

Structure and self-assembly of viruses

Reidun Twarock

Departments of Mathematics and Biology

University of York

KITP, June 2006

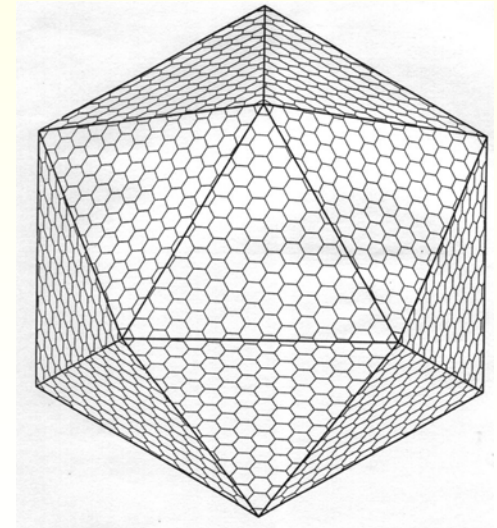
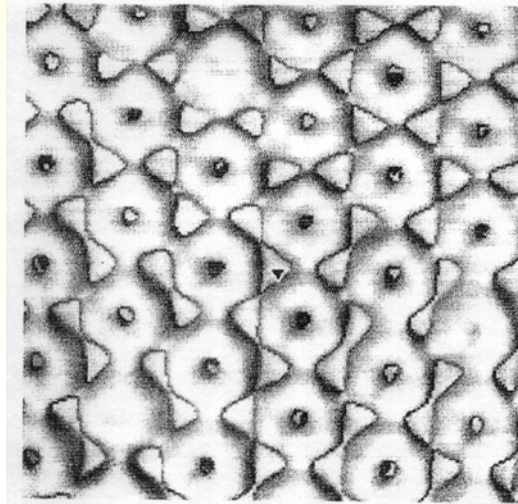
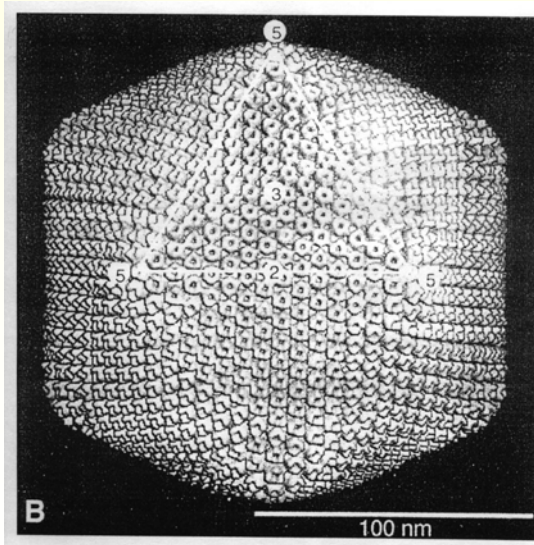
Overview:

- Part I: Virus structure
 - Caspar-Klug Models
 - Viral Tiling Theory: virus architecture via generalised lattices
 - Tubular malformations and crosslinking structures
- Part II: Assembly models
 - Prototype assembly models based on VTT
 - The role of RNA for assembly of MS2.
- Part III: Implications for genome organisation
 - Models for the structure of RNA and DNA in viral capsids

Part I: Virus structure

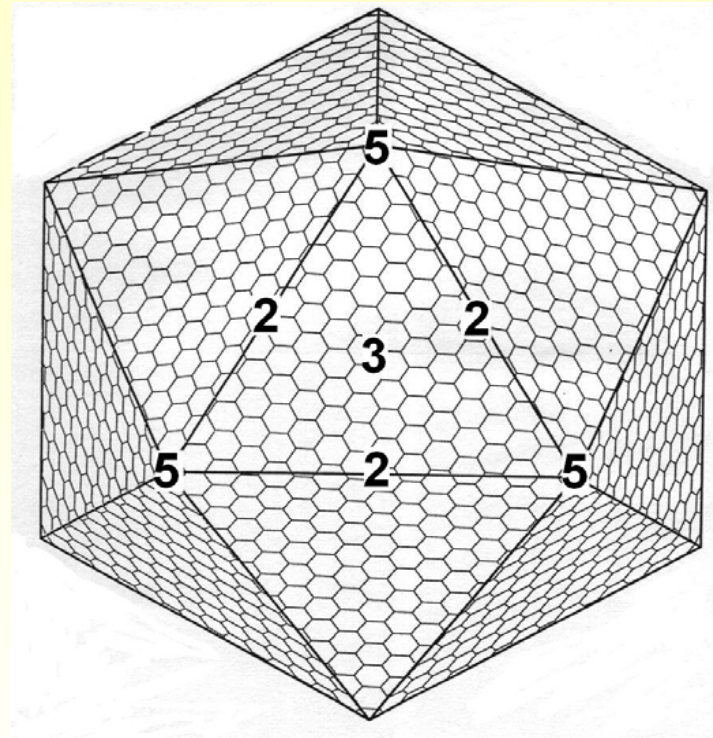
Viruses have a shell formed from proteins (called viral capsid) that encapsulates and hence provides protection for the viral genome.

Example:



Peculiarities:

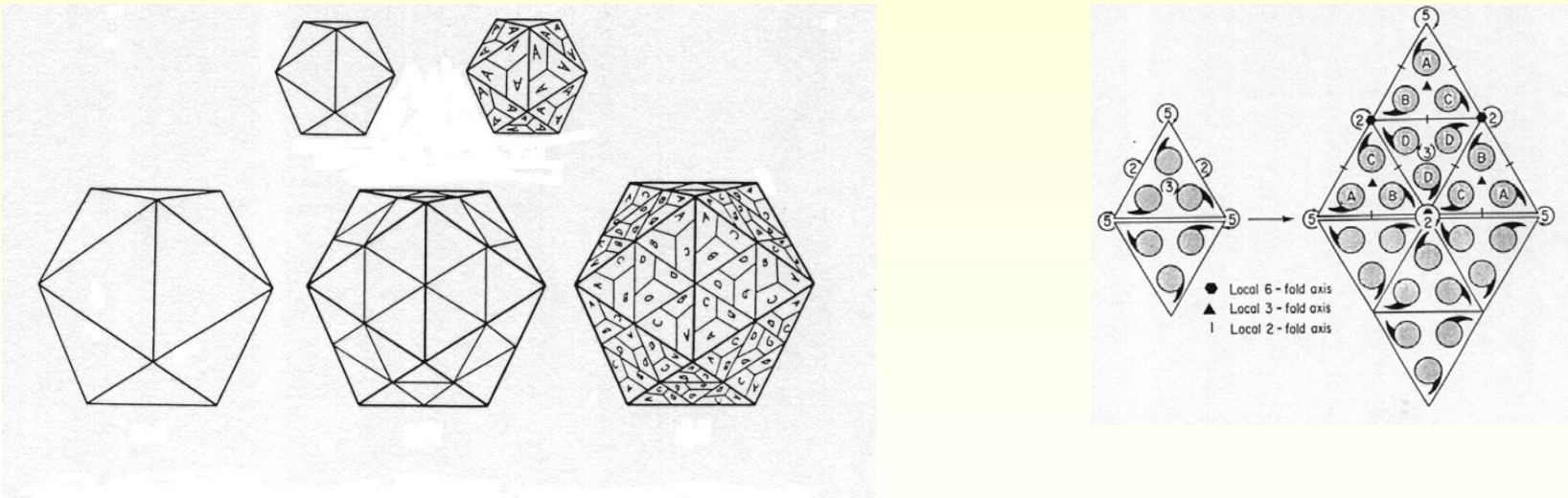
Over 50% of the known viruses have icosahedral symmetry:



There should be a general organisational principle that can be formulated based on group theory

Caspar-Klug Theory

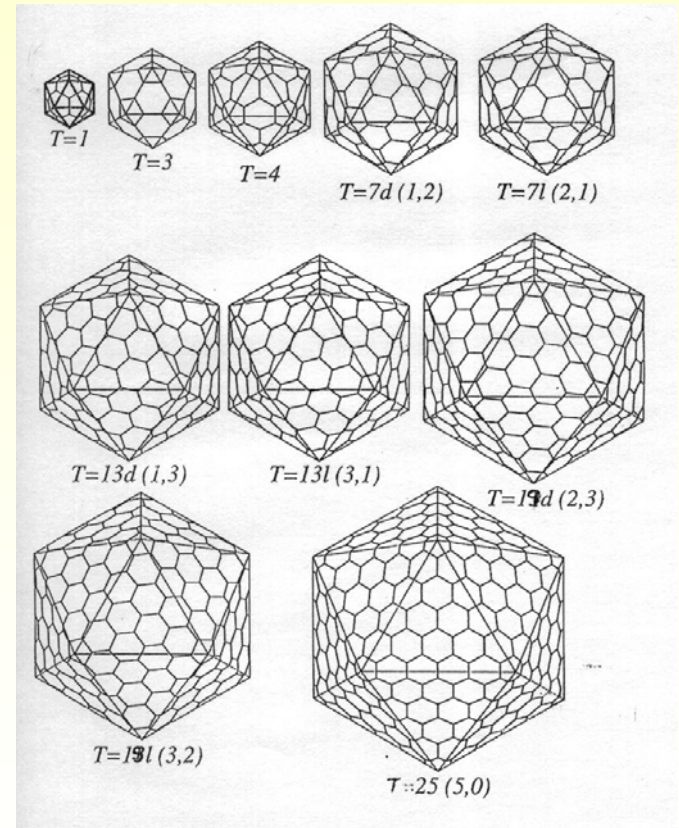
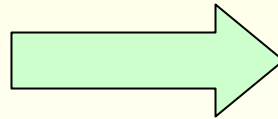
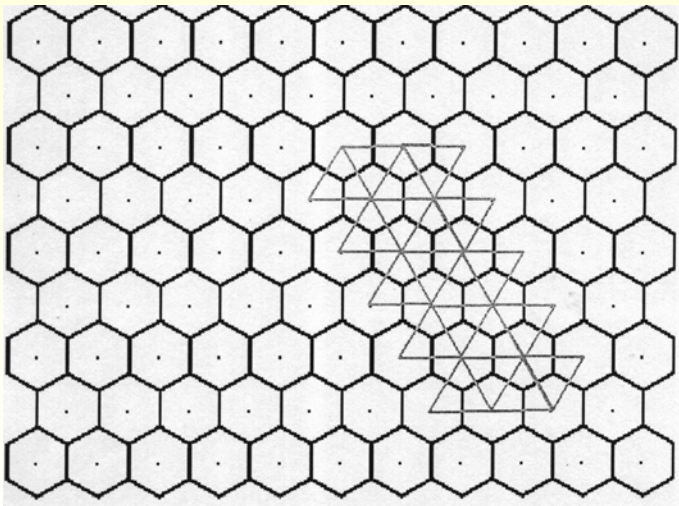
Viral capsids are modelled as icosahedral triangulations:



Locations of the proteins indicated schematically as circles.

