Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 18 April 2006
Accepted 10 June 2006

Towards a classification of icosahedral viruses in terms of indexed polyhedra

A. Janner

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

The standard Caspar & Klug classification of icosahedral viruses by means of triangulation numbers and the more recent novel characterization of Twarock leading to a Penrose-like tessellation of the capsid of viruses not obeying the Caspar–Klug rules can be obtained as a special case in a new approach to the morphology of icosahedral viruses. Considered are polyhedra with icosahedral symmetry and rational indices. The law of rational indices, fundamental for crystals, implies vertices at points of a lattice (here icosahedral). In the present approach, in addition to the rotations of the icosahedral group 235, crystallographic scalings play an important rôle. Crystallographic means that the scalings leave the icosahedral lattice invariant or transform it to a sublattice (or to a superlattice). The combination of the rotations with these scalings (linear, planar and radial) permits edge, face and vertex decoration of the polyhedra. In the last case, satellite polyhedra are attached to the vertices of a central polyhedron, the whole being generated by the icosahedral group from a finite set of points with integer indices. Three viruses with a polyhedral enclosing form given by an icosahedron, a dodecahedron and a triacontahedron, respectively, are presented as illustration. Their cores share the same polyhedron as the capsid, both being in a crystallographic scaling relation.

1. Introduction

In the celebrated paper of 1962, *Physical principles in the construction of regular viruses*, Caspar & Klug (1962) formulated the principles of virus architecture. Inspired by crystallography ('virus assembly is like crystallization'), they derive polyhedra with icosahedral symmetry by folding a honeycomb net and substituting a number of hexagons with pentagons, like in the Fuller geodesic dome. The two-dimensional crystallographic condition for the position of the pentagons in the hexagonal lattice allowed a classification of the capsid of many icosahedral viruses. The incompatibility of the icosahedral symmetry with three-dimensional crystallography was compensated by replacing strict equivalence by *quasi-equivalence*. In 1984, the discovery of icosahedral quasicrystals (Shechtman *et al.*, 1984) promoted an alternative crystallographic characterization based on the embedding of the structure in a six-dimensional Euclidean space, where a lattice with icosahedral symmetry is possible. Equivalently, one can consider the projection of this lattice in the three-dimensional space defined by the integral linear combinations of the vectors pointing to the six non-aligned vertices of an icosahedron and also denoted *icosahedral lattice*. In both cases, a lattice point is labeled by a set of six integers called *indices*, which are the components of the position vector with respect to a lattice basis.

The diffraction pattern of icosahedral quasicrystals revealed a new crystallographic symmetry: *scaling invariance*. The positions of the Bragg spots were invariant with respect to a scaling by a factor τ (or of τ^3), where τ is $(1+\sqrt{5})/2$, the *golden mean* (Elser, 1985). In general, a crystallographic scaling expressed in the basis of a lattice is represented by an invertible matrix of infinite order with rational coefficients. By applying crystallographic scalings to the vertices of an icosahedron, one gets a whole variety of forms with vertices at icosahedral lattice points. The Caspar–Klug configurations are included as a special case. It is then possible to express the Euclidean quasi-equivalence as equivalence by scale-rotational transformations. These allow one to relate points situated at different radial distances. The structural relevance of crystallographic scale rotations in biomacromolecules has already been demonstrated for enclosing forms of axial-symmetric proteins (Janner, 2005*a*,*b*,*c*) and indexed forms of morphological units (monomers, dimers, trimers, pentamers, hexamers and 60mers) in the capsid of the icosahedral rhinoviruses (Janner, 2006*a*).

The aim of the present paper is to get a first insight into the possible icosahedral forms given by points equivalent under

scale-rotational transformations with rational entries. Eventually, one then gets a classification scheme for icosahedral viruses in terms of polyhedral forms with vertices indexed by rational numbers. One should also be able to apply this geometric knowledge to a symmetry-adapted characterization of more complex viruses than the rhinovirus.

2. Definitions and notation

Considered are the six vectors $a_1, a_2, \ldots, a_6$ pointing from the center to the non-aligned vertices of an icosahedron. This set is linearly independent of the rationals $\mathbb{Q}$, therefore of the integers $\mathbb{Z}$ as well, and forms a basis $a = \{a_1, a_2, \ldots, a_6\}$ of dimension 3 and rank 6. The integral linear combinations of the a_i define the points of an *icosahedral lattice* with basis a:

$$\Lambda_{\rm ico} = \left\{ \sum_{i=1}^{i=6} n_i a_i \,\middle|\, n_i \in \mathbb{Z} \right\}. \tag{1}$$

In the orthonormal basis $e = \{e_1, e_2, e_3\}$, the basis vectors a_i can be chosen as

$$\begin{gathered} a_1 = a_0(e_1 + \tau e_3),\ \ a_2 = a_0(\tau e_1 + e_2),\ \ a_3 = a_0(\tau e_2 + e_3), \\ a_4 = a_0(-e_1 + \tau e_3),\ \ a_5 = a_0(-\tau e_2 + e_3),\ \ a_6 = a_0(\tau e_1 - e_2), \end{gathered} \tag{2}$$

where a_0 is the icosahedral lattice parameter and $\tau = (1 + \sqrt{5})/2$. In addition, the τ-cubic basis $c = \{c_1, c_2, \ldots, c_6\}$, also of dimension 3 and rank 6 on $\mathbb{Q}$, is defined by

$$c_i = c_0\, e_i, \quad c_{i+3} = c_0 \tau\, e_i, \quad i = 1, 2, 3, \tag{3}$$

with cubic parameter c_0. The components of a vector r with respect to the bases e, a and c, respectively, are indicated as

$$r = (x, y, z)_e = [n_1 n_2 \ldots n_6]_a = [m_1 m_2 \ldots m_6]_c. \tag{4}$$

In what follows, the labels e and a are omitted.

The *indices* of a point are the components of its position vector r with respect to a given lattice basis. In this paper, the icosahedral indices of a point are considered as the ones in the basis a and a point is identified with its position vector:

$$P = [n_1 n_2 \ldots n_6]. \tag{5}$$

Points with rational indices are indicated in this way because their indices are uniquely determined. After multiplication by a constant factor, rational indices can be converted into integral ones.

The *icosahedral group* $K = 235$ is the group of (proper) rotations leaving the icosahedron invariant. It is defined by

$$235 = \{ R_5, R_3 \,|\, R_5^5 = R_3^3 = (R_5 R_3)^2 = 1 \}. \tag{6}$$

Chosen for R_5 is the fivefold rotation around a_1 and for R_3 the threefold rotation around $a_1 + a_2 + a_3$. The twofold rotation around e_3 is given by $R_5^2 R_3 R_5^{-1}$. In the basis a, these rotations have a six-dimensional integer matrix representation:

$$\begin{aligned} 5(a) = R_5(a) &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \\ 3(a) = R_3(a) &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bar{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \bar{1} \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \\ 2_z(a) = R_2(a) &= \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \bar{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bar{1} \end{pmatrix}. \end{aligned} \tag{7}$$

A three-dimensional scaling S_λ with *scaling factor* λ transforms by λ-multiplication all the components of a vector. A one-dimensional scaling $S_{b,\lambda}$ scales by λ the components parallel to the given vector b and leaves invariant the perpendicular components. For a two-dimensional scaling $S_{\perp b,\lambda}$ it is the other way round: the parallel components are left invariant and the perpendicular ones are scaled by λ. Indicating by $r_\parallel$ and $r_\perp$ the parallel and perpendicular components of r with respect to b ($r = r_\parallel + r_\perp, r_\perp b = 0$), one gets the relations

$$S_\lambda r = \lambda r, \quad S_{b,\lambda} r = \lambda r_\parallel + r_\perp, \quad S_{\perp b,\lambda} r = r_\parallel + \lambda r_\perp. \tag{8}$$

This implies

$$S_\lambda = S_{b,\lambda} S_{\perp b,\lambda}. \tag{9}$$

With $b = e_i$ and $X_\lambda = S_{e_1,\lambda}$, $Y_\lambda = S_{e_2,\lambda}$ and $Z_\lambda = S_{e_3,\lambda}$, one has:

$$\begin{gathered} S_\lambda(x, y, z) = (\lambda x, \lambda y, \lambda z), \quad X_\lambda(x, y, z) = (\lambda x, y, z), \\ S_\lambda X_\lambda^{-1}(x, y, z) = (x, \lambda y, \lambda z) = Y_\lambda Z_\lambda(x, y, z). \end{gathered} \tag{10}$$

In the icosahedral basis a, the scaling transformations S_τ and X_τ are represented by six-dimensional invertible matrices with rational coefficients:

$$S_\tau(a) = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & \bar{1} & \bar{1} & 1 \\ 1 & 1 & 1 & 1 & \bar{1} & \bar{1} \\ 1 & \bar{1} & 1 & 1 & 1 & \bar{1} \\ 1 & \bar{1} & \bar{1} & 1 & 1 & 1 \\ 1 & 1 & \bar{1} & \bar{1} & 1 & 1 \end{pmatrix},$$

$$X_\tau(a) = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 2 & 0 & \bar{1} & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 1 & \bar{1} & 0 & 1 & 0 & \bar{1} \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & \bar{1} & 0 & 2 \end{pmatrix}. \qquad (11)$$

S_τ, X_τ, Y_τ and Z_τ leave invariant the set of all points with rational icosahedral indices. After multiplication by 2, these matrices become integral and transform integer indices into integer indices. This situation is generalized to *rational icosahedral scalings* with an *integer factor* f_0, where f_0 is the smallest integer factor that transforms the rational matrix into one with integer entries.

