

A new series of polyhedra as blueprints for viral capsids in the family of Papovaviridae

T. Keef¹ and R. Twarock^{1,2}

¹*Department of Mathematics
University of York*

³*Department of Biology
University of York
York YO10 5DD, U.K.*

Abstract

In a seminal paper [1] Caspar and Klug established a theory that provides a family of polyhedra as blueprints for the structural organisation of viral capsids. In particular, they encode the locations of the proteins in the shells that encapsulate, and hence provide protection for, the viral genome. Despite of its huge success and numerous applications in virology experimental results have provided evidence for the fact that the theory is too restrictive to describe all known viruses [2]. Especially, the family of Papovaviridae, which contains cancer causing viruses, falls out of the scope of this theory.

In [3] we have shown that certain members of the family of Papovaviridae can be described via tilings. In this paper, we develop a comprehensive mathematical framework for the derivation of *all* surface structures of viral particles in this family. We show that this formalism fixes the structure and relative sizes of all particles collectively so that there exists only one scaling factor that relates the sizes of *all* particles with their biological counterparts.

The series of polyhedra derived here complements the Caspar-Klug family of polyhedra. It is the first mathematical result that provides a common organisational principle for different types of viral particles in the family of Papovaviridae and paves the way for an understanding of Papovaviridae polymorphism. Moreover, it provides crucial input for the construction of assembly models along the lines of [4, 5].

1 Introduction

Icosahedral symmetry plays a fundamental role for the structure of viruses because it constrains the organisation of the proteins in the viral capsids that protect the viral genome. Based on this observation, Caspar and Klug have developed a landmark theory [1] in which they derive a family of icosadeltahedra, i.e. polyhedra with icosahedral symmetry and faces given by equilateral triangles, that act as blueprints for the organisation of viral capsids. In particular, the corners of the triangular faces of these polyhedra encode the locations of the protein subunits in the capsids. This implies that proteins are organised in clusters of 5 or 6 protein subunits, called *pentamers* and *hexamers*, respectively. The polyhedra consist of $20T$ triangular facets, where $T = h^2 + hk + k^2$ with $h \in \mathbb{N} \cup \{0\}$, $k \in \mathbb{N}$ denotes the *triangulation*

¹E-mail: tk506@york.ac.uk

²E-mail: rt507@york.ac.uk

number that parameterises the family of polyhedra. As a consequence, there are precisely 12 pentamers and $10(T - 1)$ hexamers in a viral capsid corresponding to the polyhedron of triangulation number T in this family, and the corresponding capsid contains $60T$ protein subunits.

The family of polyhedra established in Caspar-Klug Theory is fundamental in virology and has a plethora of applications ranging from image analysis of experimental data to the construction of assembly models. Despite its huge success experimental evidence has shown that it is too restrictive to account for all known viruses. In particular, viruses in the family of Papovaviridae fall out of the scope of Caspar-Klug Theory and their organisation has therefore been a long-standing open problem in virology [6, 7], which has been formulated by Liddington et al. in 1991 as follows: “The puzzle is how do the coloured pentamers fit into the hexavalent holes?”. In particular, the experiments show that also the protein clusters located off the global 5-fold axes of icosahedral symmetry can be pentamers, which is by construction excluded by the approach of Caspar and Klug because they are working with a hexagonal lattice.

We have provided a solution to this puzzle based on tiling theory in [3, 8], where we have constructed a polyhedron that describes the surface structure of the viral particles observed in [6, 7]. By construction, these models not only encode the locations of the protein subunits but also those of the inter-subunit bonds that connect different pentamers in the capsid. These results have provided the basis for the construction of assembly models for Papovaviridae in [4, 5], and the tiling approach has moreover paved the way for the study of crosslinking structures [9].

It is the purpose of this work to **develop a mathematical framework for the systematic construction of polyhedra associated with viral capsids composed of pentamers throughout, and hence to derive a family of polyhedra that constitute the exceptional cases needed to complement the Caspar-Klug series of polyhedra for the description of viruses.**

Since we are seeking polyhedra that encode the locations of pentamers off the 5-fold axes of the icosahedral group, it is not possible to work with a hexagonal lattice as in the case of the Caspar-Klug construction, and a completely different mathematical approach is hence required. In particular, the coordinates of the vertices have to be retrieved from an icosahedral lattice, i.e. a lattice that is invariant under the action of the icosahedral group. Such a lattice does not exist in three dimensions, but can be inferred from a six-dimensional space via projection. This procedure is known from the study of aperiodic structures [10], which describe the locations of atoms in alloys called quasicrystals [11]. An additional complication arises due to the fact that such projections only lead to discrete point sets if the number of lattice points projected from the higher-dimensional space is restricted in an appropriate way. We use here a restriction that is rooted in the structure of the icosahedral group itself, and can be obtained via a so-called affine extension of that group. It can be used to construct a family of finite dimensional point sets that act as blueprints for the coordinates of the desired polyhedra.

We obtain three types of polyhedra in this way that correspond to the three species of particles observed in the family of Papovaviridae. The ratios of their radii are completely determined by the mathematical formalism, so that there exists only one scaling parameter that relates the geometries of all polyhedra collectively to the biological setting. It can be used to test the predictions of our theory as we discuss in section 6.

The smallest polyhedron in the series is a triacontahedron, and we therefore call our series of exceptional polyhedra the *triacontahedral series*. The 30 faces of the triacontahedron are rhombs. By subdividing each rhomb into two triangles, it can be seen that the triacontahedron is structurally similar to the $T = 3$ capsid in Caspar-Klug Theory. However, the coordinates of the vertices are of a different type: while the coordinates of the triacontahedron can be obtained from a higher-dimensional icosahedral lattice via projection, this is not the case for the $T = 3$ Caspar-Klug structure. Moreover, the vertices of the polyhedra in the triacontahedral series are located on nested spherical shells rather than a single spherical shell, so that these particles are only approximately spherical. For example, the triacontahedron has 12 vertices (corresponding to an icosahedron) located on one shell and 20 vertices (corresponding to a dodecahedron) located on a different shell within.

The polyhedra corresponding to the medium and large sized particles have vertices on three different nested shells. The vertices at which 5 faces meet mark the centres of the pentamers, and the locations of the protein subunits correspond to the angles of the corresponding faces. In combination with the tiling approach [citeTwarock:2004a,Twarock:2005a](#) it is moreover possible to deduce the bonding structure of the viral capsids they represent as we discuss in section 6.

Our mathematical set-up moreover provides a framework for the study of the scaling transformations and rotations that map the coordinates of the polyhedra on other points induced from the higher-dimensional icosahedral lattice by projection. This is important because viral capsids are three- rather than two-dimensional objects, and these transformations can be used to associate a three-dimensional structure with the two-dimensional surfaces of the polyhedron that provides the blueprint for the capsid. This has been demonstrated in [12] for the case of Human Rhinovirus. The viral capsid of this virus follows the blueprint of our small shell, the triacontahedron. It has been shown in this reference that the viral capsid is contained between two copies of the triacontahedron that are related by a scaling by the irrational number $\tau := \frac{1}{2}(1 + \sqrt{5})$, which corresponds to a transformation of the type we are considering here. In section 5 we show how these transformations can be derived within our formalism for all viruses in the triacontahedral series. This paves the way for a construction of three-dimensional structures associated with the blueprints provided by the polyhedra for all viruses in this series, and hence to derive the locations also of the proteins that are located within the capsid and delimit the cavity filled by the genetic material. Such information is invaluable for example for the study of scaffold mediated assembly.

