# HIDDEN SYMMETRY AND PRINCIPLES OF STRUCTURAL ORGANIZATION IN SMALL ICOSAHEDRAL 'ANOMALOUS' AND DOUBLE-SHELLED CAPSIDS

S. B. Rochal[1], A.E. Myasnikova[1], O.V. Konevtsova[1]  and  V.L. Lorman[2]

[1]Physics Faculty, Southern Federal University, Rostov-on-Don, Russia
[2]Laboratoire Charles Coulomb, UMR 5221 CNRS - Université de Montpellier,
Place E. Bataillon, F-34095 Montpellier Cedex 5, France.

# Outline

-Introduction

- Hidden symmetry in the Caspar and Klug model. Quasi-equivalence theory

- Our modification of the Caspar and Klug model.

-Density waves approach and Landau theory

-Main results

# Virus Structure

- Size
  - 17 nm – 3000 nm diameter

- Basic shape
  - Rod-like
  - "Spherical"

- Protective Shell - Capsid
  - Made of many identical subunits
  - <u>Symmetrically</u> organized
  - 50% of weight
  - Enveloped or non-enveloped

- Genomic material
  - DNA or RNA
  - Single- or double-stranded
  - No unique structure
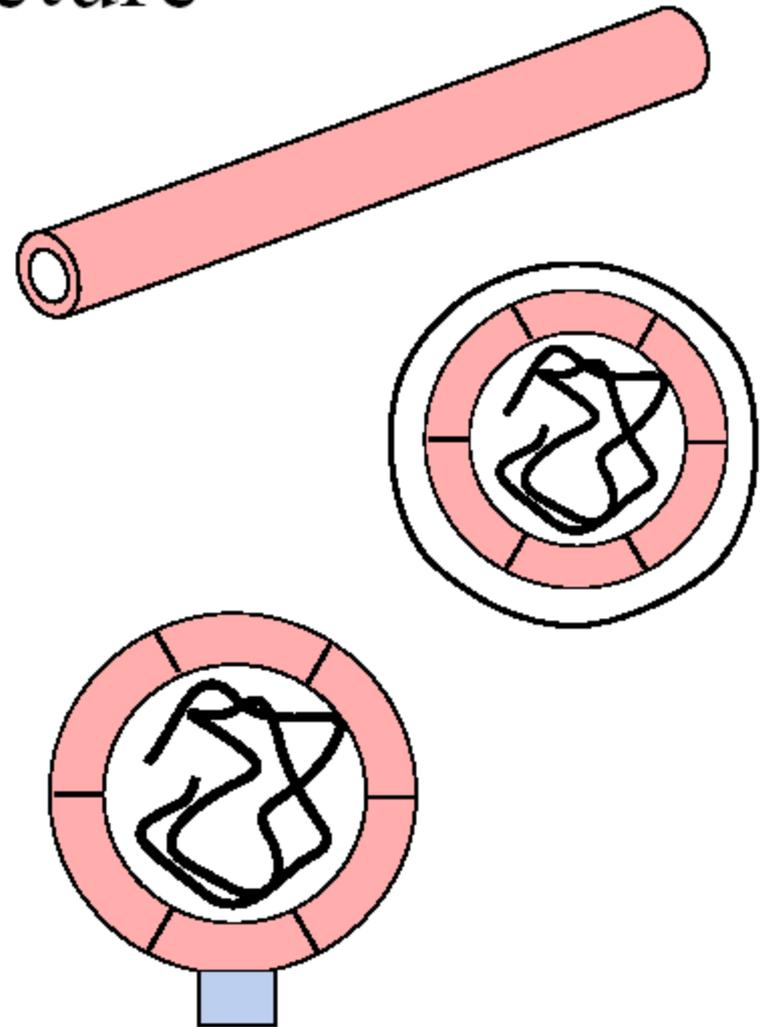
# Why it  is important to study the organization principles of viral shells

- The highly ordered viral capsid contains  a genome and therefore both mechanisms of host cell infection as well as virulence of viruses are strongly dependent on the structural organization of capsids.

- The obtained organization principles and the relation of the revealed structural peculiarities with the assembly thermodynamics can be easily generalized to the case of abiotic materials for nanotechnology.

