

Viruses and Geometry: Group, Graph and Tiling Theory Open Up Novel Avenues for Anti-Viral Therapy

Reidun Twarock

Viruses are responsible for a wide range of devastating illnesses in humans, animals and plants, yet options for treatment or prevention are limited. This is in part due to the occurrence of escape mutants that exhibit changes in the structures of the drug targets. New insights into evolutionary conserved features and constraints on virus structure are therefore key for the development of novel, more stable forms of anti-viral therapy. The area of Mathematical Virology, that focuses on the development and applications of mathematical tools to tackle open questions in virology, provides new opportunities to address this. Based on our unique interdisciplinary approach in which mathematical techniques from group, graph, tiling and lattice theory play key roles, we demonstrated that virus structure is much more constrained than previously appreciated. In particular, we identified structural constraints acting simultaneously at the level of the viral protein containers that encapsulate the viral genomes, and at the level of the packaged genomes. These new insights into virus structure have consequences for how viruses form, evolve and infect their hosts, and we developed mathematical techniques and models to quantify this. We derived new ways of characterizing genome organization within the viral capsids via graph theory, and used this to elucidate the mechanisms underpinning virus assembly. This resulted in the identification of previously unappreciated cooperative roles of single-stranded RNA genomes in virus assembly, and led to the discovery of a new anti-viral strategy against single-stranded RNA viruses, a major group of viral pathogens including HIV and Hepatitis C, in collaboration with experimental collaborators at the Universities of Leeds and Helsinki.

The following provides a summary of the key mathematical results and a discussion of their implications for our understanding of viruses and anti-viral therapy.

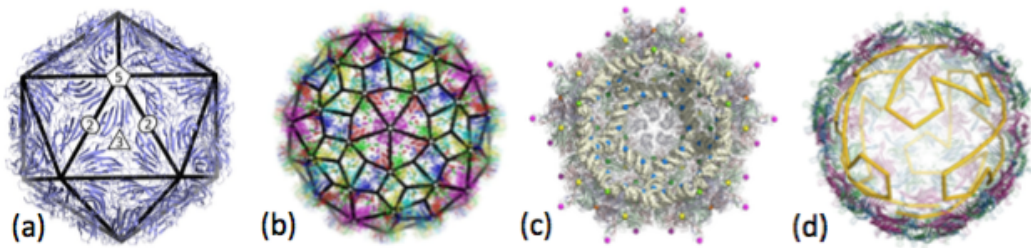


FIGURE 1. Overview of a range of mathematical tools for the modeling of virus structure: (a) Viral capsids have the characteristic 5-, 3- and 2-fold symmetry axes of icosahedral symmetry; (b) Viral Tiling theory models capsid structure in terms of tessellations indicating the positions and relative orientations of the capsid proteins; (c) point arrays representing affine extensions of icosahedral symmetry provide 3D structural constraints on the organization of material in a virus; and (d) a Hamiltonian path (yellow) modeling genome organization in proximity to capsid.

1. Affine extensions of noncrystallographic Coxeter groups and virus geometry

Protein containers encapsulating viral genomes are salient features of virus architecture. In most viruses, these containers are organized with icosahedral symmetry (cf. Fig. 1a) for reasons of genetic economy, and group theory can therefore be used to better understand virus geometry. We developed affine extensions of icosahedral symmetry to derive predictive information on the organization of viruses at different radial levels [28,32,35,45], revealing a previously unappreciated scaling principle in the overall organization of viral particles [14] (cf. Fig. 1c). Since icosahedral symmetry is non-crystallographic in three dimensions, i.e. is not compatible with periodic lattices, standard techniques for affine extensions do not apply in this case. We therefore developed a new framework for the construction of such affine extensions in the context of non-crystallographic Coxeter groups [3,8,16]. We also demonstrated that these mathematical structures, originally developed for applications in virology, can account for the atomic positions in nested fullerenes, carbon cage structures called carbon onions [7], demonstrating that they are more widely applicable in science.