Classification

Superposition of the surface of an icosahedron on a hexagonal grid:



Problems (1):

Caspar-Klug Theory is a fundamental concept in virology with wide-ranging applications. However:

(1) It does not explain all experimentally observed virus structures

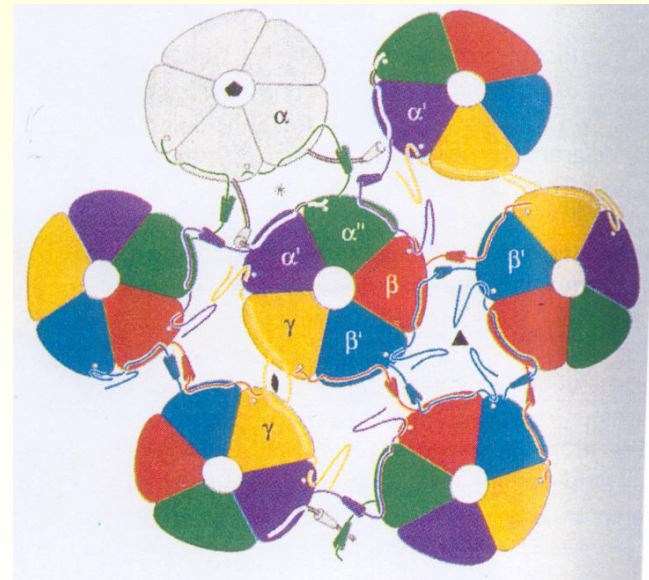


Liddington et al. have shown in Nature in 1991 that polyoma virus falls out of the scope of Caspar-Klug Theory

It has been an open problem in virology for over 10 years

Problems (2):

(2) The structure of the inter-subunit bonds cannot be explained in the framework of Caspar-Klug Theory



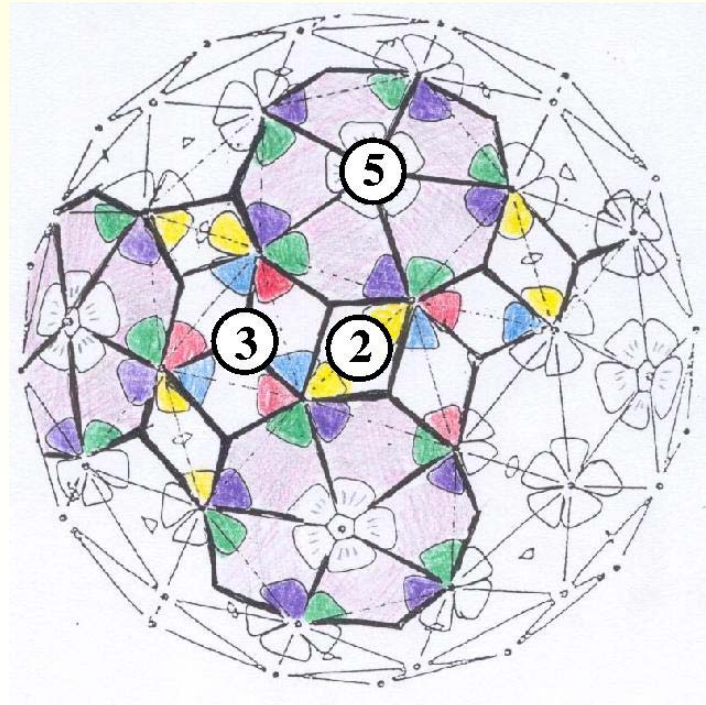
Viral Tiling Theory (VTT)

A new mathematical framework that solves these issues.

The approach is based on group theory (Coxeter groups) and leads to spherical tilings that encode the surface structures of viruses.

Example:

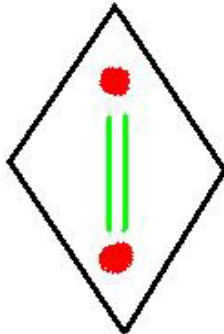
R.Twarock, *A tiling approach to virus capsid assembly explaining a structural puzzle in virology*, J. Theor. Biol. 226, 477 - 482 (2004).



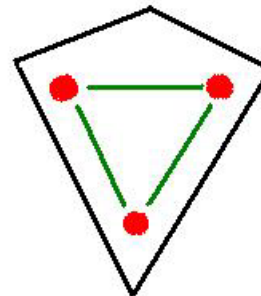
Special feature:

The building blocks of the theory, called tiles, model interactions between the proteins they represent.

Dimer interaction:



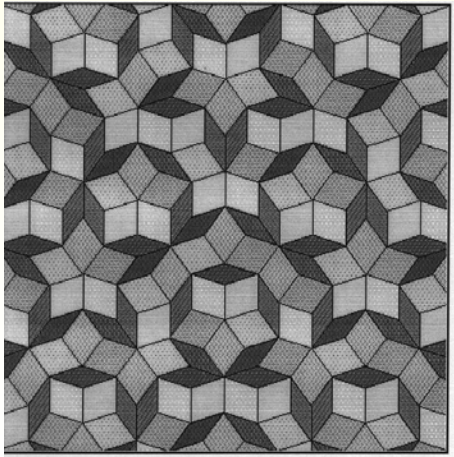
Trimer interaction:



Idea:

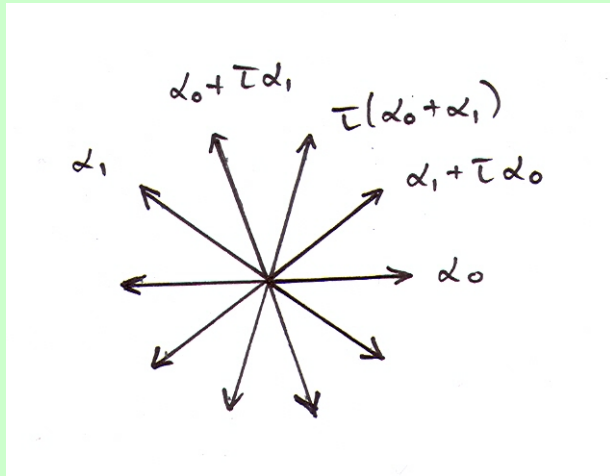
In order to derive the tilings from first mathematical principles, one uses the root system of a non-crystallographic Coxeter group to generate point sets that contain the vertex sets of the tilings.

Remark: This is similar to the construction of Penrose tilings

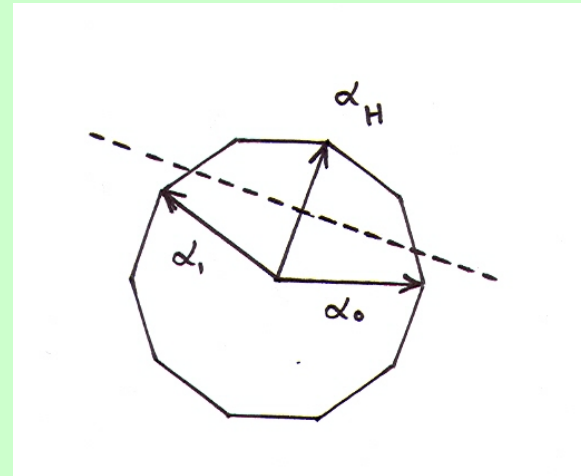


Demonstration for a 2d toy model

The root system encoding reflections:



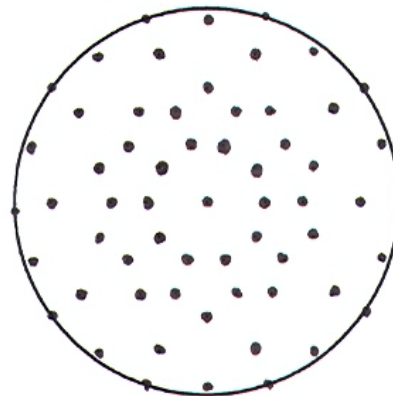
The highest root defining a translation:



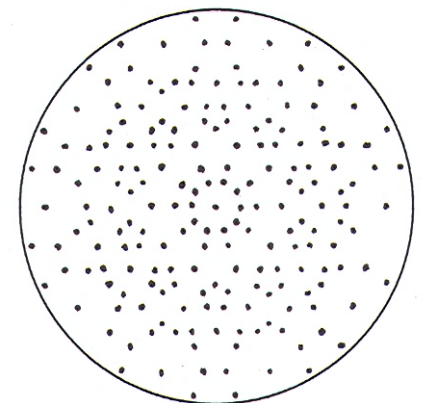
An iteration of these reflections and translation generates point sets that are subsets of Penrose tilings:

N: number of translations

N=3

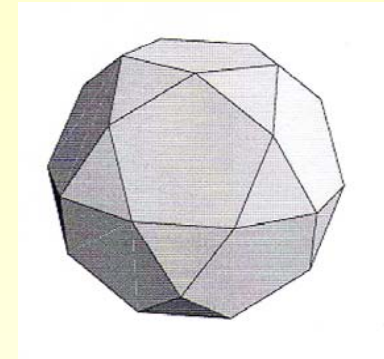


N=4

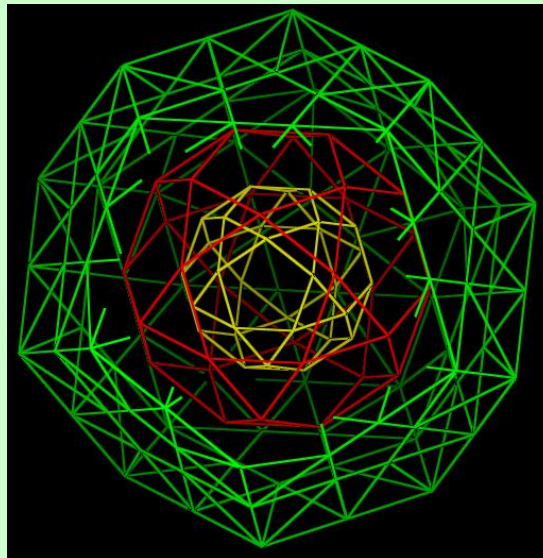


The 3d analog

The root system of the Coxeter group that contains the rotational symmetries of the icosahedron is given by the icosidodecahedron:



A similar procedure as before leads to three dimensional point sets that define nested shell structures:

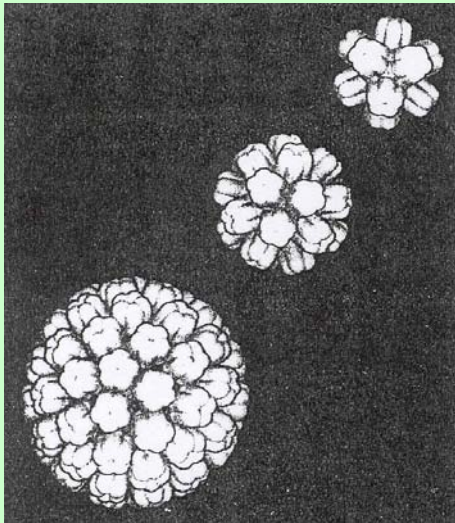


Classification

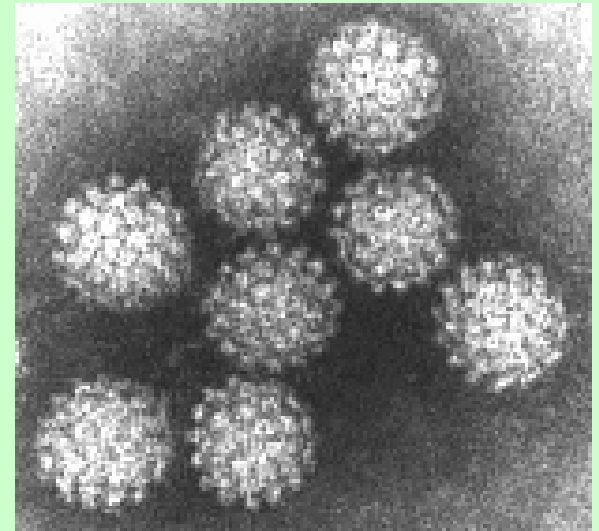
The surface structures of viruses can be obtained via this method and have been classified.

[T.Keef and R.Twarock, *A novel series of polyhedra as blueprints for viral capsids in the family of Papovaviridae*, q-bio.BM/0512047].

There are three different types of particles with all-pentamer capsids

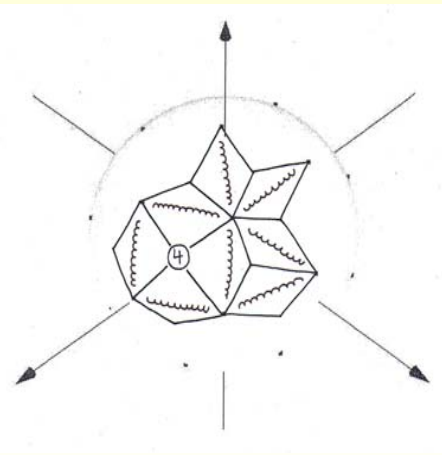
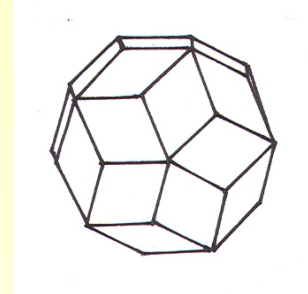


Cryo-em micrograph:



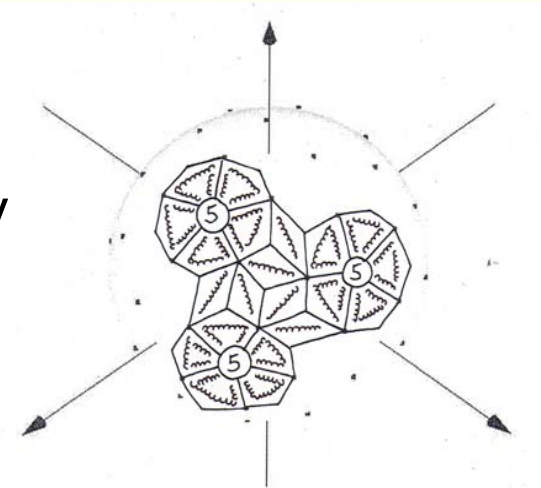
Example: all-pentamer capsids

(1) The small particle corresponds to a triacontahedron or an icosahedron



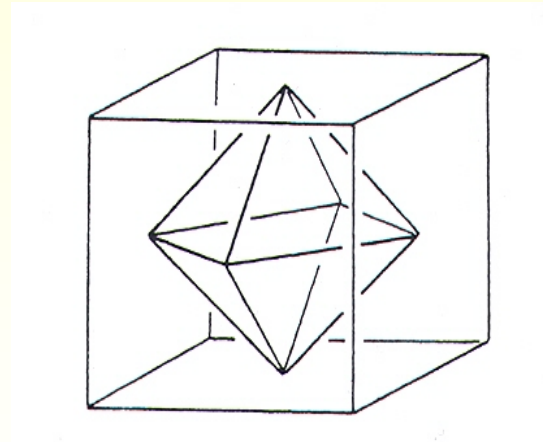
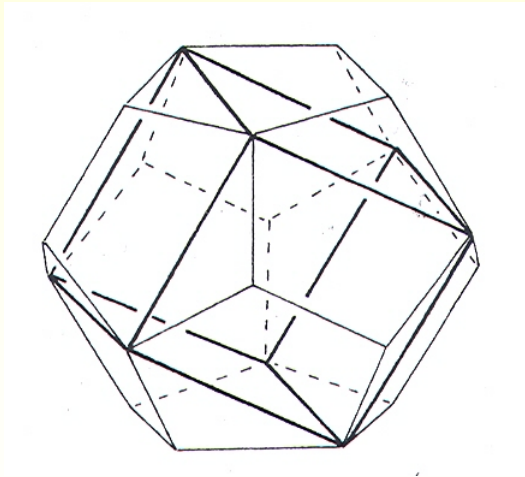
(2) The medium sized particle corresponds to the tiling shown on the left. It has octahedral symmetry.

(3) The large particle has icosahedral symmetry and corresponds to the tiling shown earlier.



Remark:

Note that it is possible to determine structures with octahedral symmetry with our approach due to the following geometric inclusions:



Comparison with experiments

(1) The relative radii of the three particles are predicted by our theory.

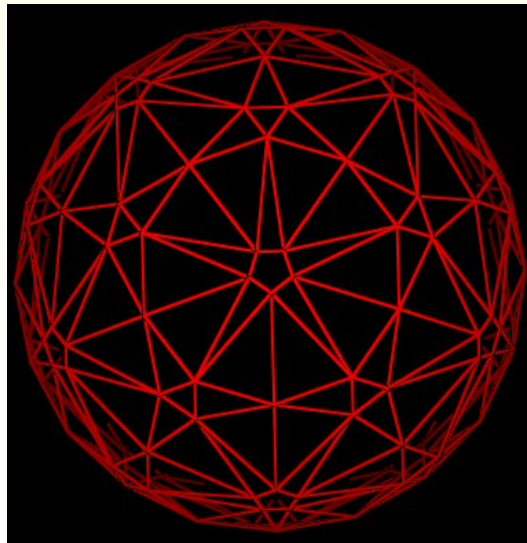
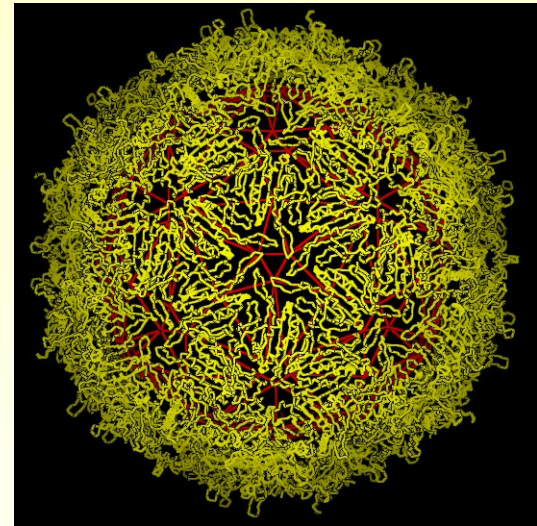
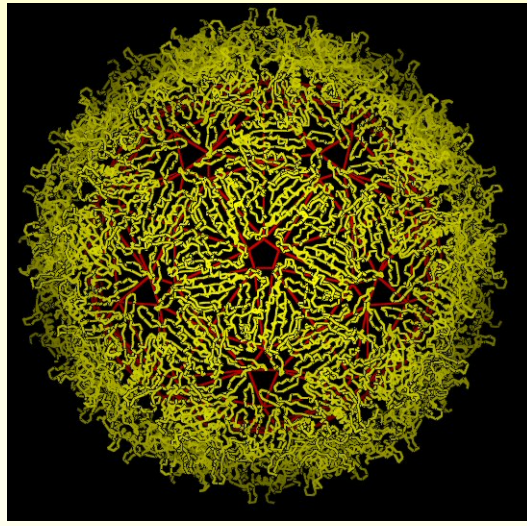
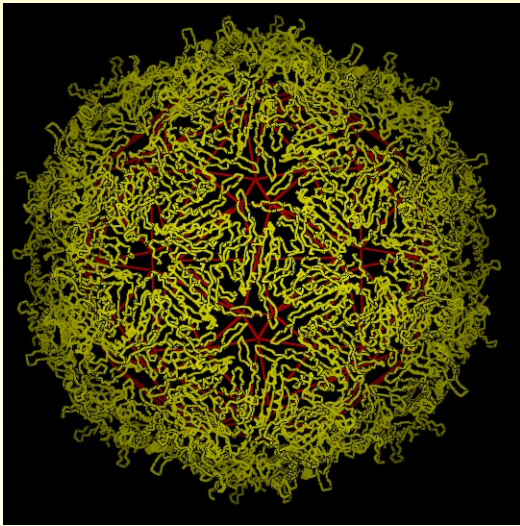
(2) Since tiles encode inter-subunit bonds, the bonding structure in the capsids is predicted by the tilings.

Our predictions in (1) and (2) agree well with experimental observations.

For experimental data on SV40, see e.g. Kaneshashi et al.

