

# The architecture of viral capsids based on tiling theory

R. TWAROCK\*

Departments of Mathematics and Biology, University of York, York YO10 5DD, UK

A vital constituent of a virus is its protein shell, called the viral capsid, that encapsulates and hence protects the viral genome. The surface structures of a large number of icosahedral viruses can be modelled via Caspar-Klug Theory, which has hence become one of the fundamental concepts in virology. However, growing experimental evidence have shown that a significant fraction of viruses falls out of the remit of this theory. Among them are the Papovaviridae, which are of particular interest for the medical sector as they contain cancer causing viruses. A novel approach for the prediction of the protein stoichiometry and bonding structure of icosahedral viruses based on tiling theory is discussed here. It generalises Caspar-Klug Theory, and is in particular applicable also to Papovaviridae. Besides describing the surface structures of the viruses, this approach also provides a tool for the classification of cross-linking structures and the construction of assembly models.

*Keywords:* Viral capsids; Tiling theory; Papovaviridae; Icosahedral symmetry

*Msc: Mathematics Subject Classifications:* 00A71; 62P10

## 1. Introduction

Mathematical models for the surface structures of viral capsids are important for the analysis of experimental data, and are a crucial step for the understanding of the viral replication cycle. In a landmark paper [1], have developed a theory for the classification of icosahedral viral capsids based on the principle of quasi-equivalence. From a mathematical point of view, quasi-equivalence implies that the architecture of the capsids corresponds to triangulations that are compatible with their icosahedral symmetry. In particular, the protein subunits in the capsids are assumed to be located in the corners of the triangular facets of these triangulations, and viral capsids corresponding to a triangulation in terms of  $T$  facets per face of the icosahedron are hence predicted to be formed from precisely  $60T$  protein subunits.

Caspar-Klug Theory models the locations of the protein subunits accurately for a large number of viruses, and it has hence become one of the fundamental concepts in virology. However, experimental evidence has shown that the assumptions of Caspar-Klug Theory are too restrictive to accommodate all icosahedral viruses [2]. In particular, Papovaviridae—a family of viruses linked to cancer—fall out of the remit of this theory [3,4]. Moreover, Caspar-Klug Theory does not make any predictions about the locations of the inter-subunit bonds, which is important in particular for the construction of assembly models.

A generalisation of Caspar-Klug Theory has been suggested based on tiling theory [5]. This approach solves a structural puzzle concerning the surface structure of the (pseudo-)  $T = 7$  particles in the family of Papovaviridae [6], which cannot be explained in the framework of Caspar-Klug Theory. Moreover, the tiling approach has been applied to tubular malformations in this family [7]. The tiling approach also predicts, besides the locations of the protein subunits, the locations of the inter-subunit bonds that stabilize the capsids. Therefore, it also explains cross-linking structures [8,9] and provides a basis for the study of assembly models [10].

## 2. Tiling theory for viral capsids

The use of tiling theory for the description of viral capsids is a natural generalisation of the principle of quasi-equivalence in Caspar-Klug Theory. Quasi-equivalence corresponds to triangulations, i.e. tessellations (or tilings) in terms of triangular facets, which encode the locations of the protein subunits. Broadening the scope to tilings in terms of other geometrical shapes (called tiles following terminology in tiling theory [11]) and allowing for tilings with more than one type of tile is a logical next step. However, without a guiding principle for the choice of these building blocks and their biological interpretation, this would be a mathematical game bare of biological relevance.

\*Corresponding author. Tel.: +44-1904-434160. Fax: +44-1904-434160. Email: rt507@york.ac.uk

The tiling approach to viral capsids, also called Viral Tiling Theory, is based on the assumptions that tiles are geometrical representations of interactions between protein subunits in the capsids. In particular, each tile corresponds to a geometrical shape on which the locations of the protein subunits are indicated schematically as dots. The shapes of the tiles are constrained by the fact that proteins are known to cluster in groups of three, five or six (corresponding to trimers, pentamers and hexamers in the capsids), so that decorations must be located at angles of size  $2\pi/3$ ,  $2\pi/5$  or  $2\pi/6$ , respectively. Moreover, they have to fulfil a set of assumptions related to their symmetries that are rooted in their interpretation in terms of inter-subunit interactions. The set of tiles compatible with these assumptions has been determined and the corresponding tilings have been classified [5].