The *icosahedral forms* considered are given in terms of a set of points with rational indices and icosahedral symmetry. The *icosahedral polyhedral forms* represent the special case where the given points are at the vertices of the polyhedron. The other forms are *decorated polyhedral forms* of various types: with *face*, *edge* or *vertex decoration*. The last case is represented by the protuberances observed in non-enveloped viruses, as in the case of the adenovirus (Stewart *et al.*, 1991), or of the spiroplasma virus (Chipman *et al.*, 1998). One can restrict the considerations to points with integer indices and specify the form in terms of *generators*. The generators are icosahedral inequivalent points which generate the whole form by applying the icosahedral group. It is convenient to indicate by (*VEF*) the number V of vertices, E of edges and F of faces of a polyhedral form. These integers satisfy the Euler relation $V - E + F = 2$.

Examples of important icosahedral polyhedral forms generated from one or two points with integer entries are:

Icosahedron: I = {235 [100000]} with (12 30 20) is generated from a_1 by 235 and has 12 vertices, 30 edges and 20 triangular faces.

Dodecahedron: D = {235 [111000]} with (20 30 12) is generated from $a_1 + a_2 + a_3$ and has 20 vertices, 30 edges and 12 pentagonal faces.

Icosidodecahedron: ID = {235 [110000]} with (30 60 32) has 20 triangular faces and 12 pentagonal ones (Coxeter, 1963).

Triacontahedron: TR = {235 [200000], [111$\bar{1}$1$\bar{1}$]} (32 60 30) has 30 rhombic faces, 12 icosahedral and 20 dodecahedral vertices. Discovered by Kepler, it is the projection in three dimensions of a six-dimensional cube.

Truncated icosahedron: TI = {235 [210000]} with (60 90 32) is one of the Archimedean solids. It is obtained by cutting an edge-decorated icosahedron, has 12 pentagonal and 20 hexagonal faces, 60 vertices and 90 edges. It corresponds to a Fuller polyhedron.

3. Fractional and golden icosahedral scalings

Fractional icosahedral crystallographic scalings are scaling transformations (one-dimensional, two-dimensional and three-dimensional in the three-dimensional space) with a fractional scaling factor and represented with respect to the icosahedral lattice basis a by invertible matrices with rational coefficients. The golden ones have a power of τ as scaling factor.

3.1. Radial scalings (three-dimensional)

In addition to the scaling S_τ with scaling factor τ already indicated, the multiplication of vector components by a fractional number m/n is trivially a three-dimensional rational scaling. It is represented by the unit matrix multiplied by m/n:

$$S_{m/n}(a) = \frac{m}{n}\mathbb{1}, \quad m, n \in \mathbb{Z}, n \neq 0. \qquad (12)$$

The integer factor is in this case $f_0 = n$.

3.2. Linear scalings (one-dimensional)

The one-dimensional scaling X_τ indicated in the previous section in the orientation fixed by the a basis is along a twofold axis of the icosahedron. The corresponding fractional scaling $X_{m/n}$ has in the τ-cubic c basis a very simple diagonal matrix representation. The one in the basis a follows directly:

$$X_{m/n}(a) = \frac{1}{2n}\begin{pmatrix} m+n & 0 & 0 & -m+n & 0 & 0 \\ 0 & m+n & 0 & 0 & 0 & m-n \\ 0 & 0 & 2n & 0 & 0 & 0 \\ -m+n & 0 & 0 & m+n & 0 & 0 \\ 0 & 0 & 0 & 0 & 2n & 0 \\ 0 & m-n & 0 & 0 & 0 & m+n \end{pmatrix}, \qquad (13)$$

with $m, n \in \mathbb{Z}, n \neq 0$.

Linear scalings by τ and by m/n, respectively, are rational in the icosahedral basis a.

In the case of a linear scaling along the fivefold axis, a_1 is chosen as the scaling direction left invariant by the rotation R_5 which permutes cyclically the other basis vectors:

$$R_5: \quad a_1 \to a_1, \quad a_2 \to a_3 \to a_4 \to a_5 \to a_6 \to a_2. \qquad (14)$$

The scaling is fixed by the transformations of the basis vectors decomposed into their parallel and perpendicular components with respect to a_1. The parallel components of $a_2, a_3, \ldots, a_6$ are all equal:

$$a_{i\parallel} = \tfrac{1}{5}(a_2 + a_3 + \ldots + a_6) = \tfrac{1}{5}[011111], \quad i = 2, \ldots, 6. \qquad (15)$$

The perpendicular components follow from $a_{i\perp} = a_i - a_{i\parallel}$. For $S_{a_1,\tau}$, one has

$$S_{a_1,\tau} a_i = \tau a_{i\parallel} + a_{i\perp} = S_\tau a_{i\parallel} + a_{i\perp}, \quad i = 1, \ldots, 6, \qquad (16)$$

with S_τ given in (11). Taking a_2 as an example, one finds

$$S_{a_1,\tau}[010000] = \tfrac{1}{10}[511111] + \tfrac{1}{5}[04\bar{1}\bar{1}\bar{1}\bar{1}] = \tfrac{1}{10}[59\bar{1}\bar{1}\bar{1}\bar{1}]. \quad (17)$$

The transformation of the other basis vectors follows by cyclic permutation. One gets

$$S_{a_1,\tau}(a) = \frac{1}{10}\begin{pmatrix} 5 & 5 & 5 & 5 & 5 & 5 \\ 5 & 9 & \bar{1} & \bar{1} & \bar{1} & \bar{1} \\ 5 & \bar{1} & 9 & \bar{1} & \bar{1} & \bar{1} \\ 5 & \bar{1} & \bar{1} & 9 & \bar{1} & \bar{1} \\ 5 & \bar{1} & \bar{1} & \bar{1} & 9 & \bar{1} \\ 5 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 9 \end{pmatrix} \quad (18)$$

with the integer factor $f_0 = 10$. In a similar way, one derives the fractional linear scaling along a_1:

$$S_{a_1,m/n}(a) = \frac{1}{5n}\begin{pmatrix} 5m & 0 & 0 & 0 & 0 & 0 \\ 0 & m+4n & m-n & m-n & m-n & m-n \\ 0 & m-n & m+4n & m-n & m-n & m-n \\ 0 & m-n & m-n & m+4n & m-n & m-n \\ 0 & m-n & m-n & m-n & m+4n & m-n \\ 0 & m-n & m-n & m-n & m-n & m+4n \end{pmatrix} \quad (19)$$

with $m, n \in \mathbb{Z}$, $n \neq 0$ and $f_0 = 5n$.

