on new geometrical and biological aspects.

7.1. Geometry

An icosahedron has vertices at points of a lattice with six basis vectors. This lattice is invariant with respect to one-, two- and three-dimensional scaling transformations which are, therefore, crystallographic. The one-dimensional scalings are of the Fibonacci type with scaling factors that are powers of the golden number τ. The two-dimensional (planar) scalings are related to polygrammal ones, characteristic in particular for the star decagons denoted by $\{10/4\}$, $\{10/3\}$ and $\{10/2\}$. The three-dimensional radial (isotropic) scalings have powers of τ as scaling factors. It is natural, but still surprising, to find relations between the decagrammal scaling factors and the icosahedron in terms of τ and other icosahedral parameters:

$$\mu_{\{10/2\}} = \frac{r_d}{2a_0}, \quad \mu_{\{10/3\}} = \frac{1}{\tau} = \frac{a_0}{a_c}, \quad \mu_{\{10/4\}} = \frac{a_0 r_d}{2a_c^2}, \qquad (36)$$

with a_0 the icosahedral lattice parameter, a_c half the edge of the square and r_d the radius of the regular decagon obtained by projecting the icosahedron along a twofold and a fivefold axis, respectively. Still open is the general case of possible connections between crystallographic regular polytopes and planar polygrammal scalings.

There is also the problem of finding all crystallographic polyhedra with icosahedral symmetry. A polyhedron with icosahedral symmetry is said to be crystallographic if it has vertices at points of an icosahedral lattice (possibly centered). The icosahedron and the dodecahedron are such. The ico-dodecahedron, defined in this paper, is a further non-trivial example. The solution of this problem is similar to the one solved by Caspar & Klug (1962) for deltahedra in terms of icosadeltahedra with triangulation numbers T. The ico-dodecahedron corresponds to the $T = 3$ case. Not all the properties of the ico-dodecahedron have been presented here. In particular, the scaling transformation relating the icosahedral vertices to the dodecahedral ones has only been mentioned. This relation shows that 8 of the 20 dodecahedral vertices form a cube and the remaining 12 are the icosahedral transformed ones.

Not every icosahedral virus obeys the rules of Caspar & Klug. This structural puzzle has been solved by Twarock for