The paper is organised as follows. In section 2 we outline the projection formalism that connects a generalised lattice (or quasi-lattice) based on the simple root vectors of the non-crystallographic Coxeter group H_3 with the root lattice of D_6 . In section 3 we derive finite dimensional subsets of this generalised lattice based on an affine extension of the non-crystallographic Coxeter group H_3 . In section 4 we use these sets for the construction of the triacontahedral series of polyhedra. The projection picture is extended in section 5 and a formalism is established that allows us to determine the range of possible scaling transformations, rotations and combinations thereof that map the vertices of the polyhedra on other points in the generalised lattice obtained via projection, because this provides the basis for the construction of three-dimensional models associated with these polyhedra along the lines of [12]. A comparison with experimental results is provided in section 6. Finally, we conclude with a discussion of the mathematical and biological implications of our results.

2 Points sets induced from an icosahedral lattices by projection

We are seeking polyhedra that are symmetric under the icosahedral group. Since the icosahedral group does not stabilize a lattice in three dimensions, a 3-dimensional generalised lattice or quasi-lattice has to be inferred from a higher dimensional crystallographic lattice via projection.

We use the fact that the icosahedral group is crystallographic in 6 dimensions, and construct our quasi-lattice from the root lattice of D_6 via projection. We remark that it would also have been possible to construct a quasi-lattice in three dimensions via a projection from \mathbb{Z}^6 rather than D_6 as in [12]. However, our approach is more general, because the root lattice of D_6 allows for an embedding with maximal symmetry [13].

Let $\mathbf{e} = \{e_j | j = 1, \dots, 6\}$ denote the standard basis in 6 dimensions with $(e_i, e_j) = \delta_{ij}$, $i, j = 1, \dots, 6$. The root lattice of D_6 corresponds to the \mathbb{Z} -linear span of the 6 simple root vectors (or roots in short) of the Coxeter group D_6 . They can be expressed in terms of the standard basis \mathbf{e} as follows:

$$\begin{aligned} a_1 &= e_2 - e_1 & a_2 &= e_1 - e_3 \\ a_3 &= e_3 - e_6 & a_4 &= e_5 + e_6 \\ a_5 &= -e_4 - e_5 & a_6 &= e_4 - e_5 \end{aligned} \tag{1}$$

We choose a projection to 3 dimensions that maps the basis \mathbf{e} on the vectors pointing to the six non-aligned vertices of an icosahedron. The coordinates of these vectors are given in terms of the 3 dimensional standard basis below, using $\tau = \frac{1}{2}(1 + \sqrt{5})$:

$$\begin{aligned} e_1 &\mapsto \frac{1}{2}(1, 0, \tau) & e_2 &\mapsto \frac{1}{2}(\tau, 1, 0) \\ e_3 &\mapsto \frac{1}{2}(0, \tau, 1) & e_4 &\mapsto \frac{1}{2}(-1, 0, \tau) \\ e_5 &\mapsto \frac{1}{2}(0, -\tau, 1) & e_6 &\mapsto \frac{1}{2}(\tau, -1, 0). \end{aligned} \tag{2}$$

Hence, by virtue of (1) and (2), the vectors a_j , $j = 1, \dots, 6$, project on the vectors in (3). Coordinates are again expressed in terms of the standard basis in 3 dimensions, with the notation $\tau' = \frac{1}{2}(1 - \sqrt{5})$ for the Galois conjugate of τ and the identity $\tau + \tau' = 1$:

$$\begin{aligned} a_1 &\mapsto \tilde{a}_1 = \frac{1}{2}(-\tau', 1, -\tau) & a_2 &\mapsto \tilde{a}_2 = \frac{1}{2}(1, -\tau, -\tau') \\ a_3 &\mapsto \tau\tilde{a}_3 = \frac{1}{2}(-\tau, -\tau^2, 1) & a_4 &\mapsto \tau\tilde{a}_2 = \frac{1}{2}(\tau, -\tau^2, 1) \\ a_5 &\mapsto \tau\tilde{a}_1 = \frac{1}{2}(1, -\tau, -\tau^2) & a_6 &\mapsto \tilde{a}_3 = \frac{1}{2}(-1, \tau, -\tau'). \end{aligned} \tag{3}$$

\tilde{a}_1 , \tilde{a}_2 and \tilde{a}_3 correspond to the root vectors of the non-crystallographic Coxeter group H_3 .

The projection π_{\parallel} in (3) is illustrated in terms of the Dynkin diagrams of D_6 and H_3 in Fig. 1. On the left, the Dynkin diagram of D_6 is shown, in a folded and hence slightly unconventional form, to illustrate how the simple root vectors of D_6 project on the root vectors of H_3 . In particular, the 5 above the link in the Coxeter diagram of H_3 on the right indicates that the angle between the corresponding root vectors is $\pi - \frac{\pi}{5}$, with the 3 over the other link omitted by convention. We remark that this procedure is similar to the projection from E_8 on H_4 considered in [14].

The root lattice of D_6 is given by \mathbb{Z} -linear combinations of the simple roots a_j , $j = 1, \dots, 6$. Correspondingly, its projection into \mathbb{R}^3 via π_{\parallel} is given by $\mathbb{Z}[\tau]$ -linear combinations of the

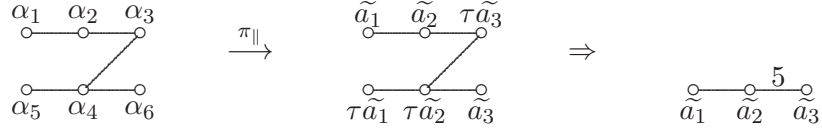


Figure 1: Figure illustrating the projection π_{\parallel} of the simple roots of D_6 on those of H_3 .

simple root vectors \tilde{a}_j , $j = 1, \dots, 3$, of H_3 , where $\mathbb{Z}[\tau]$ denotes the extended ring of integers

$$\mathbb{Z}[\tau] := \{a + \tau b \mid a, b \in \mathbb{Z}\}. \quad (4)$$

Since $\mathbb{Z}[\tau]$ is dense in \mathbb{R} , the $\mathbb{Z}[\tau]$ -linear combinations of the simple root vectors of H_3 are dense in \mathbb{R}^3 . Thus it is necessary to select subsets thereof which are suitable to guide the search for the vertex sets of the polyhedra. Such sets will be derived in the following section based on an affine extension of H_3 .

3 Nested shell structures via an affine extension of H_3

The simple root vectors \tilde{a}_j , $j = 1, \dots, 3$, in (3) form a basis of the root system of H_3 [15]. The complete root system Δ is given by [16]

$$\Delta = \left\{ \begin{array}{ll} (\pm 1, 0, 0) & \text{and all permutations} \\ \frac{1}{2}(\pm 1, \pm \tau, \pm \tau') & \text{and all even permutations} \end{array} \right\}. \quad (5)$$

The root vectors in Δ point to the vertices of an icosidodecahedron. They encode the generators of the Coxeter group H_3 as follows. Let $\alpha \in \Delta$ and let $(\cdot | \cdot)$ denote the scalar product in \mathbb{R}^3 . Then

$$r_{\alpha} : x \mapsto x - \frac{2(x|\alpha)}{(\alpha|\alpha)} \quad \text{for } x \in \mathbb{R}^3 \quad (6)$$

is the Euclidean reflection in the plane orthogonal to the root vector α . The icosidodecahedron representing the root system Δ and two of the reflection planes that are encoded by the root vectors are illustrated in Fig. 2. The intersection of the two planes in the figure corresponds to an axis of 5-fold symmetry, which intersect with the sphere at two of the 12 5-fold vertices of a (spherical) icosahedron. The other reflection planes are not shown, but their intersections with the surface of the sphere are indicated as geodesics (i.e. as spherical arcs obtained as intersections of planes through the centre of the sphere with the surface of the sphere). The intersections of the geodesics mark the locations of all symmetry axes, and one can hence reconstruct the locations of all 6 five-fold, 15 two-fold and 10 three-fold axes of rotational symmetry of the icosahedral group.