The threefold rotation R_3 transforms the basis vectors a_i according to

$$R_3: \quad a_1 \to a_2 \to a_3 \to a_1, \quad a_4 \to a_6 \to -a_5 \to a_4, \quad (20)$$

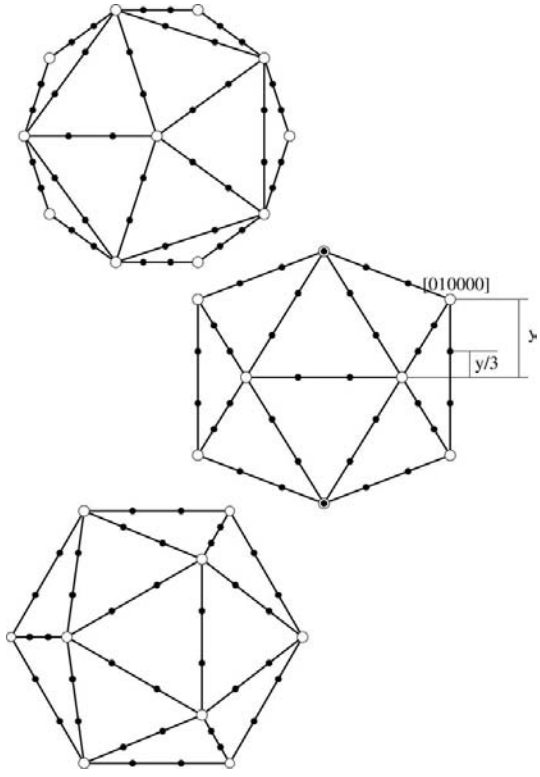

Figure 1
Edge-decorated icosahedron obtained by linear scaling from the vertices. The generating points are [300000], [0200001] = $Y_{(1/3)}$[030000], where $Y_{(1/3)}$ scales the corresponding y coordinate to $y/3$. In this figure and in the following ones, the projections along a fivefold, a threefold and a twofold axis are given at top, bottom and middle positions, respectively.

and leaves the two vectors $a_1 + a_2 + a_3$ and $a_4 - a_5 + a_6$ invariant. Both are along the threefold axis chosen. The parallel components then follow:

$$\begin{aligned} a_{1\parallel} &= a_{2\parallel} = a_{3\parallel} = \tfrac{1}{3}[111000], \\ a_{4\parallel} &= -a_{5\parallel} = a_{6\parallel} = \tfrac{1}{3}[000\bar{1}\bar{1}1]. \end{aligned} \quad (21)$$

The remaining steps are similar to those of the fivefold case. With $d_1 = [111000]$, one gets

$$S_{d_1,\tau}(a) = \frac{1}{6}\begin{pmatrix} 7 & 1 & 1 & 1 & \bar{1} & 1 \\ 1 & 7 & 1 & 1 & \bar{1} & 1 \\ 1 & 1 & 7 & 1 & \bar{1} & 1 \\ 1 & 1 & 1 & 3 & 3 & \bar{3} \\ \bar{1} & \bar{1} & \bar{1} & 3 & 3 & 3 \\ 1 & 1 & 1 & \bar{3} & 3 & 3 \end{pmatrix}, \quad (22)$$

$$S_{d_1,m/n}(a) = \frac{1}{3n}\begin{pmatrix} m+2n & m-n & m-n & 0 & 0 & 0 \\ m-n & m+2n & m-n & 0 & 0 & 0 \\ m-n & m-n & m+2n & 0 & 0 & 0 \\ 0 & 0 & 0 & 2m+n & -m+n & m-n \\ 0 & 0 & 0 & -m+n & m+2n & -m+n \\ 0 & 0 & 0 & m-n & -m+n & 2m+n \end{pmatrix}. \quad (23)$$

The linear scalings along the other threefold and fivefold directions are obtained by conjugation with elements of the icosahedral group. Edge decoration of a polyhedral form can be obtained by a linear scaling along a given edge. For example, by applying $Y_{1/3}$ to a_2, one gets the edge-decorated icosahedron of Fig. 1.

3.3. Planar scalings (two-dimensional)

The scalings in planes perpendicular to the twofold, threefold and fivefold axes are obtained by a combination of radial and linear scalings, as already mentioned. So, for example, using

$$S_{\perp a_1,\tau} = S_\tau S^{-1}_{a_1,\tau} = S_\tau S_{a_1,1/\tau} = S^{-1}_{\perp a_1,1/\tau}, \quad (24)$$

one gets from equations (11) and (18)

$$S_{\perp a_1,\tau}(a) = \frac{1}{5}\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & \bar{2} & \bar{2} & 3 \\ 0 & 3 & 3 & 3 & \bar{2} & \bar{2} \\ 0 & \bar{2} & 3 & 3 & 3 & \bar{2} \\ 0 & \bar{2} & \bar{2} & 3 & 3 & 3 \\ 0 & 3 & \bar{2} & \bar{2} & 3 & 3 \end{pmatrix}. \quad (25)$$

In addition to the cases considered, the planar pentagrammal and hexagrammal scalings deserve attention. They play an important rôle in axial symmetric proteins (Janner, 2005*a*,*b*,*c*) and in forms enclosing pentameric and trimeric morphological units of the capsid of the rhinovirus (Janner, 2006*a*). The planar pentagrammal scaling $\{5/2\}$ has the scaling factor $-1/\tau^2$, where $\{m/n\}$ is the Schäfli symbol of the corresponding star polygon (Coxeter, 1961). In the present notation, it is given by $S_{\perp a_1,-1/\tau^2}$, which is different from $-S_{\perp a_1,1/\tau^2}$ because the scaling factor $-1/\tau^2$ is associated with a planar inversion and not with a three-dimensional total inversion. For the inversion in the plane perpendicular to a_1 one has

$$I_{\perp a_1}: \quad a_1 \to a_1, \quad a_2 \to -a_2 + 2a_{2\parallel} = [0\bar{1}0000] + \tfrac{2}{5}[011111] \tag{26}$$

and so on, giving

$$I_{\perp a_1}(a) = \frac{1}{5}\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & \bar{3} & 2 & 2 & 2 & 2 \\ 0 & 2 & \bar{3} & 2 & 2 & 2 \\ 0 & 2 & 2 & \bar{3} & 2 & 2 \\ 0 & 2 & 2 & 2 & \bar{3} & 2 \\ 0 & 2 & 2 & 2 & 2 & \bar{3} \end{pmatrix}. \tag{27}$$

Finally, the pentagrammal scaling in the planes perpendicular to the fivefold axis a_1 is given by the integer matrix

$$\begin{aligned} S_{\perp a_1,-1/\tau^2}(a) &= S_{\perp a_1,\{5/2\}}(a) \\ &= I_{\perp a_1}(a) S^{-2}_{\perp a_1,\tau}(a) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \bar{1} & 1 & 0 & 0 & 1 \\ 0 & 1 & \bar{1} & 1 & 0 & 0 \\ 0 & 0 & 1 & \bar{1} & 1 & 0 \\ 0 & 0 & 0 & 1 & \bar{1} & 1 \\ 0 & 1 & 0 & 0 & 1 & \bar{1} \end{pmatrix}. \end{aligned} \tag{28}$$

The decagrammal scalings also give rise to planar scalings with rational entries. This property fits with the empirical observation of the rôle of decagrams in the enclosing forms of pentameric morphological units in the rhinovirus (Janner, 2006*a*).

For the hexagrammal scaling by $\sqrt{3}$ of the hexagonal star polygon {6/2} of triangular faces in planes perpendicular to a threefold direction, here chosen to be $d_1 = [111000]$, one finds

$$S_{\perp d_1,\{6/2\}}(a) = \frac{1}{3}\begin{pmatrix} 4 & \bar{2} & 1 & 0 & 0 & 0 \\ 1 & 4 & \bar{2} & 0 & 0 & 0 \\ \bar{2} & 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & \bar{1} & \bar{2} \\ 0 & 0 & 0 & 2 & 4 & \bar{1} \\ 0 & 0 & 0 & 1 & 2 & 4 \end{pmatrix}. \tag{29}$$

In this case, the integer factor is $f_0 = 3$ and not 1 as for the pentagrammal scaling (28). Planar scalings allow face decoration. The face decoration of Fig. 2 has been obtained by applying pentagrammal scalings to the vertices of a dodecahedron. That of Fig. 3 follows in a similar way from the planar scalings $S_{\perp a_i,1/\tau}$.

4. Indexed icosahedral forms

The forms considered (polyhedra and decorated polyhedra) are specified by a set of points with icosahedral symmetry and rational indices (in the basis of an icosahedral lattice). Vertices and faces are indexed accordingly. These forms can be given in terms of point generators with integer indices.