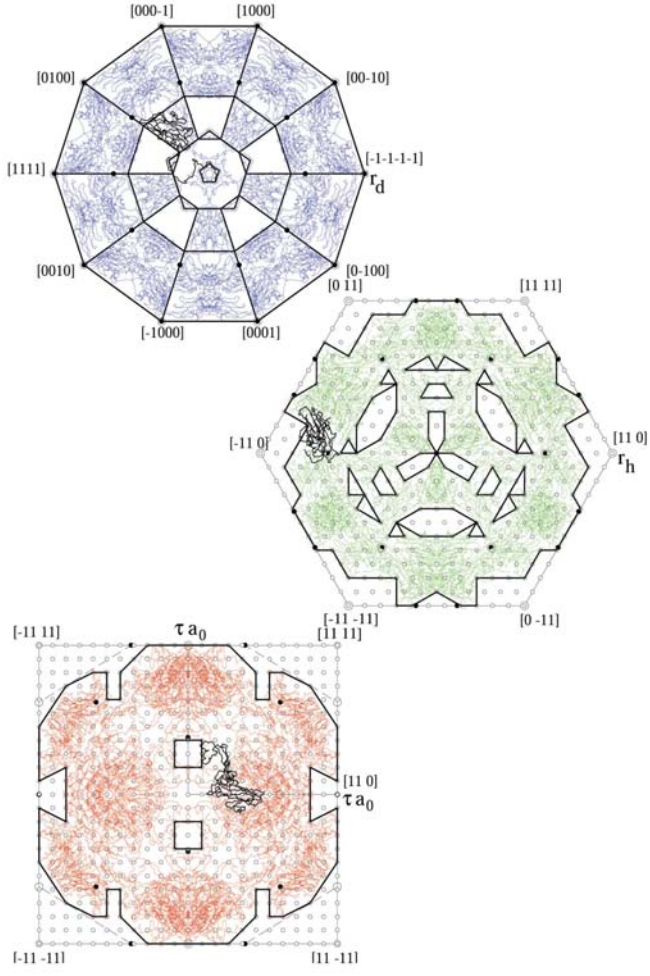

Figure 15
Icosahedral clusters of the major coat proteins. The VP1 proteins (red), viewed along the twofold axis (lower part) have a cubic form lattice with parameter $a = (\tau/11)a_0$, where a_0 is the icosahedral lattice parameter. In the direction of the threefold axis, the VP2 proteins (green) have a hexagonal form lattice with $a = \frac{1}{11} r_h$. In the fivefold orientation (upper part), the VP3 proteins (blue) reveal the decagrammal structure of the various channels. The monomeric chains VP1[0], VP2[0] and VP3[0] are plotted in black. More details are given in Table 5.

the family of *Papovaviridae* by means of a Penrose-like tessellation (Twarock, 2004). In a subsequent paper, Keef & Twarock developed a comprehensive mathematical framework for the derivation of all surface structures of the virial particles in this family (Keef & Twarock, 2005), answering for the *Papovaviridae* the two mathematical questions formulated above.

An alternative general procedure for deriving crystallographic polyhedra based on icosahedral scale rotations is presented in Janner (2006). The new approach includes as special cases the triangular net construction of Caspar & Klug and the tessellation obtained by Twarock.

The question then arises whether the morphological characterization of the rhinovirus presented here is an optimal one. The answer to this question is not trivial, because in principle all polyhedra with vertices at points of the three-dimensional projection of the six-dimensional icosahedral lattice are allowed. The set of these points (all having integral indices) is dense. Therefore, additional conditions have to be imposed. One of these conditions is that the number V of vertices should be limited by the number of morphological units considered. In the Caspar–Klug approach, the number of vertices is expressed in terms of the triangulation number T and one has $V = 10T + 2$. For the rhinovirus T is 3. This limits the enclosing form to be composed of an icosahedron ($V_i = 12$) and a dodecahedron ($V_d = 20$). The scaling ratio between the two is not determined by V and one needs a second condition which in the present crystallographic characterization is very natural: only small rational indices have a structural meaning and are allowed. In the case of crystal growth forms, this condition is well established.

The ico-dodecahedron adopted here satisfies both conditions. A look at Fig. 2 shows that the dodecahedral vertices (full circles) fix the graphical fitting, the value of the icosahedral lattice parameter a_0 and, for the indices indicated, the lattice basis as well. Moreover, the projected boundaries of the capsid along the icosahedral axes only permit very small deviations from the icosahedral vertices indicated (larger circles). Such changes always lead to high indices and have to be rejected. It should be clear that the properties pointed out are those of the enclosing polyhedral forms for the various morphological units and not directly properties of the monomers involved. For a given morphological unit, deviations from the ideal encasing form is not surprising and has to be accepted as such. In any case, it makes little sense to optimize one single case. One never finds a perfect agreement between geometric form and real structure. Sometimes the deviations are larger than what one is inclined to accept. In such a situation, one has to consider a sufficient number of different cases ensuring that the properties pointed out are significant and not purely accidental.

7.2. Biology

From the biological point of view, one of the main results is that rhinovirus has a strongly correlated structure, like the examples of axial-symmetric proteins considered in previous papers (Janner, 2005*a*,*b*,*c*). This means that the architecture of the virus is built of structural units that are all expressible in terms of one single length connecting the real structure with the geometry. Keef & Twarock arrive for the *Papovaviridae* at a similar conclusion for the surface scaling relations between morphological units. These scalings are expressible in terms of only one scaling factor (Keef & Twarock, 2005). In the present case, the single parameter can be chosen to be a_0, the icosahedral lattice parameter of the ico-dodecahedron enclosing the capsid. This is the ideal case. In the real case, one observes small variations in the value of a_0, which are of the order of a few per cent, depending on the serotype, on the coat protein and on the axial-symmetric cluster considered.