Application to other viruses

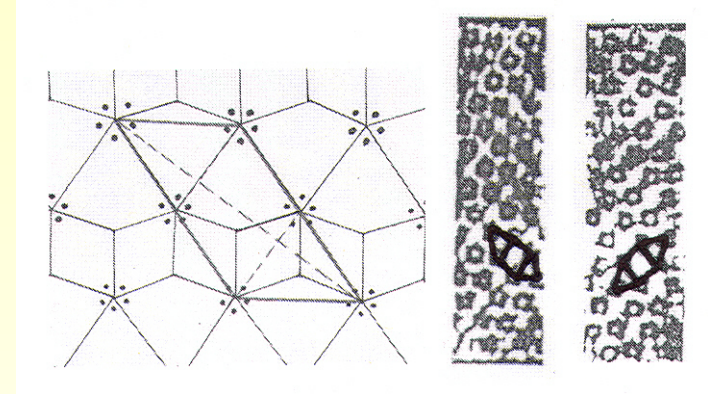
Example: MS2



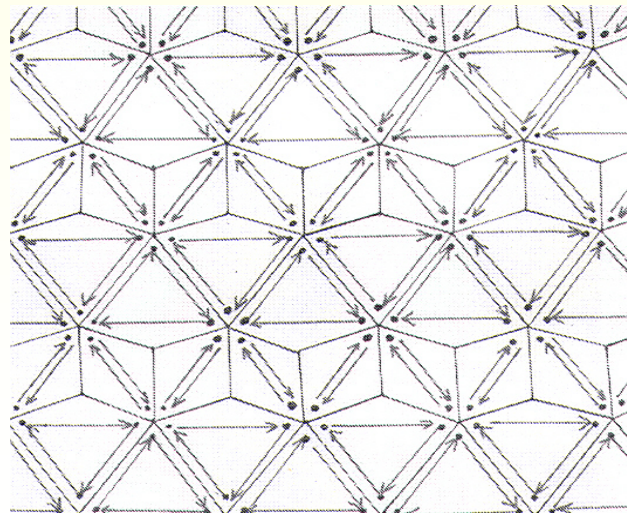
Work in progress....

Tubular malformations in VTT

With the same mathematical tools one can also describe the surface lattices of tubular malformations

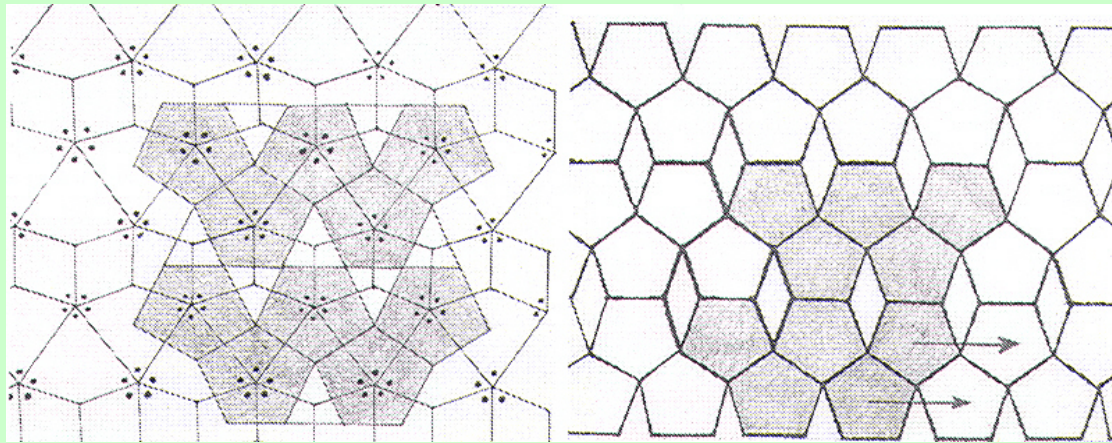


The surface lattices obtained based on group theoretical principles:



Comparison with experiments

Our surface lattices provide a better description of experimental data than those suggested previously:



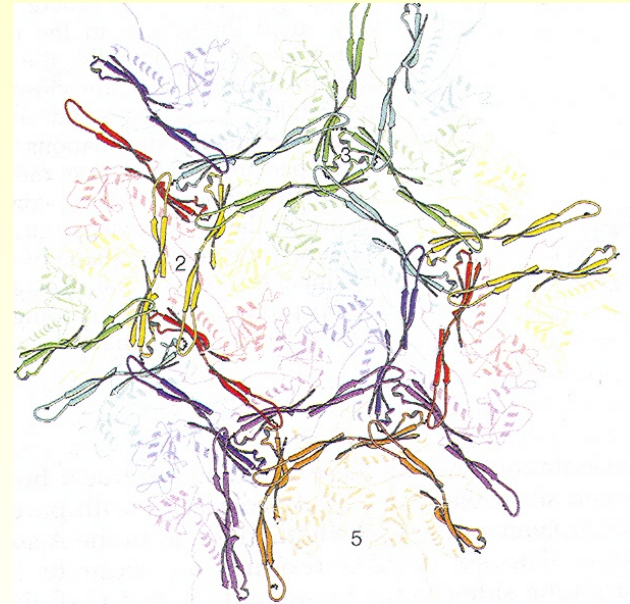
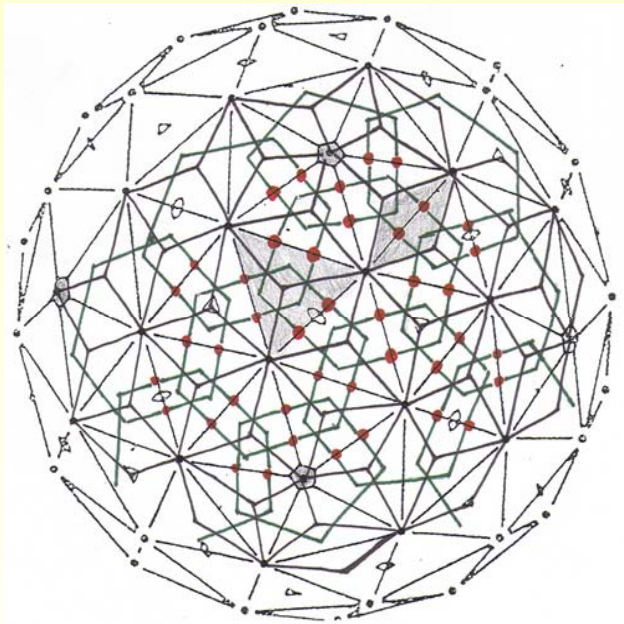
The lattice based on VTT (left) versus a lattice used earlier for the modelling of Papovaviridae tubes.

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

(b) T.Keef, A.Taormina, R.Twarock, *Classification of capped tubular viral particles in the family of Papovaviridae*, *J. Phys.: Cond. Mat.* 18, S375 (2006).

Crosslinking structures in VTT

VTT models also covalent bonds, e.g. chainmail in HK97

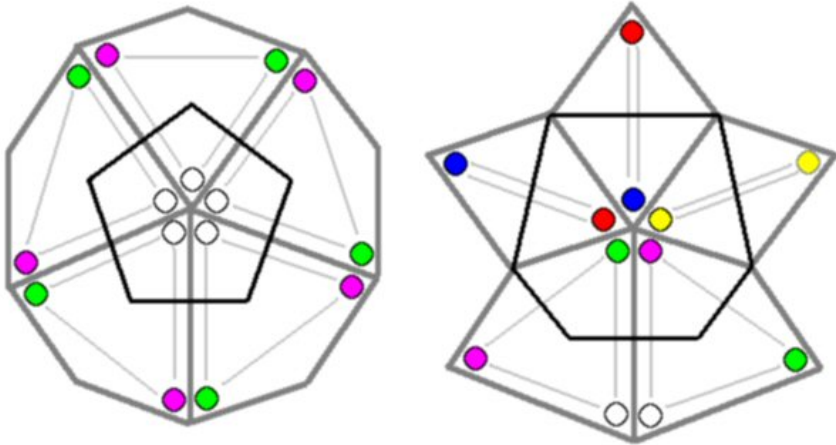


It provides a tool to decide from a theoretical point of view whether crosslinking is possible for a given viral capsid.

R. Twarock and R. Hendrix, Crosslinking in Viral Capsids via Tiling Theory, to appear in J. Theor. Biol. 2006.

Part II: Assembly models

Since tilings represent the local bonding structure in viral capsids, they can be used for the construction of assembly models.



Note:

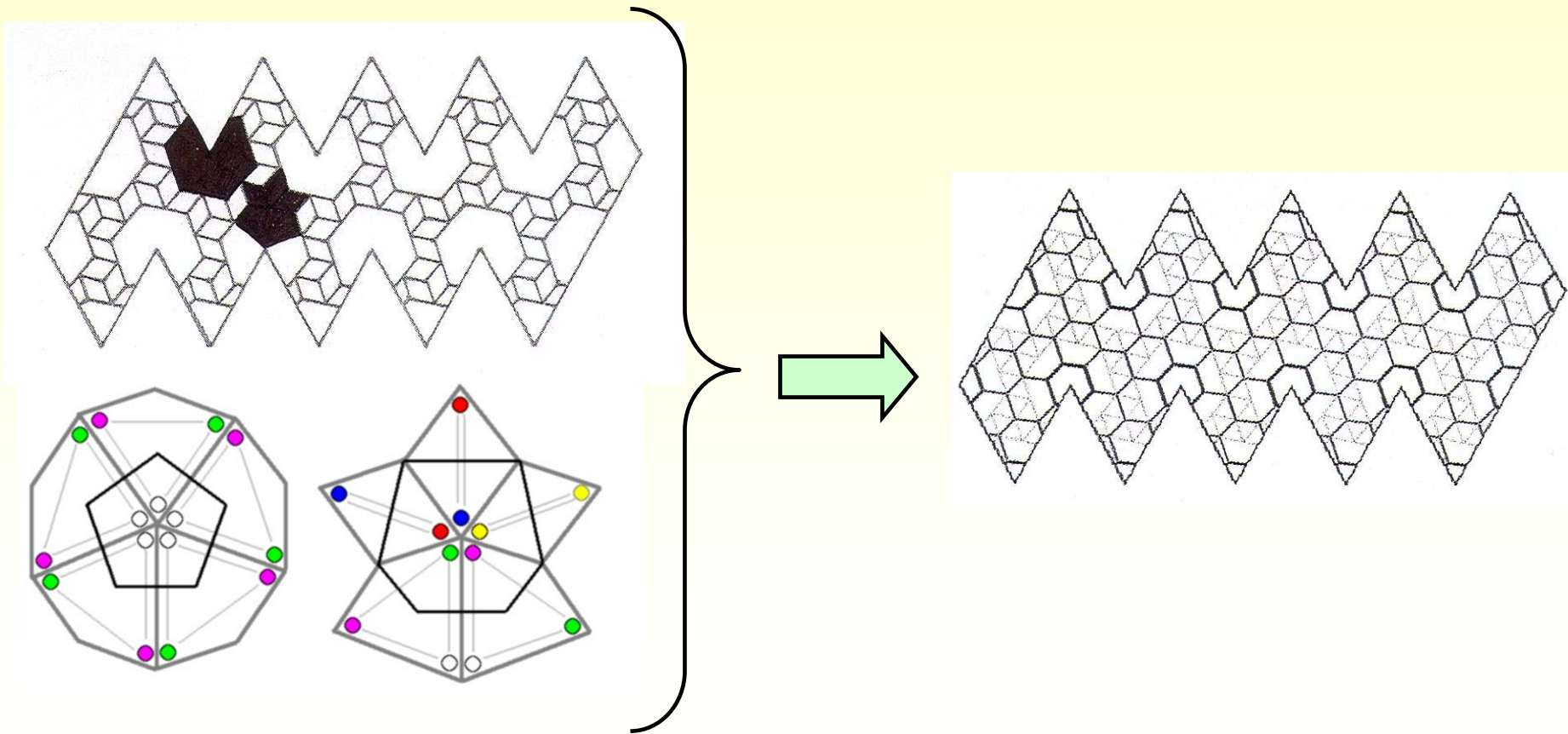
There are three different types of bonds.