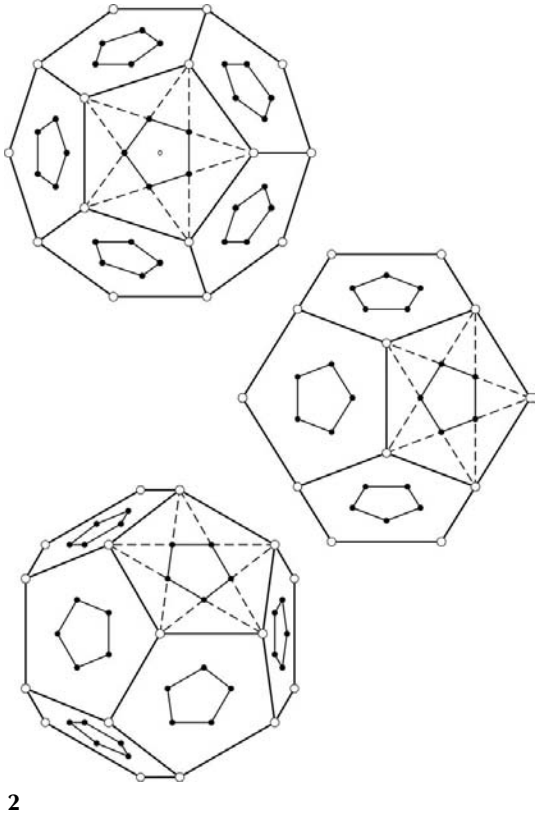

Figure 2
Face-decorated dodecahedron obtained by planar pentagrammal scalings. The generating points are [111000] and [100101] = $Y_{\perp a_1,\{5/2\}}$[111000].

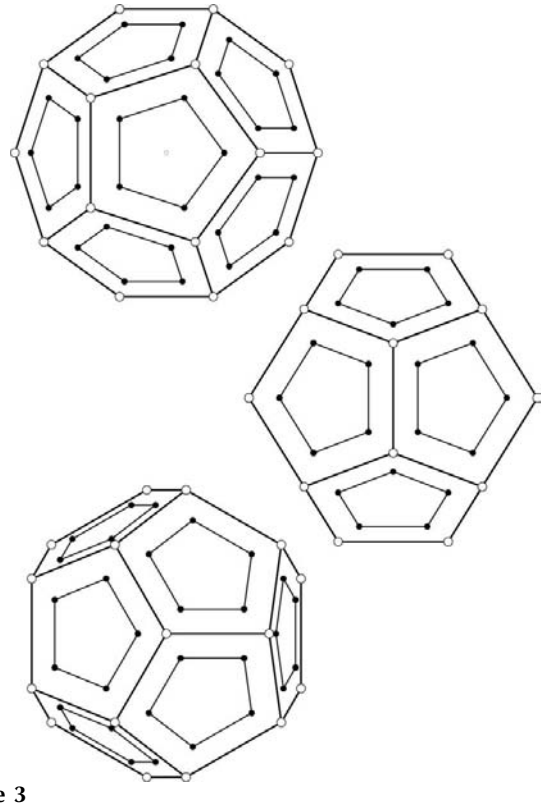

Figure 3
Pentagonal face-decorated dodecahedron obtained by the planar pentagonal scaling with scaling factor $1/\tau$. The generating points are [555000] and [533$\bar{2}$23] = $S_{\perp a_1,1/\tau}$[555000].

Table 1
Low-indices Caspar–Klug icosahedral polyhedra with triangulation number $T = Pf^2$ and $(VEF) = ([10T + 2]30T20T)$.

h	k	P	f	T	Planar net $[z_1 z_2]$	Generators $[n_1 n_2 n_3 n_4 n_5 n_6]$
1	0	1	1	1	[00]	[100000]
			2	4	[00], [10]	[200000], [110000]
			3	9	[00], [10], [11]	[300000], [210000], [111000]
			4	16	[00], [10], [20], [11]	[400000], [310000], [220000], [211000]
			5	25	[00], [10], [20], [11], [21]	[500000], [410000], [320000], [311000], [221000]
1	1	3	1	3	[00], [01]	[300000], [111000]
			2	12	[00], [01], [02], [11]	[600000], [411000], [222000], [330000]
2	1	7	1	7ℓ	[00], [01]	[700000], [412000]
1	2	7	1	$7d$	[00], [01]	[700000], [421000]

4.1. Caspar–Klug construction

Caspar & Klug start from a planar honeycomb net (Caspar & Klug, 1962). The centers of the honeycomb hexagons form a hexagonal lattice (and a net of equilateral triangles) with basis $b = \{b_1, b_2\}$ and lattice parameter b_0:

$$|b_1| = |b_2| = b_0, \quad b_1 b_2 = \tfrac{1}{2} b_0^2. \tag{30}$$

The hexagonal lattice points are indicated as

$$[z_1, z_2] = z_1 b_1 + z_2 b_2 \in \Lambda_h, \quad z_1, z_2 \in \mathbb{Z}. \tag{31}$$

The honeycomb hexagons have radius $r_{\rm hc} = b_0/\sqrt{3}$ and area $A_{\rm hc} = \sqrt{3}/2 b_0^2$. The icosahedral face has area $A = \sqrt{3} a_0^2$ and edge $|a_1 - a_2| = 2a_0$, with a_0 the icosahedral lattice parameter. The hexagonal lattice is chosen in the plane perpendicular to $d_1 = a_1 + a_2 + a_3$ and origin at a_1. The three vertices a_1, a_2, a_3 of a triangular face of the icosahedron are brought in coincidence with the lattice points $[0, 0]$, $[h, k]$ and $[-k, h + k]$, respectively. This is expressed by the set of equations

$$\begin{aligned} a_2 - a_1 &= h b_1 + k b_2 \\ a_3 - a_1 &= -k b_1 + (h + k) b_2 \end{aligned} \tag{32}$$

or, equivalently,

$$\begin{aligned} T b_1 &= (h^2 + hk + k^2) b_1 = -h a_1 + (h + k) a_2 - k a_3 \\ T b_2 &= (h^2 + hk + k^2) b_2 = -(h + k) a_1 + k a_2 + h a_3, \end{aligned} \tag{33}$$

with $T = h^2 + hk + k^2$ the *triangulation number*. The relation between the hexagonal and the icosahedral lattice parameters is given by $T b_0^2 = 4 a_0^2$. The folding of the triangular net into the triangular facets of the icosahedral form is obtained by replacing the honeycomb hexagons with center at a_1, a_2, a_3, respectively, by regular pentagons and repeating the procedure for the other icosahedral faces as well. The vertices of the *Caspar–Klug polyhedron* (CK polyhedron) are obtained from the hexagonal lattice points in the *fundamental region* of the icosahedron formed by one vertex, one side and 1/3 of the internal points of an icosahedral face. Its area is $A_0 = (1/\sqrt{3}) a_0^2$. These vertices have rational indices given by the relation

$$\begin{aligned} [z_1, z_2] = \frac{1}{T} [&T - z_1 h - z_2 (h + k), z_1 (h + k) + z_2 k, \\ &- z_1 k + z_2 h, 0, 0, 0], \end{aligned} \tag{34}$$

with $[z_1, z_2] \in A_0$ and $z_1, z_2 \in \mathbb{Z}$. The full set of vertices is obtained by applying to these points the group 235. The number of hexagons used for one face of the icosahedron is given by the ratio

$$\frac{A}{A_{\rm hc}} = \sqrt{3} a_0^2 \Big/ \left(\frac{\sqrt{3}}{2} b_0^2 \right) = \frac{T}{2}. \tag{35}$$

The hexagons substituted by pentagons represent the fraction $A_{\rm hc}/2$ of the area A. There are therefore $20(T - 1)/2$ vertices at the centers of the hexagons and 12 at the pentagonal centers, yielding for the total number of vertices $V = 10T + 2$, as indicated in Caspar & Klug (1962). The area of each triangular facet is half $A_{\rm hc}$, so that their total number F is $2A/A_{\rm hc} = 20T$. A facet is obtained from a tessellation of the triangular face of the icosahedron into smaller regular triangles. From the Euler relation follows the number of edges $E = 30T$. Thus the polyhedron obtained by the Caspar–Klug construction has

$$(VEF) = ([10T + 2]\, 30T\, 20T). \tag{36}$$

The generators of these forms are obtained from lattice points obeying (34). With f the largest cofactor of h and k and P given by $T = Pf^2$, the integer indices are obtained from (34) for h, k relatively prime, by multiplication with P after substituting T by P. Expressing R_3 in the basis b, taking into account the shift in origin, one easily finds the inequivalent points among those belonging to the icosahedral face (a_1, a_2, a_3). Expressed in the Seitz notation $\{R \,|\, t\} r = Rr + t$, one has

$$\begin{aligned} R_3(b) &= \left\{ \begin{pmatrix} \bar{1} & \bar{1} \\ 1 & 0 \end{pmatrix} \Big| \, [h, k] \right\}, \\ R_3^2(b) &= \left\{ \begin{pmatrix} 0 & 1 \\ \bar{1} & \bar{1} \end{pmatrix} \Big| \, [-k, h + k] \right\}. \end{aligned} \tag{37}$$

The low-indices CK polyhedra are listed in Table 1.

