Dechant, Pierre-Philippe ORCID logoORCID: https://orcid.org/0000-0002-4694-4010 (2018) Recent developments in mathematical virology. In: Nonlinear Algebra in Applications, 12th - 16th November 2018, Institute for Computational and Experimental Research in Mathematics (ICERM), Providence, Rhode Island. (Unpublished)

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/4017/

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repository Policy Statement

# RaY

Research at the University of York St John

For more information please contact RaY at <a href="mailto:ray@yorksj.ac.uk">ray@yorksj.ac.uk</a>





#### Recent developments in mathematical virology

Pierre-Philippe Dechant

Pro Vice Chancellor's Office, York St John University York Cross-disciplinary Centre for Systems Analysis, University of York

> Non-linear Algebra in Applications, ICERM November 15, 2018



#### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

#### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

#### What is a Virus?

- Transported piece of genetic information that e.g. can run a programme in a host cell
- Genome: RNA or DNA
- Fragile needs to be protected by a protein shell: capsid
- Gene → mRNA → protein (transcription and translation)
- Each protein = amino acid chain folds into a 3D shape: one geometric building block

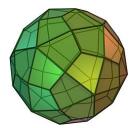








#### Watson and Crick: The Icosahedron







- Crick&Watson: Genetic economy → symmetry → icosahedral is largest
- Rotational icosahedral group is  $I = A_5$  of order 60
- Full icosahedral group is the Coxeter group H<sub>3</sub> of order 120 (including reflections/inversion); generated by the root system icosidodecahedron

# Many viruses are icosahedral



## Assembling an Icosahedron







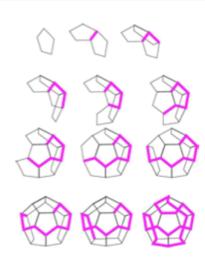
- Assemble from 20 identical triangular building blocks
- The order of addition gives a Hamiltonian path on the dual dodecahedron

#### Icosahedral symmetry

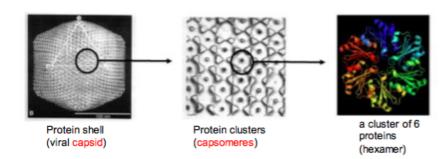
Tiling theory
Extended structures
Giant viruses

# Assembling a dodecahedron





# More than just icosahedral symmetry?

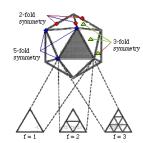


## Caspar and Klug: Triangulations

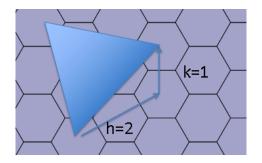
- Mathematical upper limit of 60 for equivalent subunits, but biologically want to do better!
- Gene → can already make a triangle → might as well make many!
- Caspar-Klug ideas of quasi-equivalence and triangulations







# Viruses: Caspar-Klug triangulations $T = h^2 + hk + k^2$



integer steps h and k in hexagonal directions give allowed triangulation numbers  $T = h^2 + hk + k^2$ 

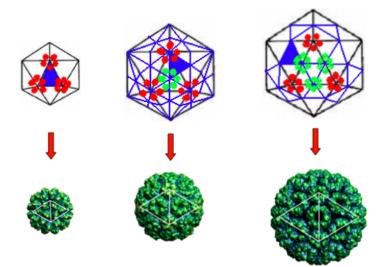
T orbits so  $60\,T$  proteins, 60 of which form 12 pentamers, and 60(T-1) form 10(T-1) hexamers



#### Icosahedral symmetry

Tiling theory
Extended structures
Giant viruses

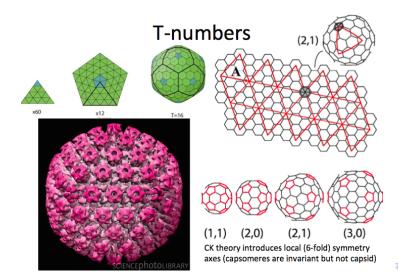
# Viruses: Caspar-Klug triangulations $T = h^2 + hk + h^2$