2. Viral Tiling Theory in virology and bio-nanotechnology

The affine extended groups are by construction related to aperiodic tilings such as the Penrose tiling. We developed Viral Tiling theory ([27,36-44], see also Fighting viruses with mathematics, a case study by the Institute of Mathematics and its Applications) to model virus architecture via spherical tilings (cf. Fig. 1b). These tilings generalize the triangulations used in Caspar-Klug theory and indicate the relative positions of the proteins in viral capsids. Viral Tiling theory solved a long-standing open problem concerning the structures of the cancer-causing papillomaviridae [44], and also delivered models for tubular malformations that can arise during self-assembly of the major capsid protein of these viruses [37]. We recently further developed our tiling approach to make it applicable to broader classes of protein containers, including protein nanoparticles used in vaccine design (the SAPN system) [4]. Our approach resulted in a classification of all possible structural blueprints consistent with the self-assembly properties of the SAPN building blocks. In combination with information from mass spectrometry on the approximate numbers of constituent building blocks in each nanoparticle, this allowed us to

identify the surface structures of the nanoparticles conclusively, which would not have been possible via experiment alone.

3. *Lattice Transitions provide insights into structural transitions important for infection*

A significant number of viruses must undergo structural rearrangements of their protein lattices in order to become infectious. Such maturation events are transient, and therefore difficult to monitor experimentally. We have developed a new mathematical framework to model such transitions based on our description of virus structure in terms of affine extended symmetry groups and surface lattices. Our approach uses the constraints encoded by the affine extensions of the icosahedral group as descriptors of capsid geometry, and exploits their relation with bases of six-dimensional crystallographic lattices to model the transition paths via projection of lattice transitions in six dimensions [5,17]. Since the descriptors derived from affine extended symmetry are by construction related to the vertex sets of tilings, this approach can also be used to model quasi-lattice transitions [18]. In order to better understand the biophysical aspects of capsid transitions, we also used an energy function to capture the interplay of different energetic contributions and describe capsid transitions in the context of a dynamical systems approach [15]. This work provides insights into the roles of asymmetric viral components in the initiation of the capsid transition, and characterizes the consequences of the resulting symmetry breaking for the expansive motion of the capsid.

4. *Novel applications of graph theory result in a paradigm shift in our understanding of virus assembly*

The assembly of viral protein containers from their component parts is a vital step in a viral life cycle. For decades, this process had been thought of as predominantly driven by protein-protein interactions, with the viral genomes at best playing minor roles via nonspecific, electrostatic interactions. Through the use of graph theory to describe the organization of the viral genomes within the capsids (see Fig. 1d), we were able to demonstrate that, by contrast, the viral genomes play important cooperative roles in virus assembly [2,12,13,20,23-26,31]. In particular, we used the concept of Hamiltonian path to formulate constraints on the organization of the packaged genomes in single-stranded RNA viruses. In combination with a novel bioinformatics approach developed by us, this revealed the existence of multiple dispersed sequence patterns/ secondary structure elements in the genomes of these viruses that specifically interact with viral capsid proteins to promote capsid formation [10,12,13,22]. We subsequently extended this work to include a number of important Human and plant pathogens, including Hepatitis B, C and HIV [46]. We moreover developed models for virus assembly to elucidate the mechanism by which these multiple dispersed contacts confer efficiency and fidelity to capsid assembly [11]. For this we used the concept of Hamiltonian paths to quantify how viruses efficiently navigate the complex network of assembly intermediates, hence effectively solving a virus assembly equivalent of the Levinthals Paradox in protein folding [6].