The concept of *icosadeltahedron*, defined in Caspar & Klug (1962) as a polyhedron with icosahedral symmetry and all faces equilateral triangles, has been avoided because, in general, it is not compatible with the folded cardboard construction. In particular, for $T = 7$ there are 12 pentagonal vertices V_p and 60 hexagonal ones V_h. Nearest-neighbor vertices all have the same distance in the plane but not in space. In space, there are two different triangular edges. For neighboring V_p and V_h, one has

$$|V_p - V_h| = |R_3V_h - V_h| \neq |R_5V_h - V_h|, \qquad (38)$$

because V_h is on the same icosahedral face as R_3V_h but on a different one to R_5V_h. So there is no icosadeltahedron for $T = 7$, only a quasi-icosadeltahedron.

Icosahedral viruses have the tendency to optimize a compact surface arrangement (a property taken into account by the Caspar & Klug construction) with an approximate spherical symmetry. Starting from a CK polyhedron by a suitable radial scaling of one or more of the generators, it is possible to realize low indices icosahedral polyhedra with a quasi-spherical shape. For the Caspar & Klug $T = 7d$ case (where d means *dextro*) with generators [700000] and [421000], see Table 1, by radial scaling the second generator is transformed into $S_{2/\tau}[421000] = [\bar{1}35315]$, yielding a quasi-spherical polyhedron with the same number of vertices, edges and faces, respectively (Fig. 4).

Another example is the *ico-dodecahedron* introduced in Janner (2006*a*) as the enclosing form of the rhinovirus capsid. By the radial scaling S_{1/τ^2} of the dodecahedral vertex [111000], one gets $S_{1/\tau^2}[111000] = \frac{1}{2}[111\bar{1}\bar{1}\bar{1}]$ in addition to the icosahedral vertex at [100000]. The corresponding integer generators of the ico-dodecahedron are [200000], $[111\bar{1}\bar{1}\bar{1}]$ with 32 vertices, 90 edges and 30 faces as the $T = 3$ CK icosahedron, generated by [3000000] and [111000], as indicated in Table 1, which yields a less good spherical approximation. The 32 vertices of the ico-dodecahedron (not those of the corresponding CK-polyhedron) define a *triacontahedron* with half as many rhombic faces, instead of 60 triangular facets (Fig. 5). The triacontahedron is the three-dimensional projection of a six-dimensional hypercube and plays an important rôle in icosahedral quasicrystals (Steinhardt & Ostlund, 1987).

4.2. Off-center (satellite) polyhedra

In the planar polygrammal case, the combination of rotational symmetries with linear scalings leads to off-center regular polygons. The transformations involved are always linear ones leaving the central origin invariant and are not around the centers of the new polygons. Off-center polygonal

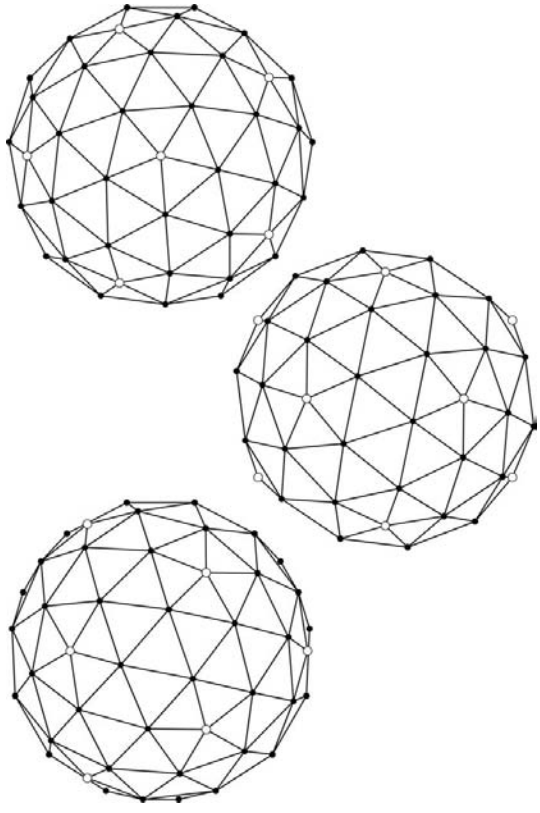

Figure 4
Quasi-spherical polyhedron with icosahedral symmetry and triangulation number $T = 7d$. The point generators [700000] and $[\bar{1}35315] = S_{2/\tau}[421000]$ from the corresponding Caspar–Klug polyhedron by radial scaling with factor $2/\tau$ of the non-icosahedral vertex [421000].

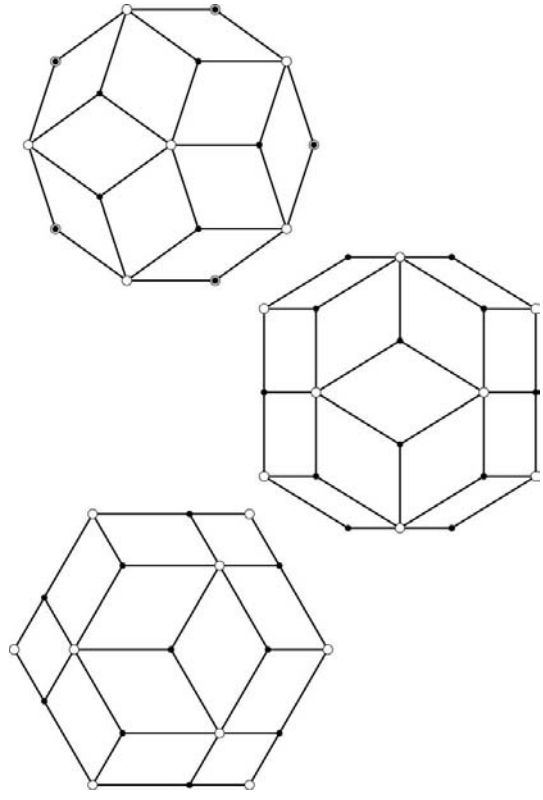

Figure 5
The triacontahedron, discovered by Kepler in 1611, which is the projection in three dimensions of the six-dimensional cube, has point generators [200000] and $[111\bar{1}\bar{1}\bar{1}] = S_{1/\tau^2}[111000]$. It has 32 vertices and 30 rhombic faces. The corresponding 60 triangular facets define the ico-dodecahedron obtained from a $T = 3$ Caspar–Klug polyhedron by rescaling of the dodecahedral vertices (Janner, 2006*a*).

boundaries (in projection) have been shown to occur in biomacromolecules with sevenfold axial symmetry (Janner, 2002) and are observed for other rotational symmetries as well.

In the icosahedral case, one can get a whole variety of off-center indexed polyhedra by combining icosahedral rotations with linear, planar and radial scalings of the type derived in the previous section. It is not possible (and even not necessary) to discuss here all these situations. It is important only to be aware of the existence of satellite polyhedra in order to be able to recognize the occurrence in the morphology of specific viruses like those with protruding proteins, as already mentioned.

As a first example, satellite icosahedra are derived, which correspond to a vertex-decorated icosahedron. One applies to the vertices of a starting central icosahedron the conjugated linear scaling $RX_\tau R^{-1}$, for all $R \in 235$, together with the radial scaling S_τ, invariant with respect to these conjugations. The result is shown in Fig. 6. Instead of a linear scaling along a twofold direction (as it is X_τ), one can consider those along threefold and fivefold axes. In particular, one then obtains 20 satellite dodecahedra attached in a way similar to the vertices of a central dodecahedron.

Another example is based on a linear scaling along the threefold direction $d_1 = a_1 + a_2 + a_3$, with a scaling factor τ. This leads to consideration of the two generators $[7111\bar{1}1]$ and $[11133\bar{3}]$ as one sees from (22). The satellite polyhedra generated by these two vector positions are truncated pentagonal pyramids attached to the 12 vertices of a central icosahedron generated from [600000] (see Fig. 7). Similar results are obtained with other scaling factors.