# Icosahedral symmetry Tiling theory Extended structures Giant viruses

## Viruses: Caspar-Klug triangulations





## The status quo for 40 years

 New insights in the 2000s from Reidun Twarock who essentially founded the field and other mathematical physicists (e.g. Anne Taormina etc)



#### **Fullerenes**

- other icosahedral objects in nature: football-shaped fullerenes
- Different shells with icosahedral symmetry: e.g.  $C_{60}$ ,  $C_{240}$ ,  $C_{540}$
- Follow Caspar-Klug-like layouts (e.g.  $T=h^2$  and  $T=3h^2$  families)

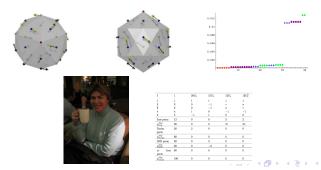






# Vibrations of capsids and fullerenes

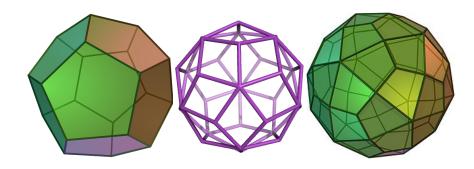
- Normal modes/vibrations of icosahedral capsids given by representation theory of the icosahedral group
- E.g.  $\Gamma_{\text{Loos}}^{\text{disp}} = \Gamma^1 + 3\Gamma^3 + \Gamma^{3'} + 2\Gamma^4 + 3\Gamma^5$
- Pioneered by Anne Taormina, Kasper Peeters and Francois Englert



#### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

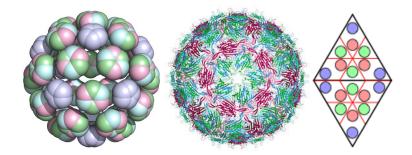
# More general icosahedral tilings



Other tile shapes can also give icosahedral tilings: pentagons (dodecahedron), rhombuses (rhombic triacontahedron), kites (deltoidal hexecontahedron)



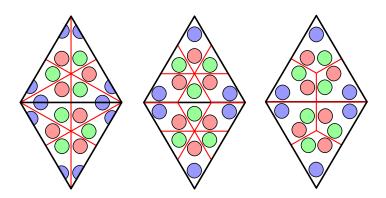
## triangulations vs other quasi-equivalent tilings



Two viral surface layouts: a T=4 triangulation (e.g. HBV) and a rhombus tiling (MS2) for a pseudo T=3 triangulation



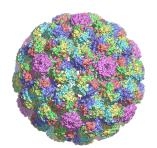
# Other quasi-equivalent tilings

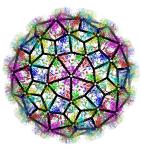


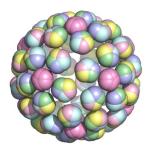
Three T = 3 capsids: Polio, MS2 and Pariacoto



## A puzzle: non-quasiequivalent tilings – Penrose







More general icosahedral tilings: Cryo-EM reconstruction of HPV, a kite-rhombus tiling and a pseudo T=7 triangulation.

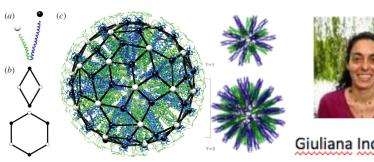


#### Architecture

- Triangulations: Buckminster Fuller geodesic domes
- kite-rhombus tiling: the new Amazon HQ



# Self-assembling protein nanoparticles





Giuliana Indelicato

Quantized e.g. in mass spec - predict units by symmetry. Particles eg. for vaccine design

# More general symmetry still?

- Improves the limit to 60 T, but only in terms of surface structures (12 pentagons and rest hexagons).
- Making the symmetry non-compact might allow more general symmetry, simultaneously constraining different 'radial levels'
- Non-compact generator is a translation motivates looking into affine extensions of icosahedral symmetry
- There is an inherent length scale in the problem given by size of nucleic acid/protein molecules

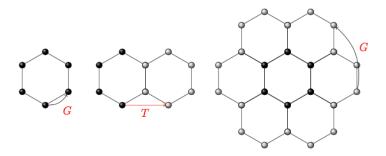


#### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

## Affine extensions - $A_2$

Unit translation along a vertex of a unit hexagon



A random translation would give 6 secondary hexagons, i.e. 36 points. Here we have degeneracies due to 'coinciding points', and building up the hexagonal lattice.

