Viral Tiling Theory: a promising insight in the structure and assembly of viral capsids

A. Taormina

Background

Viruses are fascinating nanomachines. They are simple organic systems - consisting of a very compact genome and a protective protein shell or ***capsid*** - which hijack host cells ten times their size in animals and plants to reproduce, and which have the potential to kill.

Advances in virology and the design of anti-viral therapeutics rely strongly on an understanding of the viral replication cycle [1] and, in particular, of the mechanisms which trigger capsid disassembly (allowing the release of the viral genome in the cytoplasm of the host cell) as well as the assembly of new viral capsids which encapsulate the replicated genome and proceed to infect other cells.

The importance of mathematical models for the structure of viral capsids and their assembly has been recognized for decades and is demonstrated by the significance of the 40-year old Caspar-Klug theory of quasi-equivalence [2] (**Box 1**) for the classification of viral capsids and for three-dimensional reconstructions of viral capsids from experimental data [3].

Box 1: Caspar-Klug theory of quasi-equivalence in a nutshell

A large number of known viral capsids adopt a spherical shape with icosahedral symmetry, that is, the capsids possess six 5-fold, ten 3-fold and fifteen 2-fold rotation axes. This experimental observation was anticipated in [4], where early topological and geometrical considerations led to the conclusion that viral capsids should have the symmetry of one of the five Platonic solids. Since viruses must optimize their survivability through adaptation, it is, with hindsight, no surprise their capsids choose the symmetry of the icosahedron, as the latter has the largest volume-to-surface ratio of the Platonic solids. The physical origin of icosahedral symmetry in viruses is under active investigation [5].

Small capsids are organised in clusters of protein subunits centered on the global symmetry axes of the icosahedral structure. But these symmetry axes are not sufficient to accommodate the protein subunits of larger spherical viruses, which occur in multiples of 60, the order of the rotational symmetry group of the icosahedron. The *quasi-equivalence* theory advocated by Caspar and Klug consists in organising the protein clusters around the global symmetry axes, but also around some *local* symmetry axes (6-fold) of the icosahedral capsid. The underlying rule for generating large icosahedral capsids is as follows: the building blocks are not the protein subunits themselves, but rather *pentamers* and *hexamers*, which are regular polygonal groupings of five and six individual capsid protein *asymmetric* subunits.
To build a capsid, one places a pentamer on each vertex of the icosahedron's six 5-fold symmetry axes (this accounts for $5 \times 6 \times 2 = 60$ individual subunits), and fills the space between pentamers with hexamers of the same side length. The bigger the capsid, the more hexamers one can fit, but only in the proportion of 12 to $10(T-1)$, where $T = h^2 + k^2 + hk$ with $h,\ k \in \mathbb{Z}$ is called the *triangulation number* and determines the structure of the surface lattice.

However, the existing mathematical models cannot account for important classes of viruses - including families of cancer-causing viruses of prime importance for the health sector - and are unable to explain a large number of phenomena which are vital for a successful design of anti-viral therapeutics. For instance, according to the Caspar-Klug theory (**Box 1**), quasi-equivalent capsids should contain $60T$ protein subunits, with a $T = 1, 3, 4, 7, 9, \ldots$ surface lattice (the largest viruses observed have $T = 25$). However an increasing number of viruses are being observed that are exceptions. A notorious example is the Simian Virus 40 (SV40): it has 360

subunits arranged on a $T = 7$ surface lattice [6]. Such exceptions seem to contradict quasi-equivalence and corroborate the idea that this principle might not be the leading mechanism governing capsid assembly. In particular it has been suggested that protein conformational flexibility might be more of a constraint [7] and a method has also been developed in [8] that produces a normalised score for classifying the virus reflecting the degree of quasi-equivalence by analysing the interfaces between the protein subunits.

A very promising approach to solve the classification puzzle and yield a brand new and attractive perspective on viral structure is Viral Tiling Theory (VTT). It is at the centre of the proposed research, where it will be used to model capsid assembly and genome packaging. Its mathematical foundations lie in Group Theory and Tiling Theory, and it was recently developed by Twarock [9] (**Box 2**), initially in an attempt to address some of the shortcomings of the Caspar-Klug theory regarding cancer-causing viruses in the family of Papovaviridae. The distinctive feature of the latter is the fact that the surface lattices of their large icosahedral viral capsids are composed only of pentamers while the Caspar-Klug theory predicts the presence of both pentamers and hexamers for larger triangulation numbers.

