Title
A tale of two symmetrons: Rules for construction of icosahedral capsids from trisymmetrons and pentasymmetrons

Journal
Journal of Structural Biology, 170(1)

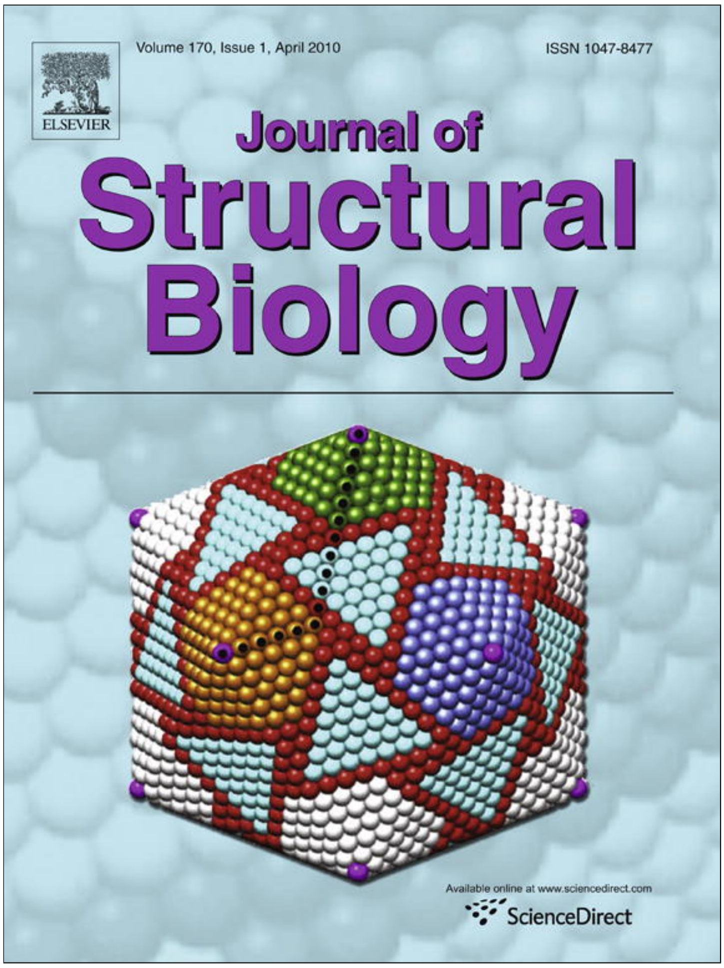
ISSN
1047-8477

Authors
Sinkovits, Robert S
Baker, Timothy S

Publication Date
2010-04-01

DOI
10.1016/j.jsb.2009.12.003

Peer reviewed


Volume 170, Issue 1, April 2010
ISSN 1047-8477
ELSEVIER
Journal of
Structural
Biology
Available online at www.sciencedirect.com
ScienceDirect

Journal of Structural Biology 170 (2010) 109–116

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

A tale of two symmetrons: Rules for construction of icosahedral capsids from trisymmetrons and pentasymmetrons

Robert S. Sinkovits [a,b], Timothy S. Baker [a,c,*]

[a] University of California, San Diego, Department of Chemistry & Biochemistry, La Jolla, CA 92093, USA
[b] University of California, San Diego, San Diego Supercomputer Center, La Jolla, CA 92093, USA
[c] University of California, San Diego, Division of Biological Sciences, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 14 October 2009
Received in revised form 1 December 2009
Accepted 2 December 2009
Available online 4 December 2009

Keywords:
Icosahedral symmetry
Trisymmetron
Pentasymmetron
Capsomer
Large DNA virus
Virus capsid structure
Triangulation number

ABSTRACT

The capsids of large, icosahedral dsDNA viruses are built from well-ordered aggregates of capsomers, known as trisymmetrons and pentasymmetrons, which are centered on the icosahedral 3-fold and 5-fold axes, respectively. We derive the complete set of rules for constructing an icosahedral structure from these symmetrons when the T lattice symmetry is odd and show that there are three classes of solutions, each of which follows from a different relationship between the size of the pentasymmetron and the values of the h and k icosahedral lattice parameters. Together, these three classes account for all possible ways of building an icosahedral structure solely from trisymmetrons and pentasymmetrons. Also, every icosahedral lattice with odd T number has solutions from exactly two of these three classes, with the set of allowed classes dependent on which of the two lattice parameters is odd. For symmetric lattices (if $h = k$ or $h = 0$), the two solutions yield the same symmetron sizes, but when the lattice parameters are equal ($h = k$) the solutions can be distinguished by the relative orientations of the symmetrons. We discuss these results in the context of known virus structures and explore the implications for viruses whose structures have not yet been solved.

1. Introduction

Early studies by Wrigley of negatively stained samples of *Sericesthis* iridescent virus (SIV)[1] suggested that the viral capsid was composed of large, well-ordered collections of capsomers (Wrigley, 1969). Capsomers were first identified in negative-stain electron microscopy studies as characteristic, morphological units in the capsids of simple, isometric viruses (Caspar and Klug, 1962; Horne and Wildy, 1961). Though they often consist of pentameric or hexameric oligomers of a major capsid protein with a β-barrel structure (Rossmann and Johnson, 1989), in adenovirus and PRD-1 (Benson et al., 2004), and many of the even larger dsDNA viruses (Yan et al., 2009) they occur, or are predicted to occur, as trimers of a coat protein with a double jelly-roll structure. The collections of capsomers seen by Wrigley, which he named "disymmetrons", "trisymmetrons", and "pentasymmetrons" were proposed to have linear, triangular, and pentagonal symmetry and be centered on the 2-, 3-, and 5-fold icosahedral axes, respectively. The latter two of these were clearly seen in both intact SIV particles and fragmented capsids. Linear fragments were also observed, but the images do not provide compelling evidence that the disymmetrons were actually incorporated into the icosahedral capsids. Starting with Goldberg's solution (Goldberg, 1937), which described how an icosahedron could be covered with a regular lattice having 5-fold symmetry at the vertices and quasi 6-fold symmetry at all other lattice points, Wrigley devised a set of rules for constructing the icosahedral lattice, and hence the viral capsid, from the three types of symmetrons. His modified Goldberg diagram (Fig. 2 of Wrigley, 1969), annotated with the allowed symmetron sizes, indicated that, when the smaller of the two lattice parameters was odd, a solution existed that involved only trisymmetrons and pentasymmetrons. Otherwise, all three types of symmetrons were necessary in order to build the icosahedral capsid. The two lattice parameters (h and k in the modern notation) are positive integers that specify the location of a 5-fold vertex relative to its neighboring vertex. The triangulation number is defined as $T = h^2 + k^2 + hk$ and the number of lattice points is equal to $12 + 10(T - 1)$.