## Affine extensions of non-crystallographic groups?

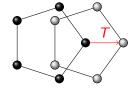
Unit translation along a vertex of a unit pentagon



## Affine extensions of non-crystallographic groups?

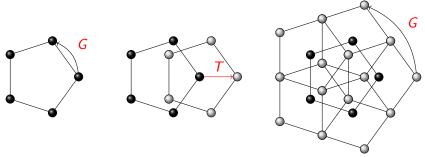
Unit translation along a vertex of a unit pentagon





## Affine extensions of non-crystallographic groups?

Unit translation along a vertex of a unit pentagon

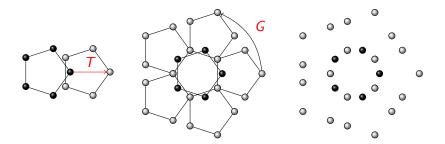


A random translation would give 5 secondary pentagons, i.e. 25 points. Here we have degeneracies due to 'coinciding points'.



## Affine extensions of non-crystallographic root systems?

Translation of length  $\tau = \frac{1}{2}(1+\sqrt{5}) \approx 1.618$  (golden ratio)

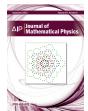


Cartoon version of a virus or carbon onion. Would there be an evolutionary benefit to have more than just compact symmetry?

The problem has an intrinsic length scale.

# Affine extensions of non-crystallographic Coxeter groups

- 2D and 3D point arrays for applications to viruses, fullerenes, quasicrystals, proteins etc
- Two complementary ways to construct these



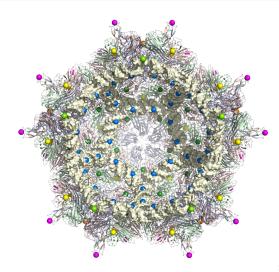




#### Other ideas

- Project symmetry orbits in 6D to get finite extended icosahedral point arrays (Emilio Zappa)
- Use projection to find 3D tiles to model viruses (David Salthouse)
- Use projection to model transitions between capsids via lattice transitions in 6D (Giuliana Indelicato)

# Use in Mathematical Virology



# New insight into RNA virus assembly

- There are specific interactions between RNA and coat protein (CP) given by icosahedral symmetry axes
- Essential for assembly, as only this RNA-CP interaction turns
   CP into right geometric shape for capsid formation
- Hamiltonian cycle visiting each RNA-CP contact once dictated by symmetry
- Even the RNA has an icosahedrally ordered component









#### Extension to fullerenes: carbon onions

- Extend idea of affine symmetry to other icosahedral objects in nature: football-shaped fullerenes (with Jess Wardman)
- Recover different shells with icosahedral symmetry from affine approach: carbon onions  $(C_{60} C_{240} C_{540})$









### Extension to fullerenes: carbon onions

- Extend idea of affine symmetry to other icosahedral objects in nature: football-shaped fullerenes
- Recover different shells with icosahedral symmetry from affine approach: carbon onions  $(C_{80} C_{180} C_{320})$

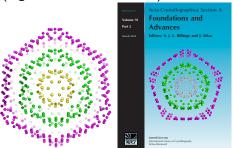


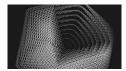




## Viruses and fullerenes – symmetry as a common thread?

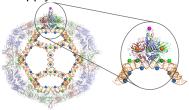
 Get nested arrangements like Russian dolls: carbon onions (e.g. Nature 510, 250253)





## Two examples

- Non-compact symmetry that relates different structural features in the same polyhedral object when there is an additional length scale
- Novel symmetry principle in Nature, shown that it seems to apply to at least fullerenes and viruses