5. *New mathematical models for virus assembly underpin the development of a novel anti-viral therapy*

Our assembly models are the first that take the cooperative roles of multiple dispersed packaging signals into account [6,11], and allow us to better understand the mechanisms underpinning packaging signal mediated assembly [1]. The models show that the importance of packaging signals in virus assembly can only be fully appreciated, if specific features of an

in vivo infection (such as the gradual build up of capsid protein, called the protein ramp) are factored into the analysis, perhaps explaining why the existence and crucial roles of the packaging signals had so long been overlooked. They also provide an in silico testing ground to probe the effects of anti-viral strategies targeting specific groups of packaging signals. The results show that such novel forms of anti-viral intervention can reduce viral load by delaying assembly and triggering miscapsidation of cellular RNAs. Moreover, an analysis of packaging signal motifs and positions across different viral strains, enabled by our novel graph theory based analysis techniques, revealed conserved features that lend themselves as drug targets for more stable anti-viral therapy [46].

Acknowledgement: This work received funding via a number of EPSRC responsive mode grants (EP/K028286/1, EP/J009059/1, GR/T26979/01) and a BBSRC grant (BB/L021803/1), an EPSRC Advanced Research Fellowship (GR/S51936/01&/02), the Wellcome Trust (via PhD students funded by the Doctoral Training Centre CIDCATS, of which Twarock is co-PI and co-director), a Leverhulme Trust Research Leadership Award (F/00224/AE), and a Royal Society Leverhulme Trust Senior Research Fellowship (LT130088). The translational component and patent exemplification work have moreover received funding from the Universities of York, Leeds and Helsinki.

References

1. N. Patel, E.C. Dykeman, R.H.A. Coutts, G. Lomonosoff, D.J. Rowlands, S.E.V. Phillips, N.A. Ranson, R. Twarock, R. Tuma & P.G. Stockley (2015) Revealing the density of encoded functions in a viral RNA, *PNAS* **112**, 2227-2232
2. J.A. Geraets, E.C. Dykeman, P.G. Stockley, N.A. Ranson & R. Twarock (2015) Asymmetric genome organization in an RNA virus revealed via graph-theoretical analysis of tomographic data, *PLoS Computational Biology*, **11**(3): e1004146
3. R. Twarock, M. Valiunas, & E. Zappa (2015) Orbits of crystallographic embeddings of non-crystallographic groups and applications to virology. *Acta Cryst. A* **71**, 569-582
4. G. Indelicato, N. Wahome, P. Ringler, S.A. Miller, M.-P. Nieh, P. Burkhard & R. Twarock (2015) Principles Governing the Self-Assembly of Coiled-Coil Protein Nanoparticles, to be submitted to *Biophys. J.* **110**, 646-660.
5. E. Zappa, E.C. Dykeman & R. Twarock (2014) On the subgroup structure of the hyperoctahedral group in six dimensions, *Acta Cryst A* **70**, 417-428.
6. E.C. Dykeman, P.G. Stockley & R. Twarock (2014) Solving a Leinthal's Paradox for Virus Assembly suggests a novel anti-viral therapy, *PNAS* **111**, 5361-5366.
7. P. Dechant, J. Wardman, T. Keef & R. Twarock (2014) Viruses and fullerenes - symmetry as a common thread? *Acta Cryst A* **70**:162-7 cover article; article has been featured in *Nature Physics* (Know your onion, Vol 10, p. 244, April 2014)
8. P.-P. Dechant, C. Bhm & R. Twarock, Affine extensions of non-crystallographic Coxeter groups induced by projection, *J. Math. Phys.* **54**, 093508 (2013) - cover article
9. G. Polles, G. Indelicato, R. Potestio, P. Cermelli, R. Twarock & C. Micheletti (2013) Mechanical and assembly units of viral shells identified via quasi-rigid decompositions of protein capsids, *PLOS Computational Biology* **9**, e1003331- cover article; special recommendation by the Faculty of 1000 (F1000Prime)
10. P.G. Stockley, R. Twarock, S.E. Bakker, A.M. Barker, A. Borodavka, E.C. Dykeman, R.J. Ford, A.R. Pearson, S.E. Phillips, N.A. Ranson & R. Tuma, Packaging signals in single-stranded RNA viruses: nature's alternative to a purely electrostatic assembly mechanism. *J. Biol. Phys.* **39**, 277-87 (2013)
11. E.C. Dykeman, P.G. Stockley & R. Twarock, Building a viral capsid in the presence of genomic RNA. *Phys. Rev. E* **87**, 022717 (2013)
12. E.C. Dykeman, P.G. Stockley & R. Twarock, Packaging signals in two single-stranded RNA viruses imply a conserved assembly mechanism and geometry of the packaged genome. *J. Mol. Biol.* **425**, 3235-49 (2013)
13. P.G. Stockley, N.A. Ranson & R. Twarock, A new paradigm for the roles of the genome in ssRNA viruses, *Future Virology* **8**, 531-543 (2013)
14. T. Keef, J.P. Wardman, N.A. Ranson, P.G. Stockley & R. Twarock, Structural constraints on the three-dimensional geometry of simple viruses: case studies of a new predictive tool. *Acta Crystallogr A.* **69**, 140-50 (2013) also received special mention in the Newsletter of the International Union of Crystallography
15. P. Cermelli, G. Indelicato & R. Twarock, Nonicosahedral pathways for capsid expansion, *Phys. Rev. E* **88**, 032710 (2013)