Note that in contrast to the case of Weyl groups $\frac{2(\alpha|\beta)}{(\alpha|\alpha)} \in \mathbb{Z}[\tau]$ (instead of \mathbb{Z}) for all $\alpha, \beta \in \Delta$. Therefore, \mathbb{Z} -linear combinations of root vectors in Δ do not form a crystallographic lattice, and Δ is therefore called *non-crystallographic*. However, it is possible to select subsets which form generalised lattices such as the quasi-lattices known from quasicrystals [10].

In order to extend the root system in a way compatible with overall icosahedral symmetry the basis of simple root vectors has to be extended. For this, we use the fact that the relations

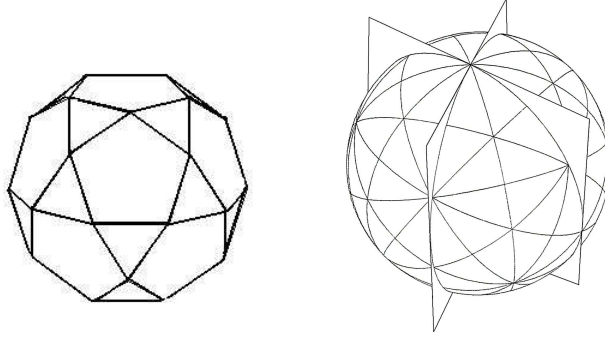


Figure 2: The root polytope (left) encoding the locations of the planes of reflection (right).

between the simple root vectors are encoded in the Cartan matrix C ,

$$C := \left(\frac{2(\alpha_i | \alpha_j)}{(\alpha_j | \alpha_j)} \right)_{ij} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -\tau \\ 0 & -\tau & 2 \end{pmatrix}. \quad (7)$$

We extend this matrix by an additional row and column via a formalism known as affine extensions in the theory of Kac-Moody algebras. The only difference here is that entries stem from the set $\mathbb{Z}[\tau]$ rather than \mathbb{Z} . We have shown in [17] that the affine extended Cartan matrix of H_3 is given by

$$\hat{C} = \begin{pmatrix} 2 & 0 & \tau' & 0 \\ 0 & 2 & -1 & 0 \\ \tau' & -1 & 2 & -\tau \\ 0 & 0 & -\tau & 2 \end{pmatrix}. \quad (8)$$

The additional row and column encode a further root vector corresponding to an affine reflection that acts as a translation T by the highest root vector $\alpha_H = \tau\alpha_1 + 2\tau\alpha_2 + \tau^2\alpha_3$. The three other reflections are cyclic operations of order two and the products of any two of them correspond to rotations around the origin as follows:

$$(r_j r_k)^M = 1 \quad \text{where} \quad \begin{cases} M = 1 & \text{if } c_{jk} = 2 \\ M = 2 & \text{if } c_{jk} = 0 \\ M = 3 & \text{if } c_{jk} = -1 \\ M = 5 & \text{if } c_{jk} = -\tau. \end{cases} \quad (9)$$

The affine extended group is hence generated by the three reflections r_1, r_2, r_3 and the translation T .

It is now possible to construct, based on these operations, finite point sets with icosahedral symmetry such that all points in the sets have a counterpart in the root lattice of D_6 . In this sense, they are finite subsets of generalised lattices induced by projection from D_6 and correspond to potential candidates for the locations of the polyhedra we are seeking.

The point sets are generated iteratively via the action of the generators of the extended group on the origin 0. If the action of the translation operator T is not restricted, one obtains an infinite point sets that densely fills \mathbb{R}^3 . However, if T acts only a finite number of times, N

say, while the action of all other operations is not restricted in order to guarantee icosahedral symmetry of the overall configuration, finite dimensional point sets $\mathcal{S}(N)$ are obtained which become larger and more dense with increasing N .

Due to the fact that T acts as a translation by the highest root vector it has been possible to derive a simple expression for the point sets $\mathcal{S}(N)$ in terms of the root system Δ [18]:

$$\mathcal{S}(N) = \left\{ \sum_{\alpha \in \Delta} n_{\alpha} \alpha \mid n_{\alpha} \in \mathbb{N} \cup \{0\}, \sum_{\alpha \in \Delta} n_{\alpha} \leq N \right\}. \quad (10)$$

In particular, $\mathcal{S}(N)$ consists of all $\mathbb{N} \cup \{0\}$ -linear combinations of up to N root vectors in Δ . N is called the *cut-off parameter* because it limits the number of points in the set.

The sets $\mathcal{S}(N)$ play a crucial role in the following for the construction of the polyhedra.

4 Construction of the triacontahedral series of polyhedra from H_3 -induced shell models

By construction, the sets $\mathcal{S}(N)$ contain the vertices of the polyhedra in the triacontahedral series. Due to the fact that we are looking for polyhedra that represent viral capsids given in terms of pentamers that are (with a certain experimental error) equidistant from their neighbouring pentamers, the vertices of the desired polyhedra follow (in a good approximation) the vertices of the Platonic and Archimedean solids, because they correspond to the polyhedra with equidistant edge lengths. They can hence be used as templates for the search of the coordinates of the polyhedra within the sets $\mathcal{S}(N)$. Moreover, since we are looking for polyhedra that mark the locations of pentamers, only those Platonic and Archimedean solids are relevant that have 5-coordinated vertices, i.e. vertices at which 5 edges meet. The three solids that fulfil this criterion are the icosahedron, the snub cube and the snub dodecahedron.

We have computed the sets $\mathcal{S}(N)$ up to $N = 5$ explicitly. The points are organised on 181 nested shells of radii between $R \approx 0.2361$ and $R = 5$. $\mathcal{S}(5)$ contains all point sets $\mathcal{S}(N)$ with $N \leq 5$ by construction, and has been chosen because it is sufficiently dense to provide coordinates for the polyhedra. We have used the icosahedron, the snub cube and the snub dodecahedron, respectively, as a template to search for subsets of $\mathcal{S}(5)$ with the desired distribution of points. We have thus obtained those coordinates of our polyhedra, that mark the locations of 5-coordinated vertices. We have then determined further vertices (not necessarily on a shell of the same radius) which together with these vertices from the vertex set of the polyhedron. The procedure is discussed explicitly in the following subsections for the individual cases and the coordinates of the vertices corresponding to centres of pentamers are provided in the Appendix.

4.1 The small species

The vertex set of the icosahedron occurs for the first time in the set $\mathcal{S}(3)$ on a shell of radius $R \approx 1.1756$. There are two ways of obtaining polyhedra that match this vertex set. One of them is the icosahedron itself and corresponds to the start of the Caspar-Klug series. The other one is the rhombic triacontahedron with 30 rhombic faces. The remaining vertices of

this polyhedron correspond to the vertices of a dodecahedron located on shell of a radius $R \approx 1.0705$ (i.e. there is a scaling factor of about 1.098 between the shells that contain the two different types of vertices). The triacontahedron is shown in Fig. 3.

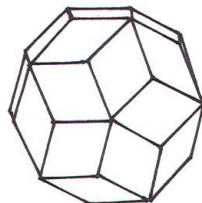


Figure 3: The triacontahedron corresponding to the polyhedron of the small species.

Since only those vertices at which 5 faces meet mark the locations of pentamers, both cases correspond to capsids with 12 pentamers located on the axes of 5-fold symmetry of the icosahedral group. However, only the centres of the pentamers coincide in both cases, and the *orientations* of the pentamers are *different*: because the proteins are located in the corners of the faces (shown schematically as dots in Fig. 4) they differ by a rotation by an angle of $\frac{2\pi}{10}$ as illustrated in this figure. It shows a face of the icosahedron with three protein subunits on the left versus three faces of the triacontahedron on the right with 2 protein subunits each. A superposition of both figures shows that the respective proteins are rotated with respect to each other.