Limitations in the quality of the SIV images led to ambiguities in their interpretation. The capsid was surmised to have a $T = 156$ lattice with di-, tri-, and pentasymmetrons containing 9, 55, and 16 capsomers, respectively, although the possibility of a $T = 129$ or $T = 147$ lattice was not excluded. Subsequent work on other large icosahedral dsDNA viruses is compatible with Wrigley's set of

* Corresponding author. Address: University of California, San Diego, Department of Chemistry & Biochemistry, La Jolla, CA 92093, USA. Fax: +1 858 534 5846.
E-mail addresses: tsb@ucsd.edu, tsb@chem.ucsd.edu (T.S. Baker).

[1] *Abbreviations:* SIV, *Sericesthis* iridescent virus; TIV, *Tipula* iridescent virus; CIV, *Chilo* iridescent virus; PBCV-1, *Paramecium bursaria chlorella* virus; PpV01, *Phaeocystis pouchetii* virus; FV3, Frog virus 3.

1047-8477/$ - see front matter
doi:10.1016/j.jsb.2009.12.003

rules, but no corroborating evidence has been seen yet for the existence of disymmetrons. An analysis of images of negatively stained samples of *Tipula* iridescent virus (TIV) (Manyakov, 1977) and, more convincingly, cryo-reconstructions of *Chilo* iridescent virus (CIV) (Yan et al., 2009), PBCV-1 (Yan et al., 2000), PpV01 (Yan et al., 2005), and Frog virus 3 (FV3) (Yan et al., unpublished results) unambiguously reveal capsids that are composed solely of tri- and pentasymmetrons. Interestingly, while the trisymmetron sizes vary for these viruses, the pentasymmetrons always contain 31 capsomers (Table 1). Given the close similarity between SIV and TIV, taken together with these other results, it is highly plausible that a modern cryo-reconstruction of SIV would yield a structure that does not contain disymmetrons.

While the large icosahedral virus structures solved to date are consistent with Wrigley's formulation, the prediction that only two-thirds of the lattices with odd T numbers can be built from tri- and pentasymmetrons is intriguing. This compelled us to suspect that one or more classes of solutions had been overlooked. We show below that this is in fact the case and that every icosahedral lattice with odd T number can be built in either one, in the case of symmetric lattices, or two distinct ways without a requirement to incorporate disymmetrons.

2. Results

2.1. Construction of icosahedral lattice from trisymmetrons and pentasymmetrons is only possible for odd T numbers

The lack of compelling experimental evidence for disymmetrons led us to investigate the rules for the construction of an icosahedral capsid solely from tri- and pentasymmetrons. A necessary, but not sufficient, condition is that the numbers of capsomers in the trisymmetrons (N_{TS}) and pentasymmetrons (N_{PS}) be consistent with the total number of capsomers forming the capsid (N_{cap}).

$$N_{cap} = 12 + 10(T - 1) = 20N_{TS} + 12N_{PS} \quad (1)$$

The shapes of the symmetrons place further restrictions on the values for N_{TS} and N_{PS} (Fig. 1). Ignoring the presence of special vertices in otherwise isometric capsids such as in PBCV-1 (Cherrier et al., 2009), the pentasymmetron must contain a central capsomer, surrounded by zero or more layers of capsomers. This leads to the result that N_{PS} is given by:

$$N_{PS} = 1 + 5e_{PS}(e_{PS} - 1)/2 \quad (2)$$

Table 1
Lattice properties for icosahedral capsids of all large dsDNA viruses with known structures. In each instance, the capsomers located at the vertices contain five subunits whereas all other capsomers contain three subunits. As a result, each complete tri- and pentasymmetron contains $3N_{TS}$ and $5 + 3(N_{PS} - 1)$ subunits, respectively.

Virus	T	h	k	N_{DS}	N_{TS}	N_{PS}	Method
SIV[a]	156	4	10	9	55	16	Neg Stain
SIV[b]	129	5	8	0	55	16	Neg Stain
SIV[b]	147	7	7	0	55	31	Neg Stain
TIV	147	7	7	0	55	31	Neg Stain
CIV	147	7	7	0	55	31	Cryo
FV3	169	7	8	0	66	31	Cryo
PBCV-1	169	7	8	0	66	31	Cryo
PpV01	219	7	10	0	91	31	Cryo

Abbreviations: T: triangulation or T number; h and k: icosahedral lattice parameters; N_{DS}, N_{TS}, and N_{PS}: number of capsomers in di-, tri-, and pentasymmetrons, respectively; Neg Stain: negatively stained sample; Cryo: unstained, vitrified sample.
[a] Wrigley's proposed interpretation of SIV images.
[b] Wrigley's alternative interpretations of SIV images.

where e_{PS} is the number of capsomers forming one edge of the pentasymmetron. Likewise, the capsomers in a trisymmetron must form an equilateral triangle, with the consequence that N_{TS} is given by:

$$N_{TS} = e_{TS}(e_{TS} + 1)/2 \quad (3)$$

where e_{TS} is the edge length of the trisymmetron.

If the forms for N_{PS} and N_{TS} are inserted into Eq. (1), we can easily show that capsids built solely from tri- and pentasymmetrons cannot have even T numbers.

$$\frac{e_{TS}(e_{TS} + 1)}{2} = \frac{(T - 1)}{2} - \frac{3e_{PS}(e_{PS} - 1)}{2} \quad (4)$$

The requirement that N_{TS} has an integral value can only be satisfied if the T number is odd, a condition that holds if at least one of the lattice parameters h or k is odd. It also follows that no capsomer will lie on a 2-fold axis since a lattice point is found at (h/2, k/2) only if both h and k are even.

We obtain an additional restriction on the symmetron sizes by considering the locations of symmetrons relative to the icosahedral symmetry axes. Requiring pentasymmetrons to be centered on the 5-fold axes ensures that their central capsomers also lie directly on these axes. Also, since the trisymmetrons cannot contain capsomers that lie on the 5-fold axes, e_{PS} must be at least equal to one. The situation for trisymmetrons is more complicated since a capsomer will be centrally located at the 3-fold axis if, and only if, the T number is divisible by three (Appendix), in which case the trisymmetron edge length must be equal to 1, 4, 7, 10, ..., $3m + 1$. For all others (i.e. T numbers not divisible by three), the edge length must be equal to 0, 2, 3, 5, 6, ... Note that there is no lower limit on e_{TS} since it is possible to construct an icosahedral lattice with points only on the vertices (e.g. $T = 1$).