For example, one of the tilings in this classification is the tiling corresponding to the (pseudo-)  $T = 7$  particles in the family of Papovaviridae (shown in figure 1 on the left).

The set of tiles consists of the rhombs and kites shown on the right. The locations of the protein subunits are shown schematically as dots on the tiles, and the bonds are indicated as spiral arms in the tiling. One observes two different types of dimer bonds, located on the global and local 2-fold axes, respectively, as well as trimer bonds represented by the kite tiles. This is consistent with experimental observations [12].

An important feature of the tiling approach is the fact that it provides geometrical building blocks with a biological

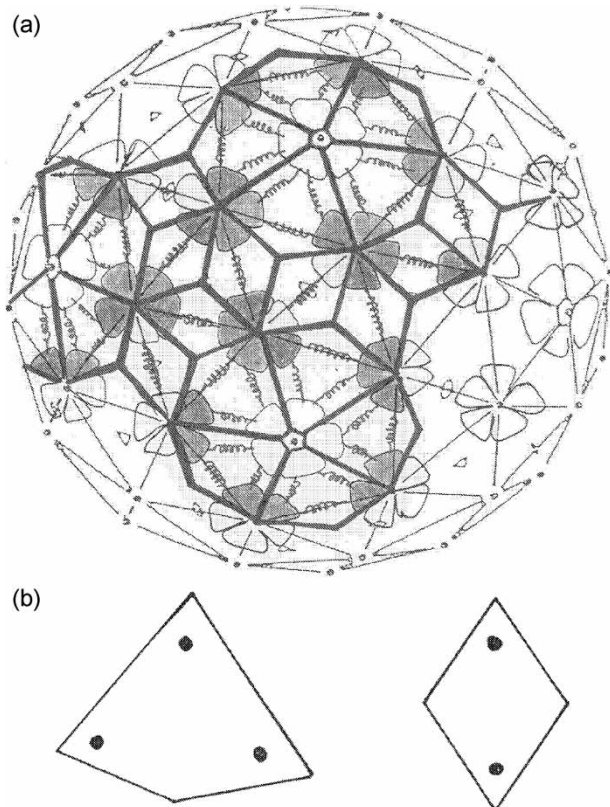


Figure 1. (a) The tiling and (b) tiles representing the surface structure of (pseudo-)  $T = 7$  particles, adapted from [7].

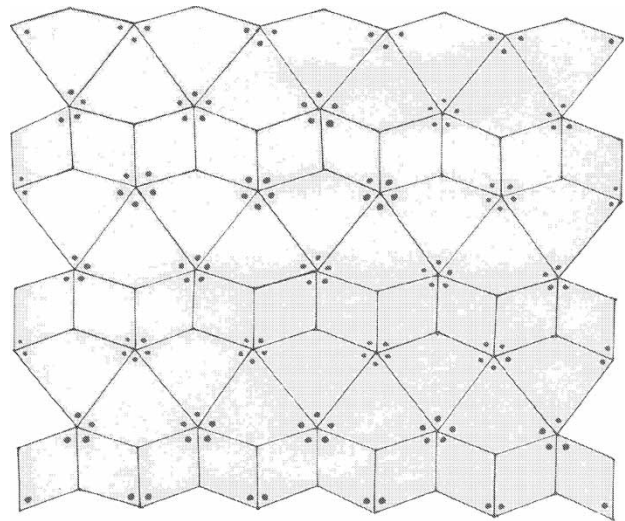


Figure 2. Planar tiling based on the tiles in figure 1.