# Why the subject we deal with is related to physics

Several steps of the capsid self-assembly demonstrate properties
   typical of ordering in passive physical systems
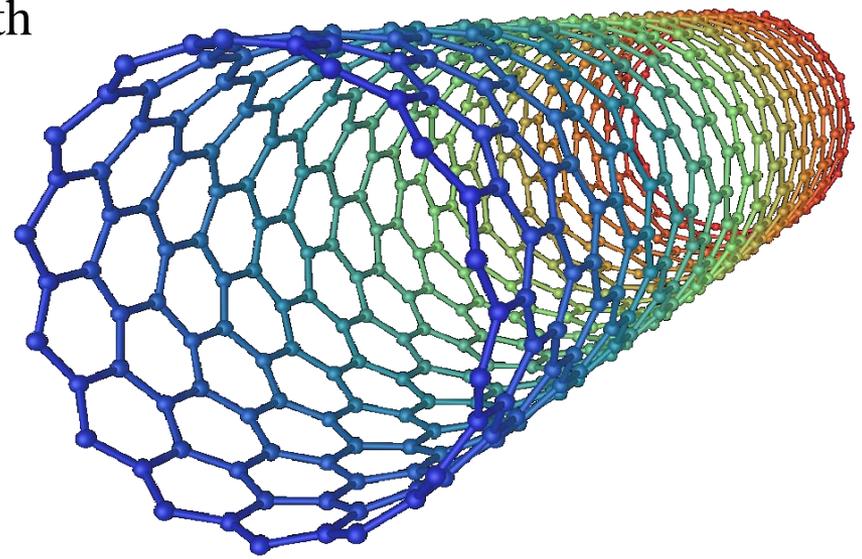
For the *capsid shell self-assembly* :
   - host cell is not necessary
   - no local energy consumption like ATP hydrolisis is needed
   - process can be reversible
   - in many cases capsid assembly does not need genome
   - for some capsids the assembly can be proceeded *in vitro* in
   purified protein solutions

 => Principles of capsid structure formation can be related to
   physics

# Physics, symmetry, and viruses

- By the middle of the last century, symmetry became the robust basis for the exploration and formulation of the fundamental principles of nonliving nature. Symmetry determines the structural organization and dictates the dynamics of relatively simple physical and chemical nanoscale systems. In living organisms, which are incommensurably more complex than the classical objects studied by physics and chemistry, the role of symmetry appears to be less significant. Nevertheless, symmetry in its different forms remains extremely important for viruses representing relatively simple systems that are intermediate between living and nonliving matter. In particular the highly ordered viral capsids have both
  conventional and hidden symmetries

Hidden symmetry can be detected only as traces of parent planar order, that covers locally the surface of nanoassembly

# Origin of the hidden symmetry in capsids

Ordinarily. viral shells self-assemble from identical proteins, which tend to form equivalent environments in the resulting assembly. However, in icosahedral capsids containing more than 60 proteins, they are enforced to occupy not only the symmetrically equivalent locations but also the quasi-equivalent ones. Due to this important fact, the symmetry of viral shells can include additional hidden components.

# Theory of Quasi-Equivalence
## D.L.D. Caspar, A. Klug, 1962
### One type of proteins icosahedral symmetry *I*

- 12 vertices

- 20 faces
  (equilateral triangles)

- 5-3-2 symmetry axes

- 60 identical* subunits
  in identical environments
  can form icosahedral shell
  * asymmetric