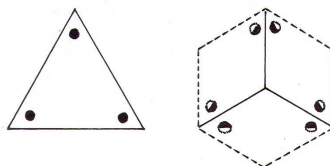


Figure 4: The locations of proteins on the triangular faces of the icosahedron (left) are rotated by an angle of $\frac{2\pi}{10}$ with respect to the locations of the proteins on the rhombs (right).

This difference has far-reaching consequences for the structure of the capsids. In particular, in one of the cases crosslinking structures are possible while they cannot occur for geometrical reasons in the other case, as has been demonstrated in [9].

The case of the triacontahedron is hence essentially different from the $T = 1$ case in the Caspar-Klug family and marks the start of a different series of polyhedra.

4.2 The intermediate species

Due to the fact that the snub cube has octahedral rather than icosahedral symmetry, we need to use the geometric inclusions illustrated in Fig. 5 in order to locate its vertices within the point set $S(5)$.

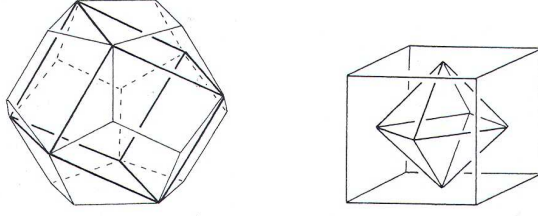


Figure 5: A cube inscribed into a dodecahedron (left) and an octahedron into a cube (right).

The points of $S(5)$ that are closest to the vertex set of the snub cube occur on a shell of radius $R \approx 2,3199$, which appears in $S(4)$ for the first time (i.e. it is not contained in a set $S(N)$ with $N < 4$). The edges are not equidistant in the case of the snub cube as expected, but the deviations are small. In particular, one obtains the following values for the three different distances A , B and C between neighbouring points in the set.

$$A \approx 1,7481, \quad B \approx 1,6625, \quad C \approx 1,8783. \quad (11)$$

The point set is illustrated in Fig. 6 (using the 3d-Grapher software) in relation to the locations of the 2-, 3- and 4-fold symmetry axes of the octahedral group. Moreover, the triangle Δ_{ABC} formed from the distances A , B and C has been superimposed.

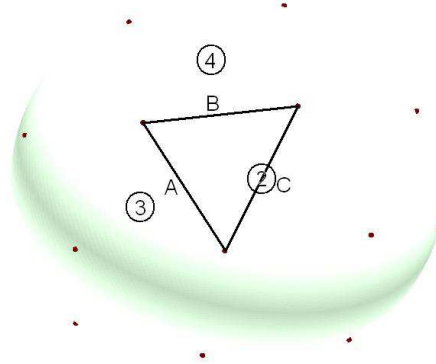


Figure 6: The triangle Δ_{ABC} in the octahedral case.

From the triangle Δ_{ABC} the polyhedron is constructed via the following argument. The line denoted as C is centred on a global two-fold axis of the symmetry group. The face containing that edge is hence 2-fold symmetric about that axis. We compute the intersection of this axis with the edge C , and determine a line that goes through that intersection point, is perpendicular to C and intersects the sphere in two points equidistant from the intersection point. We then determine the intersection of this line with the line through the global 4-fold axis (marked as an encircled 4 in Fig. 6) and the centre of edge B . The intersection is located within the sector of the sphere that is characterised by the triangle Δ_{ABC} and determines one further type of vertex of the polyhedron that does not correspond to the centre of a

pentamer. A third type of vertex is located on the global 3-fold axis, that is indicated as an encircled 3 in Fig. 6. There is flexibility in the construction in that the vertices on the global 3- and 4-fold axes do not need to be located on a shell of the same radius as the vertices representing pentamers. However, due to the fact that they are not marking locations of pentamers, their exact location is not important for our purpose. It is important to note though that a variation of the radii of the shells on which they are located causes changes in the angles of the faces.

One hence obtains the polyhedron illustrated schematically in Fig. 7. The fundamental domain of the octahedral group corresponds to the triangle between the 2-fold, 3-fold, and 4-fold axes marked as encircled 2s, 3s and 4s in Fig. 6 and the entire polyhedron can be obtained from it via the action of the octahedral group. Therefore, only a part of the polyhedron overlapping with the fundamental domain is shown, because the rest of the polyhedron is implied by symmetry. It is a schematic representation in that the vertices not marking the locations of the pentamers have been drawn on the same shell. However, this illustration shows all features of the polyhedron that are needed to read off the locations of the protein subunits, which are located in those corners of the faces that meet in multiples of 5 at a vertex.

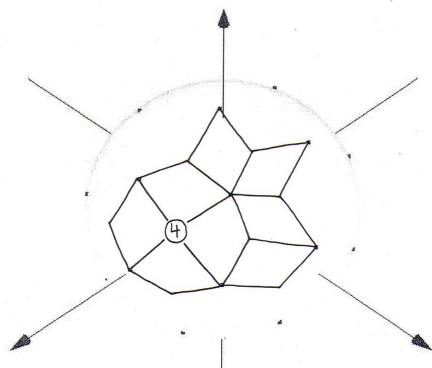


Figure 7: Schematic representation of the polyhedron corresponding to the medium species.

It is a polyhedron in terms of rhomb- and kite-shaped faces. The coordinates of the vertices that mark the centres of the pentamers are given in the Appendix. We remark that the angles of the tiles around these vertices are not equal, but the deviations are small. In particular, the angles on the kites are larger than those on the rhombs. This fact may account for the experimental observation that viral particles corresponding to such a polyhedron have only been obtained in vitro, but have never been observed in vivo.

4.3 The large species

The construction of the polyhedron corresponding to the large species is analogous to the procedure used before. In this case, we use the vertex set of the snub dodecahedron as a template. The subset of $S(5)$ closest to it is located on a shell of radius $R \approx 3,1247$ and appears in $S(5)$ for the first time.

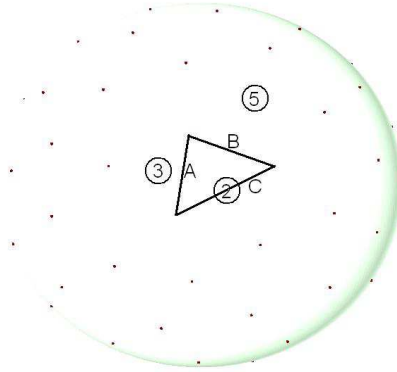


Figure 8: The triangle Δ_{ABC} in the octahedral case.

The three different distances between neighbouring points are given in (12).

$$A \approx 1,3708, \quad B \approx 1,7520, \quad C \approx 1,4364. \quad (12)$$

They define the triangle Δ_{ABC} in Fig. 8. We start again by fixing the rhomb on the 2-fold axis containing C, as well as the angles of the kite-shaped faces around the 5-fold axes. This determines the vertex within the sector corresponding to the triangle Δ_{ABC} as before. The radius of the shell that determines the locations of the global 5-fold vertices can be computed such that the angles on the kite-shaped faces at the vertices meeting rhombs match those of the rhombs containing C. Finally, the angles of the rhombs around the global 3-fold axes can be adjusted by varying the radius of the shell on which they are located. In this way, the polyhedron in Fig. 9 is obtained. As before, it is a schematic representation that encodes all important information about the polyhedron. The fundamental domain of the icosahedral group corresponds to the triangle between the 2-fold, 3-fold, and 5-fold axes marked as encircled 2s, 3s and 5s in Fig. 8 and the entire polyhedron can be obtained from it via the action of the icosahedral group. It is hence sufficient to represent only part of the polyhedron, provided that this contains the fundamental domain, in order to specify the polyhedron completely.

The coordinates of the vertices that mark the locations of the pentamers are given in the Appendix. They correspond to 12 vertices located on the 5-fold axes of icosahedral symmetry, as well as further 60 vertices off these axes.