Even more variations follow from alternative combinations of icosahedral rotations with planar scalings.

5. Morphology of icosahedral viruses

Considered are models for the capsid of virions with icosahedral symmetry. A geometric characterization can be given in terms of positions or of arrangement of architectural elements.

In the PDB files, one finds the full set of atomic positions. For a geometric morphological model, it is enough to specify the C_α positions of the polypeptide chains of the monomers involved. A further approximation of the structure is represented by the positions of monomers or of clusters of monomers organized in morphological units like capsomers, trimers, pentamers, hexamers and so on.

5.1. Tiling models

Instead of giving positions, one can partition the virial surface in *polygonal tiles*. In the Caspar–Klug construction,

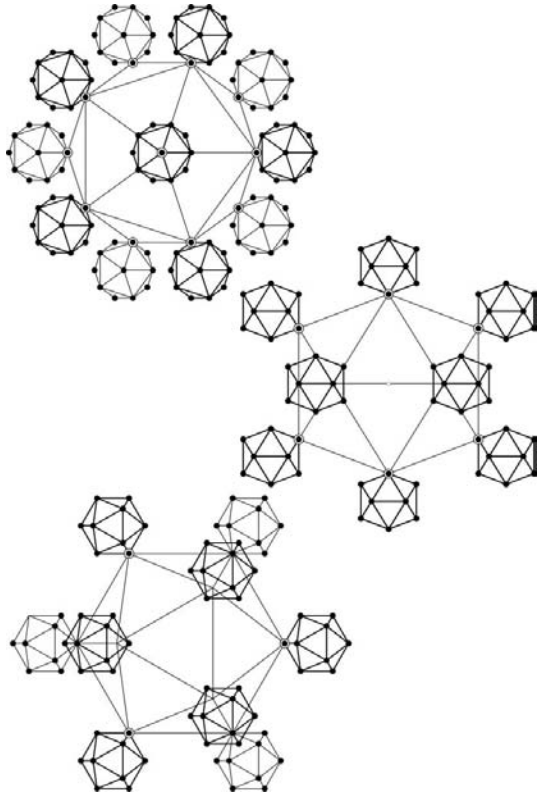

Figure 6
The 12 off-center (satellite) icosahedra attached to the vertices of a central icosahedron are obtained by combining the rotations of the icosahedral group 235 with linear and radial scalings. In the present case, the scaling factor is τ and the generating points are [200000], [110101], $[120\bar{1}00]$ and [111111].

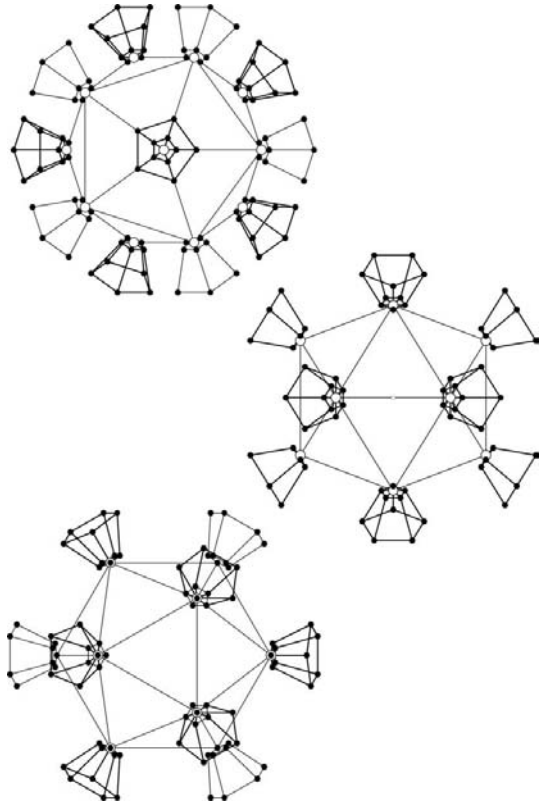

Figure 7
Icosahedron with vertices decorated by truncated pentagonal pyramids. The generating points are at [600000], $[7111\bar{1}1]$ and $[111333\bar{3}]$. This decoration could be a model for viruses with protruding proteins like the spiroplasma virus.

the tiles are triangular, possibly further subdivided to take into account the different ternary structures of the monomers. The last case is exemplified by a trapezoidal partition of the triangular facets of the CK polyhedron with triangulation number $T = 3$ defining the arrangement of the major capsid proteins VP1, VP2 and VP3 of the common cold virus (rhinovirus) (Rossmann *et al.*, 1987).

In a recent paper, the structural puzzle represented by the 72 pentamers in the polyoma virus and the simian virus 40 has been solved by considering a Penrose-like tessellation of the icosahedron in rhombus and kite tiles (Twarock, 2004). In doing so, the author R. Twarock generalizes the quasi-equivalence of Caspar & Klug.

Alternatively, the structural organization of the capsid of several serotypes of the human rhinovirus has been approached in terms of three-dimensional polyhedral forms instead of two-dimensional polygonal tiles (Janner, 2006*a*). These polyhedra enclose structural units, ranging from individual monomers to the full molecular capsid. The quasi-equivalence of Caspar & Klug has been replaced by crystallographic scale rotations, like those considered in the previous sections, relating polyhedral vertices at points of an icosahedral lattice. As has been shown in §4, the Caspar–Klug construction appears as a special case. The trapezoidal tiling of Rossmann *et al.* (1987) is consistent with the crystallographic approach applied to the ico-dodecahedron (Fig. 8*a*), which is a better polyhedral enclosing form for the rhinovirus than the icosahedron of Caspar & Klug for $T = 3$ adopted in Rossmann *et al.* (1987).

The Penrose-like icosahedral tiling of Twarock for $T = 7d$ can also be obtained from polyhedral vertices at icosahedral lattice points (Fig. 8*b*). Both tessellations already appear in the case of 60 equal monomers, as illustrated in Figs. 9(*a*),(*b*). While the trapezoidal tessellation directly follows from the Caspar–Klug construction, this is not the case for the kite tiling of the dodecahedron.

In all these cases, the vertices are related by crystallographic scale rotations. In the nomenclature adopted in Janner (2005*a*,*b*,*c*), one can say the capsid of these viruses represents a strongly correlated structure. It means that only one parameter suffices for expressing the metrical properties of the enclosing polyhedron.

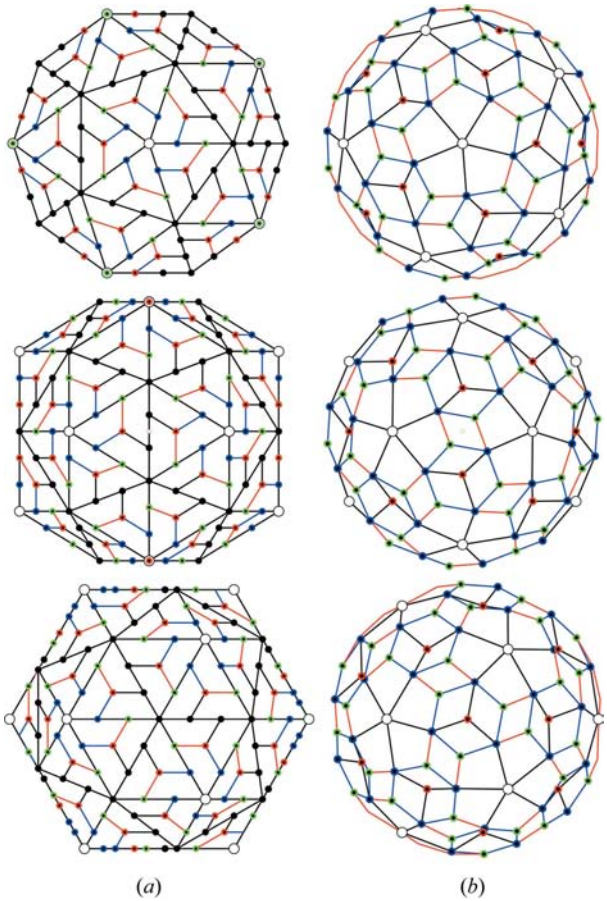

Figure 8
(*a*) Trapezoidal tessellation of the $T = 3$ ico-dodecahedron generated from [200000] and $[111\bar{1}1\bar{1}]$, as in Fig. 5. Tessellation points are $\frac{1}{3}[5111\bar{1}1\bar{1}]$, $\frac{1}{3}[313\bar{1}1\bar{1}]$, $\frac{2}{3}[211\bar{1}1\bar{1}]$ and $\frac{2}{3}[201000]$. (*b*) Rhombus and kite tessellation fitting to the $T = 7d$ polyoma virus in the way indicated by Twarock (2004). The generating points are 7[100000], $7[111\bar{1}1\bar{1}]$, $[\bar{1}35315]$ and $\frac{1}{3}[17, 5, \bar{1}, 5, \bar{17}, \bar{1}]$.