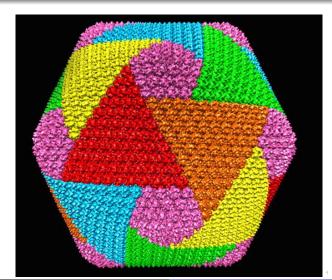




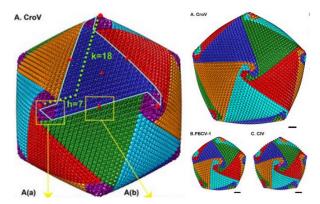
#### Overview

- 1 Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

### Giant viruses



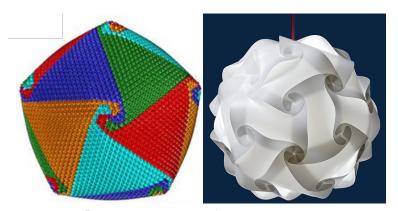
## A common approach – little hooks



Pentasymmetrons and trisymmetrons



## A common approach – little hooks



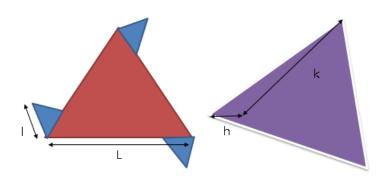
Pentasymmetrons and trisymmetrons



# A family of solutions: h = 7 – and some gaps

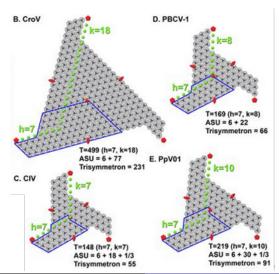
- Chilo iridescent virus: T = 147, h = 7 and k = 7
- Paramecium bursaria Chlorella virus 1: T = 169, h = 7 and k = 8
- Phaeocystis pouchetti virus: T = 219, h = 7 and k = 10
- Faustovirus: T = 277, h = 7 and k = 12
- Pacman virus: T = 309, h = 7 and k = 13
- Cafeteria roenbergensis: T = 499, h = 7 and k = 18

#### Count areas?

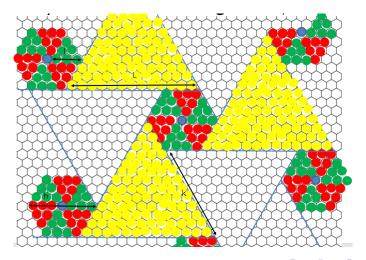


When is there an equivalent description in terms of Caspar-Klug *h* and *k*?

# Count hexagons



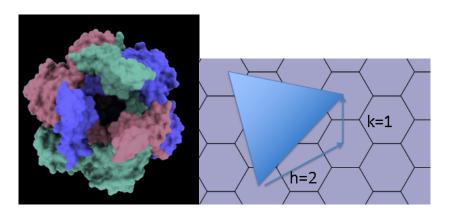
## Trisymmetrons and Pentasymmetrons – how to count



## Trisymmetrons and Pentasymmetrons – how to count

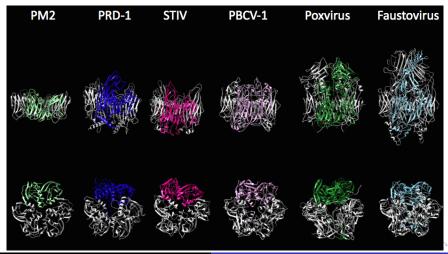
- Each has I(I+1)/2 or L(L+1)/2 hexamers
- So in total 20(3/(l+1)/2 + L(L+1)/2)
- Caspar-Klug theory predicts  $10(T-1) = 10(h^2 + hk + k^2 1)$
- So for which h, k, l, L can there be equality?
- Turns out for h = 2l + 1 and k = L l this holds identically
- So can have odd h and any k

# Major capsid protein



T is an area, so  $\sqrt{T}$  gives size of triangle and thus also particle diameter

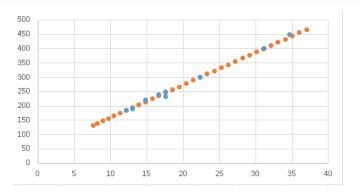
## Major capsid protein – evolutionary conservation