5-fold axis

3-fold axis

2-fold axis
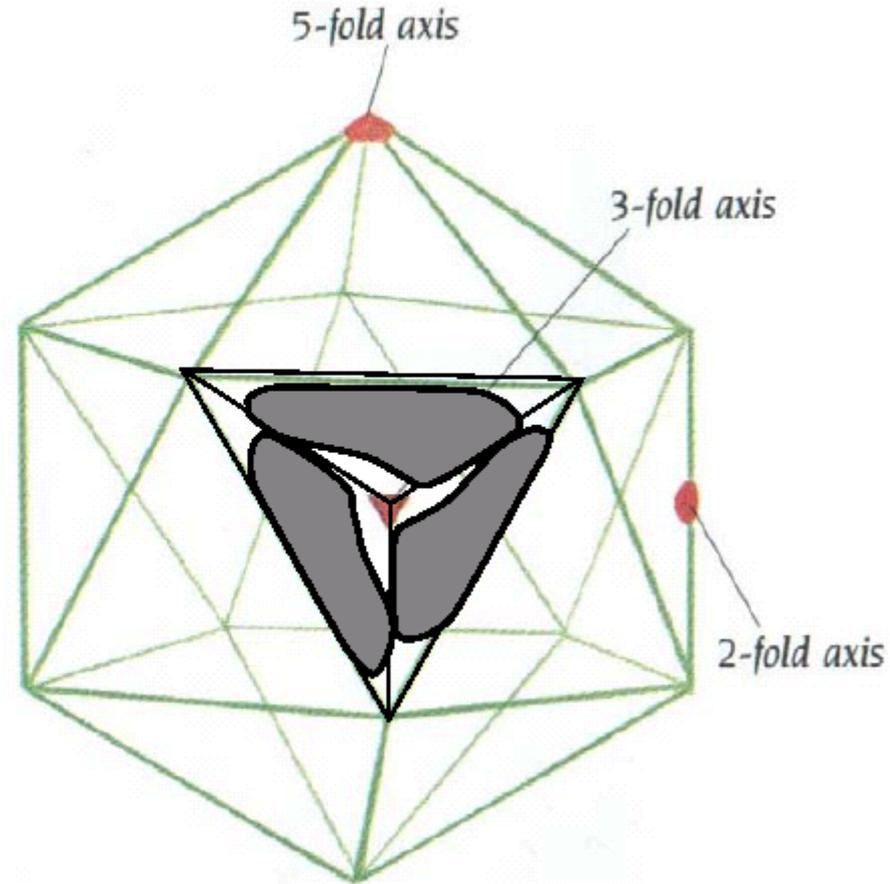
# One type of proteins in one general crystallographic position
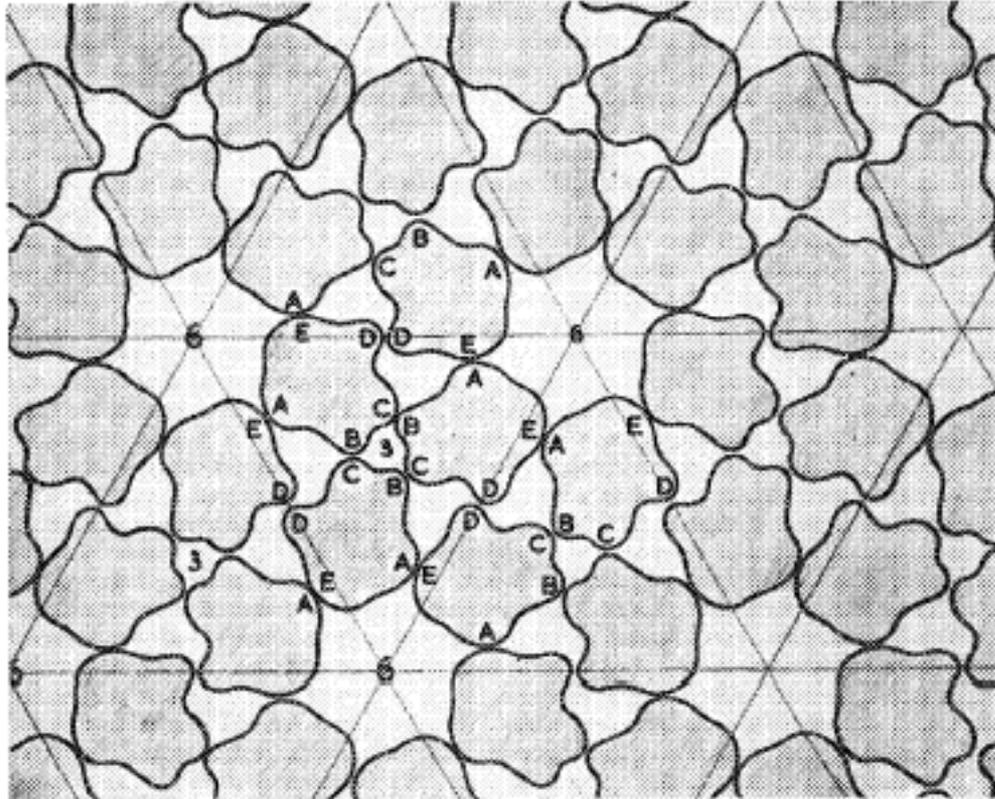