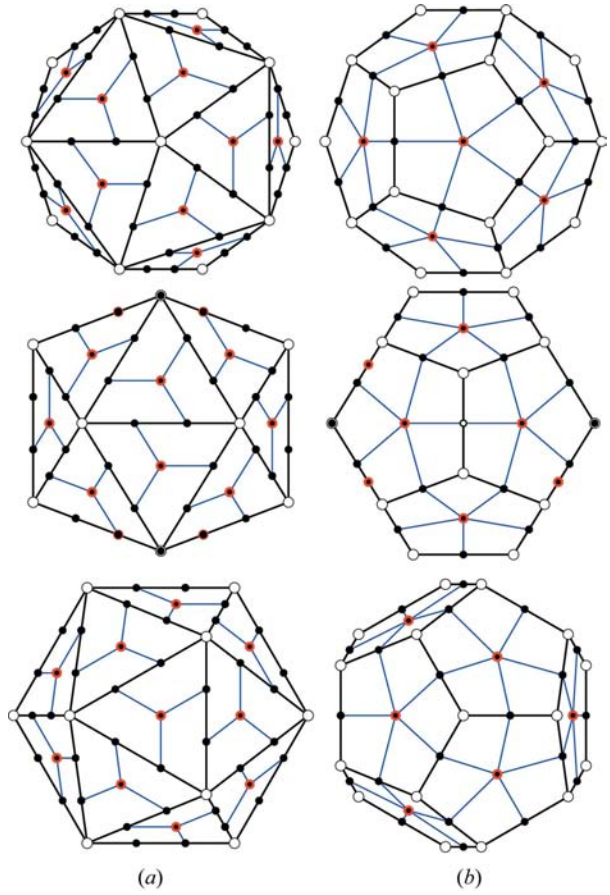

Figure 9
(*a*) Trapezoidal tessellation of the icosahedron into 60 facets with generating points [300000], [111000] and [210000]. (*b*) Corresponding kite tessellation of the dodecahedron.

5.2. Form polyhedra

The standard classification of icosahedral viruses according to the triangulation number T of Caspar & Klug is adapted to the icosahedral point symmetry of the capsid, but not always to its shape. The virial morphology is better characterized by a *form polyhedron*, as in crystal growth forms.

The capsid is delimited (within a reasonable accuracy) by two or more such polyhedra: an external one enclosing the virion, an internal one delimited by the core and possibly additional ones for layers of protein subunits (Caspar & Klug, 1962). Three illustrative examples are presented with form polyhedra given by an icosahedron, a dodecahedron and an ico-dodecahedron (triacontahedron), respectively.

The first example is one of the smallest icosahedral viruses, the satellite panicum mosaic virus, with structure determined by Ban & McPherson (1995) (PDB 1stm). It has a $T = 1$ icosahedron as enclosing form (Fig. 10, upper part). The form polyhedron of the core is also an icosahedron in a radial scaling relation by a factor $1/\tau$ with the external icosahedron (Fig. 11). One sees that free ends of the coat proteins cross the core boundary, mostly in a plane perpendicular to a twofold axis. Nevertheless, these form icosahedra keep their structural meaning.

The second example, again a $T = 1$ virus, is the canine parvovirus, with a structure determined by Xie & Chapman (1996) (PDB 4dpv). The form polyhedron is a dodecahedron (Fig. 10, middle part), so that the natural tiling is not a triangle but a kite (not shown in Fig. 10, see Fig. 9*b*). Only a fraction of the coat proteins involved in the various equatorial regions of the capsid are delimited by two dodecahedra in a $1/\tau$-scaling relation, one external and one internal (Fig. 12, left-hand side). The plot of the remaining equatorial coat proteins, in axial projection on the right-hand side of Fig. 12, suggests the 1:2 scaling ratio indicated. This is, however, only an effect of the projection of chain segments which extend far beyond the corresponding equatorial plane.

The bacteriophage MS2 is the third example. It is a $T = 3$ virus with structure determined by Golmohammadi *et al.* (1993) (PDB 2ms2). This selection has been suggested by G. Vriend because the coat protein has a fold that is different from the fold of many other virial coat proteins and, in particular, to that of another $T = 3$ virus, the rhinovirus whose

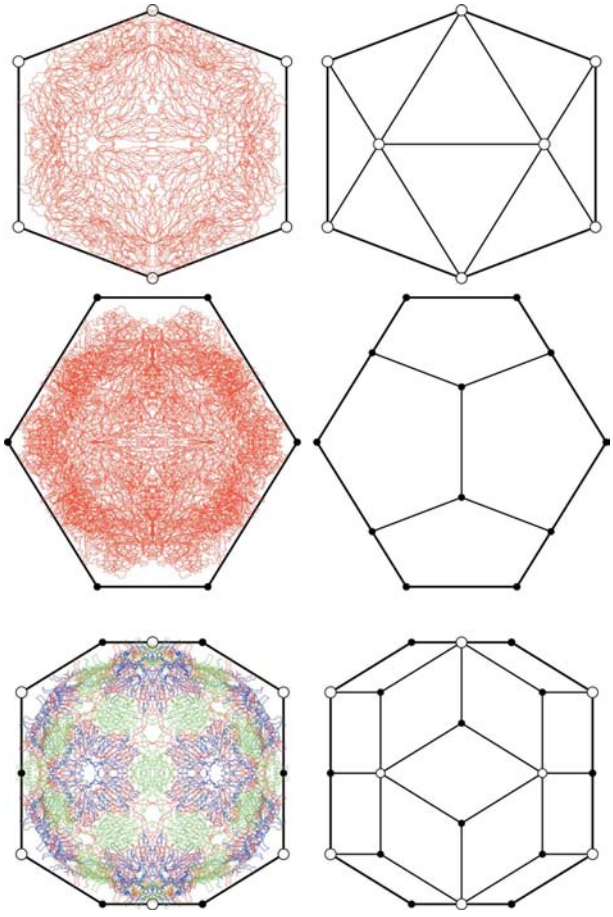

Figure 10
Three different form polyhedra enclosing icosahedral viruses are shown in a projection along the twofold axis: an icosahedron (upper part), a dodecahedron (middle part) and a triacontahedron (lower part) delimiting the capsid of the $T = 1$ satellite panicum mosaic virus, the $T = 1$ canine parvovirus and the $T = 3$ bacteriophage MS2, respectively. The icosahedral vertices are indicated by open circles, the dodecahedral ones by filled circles.

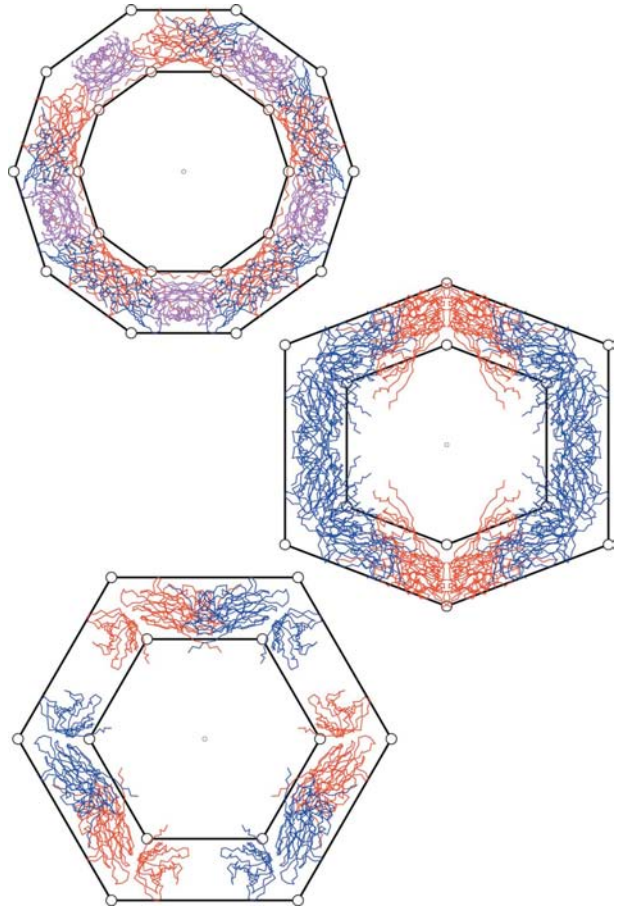

Figure 11
Two icosahedra, related by a radial scaling with factor τ, delimit the capsid of the $T = 1$ satellite panicum mosaic virus. Free ends of the monomers penetrate in projection into the core region, mainly along a plane perpendicular to the twofold axis. Only the monomers of the corresponding equatorial regions are plotted.