Pierre-Philippe Dechant

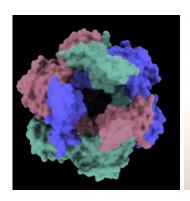
Recent developments in mathematical virology

## Scaling



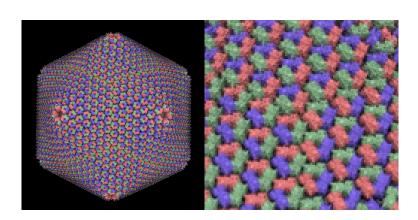
Missing points allowed geometrically but less stable? Or just not yet discovered? Predict Tetraselmis virus 1 TetV-1 of  $257nm \pm 9nm$  is exactly T=343. Predict holes in family exist and sizes given by this scaling

## Major capsid protein – trimer / pseudohexamer

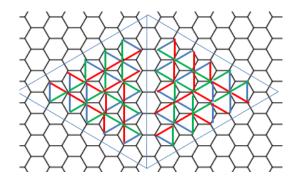




## Decorations – threefold and giants



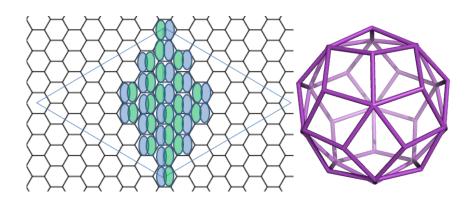
### Decorations – threefold and giants



Use hexagonal tiling unit with decorations that partially break the symmetry: 6, 3, 2, 1

If respect the 3-fold axis, then still have partial lattice symmetry! Defects/domain walls at the other symmetry axes

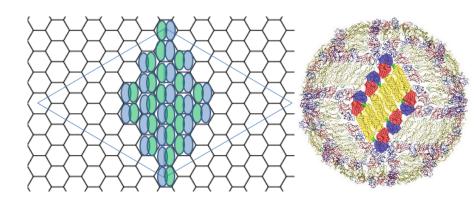
### Decorations – twofold case



Symmetric with respect to 2-fold axis – quasi-lattice symmetry Looks like a rhombic triacontahedron all over again



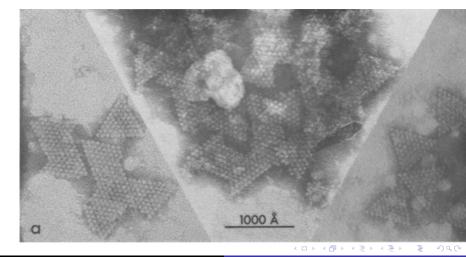
### Decorations – twofold and Zika



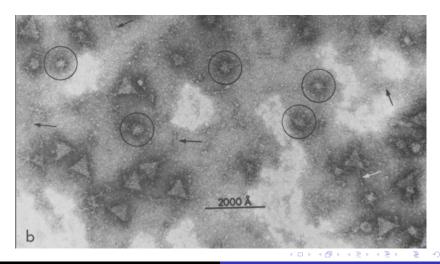
This is exactly what Zika looks like



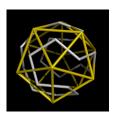
## Trisymmetrons and Pentasymmetrons



## Major capsid protein - trimer, pseudohexamer



### Build from prearranged blocks? Back to Hamiltonian paths







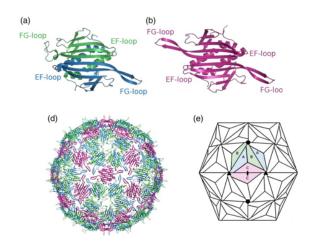
- Are the trisymmetrons and pentasymmetrons preformed? (or is that just what virions fall apart into?)
- If trisymmetrons are assembled then we're back to a Hamiltonian path for the icosahedron
- If pentasymmetrons then get a slightly new polyhedron



#### Overview

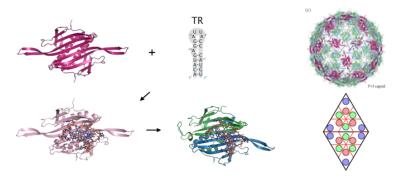
- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