An exhaustive search for compatible values of T, e_{PS}, and e_{TS} that satisfy Eq. (4) indicates that most T numbers have multiple corresponding combinations of symmetron sizes (Table 2). We will prove below that all positive integral solutions of Eq. (4) for T numbers of the form $h^2 + k^2 + hk$ and with e_{PS} greater than zero correspond to valid ways of constructing an icosahedral lattice from trisymmetrons and pentasymmetrons. It should be noted though that not all sets of lattice parameters consistent with a given T number are compatible with the solutions of Eq. (4). Although not stated explicitly, Wrigley obviously came to this same conclusion. For example, a $T = 49$ ($h = 7$, $k = 0$) capsid can be built from symmetrons of size $N_{PS} = 31$ and $N_{TS} = 6$, but not $N_{PS} = 16$ and $N_{TS} = 15$ since in the latter case it is impossible to fit the symmetrons together without resulting in both overlapping capsomers and gaps in the coverage of the lattice (Fig. 2).

2.2. Geometrical constraints on symmetron packing lead to three classes of solutions

Given the geometrical constraints between the symmetrons, the solutions of Eq. (4) can be narrowed to just those that correspond to valid ways of covering the icosahedral lattice with trisymmetrons and pentasymmetrons. This leads to three classes of solutions (Table 3), each of which follows from a different relationship between the pentasymmetron edge length and the icosahedral lattice parameters. In all of the derivations that follow we assume a right-handed ($h < k$) or symmetric ($h = k$ or $h = 0$) lattice, but analogous results can be obtained for left-handed lattices by reversing the relative orientations of the symmetrons (explained below) and interchanging the roles of h and k. To help distinguish the results for the three classes of solutions, the symmetron sizes and edge lengths are hereafter labeled to designate the class (e.g. e_{PS1}, e_{PS2}, e_{PS3}).

In the first class of solutions, each trisymmetron borders three tri- and three pentasymmetrons, whereas pentasymmetrons only

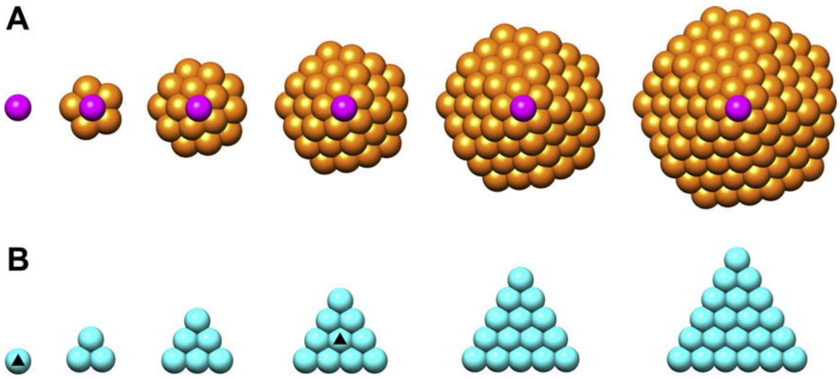


Fig. 1. Schematic models of pentasymmetrons and trisymmetrons. (A) Smallest six pentasymmetrons with capsomers modeled as magenta (at vertices) and orange spheres. The number of capsomers in a pentasymmetron, N_{PS}, is related to the edge length, e_{PS}, as given by the formula $N_{PS} = 1 + 5e_{PS}(e_{PS} - 1)/2$ (Eq. (2)). (B) Smallest six trisymmetrons with capsomers modeled as cyan spheres. The number of capsomers in a trisymmetron, N_{TS}, is related to the edge length (e_{TS}) as given by the formula $N_{TS} = e_{TS}(e_{TS} + 1)/2$ (Eq. (3)). When trisymmetrons are incorporated into an icosahedral lattice, the central capsomer (marked with black triangle) will lie on a 3-fold axis only when the value of $(e_{TS} - 1)$ is divisible by three.

Table 2
List of first 15 odd T numbers together with lattice parameters (h, k) and symmetron sizes $[N_{PS}, N_{TS}]$ corresponding to solutions of Eq. (4). Note that there is only one solution for each symmetric lattice ($h = k$ or $h = 0$), but there are always two solutions for asymmetric lattices ($h \neq k$). Although the table is limited to small T numbers, these trends persist for T numbers at least as large as 2000. $T = 49$ is the first T number for which multiple (h, k) are possible and the total number of solutions equals the number expected from the symmetric ($h = 0$, $k = 7$) and asymmetric ($h = 3$, $k = 5$) lattices. Without imposing geometrical constraints on the packing of symmetrons, there is no way to decide which solutions are associated with which lattice.

T	(h, k)	$[N_{PS}, N_{TS}]$
1	(0, 1)	[1, 0]
3	(1, 1)	[1, 1]
7	(1, 2)	[1, 3], [6, 0]
9	(0, 3)	[6, 1]
13	(1, 3)	[1, 6], [6, 3]
19	(2, 3)	[6, 6], [16, 0]
21	(1, 4)	[1, 10], [16, 1]
25	(0, 5)	[16, 3]
27	(3, 3)	[6, 10]
31	(1, 5)	[1, 15], [16, 6]
37	(3, 4)	[6, 15], [31, 0]
39	(2, 5)	[16, 10], [31, 1]
43	(1, 6)	[1, 21], [31, 3]
49	(0, 7), (3, 5)	[6, 21], [16, 15], [31, 6]
57	(1, 7)	[1, 28], [31, 10]

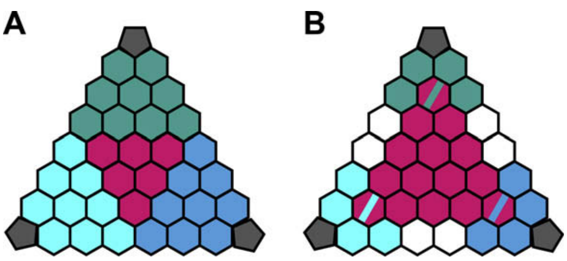


Fig. 2. Illustration of two sets of symmetron sizes permitted for a $T = 49$ ($h = 0$, $k = 7$) lattice by Eq. (2). Both panels represent a triangular face of an icosahedron with capsomers at 5-fold symmetry axes indicated by gray pentagons and all other capsomers by hexagons. (A) The portions of three pentasymmetrons ($N_{PS} = 31$) are colored in different shades of blue and the trisymmetron ($N_{TS} = 6$) in red. Although this configuration has not been observed in nature, it is a geometrically permissible packing of symmetrons. (B) Attempting to cover icosahedral lattice with symmetrons of size $N_{PS} = 16$ and $N_{TS} = 15$ results in overlap between three corner positions of the trisymmetron and three edge positions of 3-fold-related pentasymmetrons (striped hexagons) together with capsomers that do not belong to either type of symmetron (white hexagons).