Finally, we remark that the set $S(5)$ contains the vertices of all three types of polyhedra. Hence, the geometrical formalism fixes the sizes of the three different species of particles in relation to each other, so that there is indeed only one free parameter that relates the overall mathematical structure, i.e. the small, medium and large polyhedron collectively, to the biological setting. We determine the corresponding scaling factor in section 6.

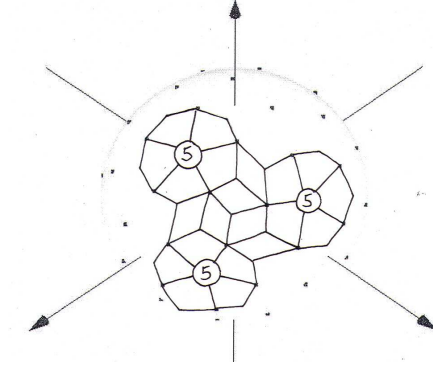


Figure 9: Schematic representation of the polyhedron corresponding to the large species.

5 Scaling transformations and rotations

In order to determine the range of possible scaling transformations, rotations and combinations thereof that map the vertices of the polyhedra onto other points in a set $S(N)$ (possibly with different N) one needs to extend the projection picture presented in section 2.

5.1 The extended projection picture

Under the action of the icosahedral group, \mathbb{R}^6 decomposes into two irreducible orthogonal subspaces, which we denote as E_\perp and E_\parallel . This decomposition induces a projection of the root lattice of D_6 onto two different copies of the root system of H_3 , which are located in E_\perp and E_\parallel , respectively.

The projection of the simple roots of D_6 onto E_\parallel has been considered in Fig. 1 in section 2. Analogously, we obtain a second projection, denoted as π_\perp , which acts as shown in Fig. 10.

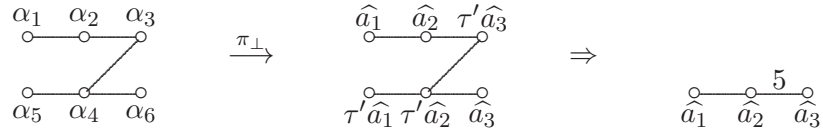


Figure 10: Figure illustrating the projection π_\perp of the simple roots of D_6 on those of H_3 .

The coordinates of the vectors in E_\perp are given in (13).

$$\begin{aligned}
 \hat{a}_1 &= \frac{1}{2}(-\tau, -\tau', 1), & \tau' \hat{a}_1 &= \frac{1}{2}(1, -\tau'^2, -\tau'), \\
 \hat{a}_2 &= \frac{1}{2}(1, -\tau, -\tau'), & \tau' \hat{a}_2 &= \frac{1}{2}(\tau', 1, -\tau'^2), \\
 \hat{a}_3 &= \frac{1}{2}(-1, -\tau, \tau'), & \tau' \hat{a}_3 &= \frac{1}{2}(-\tau', 1, \tau'^2).
 \end{aligned} \tag{13}$$

A comparison of (13) and (3) shows that the copies of H_3 in E_\parallel and E_\perp are related by an interchange of τ and τ' in the coordinates of the root vectors. Hence, each simple root vector \tilde{a}_j , $j = 1, \dots, 3$, of the H_3 -copy in E_\parallel has a counterpart \hat{a}_j , $j = 1, \dots, 3$, in one of the simple root vectors of the H_3 -copy in E_\perp and vice versa, related by an exchange of τ and τ' in their coordinates. This applies also to the root system Δ in (5). The root systems of the

two H_3 -copies in E_{\parallel} and E_{\perp} are shown superimposed onto each other in Fig. 11 as dark and light dots, respectively, to demonstrate the geometric implications of this relation.

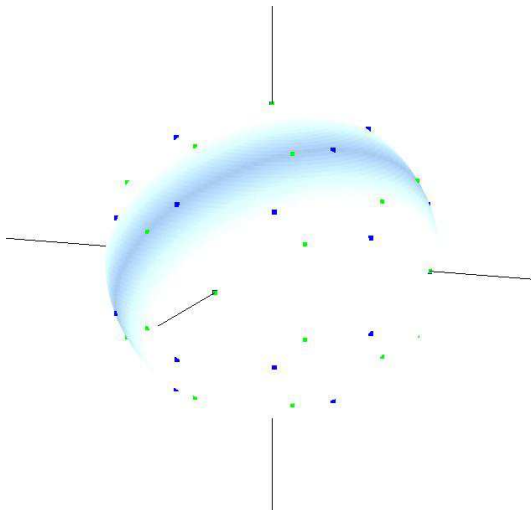


Figure 11: The root systems of the two H_3 copies on E_{\parallel} and E_{\perp} .

As mentioned in section 2, the root lattice of D_6 is given by \mathbb{Z} -linear combinations of the simple roots a_j , $j = 1, \dots, 6$, and its projection $\pi_{\parallel}(D_6)$ onto E_{\parallel} corresponds to the $\mathbb{Z}[\tau]$ -linear combinations of the simple root vectors \tilde{a}_j , where $\mathbb{Z}[\tau]$ is the extended ring of integers defined in (4). It is a dense set in \mathbb{R}^3 due to the properties of $\mathbb{Z}[\tau]$. Hence, a formalism is required to select a discrete subset of $\pi_{\parallel}(D_6)$.

This is possible via the cut-and-project method known from the study of quasicrystal [10]. In particular, we define an automorphism $*$ as a mapping from $\pi_{\parallel}(D_6)$ to $\pi_{\perp}(D_6)$ that acts on a vector $x \in \pi_{\parallel}(D_6)$ by exchanging τ and τ' in all coordinates. For any connected, bounded area $\Omega \subset E_{\perp}$ we define point sets $\Sigma(\Omega)$ by

$$\Sigma(\Omega) = \left\{ x \in \pi_{\parallel}(D_6) \mid x^* \in \Omega \right\} \subset E_{\parallel}. \quad (14)$$

The construction is illustrated in (15).

$$\begin{array}{ccccc} \Omega \subset E_{\perp} & \xleftarrow{\pi_{\perp}} & \mathbb{R}^6 & \xrightarrow{\pi_{\parallel}} & \Sigma(\Omega) \subset E_{\parallel} \\ & & \cup & & \\ & & D_6 & & \end{array} \quad (15)$$

The point sets $\Sigma(\Omega)$ have been extensively studied in the quasicrystal literature, where they are used to model the locations of atoms in alloys [11, 10].

5.2 Scaling transformations and rotations via the projection picture

The sets $S(N)$ in (10) are by construction finite dimensional subsets of the cut-and-project quasicrystals $\Sigma(\Omega)$ in (14), i.e. one has the set inclusion $S(N) \subset \Sigma(\Omega)$ for suitably chosen areas Ω . A canonical choice for Ω is the Voronoi domain of the root lattice of D_6 projected

to E_\perp [19]. It corresponds to a rhombic triacontahedron centred about the origin with edge length $1/\sqrt{2}$ and inradius $R \approx 5.32$.

The coordinates of the vertices of our polyhedra that mark the locations of pentamers are provided in the Appendix. They can be used to check that this choice of Ω is large enough to contain their counterparts in E_\perp under the mapping $*$. Let \mathcal{V}_j , $j \in \{1, 2, 3\}$, denote these vertex sets for the polyhedra corresponding the large, medium and small particle, respectively, and let $\mathcal{V}_j^* := \{x^* \mid x \in \mathcal{V}_j\}$ be their counterpart in E_\perp . Then one obtains the following radii for the shells in E_\perp on which the sets \mathcal{V}_j^* are located:

$$\begin{aligned} \mathcal{V}_1^* &: R \approx 3.77, & R \approx 4.16 \\ \mathcal{V}_2^* &: R \approx 2.76 \\ \mathcal{V}_3^* &: R \approx 1.9. \end{aligned} \tag{16}$$

Since these are all smaller than the inradius $R \approx 5.32$ of the triacontahedron that corresponds to Ω , the vertex sets of the polyhedra are indeed contained in $\Sigma(\Omega)$.