enclosing form is a triacontahedron (Janner, 2006*a*). Despite the difference in folding and in the secondary structure of the proteins involved, the form polyhedron of MS2 is also a triacontahedron (Fig. 10, lower part). Both have a triacontahedron as internal polyhedral forms as well. For the rhinovirus, the scaling ratio between the external and the internal triacontahedra is $1/\tau$ (Janner, 2006*a*). In the original version of this paper, the same ratio was also indicated for MS2, but this is actually not the case, as remarked by R. Twarock. Evidence is given in Janner (2006*b*) that for MS2 the scaling ratio is 3/4 and the internal surface of the capsid is practically spherical (this last aspect also appears in the equatorial slabs of MS2 plotted by a PhD student of Twarock). Therefore, the capsid is enclosed in the external triacontahedron and circumscribes the internal one. For the rhinovirus and for MS2, the vertices of each pair of polyhedra are at points of an icosahedral lattice.

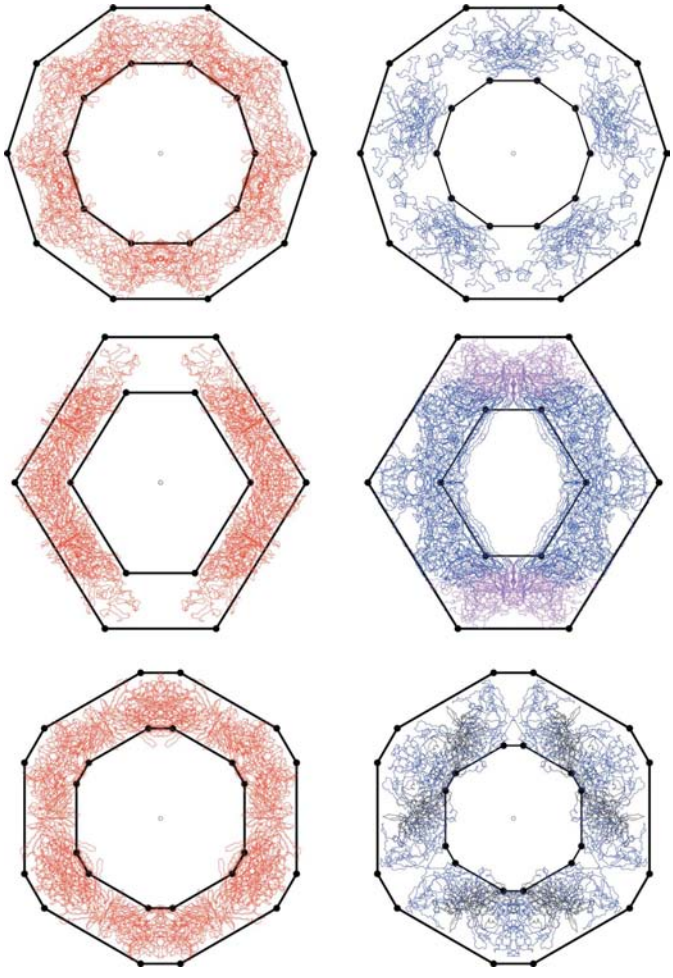

Figure 12
Part of the monomers in the equatorial regions of the capsid in the canine parvovirus is enclosed between two dodecahedra in a factor $1/\tau$ scaling relation (left-hand side). The scaling factor between the dodecahedra delimiting in projection the remaining equatorial monomers is 1/2 (right-hand side). This smaller ratio reflects a projection effect due to chain segments extending far beyond the equatorial plane and does not correspond to a core property (Janner, 2006*b*).

6. Final remarks

As has been shown, there is an incredible variety of polyhedra with icosahedral point symmetry having vertices and other decoration points with rational indices (this means at points of an icosahedral lattice). They are easily implemented by applying to few generating positions with integer indices the transformations of the icosahedral group. In all the cases considered, the generated points at different radial distance from the center are related by crystallographic scaling transformations of different type (radial, planar or linear), which have a rational matrix representation when expressed in an icosahedral lattice basis. The points of the decorated polyhedra are, therefore, related by crystallographic icosahedral scale rotations. By these linear transformations, one obtains not only edge and face decorations but also off-center (satellite) polyhedra attached to the vertices of a starting polyhedron with icosahedral symmetry (vertex decoration).

The approach already allows a geometrical classification of icosahedral viruses which includes as special case the classification of Caspar & Klug (1962) in terms of a triangular surface net and other polygonal tiling of icosahedral polyhedra, like the rhombus and kite Penrose-like tessellation considered by Twarock for solving the puzzle of viruses not obeying the Caspar–Klug rules (Twarock, 2004). One even gets a kite tessellation for the simplest dodecahedral virus case (Fig. 9*b*).

These ideas are illustrated by three simple icosahedral viruses: the bacteriophage MS2, the satellite panicum mosaic virus and the canine parvovirus, in addition to the rhinovirus already considered in a previous work (Janner, 2006*a*). In all these cases, the internal form polyhedron (delimited by the core) is related to the external one of the capsid by a crystallographic radial scaling with scaling factors $1/\tau$ or 3/4.

What has been presented is the starting point only of a more comprehensive classification of the known icosahedral viruses as listed, for example, at the web site of *Virus World* (http://virology.wisc.edu/virusworld). There, the viruses are classified according to the Caspar & Klug triangulation numbers, independently of whether the Caspar–Klug rules are obeyed or not. The flexibility of the present approach should also allow the investigation of whether individual C_αs selected by extremum conditions (like the radial distance) occur at symmetry-related positions with rational indices, in a way similar to what has been shown to be the case for the heptagonal GroEL–GroES chaperonin complex (Janner, 2003). One expects that the C_αs at these special positions play a privileged rôle in the overall structural stability and in the local conformational changes required by the transition from an empty procapsid to the capsid filled with viral genome or even in the assembly process of a viral capsid.

References

Ban, N. & McPherson, A. (1995). *Nature Struct. Biol.* **2**, 882–890.
Caspar, D. L. D. & Klug, A. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
Chipman, P. R., Agbandje-McKenna, M., Renaudin, J., Baker, T. S. & McKenna, R. (1998). *Structure*, **6**, 135–145.

Coxeter, H. S. M. (1961). *Introduction to Geometry*. New York: John Wiley.
Coxeter, H. S. M. (1963). *Regular Polytopes*, 2nd ed. New York: MacMillan.
Elser, V. (1985). *Phys. Rev. B*, **32**, 4892–4898.
Golmohammadi, R., Valegård, K., Fridborg, K. & Liljas, L. (1993). *J. Mol. Biol.* **234**, 620–639.
Janner, A. (2002). *Struct. Chem.* **13**, 277–287.
Janner, A. (2003). *Acta Cryst.* D**59**, 795–808.
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*a*). *Acta Cryst.* A**62**, 270–286.
Janner, A. (2006*b*). To be published.
Rossmann, M. G., Arnold, E., Griffith, J. P., Kramer, G., Luo, M., Smith, T. J., Vriend, G., Rueckert, R. R., Sherry, B., McKinlay, M. A., Diana, G. & Otto, M. (1987). *Trend Biochem. Sci.* **12**, 313–318.
Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. (1984). *Phys. Rev. Lett.* **53**, 1951–1953.
Steinhardt, P. J. & Ostlund, S. (1987). *The Physics of Quasicrystals.* Singapore: World Scientific.
Stewart, P. L., Burnett, R. M., Cyrklaff, M. & Fuller, S. D. (1991). *Cell*, **67**, 145–154.
Twarock, R. (2004). *J. Theor. Biol.* **226**, 477–482.
Xie, Q. & Chapman, M. S. (1996). *J. Mol. Biol.* **264**, 497–520.