FIGURE 6. Asymmetric units arrayed in an equilateral-triangular plane net. Besides having translations, here $a$ and $a$, the lattice has 6-fold rotational axes of symmetry. Although the asymmetric units are in 6 different orientations in space, they are all exactly equivalently related.

Each unit here is equipped with five "bond" sites, A, B, C, D, and E, forming three different "bonds", namely a hexamer bond AE, a trimer bond BC, and a dimer bond DD. (Note that only two of these bonds are absolutely essential for coherence of the array.)

# Classification of capsids in the frames of CK theory

Mapping of the
Honeycomb
Hexagonal Lattice
To the Surface of
an Icosahedron
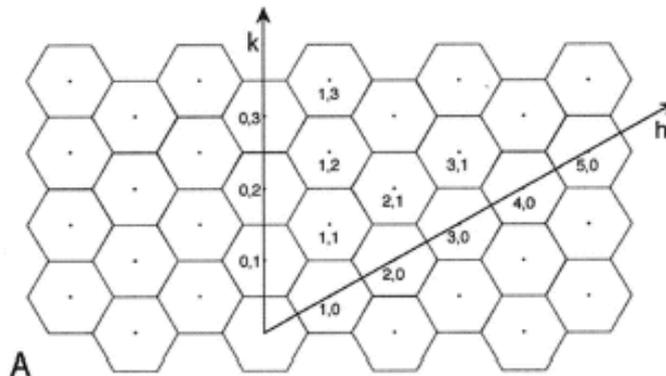


| h k | T | Example |
|---|---|---|
| 1 0 | 1 | bacteriophage φX174 |
| 1 1 | 3 | tomato bushy stunt virus |
| 2 0 | 4 | Sindbis virus |
| 1 2 | 7d | polyoma virus |
| 3 1 | 13l | reovirus |
| 1 3 | 13d | infectious bursal disease virus |
| 4 0 | 16 | herpesvirus |
| 5 0 | 25 | adenovirus |

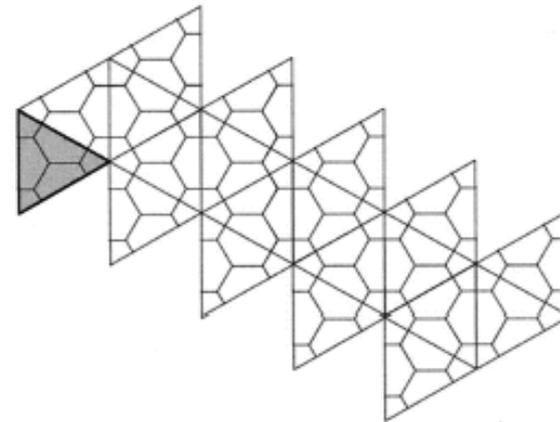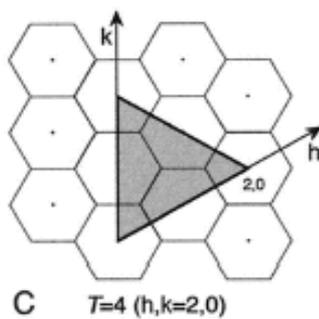Notations: d = dextro (right handed)
l = laevo (left handed)