contact trisymmetrons (Fig. 3). The relative orientations of the symmetrons can be described in various ways. For example, if a trisymmetron is viewed in an orientation such that the base of the triangle is horizontal, the bottom-left edge borders another trisymmetron and the bottom-right edge borders a pentasymmetron. Alternatively, if we consider a single pentasymmetron and its five neighboring trisymmetrons, the "dangling" edges of the trisymmetrons that extend beyond the trisymmetron–pentasymmetron boundaries point in a counter-clockwise direction. Another way of visualizing this is to note that if we step from the vertex toward the corner of the pentasymmetron and continue into the trisymmetron, making a 60° left-hand turn directs one along the dangling edge of the trisymmetron. This class corresponds to the set of solutions recognized by Wrigley that did not require disymmetrons. Noting the relationship between the pentasymmetron edge length and the smaller of the two lattice parameters, we find:

$$h = 2(e_{PS1} - 1) + 1 = 2e_{PS1} - 1 \quad (5)$$

If Eq. (5) is solved for e_{PS1} and the result is substituted into Eq. (4), the allowed edge lengths and symmetron sizes are:

$$e_{PS1} = \frac{h+1}{2}, \quad N_{PS1} = 1 + \frac{5(h^2 - 1)}{8} \quad (6)$$

$$e_{TS1} = \frac{2k + h - 1}{2}, \quad N_{TS1} = \frac{(2k + h)^2 - 1}{8} \quad (7)$$

In the above equations, h must be odd in order for e_{PS1} to have an integer value, but there are no restrictions on k other than it be greater than or equal to h. These equations, which relate symmetron sizes to the icosahedral lattice parameter, have been noted before (Simpson et al., 2003) but no derivation or geometric interpretation was provided.

The second class of solutions is obtained by considering a right-handed lattice where each trisymmetron borders three tri- and three pentasymmetrons, as in the first class, but the relative orientations of the symmetrons are reversed[2] (Fig. 3). For example, if a trisymmetron is again viewed in an orientation such that the base of the triangle is horizontal, the bottom-*right* edge borders another trisymmetron and the bottom-*left* edge borders a pentasymmetron. Similarly, the dangling edges of the trisymmetrons point in a *clockwise* direction about each pentasymmetron. Stepping from the

[2] The italicized text in this paragraph is used to emphasize the differences between the geometry of the class 1 and class 2 solutions.

Table 3
Properties of the three classes of solutions describing the coverage of a right-handed ($h < k$) and symmetric ($h = k$ or $h = 0$) icosahedral lattices solely composed of trisymmetrons and pentasymmetrons. For left-handed lattices, roles of h and k are interchanged and the relative orientations of symmetrons are reversed.

	Class 1	Class 2	Class 3
e_{PS}	$(h + 1)/2$	$(k + 1)/2$	$(h + k + 1)/2$
N_{PS}	$1 + 5(h^2 - 1)/8$	$1 + 5(k^2 - 1)/8$	$1 + 5(h^2 + k^2 + 2hk - 1)/8$
e_{TS}	$(2k + h - 1)/2$	$(2h + k - 1)/2$	$(k - h - 1)/2$
N_{TS}	$((2k + h)^2 - 1)/8$	$((2h + k)^2 - 1)/8$	$((k - h)^2 - 1)/8$
$\min(e_{TS})$	$3e_{PS} - 2$	$e_{PS} - 1$	0
$\max(e_{TS})$	∞	$3e_{PS} - 2$	$e_{PS} - 1$
Restrictions	h odd	k odd	$(h + k)$ odd
Notes	Edges of trisymmetron point counterclockwise about 5-fold axis	Edges of trisymmetron point clockwise about 5-fold axis	Edges of pentasymmetron point clockwise about 3-fold axis

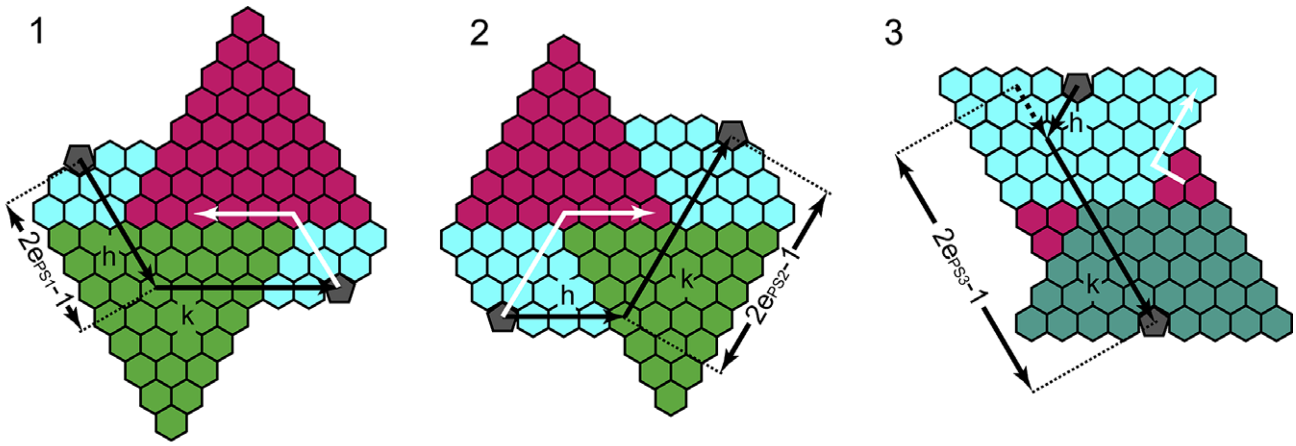


Fig. 3. Relation between pentasymmetron edge length and icosahedral lattice parameters for the three classes of solutions described in the text. Capsomers at 5-fold axes are shown as gray pentagons. Pentasymmetrons are colored in shades of blue and trisymmetrons in red or green. In the three classes, the distance ($2e_{PS} - 1$) is equal to h in class 1, k in class 2, and ($h + k$) in class 3. The white arrows illustrate the 60° and 90° turns at the boundaries of the symmetrons as described in the text. Note that the centers of the trisymmetrons do not necessarily correspond to the centers of capsomers. The T numbers and icosahedral lattice parameters in this figure, $T = 91$ ($h = 5$, $k = 6$), $T = 93$ ($h = 4$, $k = 7$), and $T = 67$ ($h = 2$, $k = 7$), do not have any special significance and were chosen simply to provide illustrative examples for the three classes.

vertex toward the corner of the pentasymmetron and continuing into the trisymmetron, then making a 60° *right*-hand turn, directs one along the dangling edge of the trisymmetron. In this case the relationship between the pentasymmetron edge length and the *larger* of the two lattice parameters is given by:

$$k = 2(e_{PS2} - 1) + 1 = 2e_{PS2} - 1 \quad (8)$$

If Eq. (8) is solved for e_{PS2} and the result is substituted into Eq. (4), the allowed edge lengths and symmetron sizes for class 2 structures are:

$$e_{PS2} = \frac{k+1}{2}, \quad N_{PS2} = 1 + \frac{5(k^2-1)}{8} \quad (9)$$

$$e_{TS2} = \frac{2h+k-1}{2}, \quad N_{TS2} = \frac{(2h+k)^2-1}{8} \quad (10)$$

As expected, Eqs. (9) and (10) have the same form as Eqs. (6) and (7), but with the roles of h and k reversed. In addition, we now have the restriction that k must be odd while h can take on arbitrary values less than or equal to k.