Task: Characterise assembly in dependence on the bond strengths

- (a) T.Keef, A.Taormina, R.Twarock, *Assembly Models for Papovavirida based on Tiling Theory*, Phys. Biol. 2, 175-188, 2005.
- (b) (b) T.Keef, C. Micheletti, R.Twarock, *Master equation approach to the assembly of viral capsids*, to appear in Theor. Biol., June 2006.

Modelling assembly geometrically

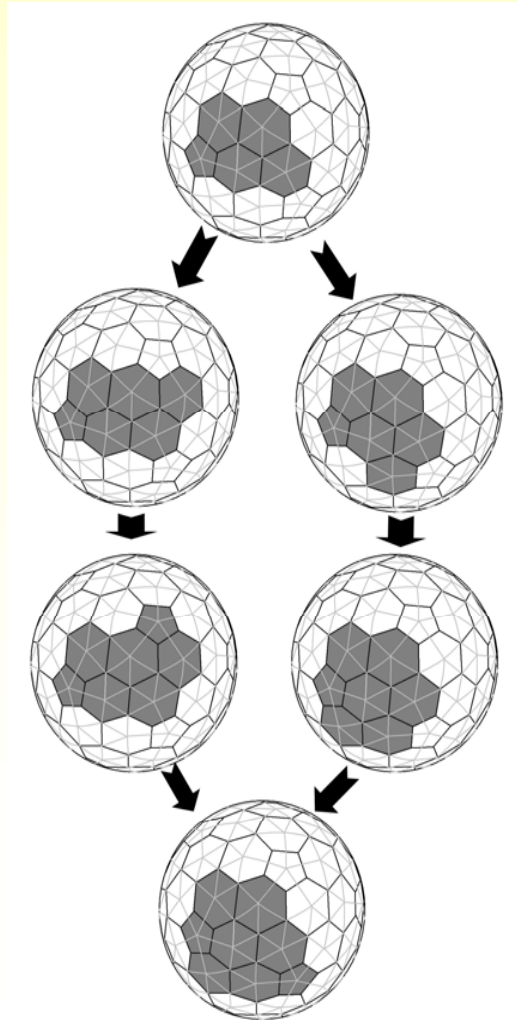
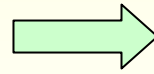
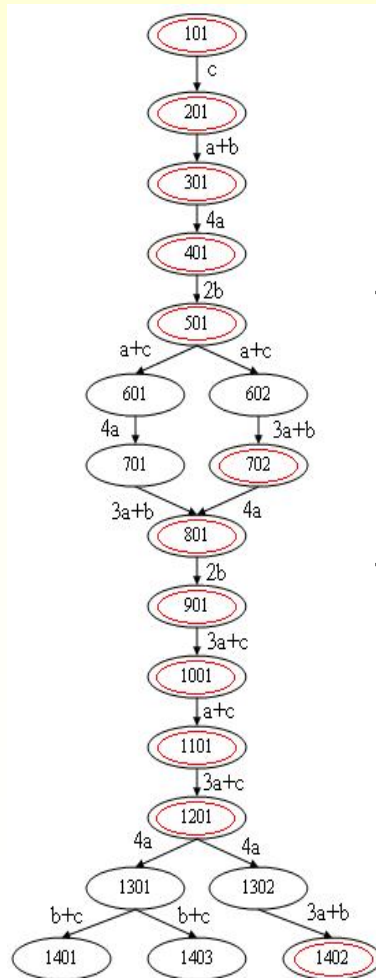
We characterise assembly intermediates in planar geometry:



Assumption: Attachment of a single building block per iteration step.

Assembly graphs

For a given choice of association (bond) energies a , b and c one obtains an assembly scenario that can be represented as a graph:



Associated assembly reactions

Association constant: $K_N := \frac{[N]}{[N-1][1]}$

Factorise: $K_N = S_{in} \cdot S_N \cdot e^{\frac{-\sum c_j \Delta G_j}{RT}}$

S_{in} : denotes the geometric degeneracy of the incoming subunit

$S_N := \frac{O_{sym}(N-1)}{O_{sym}(N)}$: denotes the degeneracy due to symmetry of the intermediates

ΔG_j : denotes the bond energies

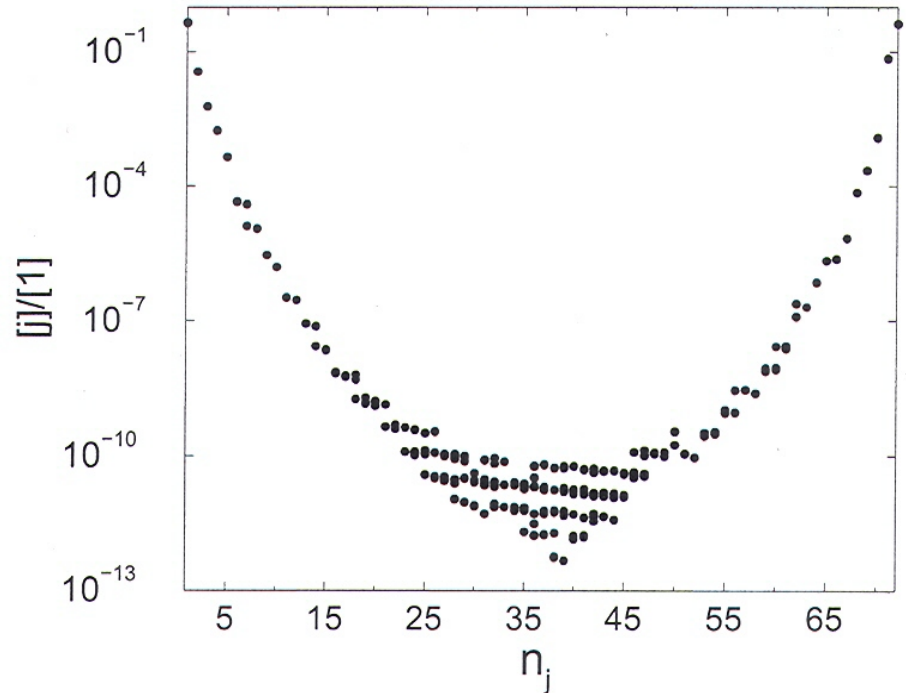
R: gas constant T: temperature

Following an approach to capsid assembly pioneered by Zlotnick.

Concentration profile

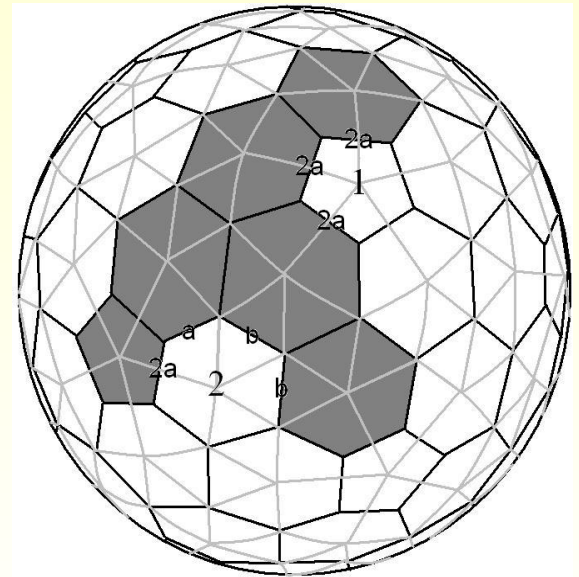
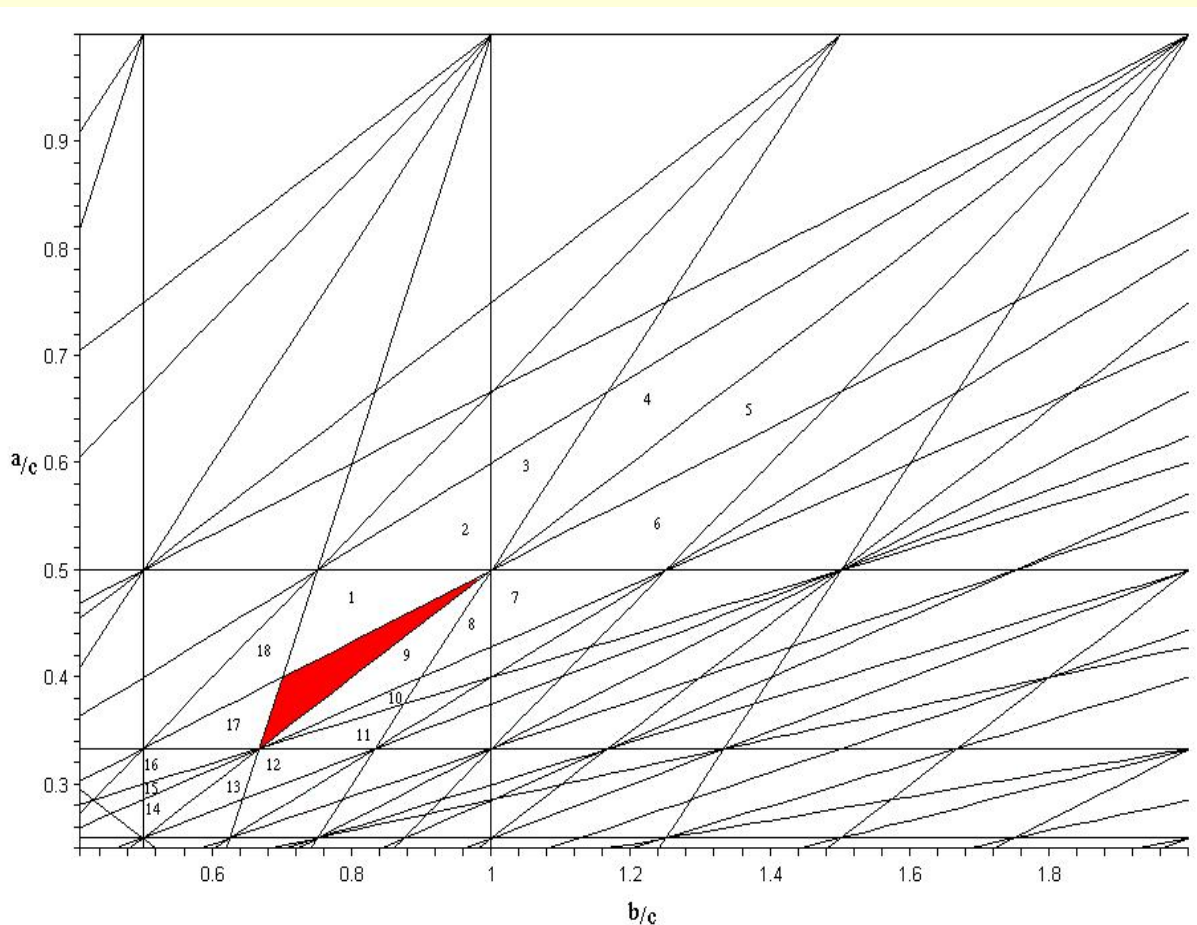
One obtains the concentrations of the assembly intermediates via a recurrence relation:

$$[n] = \Omega(N) \cdot [n - 1]$$
$$\Omega(N) = 5 \cdot [1] \cdot S_N e^{\frac{-\sum c_j \Delta G_j}{RT}}$$



Changing association energies

For different ratios a/c and b/c of the association constants, different assembly scenarios are obtained:



$$6a = 3a + 2b$$

Applications:

Via changes in the polypeptide chains of the capsid proteins, the association energies of the inter-subunit bonds can be manipulated. The diagramme tells us how this affects the assembly scenario.

Example: CCMV assembly (J.M. Johnson et al., Nano Lett.,2005)

A truncated form of the CCMV CP (lacking residues 2-36) can assemble into a mixture of shells, with 90 (native), 60 and 30 dimers.

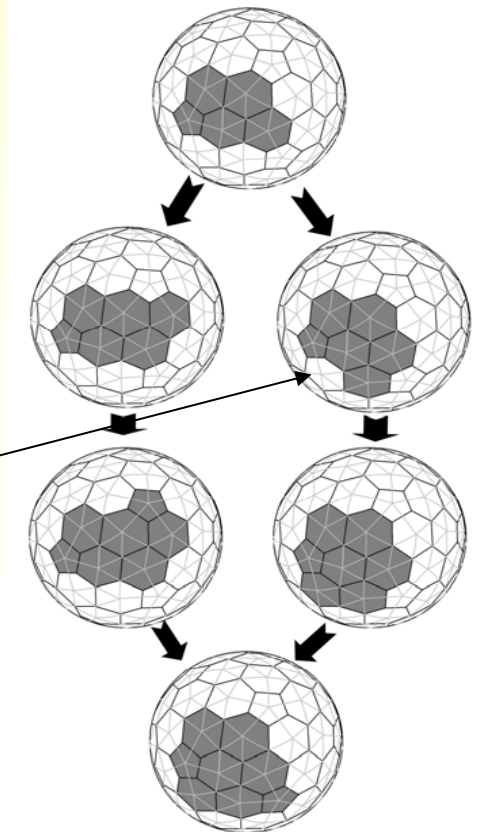
Computation of the dominant pathway

The relative probabilities of the assembly intermediates are computed via a master-equation approach.

One hence obtains the dominant pathway of assembly.

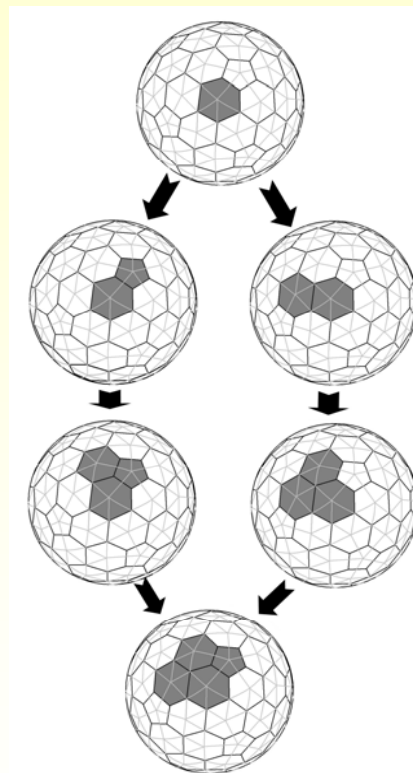
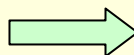
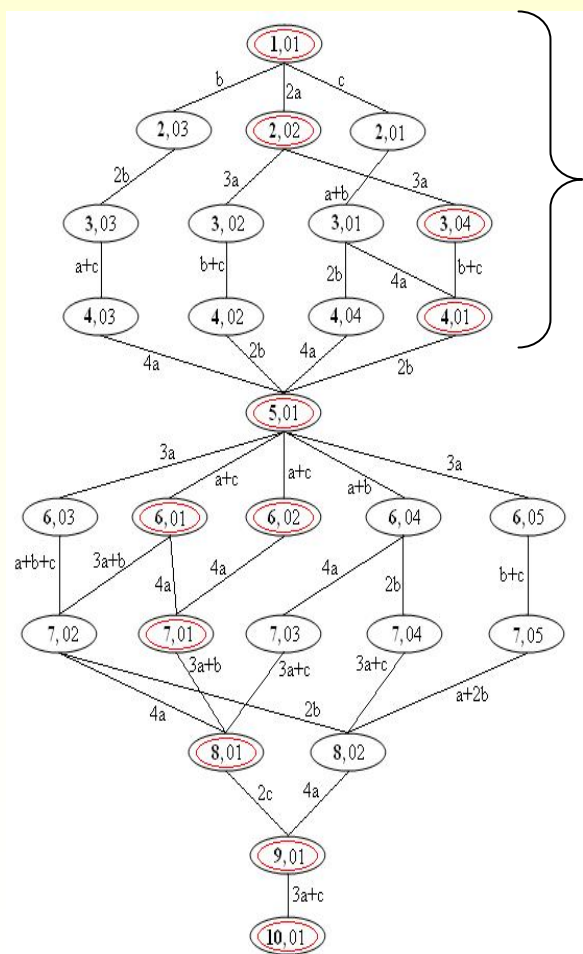
A geometric analysis of the assembly intermediates on this pathway shows a characteristic pattern:

A formation of bonds with association energies $2a$, a and b is pertinent to the selection of the dominant pathway.



Expanding the ensemble of intermediates

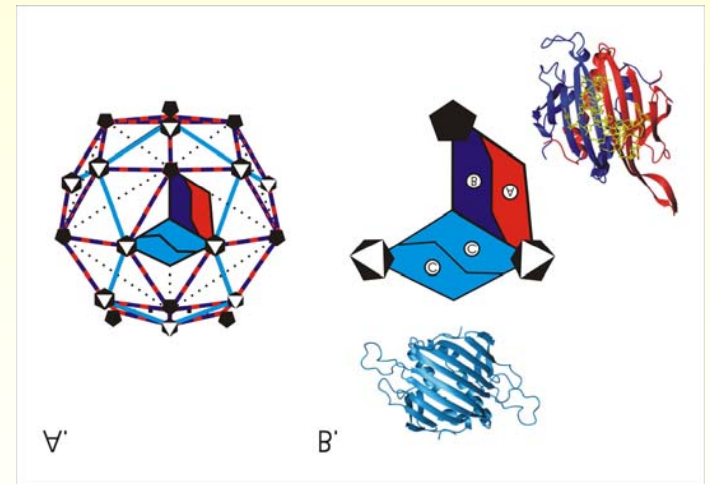
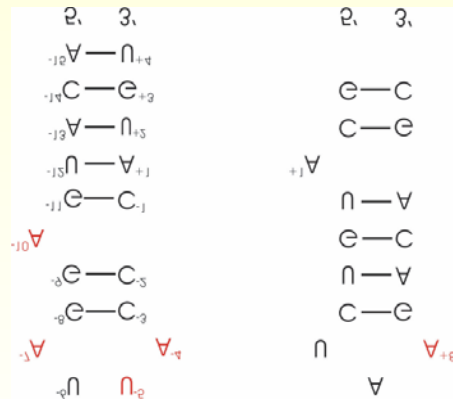
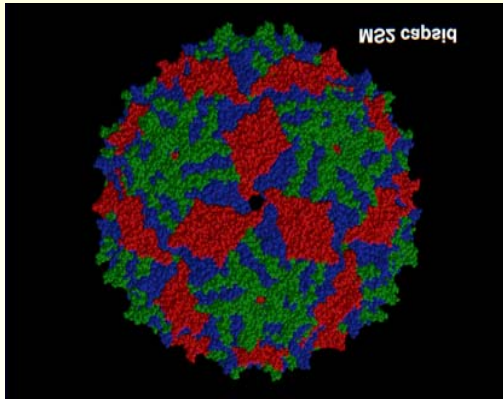
Allowing for the 20% energetically most favourable intermediates, one obtains slight changes in the dominant pathway:



A change at the start of the dominant pathway.

The role of RNA during assembly

Experimental results concerning bacteriophage MS2 suggest that the viral RNA plays an important role during capsid assembly:



In order to integrate the role of RNA into the assembly models, one needs to know more about genome organisation within viral capsids.

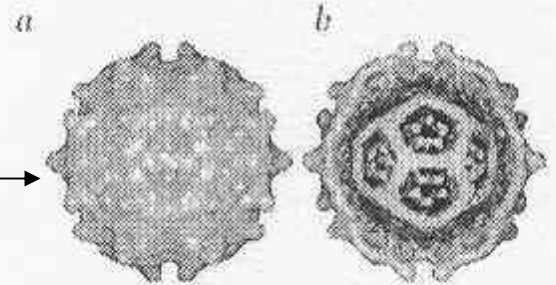
Work in progress in collaboration with the Stockley Lab.