## MS2 tiling and dimeric building blocks: A/B and C/C



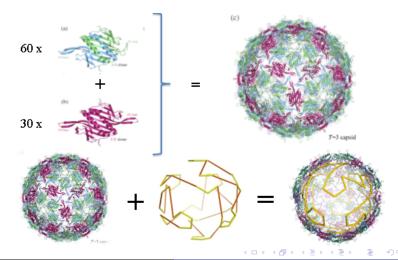
### Need to bind RNA in 60 places

The TR sequence is known to initiate assembly by associating with the maturation protein. It forms a <u>stemloop</u> and it has been shown that the <u>stemloop</u> changes the conformation of the <u>symmetric</u> C/C dimer to the <u>asymmetric</u> A/B dimer (allosteric switch).



Peter Stockley (Leeds), Neil Ranson (Leeds), Eric Dykeman (York)

## MS2 Hamiltonian path



## New insight into RNA virus assembly

- More realistic examples for MS has 60 vertices with 41,000 paths
- The RNA is actually circularised by Maturation Protein: only 66 cycles
- With thermodynamical assembly kinetics and 5-fold averaging experiments uniquely idenfied an evolutionarily conserved cycle
- Patents for new antiviral strategies and virus-like nanoparticles
   e.g. for drug delivery (Twarock group)



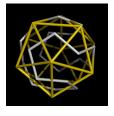






### Hamiltonian cycles on icosahedral solids







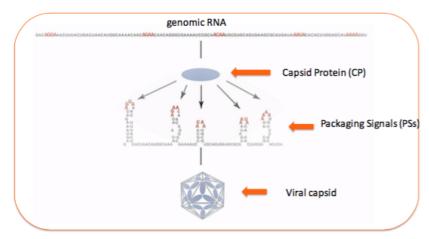
- So interaction contacts are given by the symmetry
- Orbits of the interaction points have to be visited by the RNA exactly once
- Even the RNA has an icosahedrally ordered component
- Hamiltonian cycles for dodecahedron, icosahedron and rhombic triacontahedron



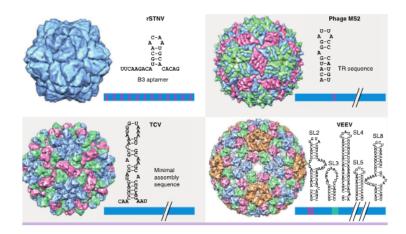
### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

# Multiple dispersed Packaging Signals paradigm



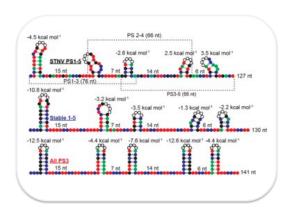
## Common Mechanism across groups of viruses



#### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

## Engineering Packaging Signals to make VLPs

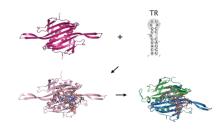


Improved sequences assemble twice as efficiently (Nikesh Patel). Potential applications to vaccines or drug delivery.

### Overview

- 1 Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

### Understanding assembly allows one to interfere



- target RNA
- target CP
- introduces competitors
- this might drive evolution due to exerting selection pressures



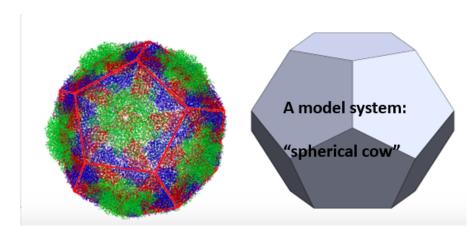
### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- Oisease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

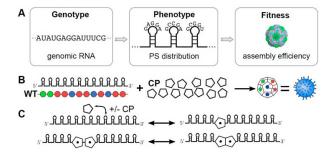
### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- Oisease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

# Dodecahedral cow (Rich Bingham)



## Viral evolution and quasispecies

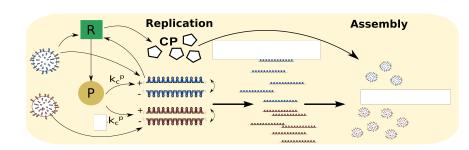


A phenomenological genome space of 12 packaging signals with 3 binding affinity bands (weak, medium, strong). Can compute the whole space explicitly in terms of assembly efficiency.