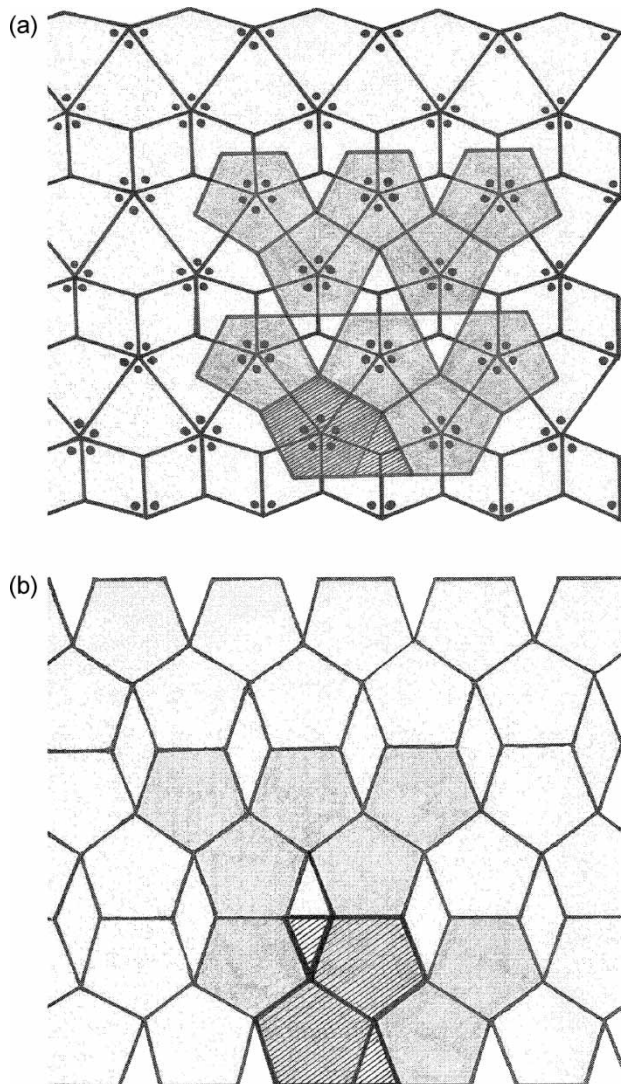


Figure 3. (a) The dual lattice compared with (b) the surface lattice in [13].

interpretation that can be used to model also other structures formed from the same capsid proteins. For example, *in vitro* experiments have shown that polyomavirus capsid protein self-assemble both into spherical particles and into non-capsid polymers [13]. Since tiles correspond to interactions between the capsid proteins, they can be used to model also the surface structures of these tubular malformations. For example, the surface lattices of the tubular structures in the family of Papovaviridae can be modelled in terms of the tiles in figure 1 (b) and the corresponding surface structures are the cylindrical compactifications of the periodic lattice in figure 2 [7].

Note that this surface lattice is difficult to derive without the insight provided by the tiling approach. Earlier models for the explanation of the surface structures of the tubular variants observed in experiments included the lattice shown in figure 3(b) (adapted from [13]). A comparison with the prediction of the tiling approach (figure 3(a)) demonstrates that the local environment around the pentamers exhibits more bonds in the case of the tiling model.

To facilitate comparison between both lattices, their fundamental domains are shown as dashed areas in the figure. In both cases, the lattices are generated by half-turns and translations from these areas, and thus the symmetry group of both lattices is  $p2$ . However, their fundamental domains are of different size. The maximally achievable bonds in the earlier model are three bonds per pentamer for every pentamer in the surface lattice. For the lattice derived via the tiling approach, the inter-subunit bonds are located both along all of the 2-fold symmetry axes (in fact there are two C-terminal arm extensions along each of the two-fold symmetry axes) as well as one further C-terminal arm extension per kite tile that is not located along a 2-fold symmetry axis. The lattice based on the tiling approach hence describes a more stable structure and is more likely to occur in nature.