Honeycomb
Hexagonal Lattice
« composed of hexamers »
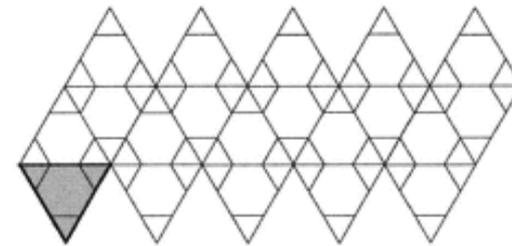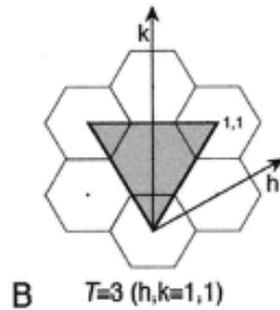
Trinagulation Number
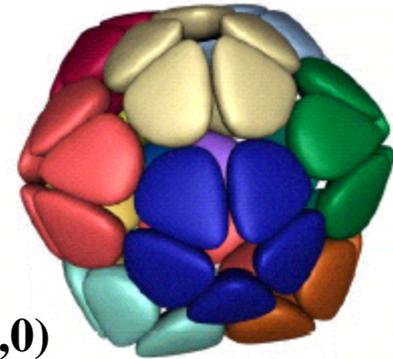$T = h^2 + hk + k^2$

Number of proteins is 60T

Selection rules for the
Triangulation Number
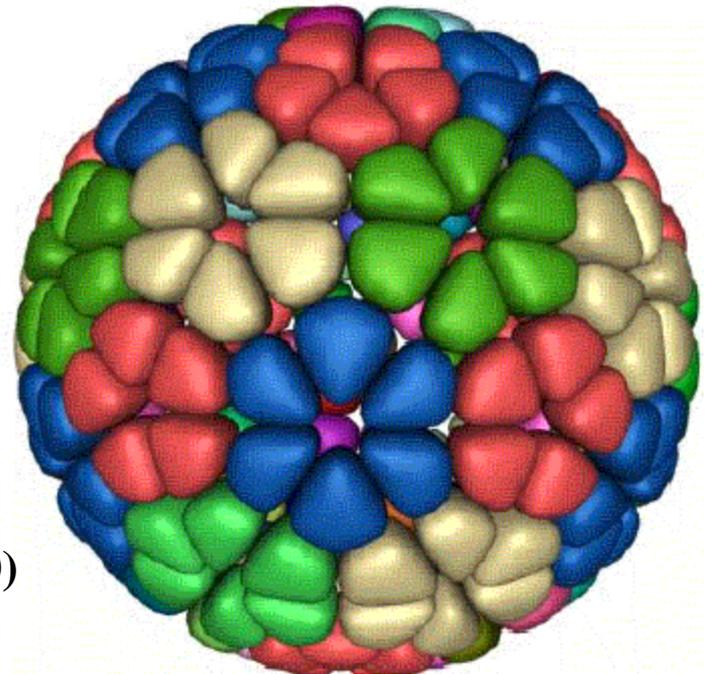T=1,3,4,7...
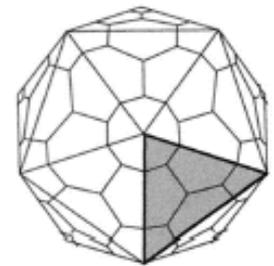
# Hidden symmetry and protein quasi-equivalence

- Subunits are in "minimally" different environments
  - Pentamers at vertices
  - Hexamers elsewhere

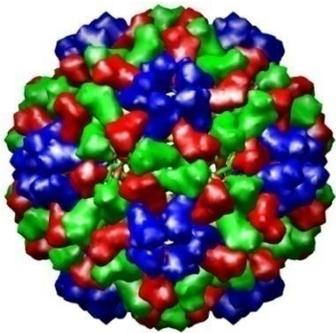- Predicts packing arrangements of larger capsids

T = 1
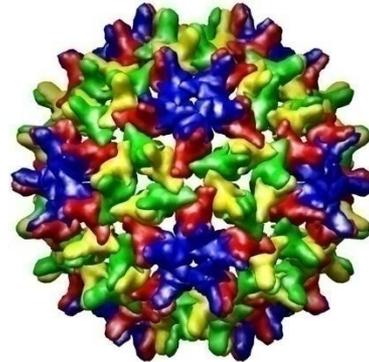(h,k) = (1,0)

T = 4
(h,k) = (2,0)