Part III: Genome organisation

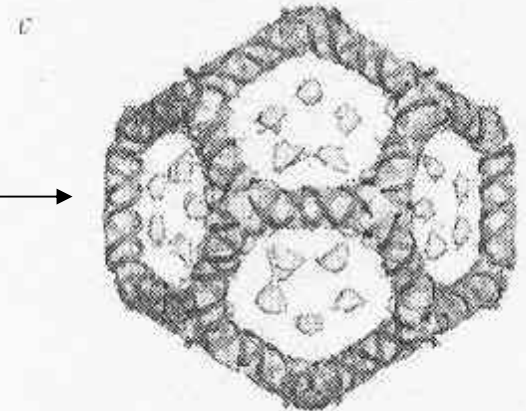
Cryo-electron micrographs show that the viral genome takes on an ordered structure within the capsid.

Example: Pariacoto virus (see work of Bruinsma and Rudnick)

The viral capsid



About 1500 of the 4322 nucleotides of the genome are confined to this dodecahedral cage, i.e. 50 nucleotides per dodecahedral edge, organised in a double-helical structure.



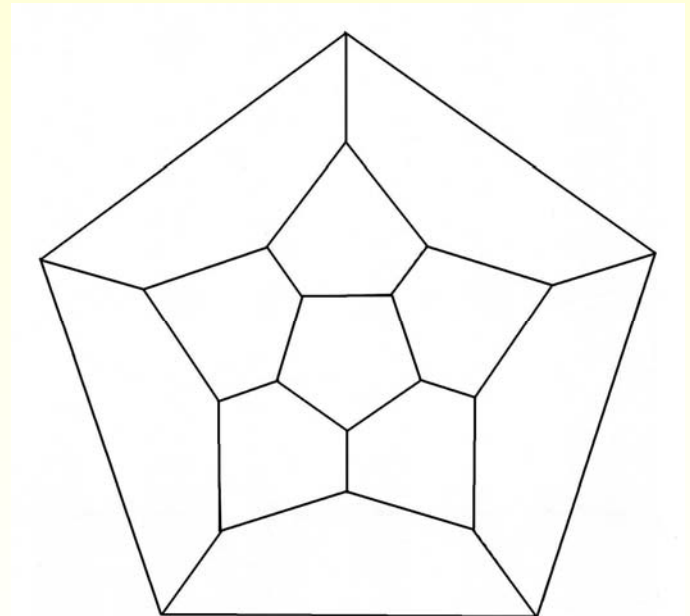
Models based on graph theory

Investigate, how the RNA is arranged on the polyhedral shell.

Consider the dodecahedral shell as a planar graph:

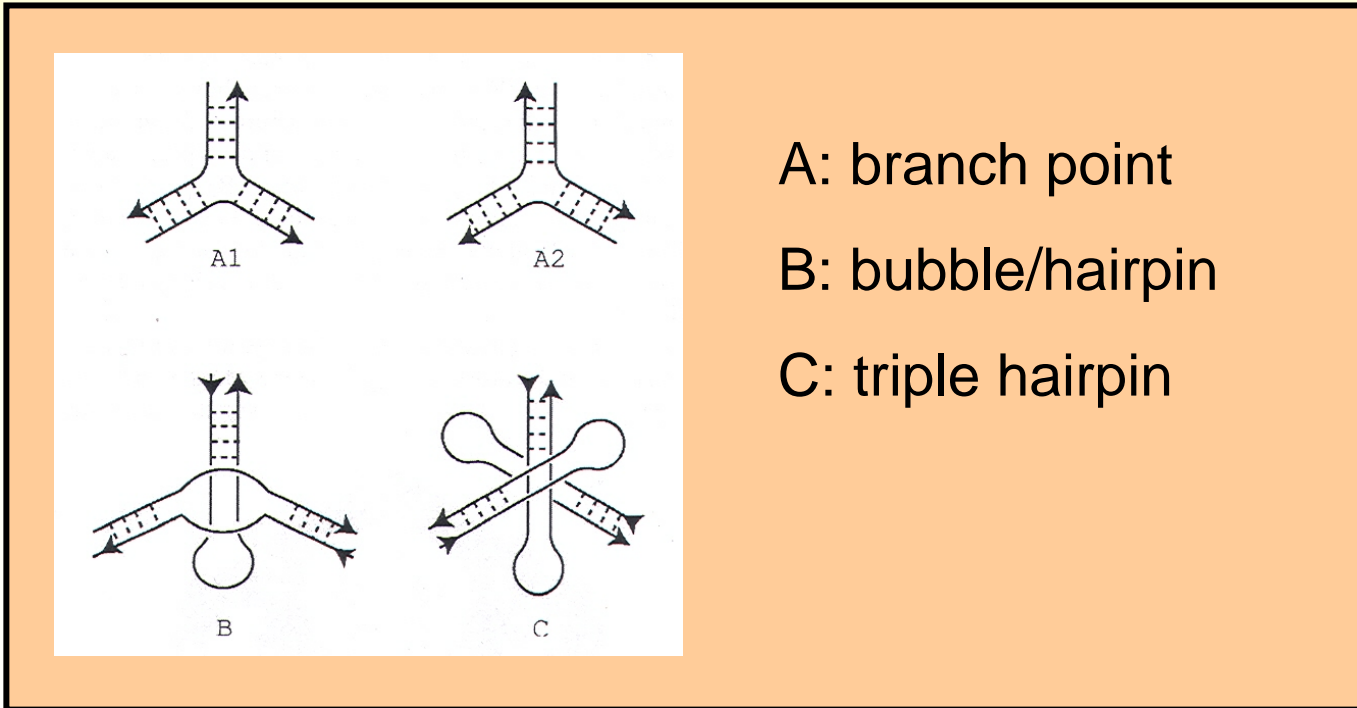
Question:

What is the energetically best arrangement of the RNA?



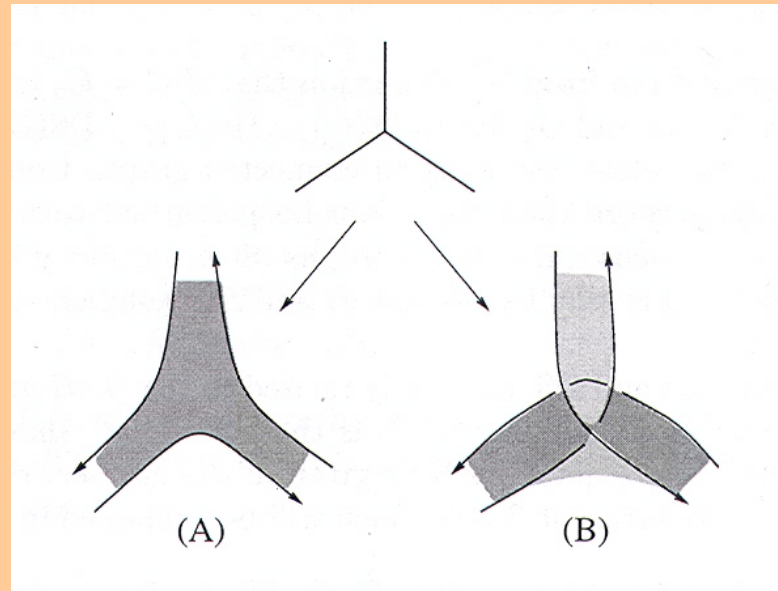
The junctions

As observed by Rudnick and Bruinsma [PRL 94, 038101, 2005], the structure of the junctions is crucial for answering this question.



Junctions (2)

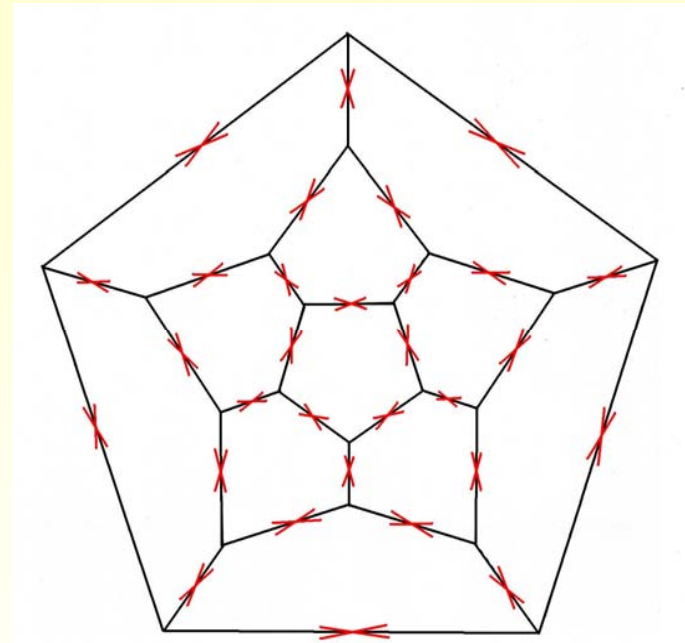
In addition, there is the alternative shown in (B) below [see N. Jonoska and M. Saito, Boundary Components of Thickened Graphs, in: DNA7, LNCS 2340]:



RNA organisation:

Based on these junction types, one obtains the energetically best RNA organisation as follows:

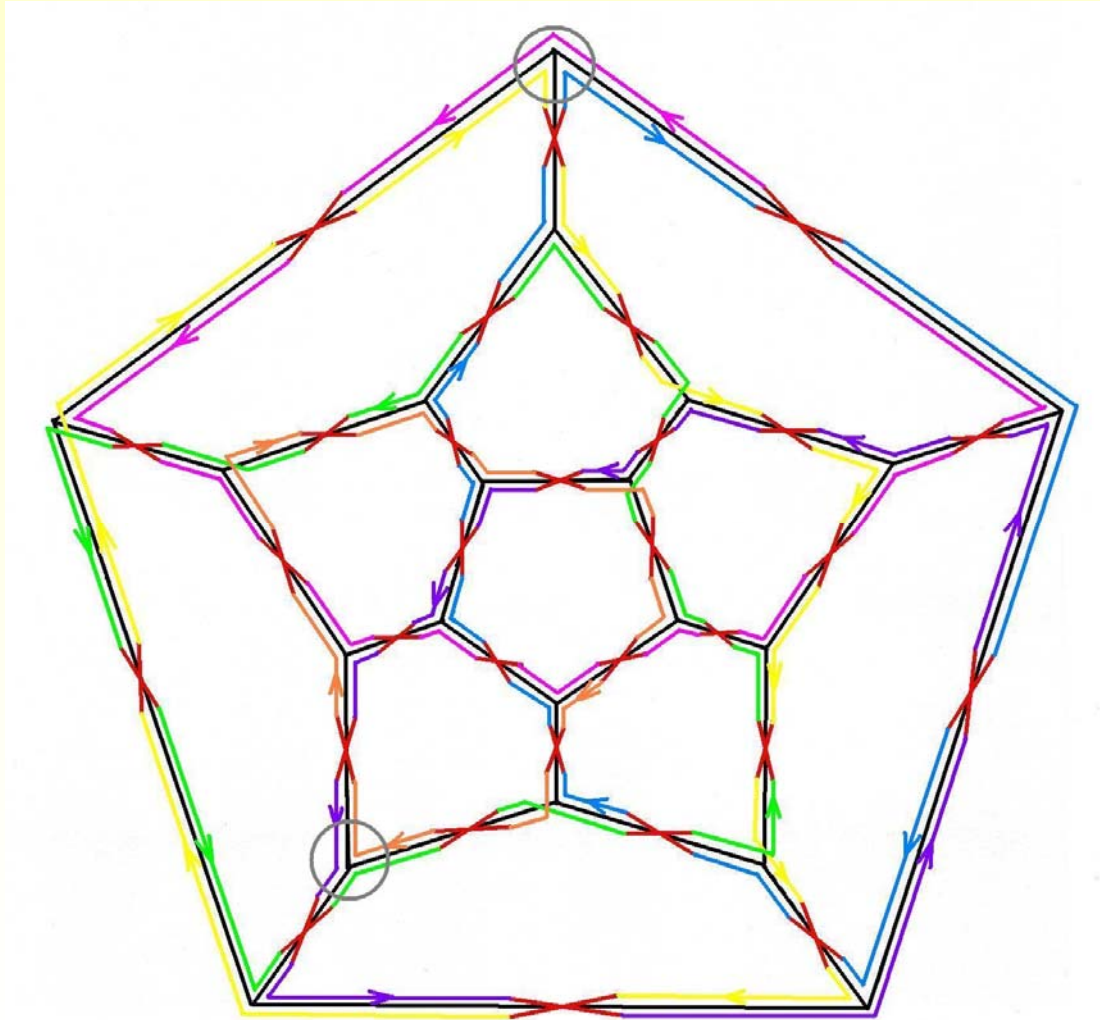
There are 2.5 helical turns of RNA per edge. This can be represented mathematically via cross-overs on the edges (shown schematically in red):



Task: Connect the cross-overs such that only one loop is obtained that meets every edge precisely once in opposite direction.

Work in collaboration with Natasha Jonoska.

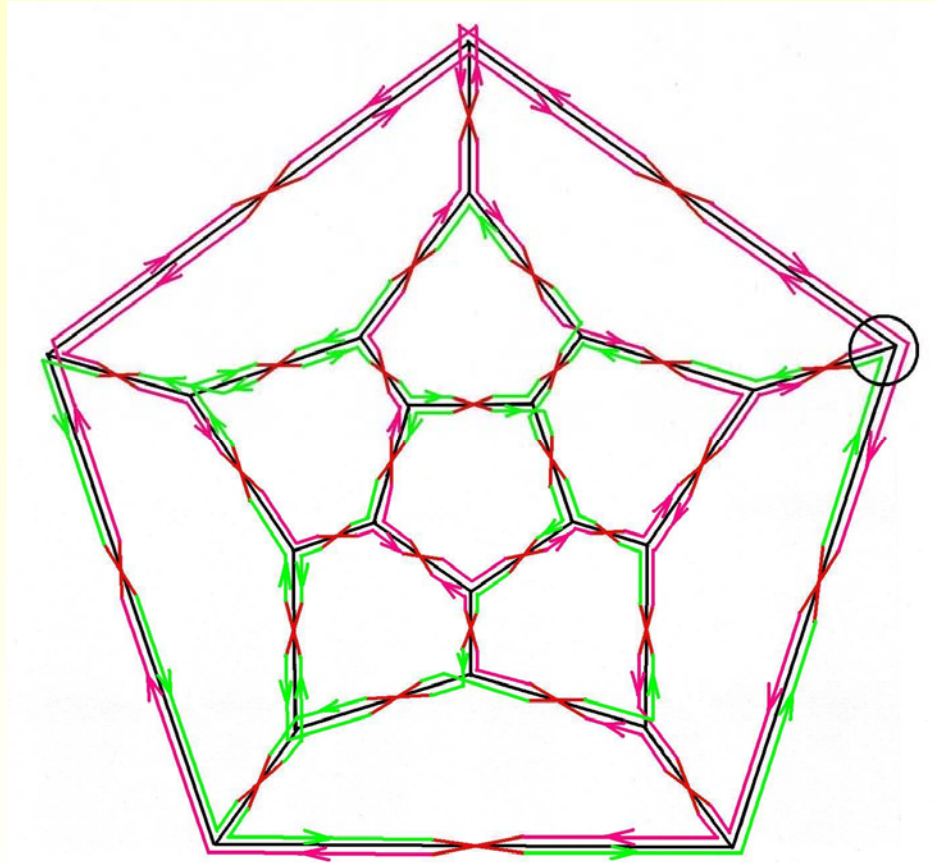
Step 1: six loops



Junctions to be replaced by other vertex types are indicated by circles.

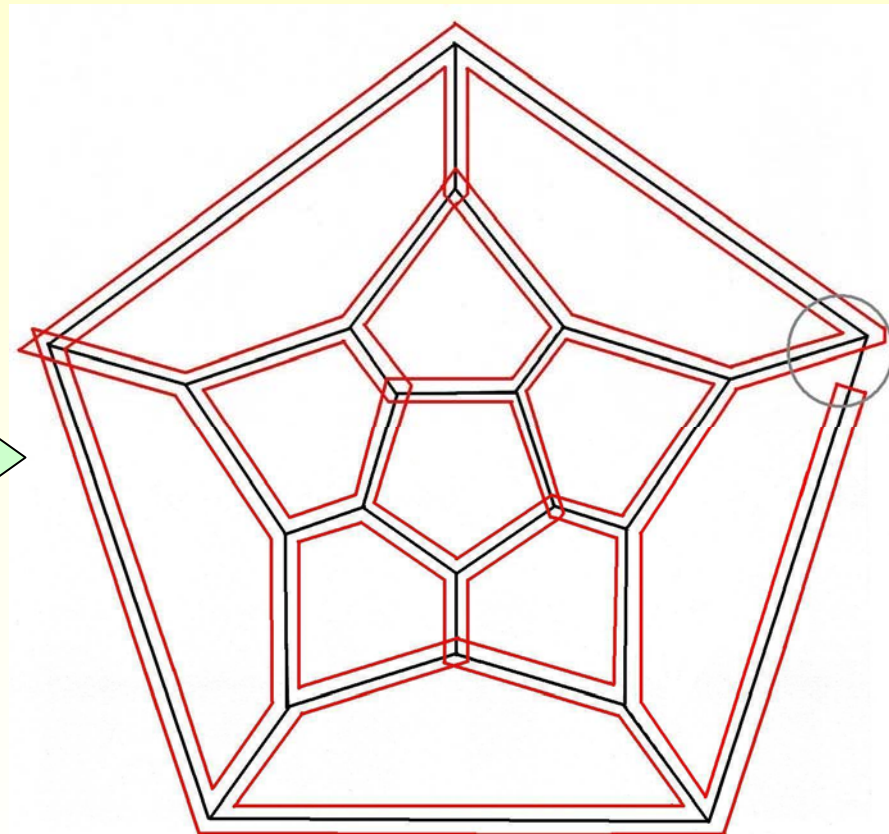
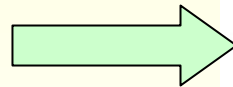
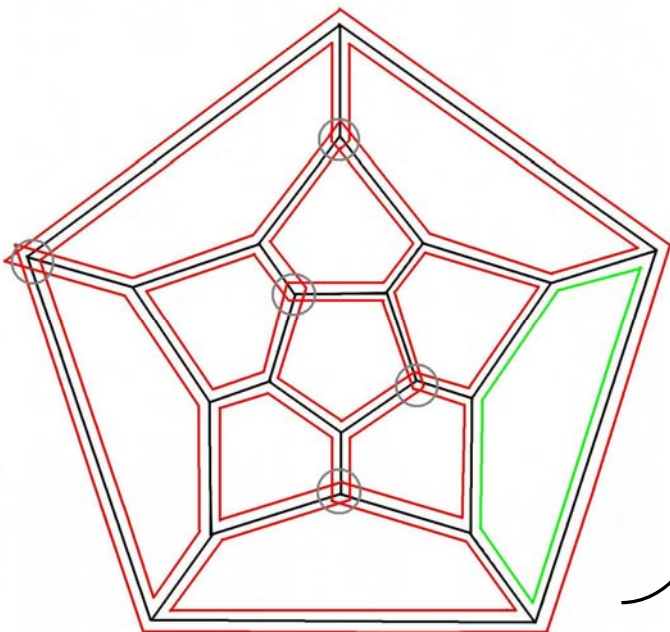
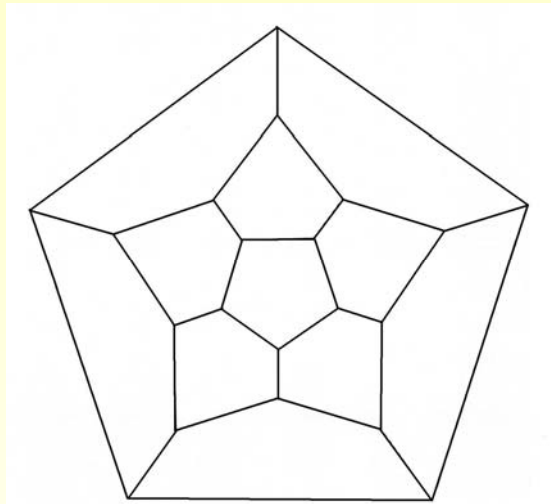
Step 2: two loops

After replacement:



Circle indicating the junction to be replaced by a hairpin/bubble.
This identifies the green and the pink loop.

Without extra helical turns:



Summary and outlook

We have developed new mathematical tools for the description of virus structure and assembly.

They are based on group theory and tiling theory, and have been used in order to model

- the structure of viral capsids in terms tilings that encode the locations of the capsid proteins and the bonds between them
- the structure of tubular malformations
- the assembly process
- the structure of the viral genome within the capsids

Outlook

Our group is currently working on assembly models that include:

A 3d representation of proteins via encasing forms.

- The role of RNA during assembly of RNA viruses.
- Assembly via agglomeration of intermediates.
- The dependence on experimental conditions (eg pH value)
- Simultaneous assembly of different species.

Applications include

- the use of capsids for drug delivery.
- interference with capsid assembly for anti-viral drug design.

Financial support by an EPSRC Advanced Research Fellowship and EPSRC grant GR/T26979/01 are gratefully acknowledged.