16. P.-P. Dechant, C. Bhm & R. Twarock, extensions of non-crystallographic Coxeter groups. *J. Phys. A* 45(28), 285202 (2012)
17. G. Indelicato, P. Cermelli, D.G. Salthouse, S. Racca, G. Zanzotto & R. Twarock, A crystallographic approach to structural transitions in icosahedral viruses. *J. Math. Biol.* 64,745-73 (2012) Novel Kac-Moody-type affine
18. G. Indelicato, T. Keef, P. Cermelli, D.G. Salthouse, R. Twarock & G. Zanzotto, Structural transitions in quasicrystals induced by higher dimensional lattice transitions, *Proc R Soc A* 468: 1452-1471 (2012)
19. K.M. ElSawy, R. Twarock, C.S. Verma & L.S. Caves, Peptide inhibitors of viral assembly: a novel route to broad-spectrum antivirals. *J Chem Inf Model* 52:770- 6 (2012)
20. E.C. Dykeman, N.E. Grayson, K. Toropova, N.A. Ranson, P.G. Stockley & R. Twarock, Simple rules for efficient assembly predict the layout of a packaged viral RNA. *J. Mol. Biol.* 408, 399-407 (2011)
21. K.M. ElSawy, L.S. Caves & R. Twarock, On the origin of order in the genome organization of ssRNA viruses. *Biophys. J.* 101, 774-80 (2011)
22. D.H. Bunka, S.W. Lane, C.L. Lane, E.C. Dykeman, R.J. Ford, A.M. Barker, R. Twarock, S.E. Phillips & P.G. Stockley, Degenerate RNA packaging signals in the genome of Satellite Tobacco Necrosis Virus: implications for the assembly of a T=1 capsid. *J. Mol. Biol.* 413, 51-65 (2011)
23. V.L. Morton, E.C. Dykeman, N.J. Stonehouse, A.E. Ashcroft, R. Twarock & P.G. Stockley, The impact of viral RNA on assembly pathway selection. *J. Mol. Biol.* 401, 298-308 (2010)
24. K.M. ElSawy, L.S. Caves & R. Twarock, The impact of viral RNA on the association rates of capsid protein assembly: bacteriophage MS2 as a case study, *J. Mol. Biol.* 400, 935-47 (2010)
25. E.C. Dykeman & R. Twarock, All-atom normal-mode analysis reveals an RNA- induced allostery in a bacteriophage coat protein. *Phys. Rev. E* 81, 031908 (2010)
26. E.C. Dykeman, P.G. Stockley & R. Twarock, Dynamic allostery controls coat protein conformer switching during MS2 phage assembly. *J Mol. Biol.* 395, 916- 23 (2010)
27. R. Twarock & T. Keef, Viruses and geometry where symmetry meets function, *Microbiology Today* 37: 24-27 (2010)
28. T. Keef & R. Twarock, Affine extensions of the icosahedral group with applications to the three-dimensional organisation of simple viruses. *J. Math. Biol.* 59, 287-313 (2009)
29. N. Jonoska, A. Taormina & R. Twarock, DNA cages with icosahedral symmetry in bionanotechnology in: *Algorithmic Bioprocesses, Natural Computing Series, Springer Verlag*, 141-158 (2009)