Following [12] we call a transformation *crystallographic* if it can be expressed as an N -dimensional matrix with integral entries in an N -dimensional vector space. All transformations that map the vertex sets of our polyhedra on points in $\Sigma(\Omega)$ are by construction crystallographic in this sense. It is hence possible to determine the corresponding scaling transformations, rotations and combinations thereof via the projection picture. In particular, they can be studied as transformations in E_\perp and correspond to those transformations that map \mathcal{V}_j^* onto points in $\pi_\parallel(D_6) \subset \Omega$.

- **Scaling transformations:**

All transformations that act as contractions in Ω , that is scalings by $z = z_1 + \tau z_2 \in \mathbb{Z}[\tau]$ (cf. (4)) with modulus $|z| < 1$ correspond to crystallographic scaling transformations. In particular, they map \mathcal{V}_j^* on a set $\mathcal{V}_{j,z}^* := \{zx^* \mid x^* \in \mathcal{V}_j^*\}$. Using $z^{**} = z$ this induces a scaling of \mathcal{V}_j by $z^* \in E_\parallel$, which acts as a stretching transformation since $|z^*| > 1$.

Hence, every $z \in \mathbb{Z}[\tau]$ with $|z| < 1$ induces a stretching transformation T_{z^*} on \mathcal{V}_j in E_\parallel . An example is $z = 1 - \tau = \tau'$ which induces a scaling of the index set \mathcal{V}_j by τ .

- **Rotations:**

Rotations can again be studied as rotations in Ω . Since the shells on which the sets \mathcal{V}_j^* are located are entirely contained within the inradius of the triacontahedron that defines Ω , any rotation in Ω that maps \mathcal{V}_j^* on points in $\pi_\perp(D_6) \cap \Omega$ again induces a transformation of \mathcal{V}_j in $\pi_\parallel(D_6)$.

The same holds for combinations of rotations and scaling transformations. In particular, also the polygrammatical transformations that play an important role in [12] can easily be studied via the projection formalism in this way.

6 Validation against experimental results

Our approach can be tested based on the following two criteria: the predictions for the ratios of the radii of the different species of particles which are fixed completely by our mathematical

formalism, as well as the predictions concerning the locations of the inter-subunit bonds in the capsids that are implied by the structure of the polyhedra.

6.1 Predictions concerning the relative radii of particles in different species

Our construction fixes the radii of the different polyhedra with respect to each other, and there is hence only one free parameter, α say, that relates the geometries of all particles collectively with their biological counterparts. In particular, we obtain a radius of $R_L = 3.1247$ for the large particle, of $R_M = 2.3199$ for the medium particle and of $R_S = 1.1756$ (using the radius at the global five-fold vertices where the pentamers are located) for the small particle. The ratios of these are

$$\frac{R_M}{R_S} \approx 1.9734, \quad \frac{R_L}{R_S} \approx 2.658. \quad (17)$$

The polymorphic assemblies of the major capsid protein of Simian Virus 40, which is a member of the family of Papovaviridae, has been studied in [20]. They report the occurrence of three different types of particles: a large particle with a diameter of approximately 40 to 45 nm, a medium sized particle with a diameter of size 25 to 35 nm and a small particles with a diameter of approximately 20 nm. It has not been possible in that work to distinguish whether the medium sized particles correspond to the octahedral particles consisting of 24 pentamers reported in [21]. The radii of these particles are, taking averages, approximately

$$R_L^K \approx 21.25\text{nm}, \quad R_M^K \approx 15\text{nm}, \quad \text{and} \quad R_S^K \approx 10\text{nm}. \quad (18)$$

Since these are approximate values, we test whether any combination with small deviations from these values is compatible with the ratios obtained theoretically in (17). One finds that the combination

$$R_L^K = 21.26\text{nm}, \quad R_M^K = 15.79\text{nm}, \quad \text{and} \quad R_S^K = 8\text{nm} \quad (19)$$

corresponds to the ratios in (17). The correspondence with the values in (18) is remarkably close. It implies that the ratio between large and medium size particles are in excellent agreement with the experimental results. Moreover, given that the exact value of the small sized shell is more difficult to determine experimentally than those of the larger ones, a deviation of 2nm from the experimentally measured average can be considered as an equally good agreement with experimental results.

Based on (19), it is possible to compute the scaling factor, that relates the geometry of the shells of radii R_L , R_M and R_S in the theoretical model collectively with the biological system, as

$$\alpha = \frac{R_\lambda^K}{R_\lambda} \approx 6,80\text{nm} \quad \text{with } \lambda \in \{S, M, L\}. \quad (20)$$

6.2 Predictions concerning the bonding structure

The polyhedra provide information on the locations of intersubunit bonds between the proteins in different pentamers in the capsid. Based on the tiling approach developed earlier [3, 8, 9], the faces of the polyhedra encode the locations of these interactions as follows:

a face with three corners corresponding to centres of pentamers in the capsid represents a trimer interaction between the proteins located in these corner; respectively, a face with two such corners represents a dimer interaction between these two protein subunits. The bonding structure implied by the polyhedra of the medium and large species are illustrated in Fig. 12, where the locations of the bonds are shown schematically by spiral arms.

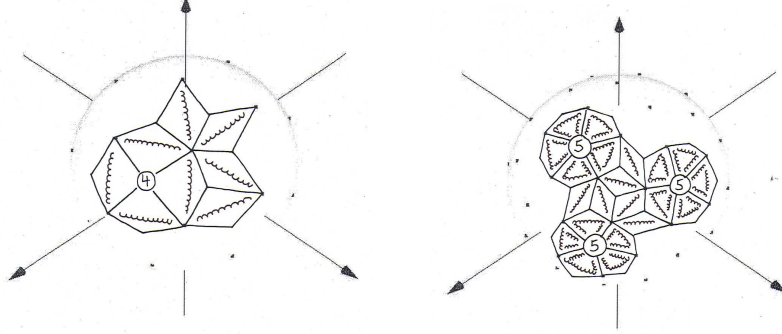


Figure 12: The bonding structure of medium and large particles implied by the tiling approach.

The combination of dimer and trimer bonds predicted by our theory for the large shell in Fig. 12 on the right coincides with the experimental results for the locations of the C-terminal arm extensions observed in [22]. It would be interesting to validate also the predictions for the bonding structure of the other species experimentally.

7 Conclusions

The results in this paper provide answers to the two open mathematical questions stated in the conclusion of [12].

Firstly, we have addressed the question of how the series of polyhedra that starts with the triacontahedron and contains all polyhedra corresponding to all-pentamer configurations, can be completed. For this, we have derived a mathematical method that combines a projection of the root lattice of D_6 with an affine extension of the non-crystallographic Coxeter group H_3 in order to determine finite subsets of generalised lattices that encode the vertices of these polyhedra. In this way, we have been able to establish the triacontahedral series that complements the family of polyhedra in Caspar-Klug Theory.

Secondly, the mathematical formalism developed here provides a framework for the systematic analysis of crystallographic scaling transformations and rotations for polyhedra with vertices in the triacontahedral series. These are important as they can be used to associate three-dimensional structures with the blueprints provided by the surface structures of the polyhedra as demonstrated in [12] for the case of Human Rhinovirus, which corresponds to the small particle in our triacontahedral series. In this respect, our results pave the way for an analysis of the interior organisation of all viral capsids corresponding to the triacontahedral series.