Finally, a third class of solutions exists where each pentasymmetron borders five tri- and five pentasymmetrons, whereas the trisymmetrons only contact adjacent pentasymmetrons (Fig. 3). If we view the pentasymmetron such that the top edge is horizontal, the top-right and top-left edges border a trisymmetron and pentasymmetron, respectively. The dangling edges of the pentasymmetrons that extend beyond the trisymmetron–pentasymmetron boundaries point clockwise about a central trisymmetron. Equivalently, stepping from the center of the trisymmetron towards the center of its edge, continuing into the pentasymmetron, and making a 90° right-hand turn directs one along the edge of the trisymmetron.

The pentasymmetron edge length is seen to depend on both lattice parameters.

$$h + k = 2(e_{PS3} - 1) + 1 = 2e_{PS3} - 1 \quad (11)$$

If Eq. (11) is solved for e_{PS3} and the result is substituted into Eq. (4), the allowed edge lengths and symmetron sizes for class 3 structures are:

$$e_{PS3} = \frac{h+k+1}{2}, \quad N_{PS3} = 1 + \frac{5(h^2+k^2+2hk-1)}{8} \quad (12)$$

$$e_{TS3} = \frac{k-h-1}{2}, \quad N_{TS3} = \frac{(k-h)^2-1}{8} \quad (13)$$

For the third class of solutions, there are no restrictions on h or k individually, but the quantity $h + k$ must be odd. This condition is only satisfied when one lattice parameter is odd and the other is even.

Every icosahedral lattice with odd T number will have solutions from exactly two of these classes. When h is odd and k is even, there will be solutions from classes 1 and 3, but when h is even and k is odd solutions from classes 2 and 3 result. Finally, h and k both odd leads to solutions from classes 1 and 2 (Fig. 4).

2.3. Three classes of solutions encompass all possible ways of building T = odd icosahedral lattices from trisymmetrons and pentasymmetrons

We began by asking which symmetron sizes are allowed for a given icosahedral lattice. It is instructive to reverse the question

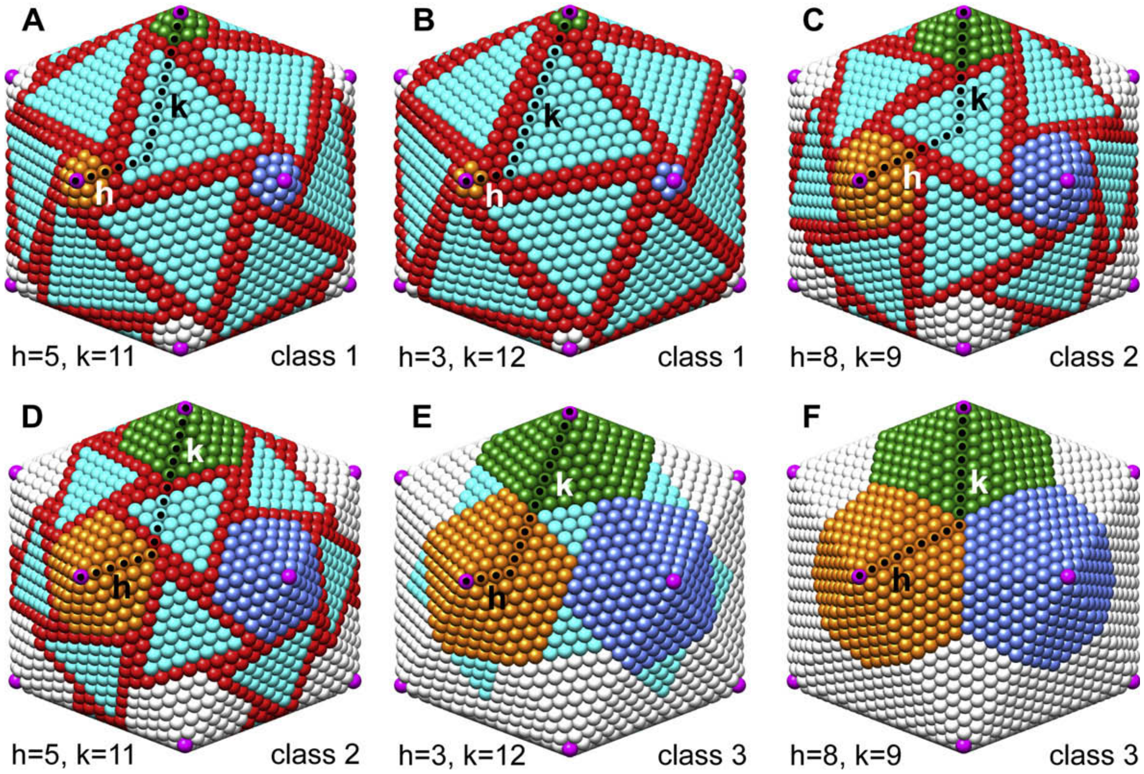


Fig. 4. Icosahedra with capsomers (modeled as spheres) arranged in six different, asymmetric lattices. In each example, capsomers at the vertices are colored magenta and the remaining ones in three adjacent pentasymmetrons are colored orange, blue, and green to distinguish them from each other and all other, uncolored pentasymmetrons. Capsomers in trisymmetrons are colored cyan and, except for (E), bordered in red to highlight symmetron boundaries. Models in the top row (A–C) share identical icosahedral lattice parameters with the models directly below them (D–F). This illustrates that a given lattice can be built from symmetrons in two different ways, and sometimes solely from pentasymmetrons (F).

and ask which combination of symmetron sizes can be used to build a given icosahedral lattice. The answer to this question provides a key to determining if another set of possible solutions has been overlooked. By deriving the allowed trisymmetron sizes relative to a given pentasymmetron size for the three classes of solutions, we can show below that they cover all possible combinations.