From a more general point of view, viruses can be classified according to their biological activity and to their geometrical structure. The relation between the two points of view is of fundamental importance. Another distinction is between the conserved and the varying structural elements in viruses. The present investigation reveals conserved geometrical features in the genus rhinovirus of the family $T = 3$ picorna viruses. It has been shown that the set of crystallographic forms enclosing clusters of coat proteins with a given axial symmetry is conserved in the five different serotypes considered. It is natural to conjecture that this is also true for all the 100 or more serotypes.

The question then arises about possible geometrical elements characteristic for a given serotype. On the basis of the results obtained for the GroEL and GroES chaperonin (Janner, 2003*a*,*b*), one can hope that enclosing forms of segments of the monomeric chains will lead to the desired goal. This same approach should give further information on the folding, because a chain reaching the form boundary has only two possibilities: to stop or to fold.

Alternative to a more detailed geometrical characterization of the conserved structural organization of HRVs is a comparison between different orderings, as expected in modified rhinoviruses like the HRV-HIV-1 chimeras. This could represent a step forward in a structure-based approach of vaccine design.

In the icosahedral case, the number of different indexed forms enclosing a single monomer is incredibly large. This number is of the order of 32 for each coat protein: 15 for the tetramers, 10 for the hexamers and 6 for the decamers and one for the icosahedral cluster, as summarized in the various tables. Additional clusters have also been investigated, but not in a systematic way. The mutual relations between these forms are sometimes evident, sometimes obscure because not familiar and difficult to visualize in three dimensions. In any case, the knowledge of the geometry of these forms should allow the use of symmetry-adapted coordinates for the local characterization of a given active site.

An explanation of the crystallographic character of the enclosing forms and of their lattices is still missing. They all fit with the geometrical boundary of the virion. This suggests the idea of a capsid as a kind of resonator, with nodes of wave-like eigenmodes at the various lattice points.

References

Arnold, E. & Rossmann, M. G. (1990). *J. Mol. Biol.* **211**, 763–801.
Caspar, D. L. D. & Klug, A. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
Coxeter, H. S. M. (1961). *Introduction to Geometry*. New York: J. Wiley.
Gelder, R. de & Janner, A. (2005*a*). *Acta Cryst.* B**61**, 287–295.
Gelder, R. de & Janner, A. (2005*b*). *Acta Cryst.* B**61**, 296–303.
Hadfield, A. T., Lee, W., Zhao, R., Oliveira, M. A., Minor, I., Rueckert, R. R. & Rossmann, M. G. (1997). *Structure*, **5**, 427–441.
Janner, A. (2003*a*). *Acta Cryst.* D**59**, 783–794.
Janner, A. (2003*b*). *Acta Cryst.* D**59**, 795–808.
Janner, A. (2004). *Acta Cryst.* A**60**, 611–620.
Janner, A. (2005*a*). *Acta Cryst.* D**61**, 247–255.
Janner, A. (2005*b*). *Acta Cryst.* D**61**, 256–268.
Janner, A. (2005*c*). *Acta Cryst.* D**61**, 269–277.
Janner A. (2006). *Acta Cryst.* A**62**. In the press.
Keef, T. & Twarock, R. (2005). *q-bio.BM/0512047*, pp. 1–21.
Kim, S., Smith, Th. J., Chapman, M. S., Rossmann, M. G., Pevear, D. C., Dutko, F. J., Felock, P. J., Diana, G. D. & McKinlay, M. A. (1989). *J. Mol. Biol.* **210**, 91–111.
Twarock, R. (2004). *J. Theor. Biol.* **226**, 477–482.
Verdaguer, N., Blaas, D. & Fita, I. (2000). *J. Mol. Biol.* **300**, 1179–1194.
Zhao, R., Pevear, D. C., Kremer, M. J., Giranda, V. L., Kofron, J. A., Kuhn, R. J. & Rossmann, M. G. (1996). *Structure*, **4**, 1205–1220.