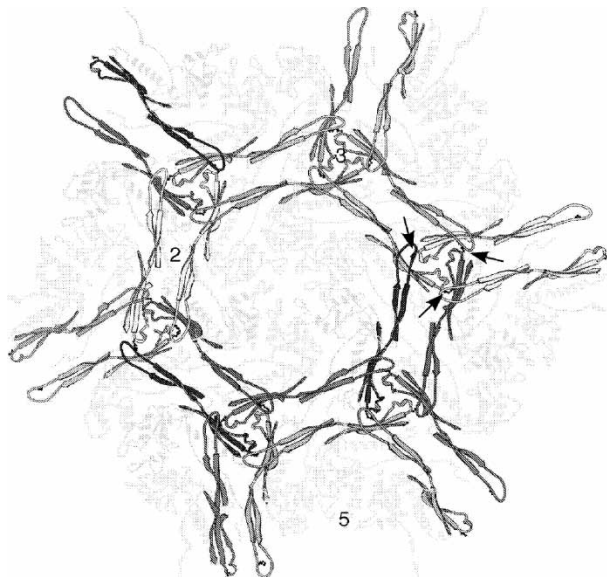


Figure 4. A cartoon showing cross-linking in HK97 bacteriophage capsids adapted from [14].

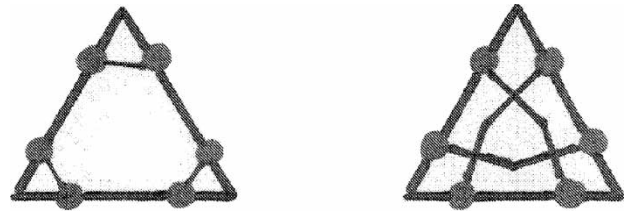


Figure 5. The two possible tiles for higher-level tilings related to head II of HK97 bacteriophage.

### 3. Extensions to higher-level tilings

By construction, the tiles in the previous section encode only dimer- and trimer-interactions. In order to model more complicated interactions such as cross-linking structures, a multi-level tiling approach is needed, in which further tilings are superimposed on the original tiling such that a set of compatibility conditions are fulfilled.

For example, the chainmail structure in head II of bacteriophage HK97 can be modelled in this way [9]. A cartoon of this cross-linking structure around a hexamer is shown in figure 4 adapted from [14].

It links monomers together in 5- or 6-membered rings, which surround each pentamer or hexamer in the capsid