For the first class of solutions, N_{PS1} and therefore e_{PS1} depend only on h and there are no restrictions on k. We can therefore obtain the minimum and maximum trisymmetron edge lengths for a fixed pentasymmetron size by allowing k to tend toward h and infinity, respectively.

$$\min(e_{TS1}) = \lim_{k \to h} \frac{2k+h-1}{2} = \frac{3h-1}{2} = 3e_{PS1} - 2 \quad (14)$$

$$\max(e_{TS1}) = \lim_{k \to \infty} \frac{2k+h-1}{2} \to \infty \quad (15)$$

For the second class of solutions, N_{PS2} and therefore e_{PS2} depend only on k and there are no restrictions on h other than $h \leqslant k$. The minimum and maximum trisymmetron edge lengths for class 2 solutions are therefore obtained as h tends to 0 and k, respectively.

$$\min(e_{TS2}) = \lim_{h \to 0} \frac{2h+k-1}{2} = \frac{k-1}{2} = e_{PS2} - 1 \quad (16)$$

$$\max(e_{TS2}) = \lim_{h \to k} \frac{2h+k-1}{2} = \frac{3k-1}{2} = 3e_{PS2} - 2 \quad (17)$$

For the third class of solutions, N_{PS3} and e_{PS3} depend on $h + k$, whose sum must be odd. We can obtain the range of allowed trisymmetron size by holding this sum constant and considering the limiting values on h.

$$\min(e_{TS3}) = \lim_{h \to k-1} \frac{k-h-1}{2} = 0 \quad (18)$$

$$\max(e_{TS3}) = \lim_{h \to 0} \frac{k-h-1}{2} = \frac{k-1}{2} = e_{PS3} - 1 \quad (19)$$

An unexpected result (Eq. (18)) is that it is possible to construct an icosahedral structure solely from pentasymmetrons when the condition $h = k - 1$ is satisfied. One restriction on an all-pentasymmetron structure is that no capsomers lie on the 3-fold axes. We can show that this is true since the T number will be equal to $3(k^2 - k) + 1$, which is guaranteed not to be divisible by three (see Appendix).

An important consequence of these limits (Eqs. (14)–(19)) is that, for a given non-zero pentasymmetron size, trisymmetrons of all sizes are allowed. The allowed values for e_{TS} range from zero (class 3) to infinity (class 1), and the overlap between the classes guarantees that all possibilities between these extremes are covered. This result also confirms our earlier assertion that all solutions of Eq. (4), with T numbers of the form $h^2 + k^2 + hk$, correspond to valid ways of constructing an icosahedral lattice from symmetrons. Since any set of symmetron sizes can be combined to create an icosahedral lattice, and their sizes determine the T number, every solution of Eq. (4) must be permissible. The key restriction though is that the solutions are associated with only one set of lattice parameters if two or more sets of lattice parameters result in the same T number. For example, when T equals 49, the solution $[N_{PS} = 31,\ N_{TS} = 6]$ is valid only for the $(h = 0, k = 7)$ lattice whereas the solutions $[N_{PS} = 16,\ N_{TS} = 15]$ and $[N_{PS} = 6, N_{TS} = 21]$ are valid only for the $(h = 3, k = 5)$ lattice.

If we assume that there is only one way, or two if we consider handedness, to pack together symmetrons of specified sizes, then

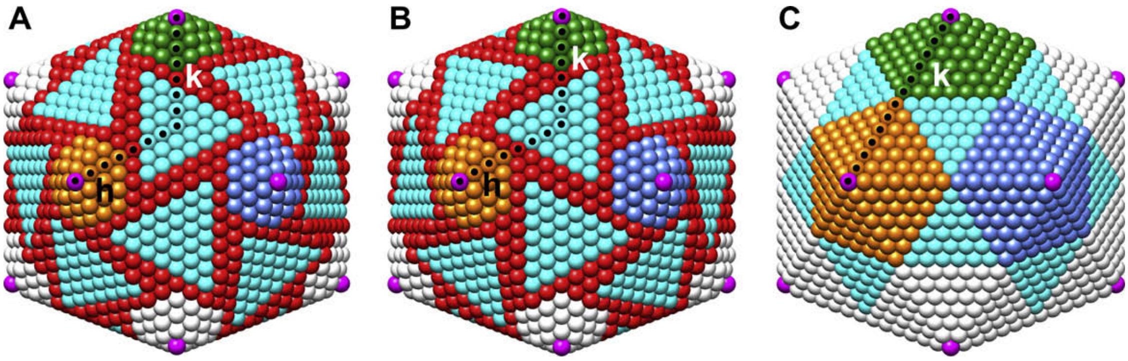


Fig. 5. Icosahedra with capsomers arranged in three different, symmetric lattices. Capsomers are rendered using the same coloring scheme as in Fig. 4. Models in (A) and (B), both with $T = 147$ ($h = 7$, $k = 7$) lattices, illustrate solutions from classes 1 and 2. Although a $T = 147$ lattice is not handed, the two allowed ways to pack symmetrons result in trisymmetron arrangements that are mirror images of each other. (C) Model with a $T = 169$ ($h = 0$, $k = 13$) lattice, where a trisymmetron can adopt only one orientation relative to the neighboring pentasymmetrons and class 2 and 3 solutions are indistinguishable.

our three classes of solutions are complete and describe all ways of covering an icosahedral lattice with tri- and pentasymmetrons. This assumption is reasonable since a consistent feature of all three classes is that at least one edge of the trisymmetron is continuous with the edge of a neighboring pentasymmetron. Without this constraint, we would end up with gaps in the coverage of the lattice that could not be filled without resorting to disymmetrons.

These results also provide additional insights regarding the symmetric lattices, which occur when $h = k$ or when $h = 0$. In the first instance ($h = k$), the symmetron edge lengths for class 1 match those for class 2 and are given by $e_{PS} = (k + 1)/2$ and $e_{TS} = (3k - 1)/2$. The two solutions can still be distinguished though by the orientation of the trisymmetrons relative to the pentasymmetrons (Fig. 5A and B). When h is equal to zero, the symmetron edge lengths for classes 2 and 3 are the same and given by $e_{PS} = (k + 1)/2$ and $e_{TS} = (k - 1)/2$. In this case the two solutions are indistinguishable since a trisymmetron with an edge length equal to $e_{PS} - 1$ can only assume one orientation relative to the pentasymmetron (Fig. 5C).

3. Discussion

We have shown that the potential ways of constructing an icosahedral lattice from trisymmetrons and pentasymmetrons is much richer than had been expected. The large dsDNA virus structures that have been solved to date are all in agreement with the previously recognized set of solutions, what we have termed class 1, but this may be fortuitous given the small number of structures and the fairly narrow range of T numbers that they span.