#### **Simulations**

- Stochastic simulations rather than ODE models because of discrete nature and low numbers
- Gillespie algorithm that selects a random reaction to occur (Eric Dykeman)
- Couple an intracellular model (replication) with an infection model (immune system)

## Intracellular model: replication



### Intracellular reactions

$$p_V^+ + R \stackrel{k_r^{on}}{\longleftrightarrow} (p_V^+, R)$$
 (Ribosome positive strand virus binding/unbinding)

$$(p_V^+,R) \xrightarrow{k_v^c} p_V^+ + R + P$$
 (Genome translation - makes P and abundant CP)

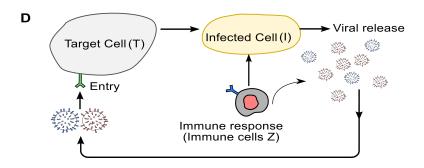
$$p_{V/S}^{\pm} + P \underset{\stackrel{k}{\leftarrow} off}{\overset{k_p^{\pm on}}{\leftarrow}} (p_{V/S}^{\pm}, P)$$
 (Polymerase positive/negative strand virus binding/unbinding)

$$(p_{V/S}^{\pm}, P) \xrightarrow{k_p^c} p_{V/S}^{\pm} + p_{V/S}^{\mp} + P$$
 (complementary strand production)

### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- Oisease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

## Infection model: immune system



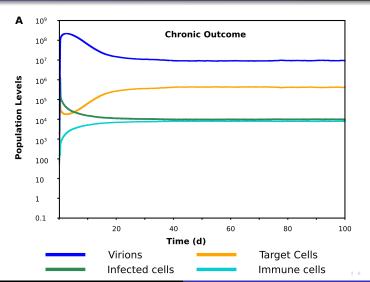
#### Immune cell reactions

$$T \xrightarrow{\lambda} 2T$$
 (Target cell birth) 
$$T \xrightarrow{d_T} 0 \text{ (Target cell death)}$$
 
$$T + pV + qS \xrightarrow{\beta} I \text{ (Infection of target cell)}$$
 
$$I \xrightarrow{a} rV + sS \text{ (Infected cell death/lysis, if } p > 0)$$
 
$$I + Z \xrightarrow{\pi} Z \text{ (Infected cell removal by immune system)}$$
 
$$V + Z \xrightarrow{u} Z \text{ (Virion removal by immune system)}$$
 
$$I + Z \xrightarrow{c} I + 2Z \text{ (Immune cell birth)}$$
 
$$Z \xrightarrow{b} 0 \text{ (Immune cell death)}$$

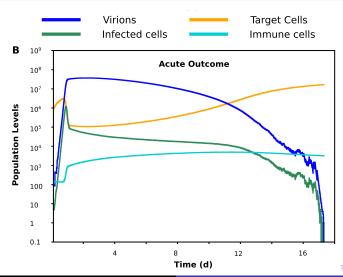
#### Overview

- Virus structure and dynamics
  - Icosahedral symmetry
  - Tiling theory
  - Extended structures
  - Giant viruses
- 2 Virus assembly
  - MS2 and Packaging Signals
  - Virus-like Particles
  - Anti-virals
- 3 Disease dynamics
  - Intracellular Replication
  - Immunological Infection Dynamics
  - Modelling