VTT successfully describes the capsid structure of Papovaviridae while still reproducing the tessellations (tilings) of the viruses in the Caspar-Klug classification. The predictive power of VTT is significantly enhanced, compared to Caspar-Klug theory, through its ability to locate the *bonds* between protein subunits, and not only the location of the protein subunits themselves. It is proving a versatile and powerful ally in tackling the puzzles of modern structural virology, one of them being the mechanisms of viral capsid assembly which we will investigate in this project.

Box 2: Viral Tiling Theory

Tiling theory describes how surfaces can be tessellated in terms of a set of building blocks called tiles [10]. It has been successfully used in the study of quasicrystals [11], i.e. alloys with a non-crystallographic rotational symmetry and long range order. An important class of examples are Penrose tilings [12], because the corresponding vertex sets, that is the collections of corner points of the tiles in the tessellations, are suitable models for a significant number of real life quasicrystals. Although viral capsids are compact 2-dimensional structures, as opposed to the planar 2-dimensional structures modelling quasicrystals, they share similar types of non-crystallographic rotational symmetries correlated to the tile shapes. Hence similar types of mathematical techniques play an important rôle in the two contexts. Viral capsids may indeed be tessellated in terms of a set of tiles similar to the ones observed in quasicrystals. The tilings encode interactions between protein subunits and both the locations of the protein subunits and of the inter-subunit bonds can be read off from the tilings.

Remarkably, VTT exploits the concept of symmetry to the full. Its group theoretical roots lie in the classification of all local symmetry axes of icosahedral structures, which are determined via the affinisation of the non-crystallographic Coxeter group H_3 using a method inspired by the projection formalism known from the theory of quasicrystals and Penrose tilings [13, 14]. Protein subunits are located in the corners of the tiles that meet at global and local symmetry axes, as illustrated in Fig.2 of [15] for viruses in the family of Papovaviridae such as SV40.

Given the important rôle played by the viral capsid in the life-cycle of a virus, interference with its assembly in an attempt to reduce the infectiousness of the virus has been studied for some time. Recent simulations of virus infection have confirmed that assembly is a viable target in the fight against diseases [16]. Virus capsid assembly has been considered from various points of view in the literature. Besides the approach of molecular dynamics [17] and of thermodynamics as self-organisation of disks on a sphere [18], combinatorial optimisation studies have been performed in [19]. Two additional methods are the local rules approach [20], where assembly is constrained by a set of local rules that indicate possible locally allowed configurations for the

protein subunits, and an equilibrium approach due to Zlotnick [21], where kinetic rate equations determine the concentrations of the assembly intermediates. A distinctive feature of the latter approach is the fact that the local bonding structure of the capsomers is *not* taken into account when constructing the assembly models. While this is an appropriate simplification for the viruses studied in [21], it does not accurately reflect the fact that the capsids of Papovaviridae are formed from pentamers differing by the structure of the inter-subunit bonds surrounding them. In two very recent papers [15], the information on local environments around pentamers encoded in VTT was combined with the equilibrium approach of Zlotnick to obtain a capsid assembly model for viruses of the Papovaviridae family.

It is known that besides isometric viral particles, non-isometric particles and tubular structures [22] can also arise during assembly. While spherical particles and closed non-isometric particles can package the viral genome and are hence infectious, the tubular variants are not infectious. It is therefore of interest to predict the concentration of the tubular variants during assembly and to describe the mechanisms which trigger their formation.

The infectiousness of a virus relies on both a correct capsid assembly and the presence of RNA or DNA genome within it. It is therefore desirable to unravel the mechanisms of genome packaging - i.e. how it is inserted and folded in the capsid - both to stop the infection process and to devise efficient packaging of non-native nucleic acids for gene therapy. There are essentially two classes of models of genome packaging [23, 24]. Either the capsid proteins are attached to a specific site at the nucleic acid and the protein subunits assemble to form a complete capsid, or a nearly completed capsid is assembled and the genome is inserted before the final assembly step [25]. It has recently been observed that certain single stranded RNA viruses have their genome packaged in a way that mirrors the icosahedral symmetry of the capsid so that the RNA is acting as a template for the capsid assembly [26]. This intimate link between capsid assembly and genome packaging indicates that VTT could be further exploited to develop a model of genome packaging.