# Experimental Confirmation

The capsids of many « spherical » viruses exhibit spatial organization consistent with the quasi-equivalence principle
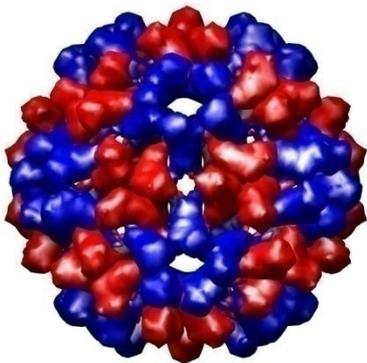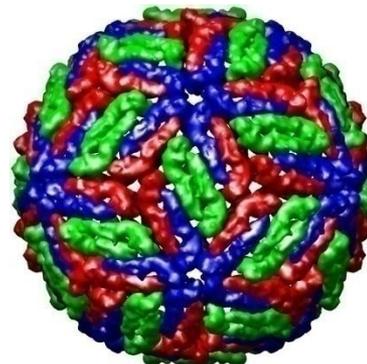


Cowpea Chlorotic Mottle Virus (CCMV)
T = 3



Hepatitis B Virus (HBV)
T = 4

**However, some don't**



L-A Virus
T = 2
forbidden by Caspar-Klug selection rules



Dengue Virus
T = 3
but without Caspar-Klug hexamers

**The main idea**: Transfer of the primitive hexagonal lattice onto the icosahedron's surface

Chiral SL with the indices <4,1>, the triangulation number T=21 has the rotational icosahedral symmetry group *I*. Among 212 nodes of the SL there are 180 nodes (full circles) which have the trivial local symmetry and are compatible with the protein asymmetry. The nodes with the non-trivial local symmetry (open circles) cannot be occupied by the asymmetric proteins. They are located at icosahedral 5-fold and 3-fold axes.
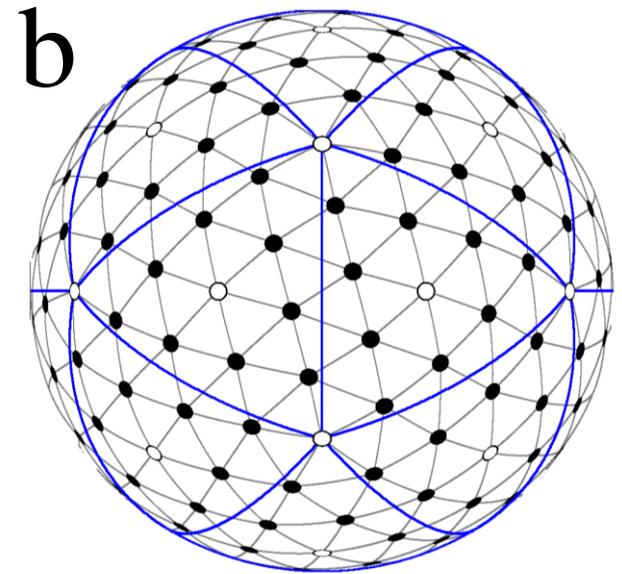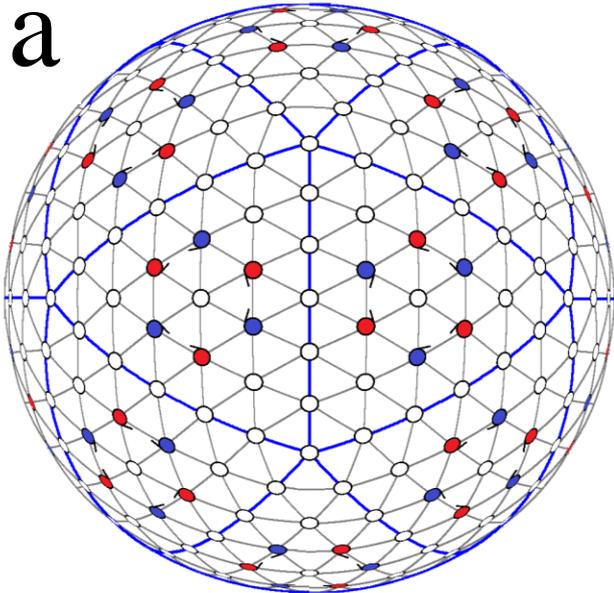
Achiral SL with the indices <6,0>, the triangulation number T=36 and the full icosahedral symmetry group $I_h$. Among its 362 nodes only 120 nodes (colored circles) belong to 2 orbits of general positions with the trivial local symmetry. In addition, these general nodes have to contain both left-handed (red circles) and right-handed (blue circles) SUs, but this constraint is incompatible with the fixed protein handedness. Smaller achiral SLs do not contain nodes with the non-trivial symmetry.