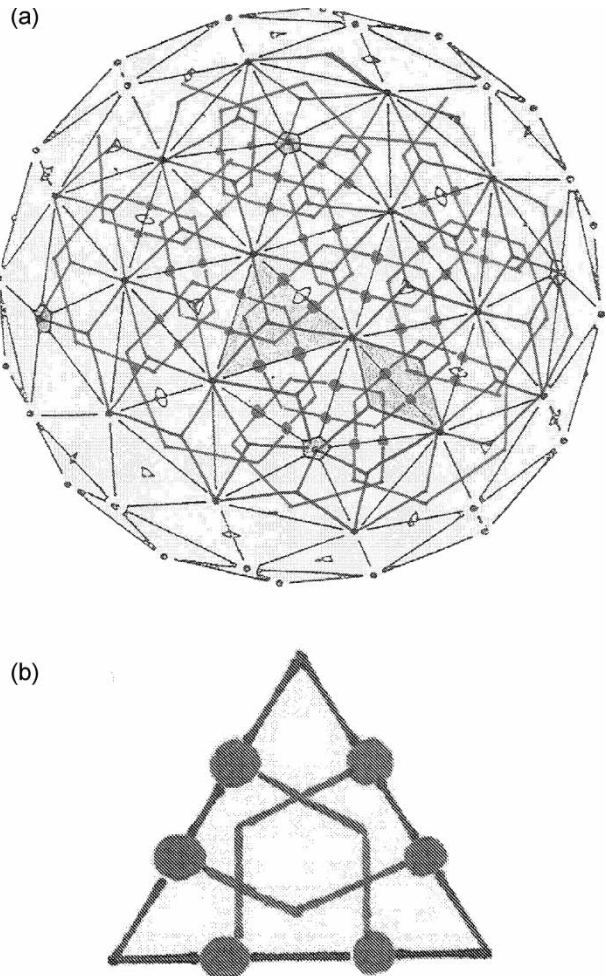


Figure 6. (a) A multi-level tiling and (b) tiles representing the cross-linking structure of head II of HK97 bacteriophage capsids.

and interlock to cover the whole capsid in a chainmail organisation that provides particular stability.

By construction, a single tiling can only represent a subset of the interactions such that each protein subunit participates in precisely one interaction. In the case of head II, the protein stoichiometry is represented as a rhomb tiling, where each rhomb represents a dimer interaction. Hence, further tilings are needed to represent the cross-linking structure. It is shown in [9] that the only tilings that qualify as higher-level tilings under a set of consistency conditions are triangulations in terms of the tiles shown in figure 5.

The triangulation corresponding to the tile shown on the right represents the cross-linking structure shown in figure 6. Both the triangular tile in the higher-level tiling and the rhomb in the original tiling are highlighted in grey in order to show the two different levels of tilings.

Bonds within the capsomers are represented by a further triangulation in terms of the tile in figure 3 on the left.

Similarly, cross-linking structures can be determined for other families of viruses [9], and the tiling approach hence provides a geometrical means to test for the possibility of cross-linking.

### Acknowledgements

Financial support by an EPSRC Advanced Research Fellowship is gratefully acknowledged.

### References

- [1] Caspar, D.L. and Klug, A., 1962, Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.*, **27**, 1–24.
- [2] Casjens, S., 1985 *Virus Structure and Assembly* (Boston, Massachusetts: Jones and Bartlett).
- [3] Liddington, R.C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T.L. and Harrison, S.C., 1991, Structure of Simian Virus 40 at 3.8 Å resolution. *Nature*, **354**, 278–284.
- [4] Rayment, I., Baker, T.S., Caspar, D.L.D. and Murakami, W.T., 1982, Polyoma virus capsid structure at 22.5 Å resolution. *Nature*, **295**, 110–115.
- [5] Twarock, R., 2004, Viral tiling theory: Protein stoichiometry and bonding structure for viral capsids based on tiling theory. Preprint, submitted to *J. Math. Biol.*
- [6] Twarock, R.J., 2004, *Theor. Biol.*, **226**, 477.
- [7] Twarock, R., 2005, Mathematical models for tubular structures in the family of Papovaviridae. *Bull. Math. Biol.*, **68** in press.
- [8] Twarock, R., 2005, *Classification of Higher-level Tilings in Viral Tiling Theory*, in preparation.
- [9] Twarock, R. and Hendrix, R., 2005, Cross-linking in HK97 bacteriophage capsids via tiling theory. Preprint, submitted to *J. Theo. Biol.*
- [10] Keef, T., Taormina, A. and Twarock, R., 2005, Assembly models for papovaviridae based on tiling theory. Submitted to *Phys. Biol.*
- [11] Senechal, M., 1996 *Quasicrystals and Geometry* (Cambridge University Press).
- [12] Modis, Y., Trus, B.L. and Harrison, S.C., 2002, Atomic model of the papillomavirus capsid. *EMBO J.*, **21**, 4754–4762.
- [13] Kiselev, N.A. and Klug, A., 1969, *J. Mol. Biol.*, **40**, 155.
- [14] Helgstrand, Ch., Wikoff, W.R., Duda, R.L., Hendrix, R.W., Johnson, J.E. and Liljas, L., 2003, The Refined Structure of a Protein Catenane: The HK97 Bacteriophage Capsid at 3.44 Å Resolution. *J. Mol. Biol.*, **334**, 885–899.