We have proven that any symmetric or skewed icosahedral lattice with odd T number can be constructed from one or two sets of symmetron sizes, respectively, but several open questions remain for real biological systems. For example, what are the smallest collections of capsomers that form recognizable symmetrons? Cryo-reconstructions indicate that the capsomers within a trisymmetron all have the same orientation, while those within a pentasymmetron take on one of two possible orientations (Yan et al., 2000, 2005, 2009). In addition, the higher resolution that has been achieved for CIV reveals that minor capsid proteins are located at the symmetron boundaries (Yan et al., 2009). $T = 1$ and $T = 3$ lattices are mathematically consistent with symmetrons having edge lengths equal to one, but viruses with these triangulation symmetries do not exhibit the higher levels of organization that have been observed in much larger viruses.

A related question is whether there are either optimal or minimal sizes for the symmetrons. Cryo-reconstructions of CIV, PBCV-1, FV3, and PpV01 and images of negatively stained TIV reveal that all five structures have pentasymmetrons comprised of 31 capsomers. This is consistent with $N_{PS} = 31$ being both an optimal and a minimal value, but this finding may be just a coincidence since, as mentioned earlier, these viruses span a fairly small range of T numbers. If $N_{PS} = 31$ did represent a minimum pentasymmetron size, lattices with odd T numbers ranging from $T = 37$ ($h = 3$, $k = 4$) to $T = 127$ ($h = 6$, $k = 7$) would still be allowed, but none have been observed to date. The largest odd T number less than 147 for which there is a known structure, *Sulfolobus* turreted icosahedral virus (Rice et al., 2004), is 31 ($h = 1$, $k = 5$). A lack of known icosahedral viruses in this T number range does not preclude their existence, but is suggestive that the smallest allowed trisymmetron contains 55 capsomers. The $T = 169$ (PBCV-1) and $T = 219$ (PpV01) capsids provide additional evidence in support of this view. In these instances, a solution from class 3 exists since one of the lattice parameters is even and the other odd. The $T = 169$ ($h = 7$, $k = 8$) and $T = 219$ ($h = 7$, $k = 10$) lattices could theoretically be built from symmetrons with sizes [$N_{PS} = 141$, $N_{TS} = 0$] and [$N_{PS} = 181$, $N_{TS} = 1$], respectively, but rather evolved to have the structures from class 1 that we observe. An alternative interpretation of these results is that they argue for a maximum pentasymmetron size of $N_{PS} = 31$. These hypotheses of course are merely speculative and based on limited structural data. New studies of the morphologies of additional capsids may provide more clues about rules for constructing large icosahedral viruses.

Just as there may be a lower limit on the T number of icosahedral capsids constructed from trisymmetrons and pentasymmetrons, there is evidence that there may also be an upper limit. The structure of the 5000-Å diameter (ignoring fibers) mimivirus suggests that a different capsid architecture is used (Xiao et al., 2009). In all other large dsDNA viruses that have been studied, the capsomers are arranged in a hexagonal close-packed lattice and the capsomers within a trisymmetron all have the same orientation. Mimivirus appears to be unique though in that one-third of the capsomers are missing, resulting in a regular pattern of vacancies. In addition, the capsomers within a face of the icosahedral structure appear to have alternating orientations.

Mimivirus may be unique, but it is still instructive to consider the possible symmetron sizes should a modified version of our formulation prove to be applicable. As reported (Xiao et al., 2009), if all vacancies were replaced by capsomers, the lattice parameters of mimivirus would be estimated to have values of h and k ranging between 18 and 20. Over the range of allowed T numbers, there are five possible sets of symmetron sizes (Table 4). Two of these solutions require all-pentasymmetron structures (e.g. Fig. 4F), both of which we believe are rather unlikely. This is because cryo-reconstructions of other large dsDNA viruses indicate that most capsid curvature originates from the inherent pyramidal structure of

Table 4
Properties of allowed icosahedral lattices for mimivirus based on current estimates for lattice parameters. Results listed are only provided for T = odd lattice symmetries. The column labeled "Class" corresponds to the three classes of solutions described in the text and in Table 3. For the example in which $h = k = 19$, the solutions for classes 1 and 2 are degenerate.

h	k	T	Class	N_{PS}	N_{TS}
18	19	1027	2	226	378
18	19	1027	3	856	0
19	19	1083	1/2	226	406
19	20	1141	1	226	435
19	20	1141	3	951	0

pentasymmetrons, whereas the approximately planar regions, such as those seen in mimivirus, are attributed to the presence of trisymmetrons. In addition, AFM images of defibered mimivirus virions (Xiao et al., 2009) show irregular cracks that might be interpreted as trisymmetron boundaries. The remaining three solutions all correspond to $N_{PS} = 226$ and with trisymmetrons containing 378, 406, or 435 capsomers. If these results prove to be correct, they would provide additional support for minimum symmetron sizes of $N_{PS} = 31$ and $N_{TS} = 55$.

The results derived in this paper provide a framework for understanding the morphology of large icosahedral viruses, but many outstanding questions remain. There is no compelling reason, for example, that large T = even viruses with disymmetrons should be forbidden, but is their notable absence a consequence of nature taking the simpler path of employing just two types of large scale building blocks rather than three? Also, why is there such a large gap in solved structures between $T = 31$ and $T = 147$? Is it simply that viruses in this gap have yet to be discovered or is it that the different assembly mechanisms used to build small and large capsids both fail in this intermediate size regime? Normally, there is a correlation between length of genome and capsid size, but the large dsDNA viruses tend to be much larger than is required. Could it be that capsids with T numbers of at least 147 are easier to construct compared to smaller capsids, which pack the genome more tightly? These and other questions ought to be answered as more structures are solved and higher resolutions are achieved.

Acknowledgments

We thank X. Yan for helpful discussions and K.N. Parent for a critical reading of the manuscript. The work was supported in part by NIH Grants R37 GM-033050 and R01 AI-079095 and a gift from the Agouron Foundation to T.S.B.

Appendix. Proof that a capsomer can only lie on a 3-fold axis if the *T* number is divisible by three

Consider an equilateral triangle with sides of unit length and orientated as shown in Fig. 6. If this triangle represents a face of an icosahedron, starting at the origin (lower left 5-fold vertex) and taking h and k steps along the $\boldsymbol{h}$ and $\boldsymbol{k}$ lattice vectors, respectively, corresponds to a translation from the Cartesian coordinates (0, 0) to (1, 0). Owing to the 3-fold symmetry of the equilateral triangle, the sum of the angles α and β must be equal to 60° and hence the angle between $\boldsymbol{h}$ and $\boldsymbol{k}$ is 120°. By invoking the law of cosines, it can be shown that the spacing between the lattice points d is equal to $1/\sqrt{T}$,

$$1 = h^2d^2 + k^2d^2 - 2hkd^2\cos(120^\circ) \quad \text{(A1)}$$

Using the law of sines, we find that the angle α is equal to $\sin^{-1}(\sqrt{3}h/2\sqrt{T})$.