From the biological point of view, our triacontahedral series closes an important gap because it complements the series of Caspar-Klug polyhedra. Similar as the in the case of the Caspar-Klug series, this paves the way for a broad spectrum of applications in virology. In particular, due to the fact that Papovaviridae contain cancer-causing viruses, a better understanding of their structure is highly desirable to assist the development of new anti-viral drugs. For example, based on the structure of the polyhedra it is possible to develop assembly models that explain how viral capsids self-assemble from their capsid proteins. The theory developed here provides a framework for the generalisation of earlier work [4, 5] to the case of simultaneous assembly of different species of particles. Such models are important in order to study how viral assembly can be misdirected as a means to interfere with the viral replication cycle.

Acknowledgements

RT has been supported by an EPSRC Advanced Research Fellowship. TK has been supported by the EPSRC grant GR/T26979/01.

References

- [1] Caspar, D.L.D & Klug, A. (1962) Physical Principles in the Construction of Regular Viruses. *Cold Spring Harbor Symp. Quant. Biol.* 27, 1–24.
- [2] Casjens, S. (1985). Virus structure and assembly. Jones and Bartlett, Boston, Massachusetts.
- [3] Twarock, R. (2004), A tiling approach to virus capsid assembly explaining a structural puzzle in virology, *J. Theor. Biol.* **226**, pp. 477.
- [4] Keef, T., Taormina, A. and Twarock, R. (2005) *Assembly Models for Papovaviridae based on Tiling Theory*, to appear in *J. Phys. Biol.*.
- [5] Keef, T., Micheletti, C. and Twarock, R. (2005) *Master equation approach to the assembly of viral capsids*, submitted to *J. Theor. Biol.*.
- [6] Rayment, I., et al. (1982), *Nature* **295**, pp. 110.
- [7] Liddington, R.C. et al. (1991), *Nature* **354**, pp. 278.
- [8] Twarock, R. (2005) The architecture of viral capsids based on tiling theory, *J. Theor. Medicine* **6**, 87–90.
- [9] Twarock, R. and Hendrix, R. (2005) Crosslinking in Viral Capsids via Tiling Theory, preprint, submitted to *J. Theor. Biol.*
- [10] Senechal, M. (1996) Quasicrystals and Geometry, Cam. Univ. Press.
- [11] Shechtman, D., Blech, I., Gratias, D. & Cahn, J.W. (1984) Metallic phase with long-range order and no translational symmetry. *Phys. Rev. Lett.* 53, 1951–1953.

- [12] Janner, A. (2005) Crystallographic Structural Organization of Human Rhinovirus Serotype 16, 14, 3, 2 and 1A, submitted to *Acta Crystallographica B*.
- [13] Baake, M., Joseph, D., Kramer, P. and Schlottmann, M. (1990) Root lattices and quasicrystals *J. Phys. A* **23**, L1037 – L1041.
- [14] Moody, R. V. and Patera, J (1993), The E_8 family of quasicrystals, in: *Proc. NATO ASI Noncompact Lie groups and Some of Their Applications*, San Antonio, eds. Tanner, E.A. and Wilson, F., NATO ASI Series C, Vol.429, pp.341.
- [15] Humphreys J E 1992 *Reflection Groups and Coxeter groups* Cambridge studies in advanced mathematics 29 (Cambridge Univ. Press)
- [16] Champagne B, Kjiri M, Patera J and Sharp R 1995 Description of reflection generated polytopes using decorated coxeter diagrams *Can. J. Phys.* **73** 566–584
- [17] J. Patera and R. Twarock, *Affine extensions of noncrystallographic Coxeter groups and quasicrystals*, *J. Phys. A* **35** 1551–1574, (2002).
- [18] R. Twarock, *New group structures for carbon onions and carbon nanotubes via affine extensions of noncrystallographic Coxeter groups*, *Phys. Lett. A* **300** 437–444, (2002).
- [19] Papadopolos, Z., Hohneker, C. and Kramer, P. (1999) Tiles-inflation rules for the class of canonical tilings $\mathcal{T}^{*(2F)}$ derived by the projection method, *math-ph/9909012*.
- [20] Kaneshashi, S. et al. (2003) Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro, *J. Gen. Virol.* **84**, 1899 - 1905.
- [21] Salunke, D.M., Caspar, D.L. and Garcea, R.L. (1989) Polymorphism in the assembly of polyomavirus capsid protein VP1, *Biophys. J.* **56**, 887-900.
- [22] Modis, Y., et al. (2002), *EMBO J.* **21**, pp. 4754.

Appendix

In the Appendix, we provide the coordinates of the vertices of the polyhedra that correspond to the centres of pentamers.

The vertices corresponding to the 60 pentamers off the 5-fold axes of the large particles

are located on a shell of radius $R \approx 3.1247$. Their coordinates are given in (21).

$$\begin{aligned}
& \frac{1}{2} (2\tau - \tau', 1 - 2\tau', 6 - \tau) & \frac{1}{2} (2\tau - \tau', -(1 - 2\tau'), -(6 - \tau)) \\
& \frac{1}{2} (-(2\tau - \tau'), 1 - 2\tau', 6 - \tau) & \frac{1}{2} (-(2\tau - \tau'), 1 - 2\tau', -(6 - \tau)) \\
& \frac{1}{2} (1 - 2\tau', 6 - \tau, 2\tau - \tau') & \frac{1}{2} (1 - 2\tau', -(6 - \tau), -(2\tau - \tau')) \\
& \frac{1}{2} (-(1 - 2\tau'), -(6 - \tau), 2\tau - \tau') & \frac{1}{2} (-(1 - 2\tau'), 6 - \tau, -(2\tau - \tau')) \\
& \frac{1}{2} (6 - \tau, 2\tau - \tau', 1 - 2\tau') & \frac{1}{2} (6 - \tau, -(2\tau - \tau'), -(1 - 2\tau')) \\
& \frac{1}{2} (-(6 - \tau), -(2\tau - \tau'), 1 - 2\tau') & \frac{1}{2} (-(6 - \tau), 2\tau - \tau', -(1 - 2\tau')) \\
& \frac{1}{2} (\tau' + 4\tau, -1 - 4\tau', \tau) & \frac{1}{2} (\tau' + 4\tau, -(-1 - 4\tau'), -\tau) \\
& \frac{1}{2} (-(\tau' + 4\tau), -(-1 - 4\tau'), \tau) & \frac{1}{2} (-(\tau' + 4\tau), -1 - 4\tau', -\tau) \\
& \frac{1}{2} (-1 - 4\tau', \tau, \tau' + 4\tau) & \frac{1}{2} (-1 - 4\tau', -\tau, -(\tau' + 4\tau)) \\
& \frac{1}{2} (-(-1 - 4\tau'), -\tau, \tau' + 4\tau) & \frac{1}{2} (-(-1 - 4\tau'), \tau, -(\tau' + 4\tau)) \\
& \frac{1}{2} (\tau, \tau' + 4\tau, -1 - 4\tau') & \frac{1}{2} (\tau, -(\tau' + 4\tau), -(-1 - 4\tau')) \\
& \frac{1}{2} (-\tau, -(\tau' + 4\tau), -1 - 4\tau') & \frac{1}{2} (-\tau, \tau' + 4\tau, -(-1 - 4\tau')) \\
& \frac{1}{2} (\tau - 3\tau', 1 + 2\tau', 2 + \tau) & \frac{1}{2} (\tau - 3\tau', -(1 + 2\tau'), -(2 + \tau)) \\
& \frac{1}{2} (-(\tau - 3\tau'), -(1 + 2\tau'), 2 + \tau) & \frac{1}{2} (-(\tau - 3\tau'), 1 + 2\tau', -(2 + \tau)) \\
& \frac{1}{2} (-(1 + 2\tau'), -(2 + \tau), \tau - 3\tau') & \frac{1}{2} (-(1 + 2\tau'), 2 + \tau, -(\tau - 3\tau')) \\
& \frac{1}{2} (1 + 2\tau', 2 + \tau, \tau - 3\tau') & \frac{1}{2} (1 + 2\tau', -(2 + \tau), -(\tau - 3\tau')) \\
& \frac{1}{2} (2 + \tau, -(\tau - 3\tau'), -(1 + 2\tau')) & \frac{1}{2} (2 + \tau, \tau - 3\tau', 1 + 2\tau') \\
& \frac{1}{2} (-(2 + \tau), \tau - 3\tau', -(1 + 2\tau')) & \frac{1}{2} (-(2 + \tau), -(\tau - 3\tau'), 1 + 2\tau') \\
& \frac{1}{2} (\tau' + 2 + 2\tau, -(2\tau - 2\tau' - 1), -(-2 - \tau - 2\tau')) & \frac{1}{2} (\tau' + 2 + 2\tau, 2\tau - 2\tau' - 1, -2 - \tau - 2\tau') \\
& \frac{1}{2} (-(\tau' + 2 + 2\tau), 2\tau - 2\tau' - 1, -(-2 - \tau - 2\tau')) & \frac{1}{2} (-(\tau' + 2 + 2\tau), -(2\tau - 2\tau' - 1), -2 - \tau - 2\tau') \\
& \frac{1}{2} (2\tau - 2\tau' - 1, -2 - \tau - 2\tau', \tau' + 2 + 2\tau) & \frac{1}{2} (2\tau - 2\tau' - 1, -(-2 - \tau - 2\tau'), -(\tau' + 2 + 2\tau)) \\
& \frac{1}{2} (-(2\tau - 2\tau' - 1), -(-2 - \tau - 2\tau'), \tau' + 2 + 2\tau) & \frac{1}{2} (-(2\tau - 2\tau' - 1), -2 - \tau - 2\tau', -(\tau' + 2 + 2\tau)) \\
& \frac{1}{2} (-(-2 - \tau - 2\tau'), -(\tau' + 2 + 2\tau), 2\tau - 2\tau' - 1) & \frac{1}{2} (-(-2 - \tau - 2\tau'), \tau' + 2 + 2\tau, -(2\tau - 2\tau' - 1)) \\
& \frac{1}{2} (-2 - \tau - 2\tau', \tau' + 2 + 2\tau, 2\tau - 2\tau' - 1) & \frac{1}{2} (-2 - \tau - 2\tau', -(\tau' + 2 + 2\tau), -(2\tau - 2\tau' - 1)) \\
& (3, \tau', -\tau') & (3, -\tau', \tau') \\
& (-3, -\tau', \tau') & (-3, \tau', -\tau') \\
& (-\tau', \tau', 3) & (-\tau', -\tau', -3) \\
& (\tau', -\tau', 3) & (\tau', \tau', -3) \\
& (-\tau', -3, -\tau') & (-\tau', 3, \tau') \\
& (\tau', 3, -\tau') & (\tau', -3, \tau')
\end{aligned}$$