30. N.E. Grayson, A. Taormina & R. Twarock, RNA and DNA Cages with Icosahedral Symmetry, *J. Theor. Comp. Sci.* 410, 1440 (2009)
31. K. Toropova, G. Basnak, R. Twarock, P.G. Stockley & N.A. Ranson, The three- dimensional structure of genomic RNA in bacteriophage MS2: implications for assembly, *J. Mol. Biol.* 375, 824-36 (2008)
32. T. Keef, R. Twarock & K.M. ElSawy, Blueprints for viral capsids in the family of polyomaviridae. *J. Theor. Biol.* 253, 808-16 (2008)
33. N. Jonoska & R. Twarock, Blueprints for Dodecahedral DNA Cages, *J. Phys. A* 41, 304043-304057 (2008)
34. K. M. ElSawy, L. S.D. Caves & R. Twarock, Polyomaviridae assembly polymorphism from an energy landscape perspective, *Computational and Mathematical Methods in Medicine* 9, 245-256 (2008)
35. T. Keef & R. Twarock, New insights into viral architecture via affine extended symmetry groups, *Computational and Mathematical Methods in Medicine* 9, 221- 229 (2008)
36. K.M. ElSawy, A. Taormina, R. Twarock & L. Vaughan, Dynamical Implications of Viral Tiling Theory, *J. Theor. Biol.* 252, 357-369 (2008)
37. T. Keef, A. Taormina & R.Twarock, Classification of capped tubular viral particles in the family of Papovaviridae, *J. Phys.: Condens. Matter* 18, S375- S387 (2006)
38. R. Twarock & R. Hendrix, Crosslinking in Viral Capsids via Tiling Theory, *J. Theor. Biol.* 240, 419-424 (2006)
39. T. Keef, C. Micheletti & R. Twarock, Master equation approach to the assembly of viral capsids, *J. Theor. Biol.* 242, 713-21 (2006)
40. R. Twarock, A Mathematical Physicist's Approach to the Structure and Assembly of Viruses, *Phil. Trans. A* 365, 3357-3374 (2006)
41. R. Twarock, The architecture of viral capsids based on tiling theory, *J. Theor. Medicine* 6, 87-90 (2005)
42. T. Keef, A. Taormina & R. Twarock, Assembly Models for Papovaviridae based on Tiling Theory, *Phys. Biol.* 2, 175-188 (2005)
43. R. Twarock, Mathematical models for tubular structures in the family of Papovaviridae, *Bull. Math. Biol.* 67, 973-987 (2005)
44. R. Twarock, A tiling approach to virus capsid assembly explaining a structural puzzle in virology, *J. Theor. Biol.* 226, 477-482 (2004)
45. Stockley, P. G. & Twarock, R. (Eds.) *Emerging Topics in Physical Virology*, Imperial College Press, London, 2010 (ISBN:978-1-84816-464-2)
46. Patent: A novel anti-viral strategy against RNA viruses; filed 5th September 2013 (GB1315785.4)

*Reidun Twarock,
Departments of Mathematics and Biology,
York Centre for Complex Systems
Analysis,
University of York,
York YO10 5GE,
UK
reidun.twarock@york.ac.uk*