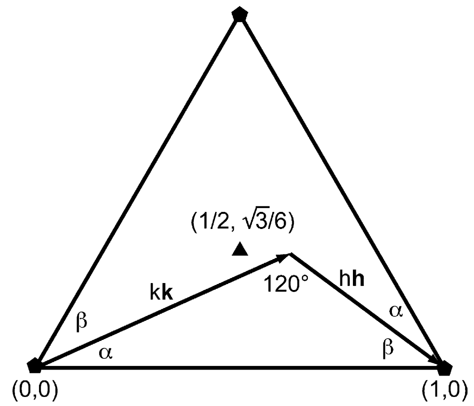


Fig. 6. Triangular face of an icosahedron with unit edge length. Fivefold and 3-fold symmetry axes shown as solid pentagons and triangle and the angles between the lattice vectors and the base of the triangle are designated as α and β. As a consequence of rotational symmetry, the sum $\alpha + \beta$ must be equal to 60° and the angle between the lattice vectors is 120°.

$$\frac{\sin\alpha}{hd} = \frac{\sin 120^\circ}{1} \rightarrow \sin\alpha = \frac{\sqrt{3}h}{2\sqrt{T}} \quad \text{(A2)}$$

With these results, we can represent the lattice vectors in terms of the Cartesian unit vectors as:

$$\boldsymbol{h} = \frac{(2h+k)}{2T}\hat{i} - \frac{\sqrt{3}k}{2T}\hat{j} \quad \text{(A3)}$$

$$\boldsymbol{k} = \frac{(2k+h)}{2T}\hat{i} + \frac{\sqrt{3}h}{2T}\hat{j} \quad \text{(A4)}$$

The icosahedral 3-fold axis lies at the center of the triangle, has the Cartesian coordinates $(1/2, \sqrt{3}/6)$, and can be written in terms of the lattice vectors as (h_3, k_3). Using Eqs. (A3) and (A4) to solve for (h_3, k_3), leads to the following pair of relations:

$$h_3(2h+k) + k_3(2k+h) = T \quad \text{(A5)}$$

$$k_3h - h_3k = \frac{T}{3} \quad \text{(A6)}$$

Since h and k have integer values, the latter can only have integral solutions for (h_3, k_3) if T is divisible by three, thereby providing a necessary, but not sufficient, condition for having a capsomer located at the 3-fold axis. Simultaneously solving Eqs. (A5) and (A6) for (h_3, k_3) leads to the following results that hold true for all T numbers:

$$h_3 = \frac{h-k}{3}, \quad k_3 = \frac{h+2k}{3} \quad \text{(A7)}$$

Although it is not obvious, we can easily show that (h_3, k_3) have integral solutions when T is a multiple of three. Expressing the squares of the numerators in Eq. (A7) in terms of h, k, and recalling that T equals $h^2 + k^2 + hk$, we obtain

$$(h-k)^2 = h^2 + k^2 - 2hk = T - 3hk \quad \text{(A8)}$$

$$(h+2k)^2 = h^2 + 4k^2 + 4hk = T + 3(k^2 + hk) \quad \text{(A9)}$$

Since h and k are integers, the condition that the squares of the numerators be divisible by three guarantees that the numerators themselves are also divisible by three. (Any integer can be expressed as a product over prime numbers. A power of an integer will contain a particular prime factor only if the integer itself contains the factor.) Together with our earlier result, this proves that a capsomer lies at the 3-fold axis if, and only if, the *T* number is divisible by three. We note that these results are in conflict with and correct the erroneous assertion that a capsomer lies at the 3-fold axis whenever *h* and *k* are both odd (Yan et al., 2009).

References

Benson, S.D., Bamford, J.K., Bamford, D.H., Burnett, R.M., 2004. Does common architecture reveal a viral lineage spanning all three domains of life? Mol. Cell 16, 673–685.

Caspar, D.L., Klug, A., 1962. Physical principles in the construction of regular viruses. Cold Spring Harb. Symp. Quant. Biol. 27, 1–24.

Cherrier, M.V., Kostyuchenko, V.A., Xiao, C., Bowman, V.D., Battisti, A.J., Yan, X., Chipman, P.R., Baker, T.S., Van Etten, J.L., Rossmann, M.G., 2009. An icosahedral algal virus has a complex unique vertex decorated by a spike. Proc. Natl. Acad. Sci. USA 106, 11085–11089.

Goldberg, M., 1937. A class of multi-symmetric polyhedra. Tohoku Math. J. 43, 104–108.

Horne, R.W., Wildy, P., 1961. Symmetry in virus architecture. Virology 15, 348–373.

Manyakov, V.F., 1977. Fine structure of the iridescent virus type I capsid. J. Gen. Virol. 36, 73–79.

Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J.E., Douglas, T., Young, M., 2004. The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. Proc. Natl. Acad. Sci. USA 101, 7716–7720.

Rossmann, M.G., Johnson, J.E., 1989. Icosahedral RNA virus structure. Annu. Rev. Biochem. 58, 533–573.

Simpson, A.A., Nandhagopal, N., Van Etten, J.L., Rossmann, M.G., 2003. Structural analyses of Phycodnaviridae and Iridoviridae. Acta Crystallogr. D Biol. Crystallogr. 59, 2053–2059.

Wrigley, N.G., 1969. An electron microscope study of the structure of *Sericesthis* iridescent virus. J. Gen. Virol. 5, 123–134.

Xiao, C., Kuznetsov, Y.G., Sun, S., Hafenstein, S.L., Kostyuchenko, V.A., Chipman, P.R., Suzan-Monti, M., Raoult, D., McPherson, A., Rossmann, M.G., 2009. Structural studies of the giant mimivirus. PLoS Biol. 7, e92.

Yan, X., Chipman, P.R., Castberg, T., Bratbak, G., Baker, T.S., 2005. The marine algal virus PpV01 has an icosahedral capsid with *T* = 219 quasisymmetry. J. Virol. 79, 9236–9243.

Yan, X., Olson, N.H., Van Etten, J.L., Bergoin, M., Rossmann, M.G., Baker, T.S., 2000. Structure and assembly of large lipid-containing dsDNA viruses. Nat. Struct. Biol. 7, 101–103.

Yan, X., Yu, Z., Zhang, P., Battisti, A.J., Holdaway, H.A., Chipman, P.R., Bajaj, C., Bergoin, M., Rossmann, M.G., Baker, T.S., 2009. The capsid proteins of a large, icosahedral dsDNA virus. J. Mol. Biol. 385, 1287–1299.