# Modified CK capsid model
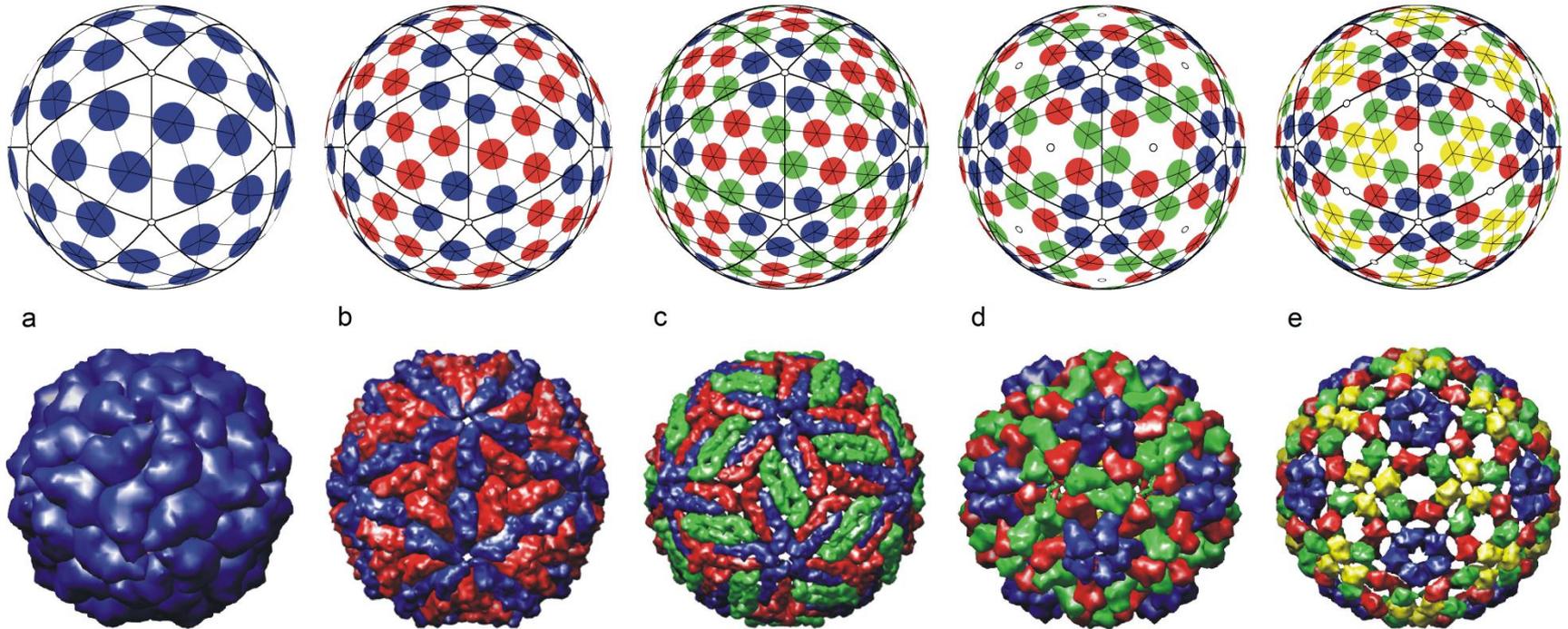


The upper line shows the first chiral spherical lattices: (a) <2,1>, (T=7, N=1); (b) <3,1>, (T=13, N=2); (c) <3,2>, (T=19, N=3); (d) <4,1>, (T=21, N=3); (e) <4,2>, (T=28, N=4). The nodes with the non-trivial local symmetry which are not suitable for occupation by the asymmetric proteins are represented by small open circles. The nodes with the trivial local symmetry occupied by the asymmetric proteins are shown by big colored circles.