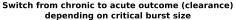
### Chronic infections

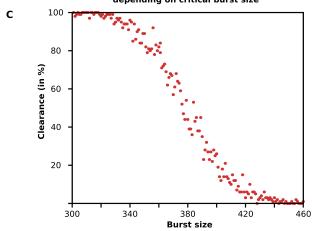


### Acute infections

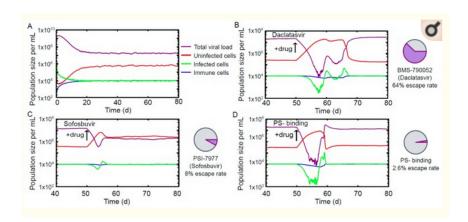


#### Transition between chronic and acute





## **Evolutionarily stable drugs**



# Summary of mathematical virology

- a truly diverse interdisciplinary endeavour: group theory, tiling theory, dynamical systems, graph theory, computational modelling, biophysics, bioinformatics, biochemistry, cell biology, structural biology, immunology
- curiosity-driven and impactful: from understanding a structural puzzle to a new generation of anti-virals – by accident

## Other algebraic interests

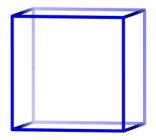
- exceptional root systems/geometries
- (reflexive) polytopes
- Clifford algebras
- ADE correspondences

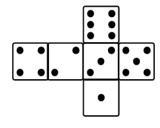
Intracellular – Replication Immunological – Infection Dynamics Modelling

Thank you! LGBTQ+ lunch at 1pm.

# Symmetry averaging

- As an example use the symmetry of a cube
- (axes of 4-fold, 3-fold and 2-fold symmetry)
- Dice have approximately cubic symmetry, but have slightly asymmetric features (the faces)



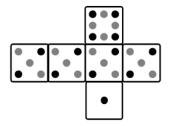


# Symmetry averaging

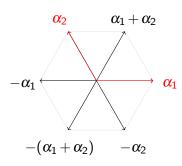
- As an example use the symmetry of a cube
- Approximately cubic symmetry, but have slightly asymmetric features (the faces)
- Experiment averages over all orientations
- Washes out all asymmetric features

# Symmetry averaging

- Assume we have a distinguished face, e.g. 1 is always at the bottom
- The 6 can be in two configurations
- Average over the other 4 faces



## Root systems



Root system  $\Phi$ : set of vectors  $\alpha$  in a vector space with an inner product such that

1. 
$$\Phi \cap \mathbb{R}\alpha = \{-\alpha, \alpha\} \ \forall \ \alpha \in \Phi$$

$$2. s_{\alpha} \Phi = \Phi \ \forall \ \alpha \in \Phi$$

Simple roots: express every element of  $\Phi$  via a  $\mathbb{Z}$ -linear combination.

reflection/Coxeter groups 
$$s_{\alpha}: v \to s_{\alpha}(v) = v - 2 \frac{(v|\alpha)}{(\alpha|\alpha)} \alpha$$

#### Affine extensions

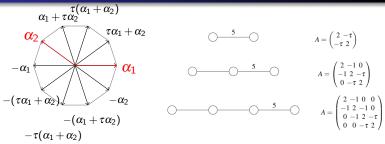
An affine Coxeter group is the extension of a Coxeter group by an affine reflection in a hyperplane not containing the origin  $s_{\alpha_0}^{aff}$  whose geometric action is given by

$$s_{\alpha_0}^{aff} v = \alpha_0 + v - \frac{2(\alpha_0|v)}{(\alpha_0|\alpha_0)} \alpha_0$$

Non-distance preserving: includes the translation generator

$$Tv = v + \alpha_0 = s_{\alpha_0}^{aff} s_{\alpha_0} v$$

# Non-crystallographic Coxeter groups $H_2 \subset H_3 \subset H_4$



 $H_2 \subset H_3 \subset H_4$ : 10, 120, 14,400 elements, the only Coxeter groups that generate rotational symmetries of order 5 linear combinations now in the extended integer ring

$$\boxed{\mathbb{Z}[ au] = \{a + au b | a, b \in \mathbb{Z}\}}$$
 golden ratio  $\boxed{ au = \frac{1}{2}(1 + \sqrt{5}) = 2\cos\frac{\pi}{5}}$ 

$$\boxed{x^2 = x + 1} \boxed{\tau' = \sigma = \frac{1}{2}(1 - \sqrt{5}) = 2\cos\frac{2\pi}{5}} \boxed{\tau + \sigma = 1, \tau\sigma = -1}$$