(21)

The locations of the coordinates in (21) on a sphere are illustrated in Fig. 13 via the 3d-Grapher software.

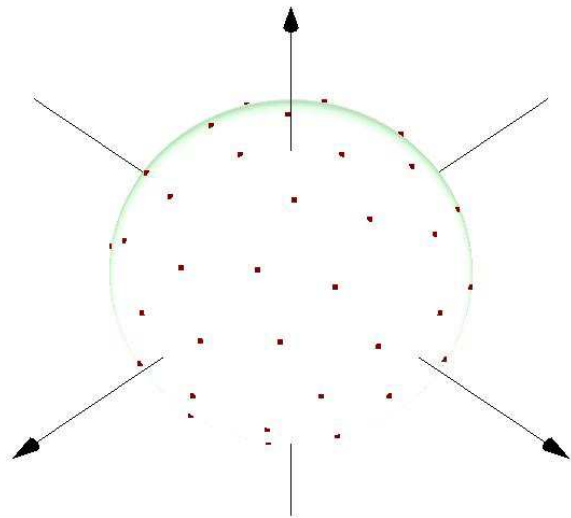


Figure 13: The 60 vertices in (21) corresponding to the centres of the pentamers off the global symmetry axes.

The vertices corresponding to the centres of the remaining pentamers of the large particles point to the vertices of an icosahedron inscribed into a sphere of radius $R = 5 \times 0.7265 \approx 3.6325$. The coordinates are given in (22).

$$\frac{1}{2}(\pm(\tau - 3 - \tau'), 0, \pm 2\tau') \text{ and all cyclic permutations.} \quad (22)$$

The vertices of the polyhedron corresponding to the centres of the pentamers in the capsids of the medium sized particles are located on a sphere of radius $R \approx 2.3199$. Their

coordinates are given in (23).

$$\begin{array}{ll}
(-1 - \tau + 2\tau', 1, -3 - \tau)/2 & (-1, 3 + \tau, -1 - \tau + 2\tau')/2 \\
(-3 - \tau, 1 + \tau - 2\tau', -1)/2 & (3\tau' - 1 - \tau, 2\tau', 0)/2 \\
(-2\tau', 0, 3\tau' - 1 - \tau)/2 & (0, -3\tau' + 1 + \tau, -2\tau')/2 \\
(\tau' - 2, 2\tau' - 2\tau + 1, -\tau)/2 & (-2\tau' + 2\tau + 1, \tau, \tau' - 2)/2 \\
(\tau, \tau, 2 - \tau) & (\tau, -\tau, 2 - \tau) \\
(\tau, -2 + \tau, \tau) & (-\tau, -2 + \tau, \tau) \\
(2 - \tau, -\tau, \tau) & (2 - \tau, -\tau, -\tau) \\
(2 + \tau, 3\tau', 2 - \tau)/2 & (-3\tau', -2 + \tau, 2 + \tau)/2 \\
(3\tau', -2 + \tau, 2 + \tau)/2 & (2 - \tau, -2 - \tau, -3\tau')/2 \\
(1 + \tau', -3 - \tau, 2 - \tau + \tau')/2 & (3 + \tau, -2 + \tau - \tau', 1 + \tau')/2 \\
(2 - \tau + \tau', -1 - \tau', 3 + \tau)/2 & (-\tau', 2, -1) \\
(-2, 1, -\tau') & (-1, \tau', -2).
\end{array} \tag{23}$$

The locations of the coordinates in (23) on a sphere are illustrated in Fig. 14 via the 3d-Grapher software.

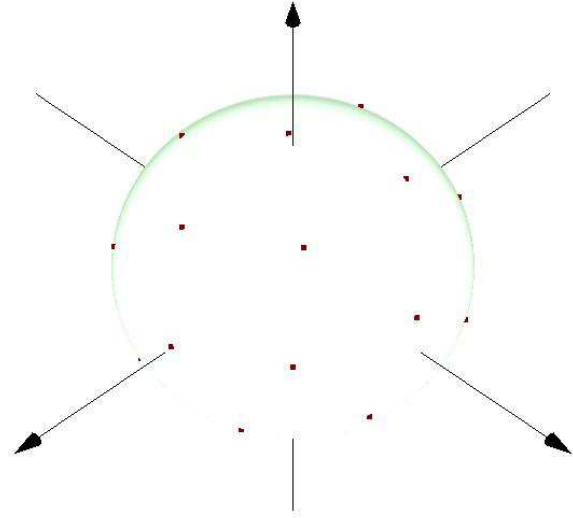


Figure 14: The 24 vertices in (23) corresponding to the centres of the pentamers.

The vertices of the triacontahedron that correspond to the centres of the pentamers of the small particles are located on a sphere of radius $R \approx 1.1756$ with coordinates given in (24).

$$(\pm 1, 0, \pm \tau') \text{ and all cyclic permutations.} \tag{24}$$