References

[1] D.D. Richman, Nature **410**, 995 (2001); H. Lodish *et al*, *Molecular Cell Biology*, W.H. Freeman & Co, 4th ed. (2000).

[2] D.L.D. Caspar and A. Klug, Cold Spring Harbor Symp. Quant. Biol. **27** 1 (1962).

[3] T.S. Baker *et al*, Microbiology and Molecular Biology Reviews **63**, 862 (1999).

[4] F. Crick and J. Watson, Nature **177**, 473 (1956).

[5] R. Zandi *et al*, Proc. Natl. Acad. Sci. **101, 44**, 15556 (2004).

[6] T. Stehle *et al*, Structure **4(2)**, 165 (1996).

[7] T. Dokland, Structure **8**, 157 (2000).

[8] K.V. Damodaran *et al*, J. Mol. Biol. **324**, 723 (2002).

[9] R. Twarock, J. Theor. Biol. **226**, 477 (2004); Bull. Math. Biol. **68** (2005).

[10] B. Grünbaum and G. C. Shephard , *Tilings and Patterns*, W. H.Freeman, New York, 1987.

[11] D. Shechtman *et al*, Phys. Rev. Lett. **53** (1984).

[12] R. Penrose, Bull. Inst. Math. Appl. **10**, 266 (1974).

[13] J. Patera and R. Twarock, J. Phys. A, Math. Gen. **35** 1551 (2002).

[14] R. Twarock, *A toolkit for the construction of icosahedral particles with local symmetry axes*,
q-bio.BM/0508015.

[15] T. Keef, A. Taormina and R. Twarock, *Assembly models for Papovaviridae based on tiling theory*, q-bio.BM/0508031, Phys. Biol. 2 (2005); T. Keef, C. Micheletti and R. Twarock, *Master equation approach to the assembly of viral capsids*, q-bio.BM/0508030.

[16] I.D. Endy and J. Yin, Antimicrob. Agents Chemother. **44**, 1097 (2000).

[17] D. Rapaport *et al*, Computer Physics Communications, 231 (1999).

[18] R.F. Bruinsma *et al*, Phys. Rev. Lett. **90**, 24 (2003).

[19] V.S. Reddy *et al*, Biophysical Journal **74**, 546 (1998); N. Horton and M. Lewis, Protein Science, **1**, 169 (1992).

[20] B. Berger *et al*, Proc. Natl. Acad. Sci. USA **91** 7732 (1994).

[21] A. Zlotnick, J. Mol. Biol. **241** 59(1994); Virology **277** 450 (2000); D. Endres and A. Zlotnick, Biophysical Journal **83**, 1217 (2002).

[22] W.R. Wikoff and J. E. Johnson, Current Biology **9** R296 (1999); J.A.G. Briggs *et al*, EMBO Journal **22**, 1707 (2003); N.A. Kiselev and A. Klug, J. Mol. Biol. **40** 155 (1969).

[23] K.E. Richards *et al*, R. J. Mol. Biol. **78**, 255 (1973); W.C. Earnshaw and S.C. Harrison, Nature **268**, 598 (1977).

[24] J. Arsuaga *et al*, Biophys. Chem. **101**, 475 (2002); J. C. LaMarque *et al*, Biopolymers **73**, 348 (2004); S. Tzlil *et al*, Biophys. J. **84**, 1616 (2003);J. Arsuaga *et al*, Proc. Natl. Acad. Sci. USA **99** 5373 (2002).

[25] W. Lucas and D.W. Knipe, http://knipelab.med.harvard.edu/pdfs/viral_capsids-Lucas.pdf (2002).

[26] J. Rudnick and R. Bruinsma, *Icosahedral packing of RNA viral genome*, cond-mat/0301305.

[27] 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).

[28] R. Twarock, *Mathematical models for tubular structures in the family of Papovaviridae*, Bull. Math. Biol. **67**, 973–987, (2005).