The experimental capsids structures[*] are shown in the bottom line: (a) Satellite Tobacco Mosaic Virus (N = 1); (b) L–A Virus (N = 2); (c) Dengue Virus (N = 3); (d) Chlorosome Vigna Virus (N = 3); (e) Sindbis Virus (N = 4). Protein centers of mass are located in the vicinity of the occupied nodes of the spherical lattices.

# Commensurate concentric nanoshells and double-shelled capsid structures of reoviridae and cystoviridae families
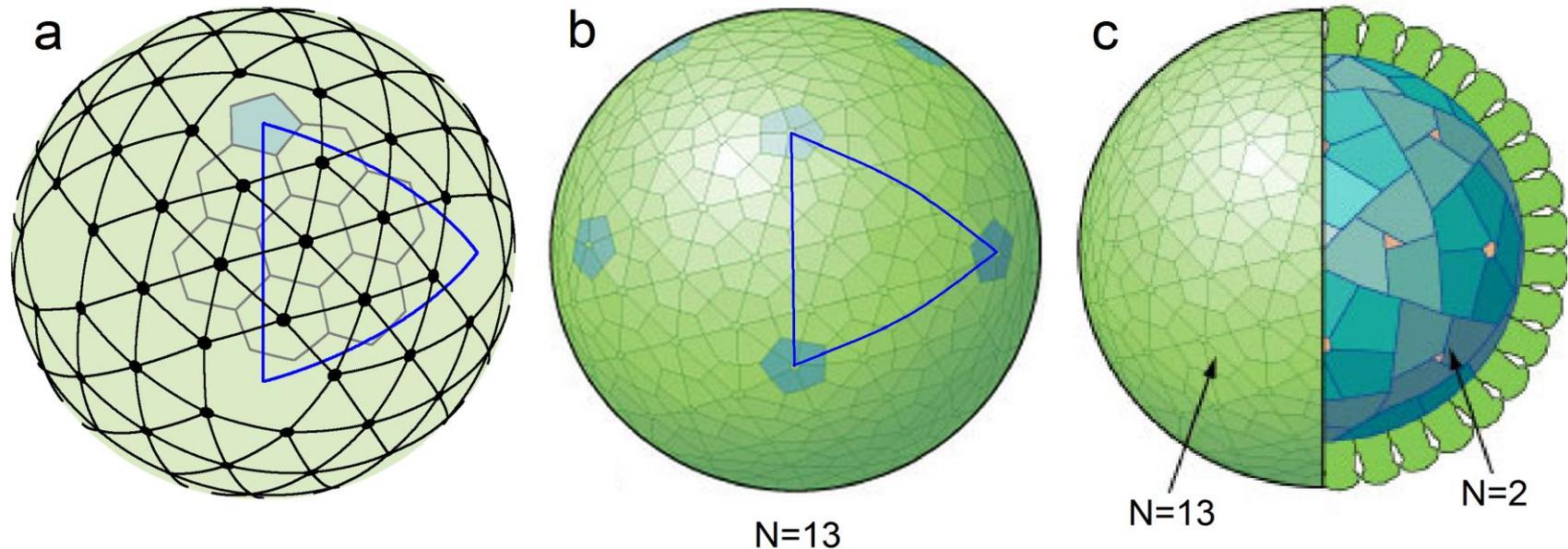


(a) Spherical tiling based on the SL with the indices <3,1>. The inner shell proteins are located in the nodes of the SL (full circles) while the outer shell proteins occupy the general positions of the underlying hexagonal lattice and form the hexamers around the SL nodes.

(b) Standard schematic representation* of the capsid with $T$=13 satisfying the original CK model requirements. It corresponds to the outer shell structure with $N$=13 in the capsids of the reoviridae and cystoviridae families.

(c) Standard schematic representation* of the inner and outer shell structures in the capsids of the reoviridae and cystoviridae families.

# The structures of icosahedral viruses with the double-shelled capsids related to the SLs
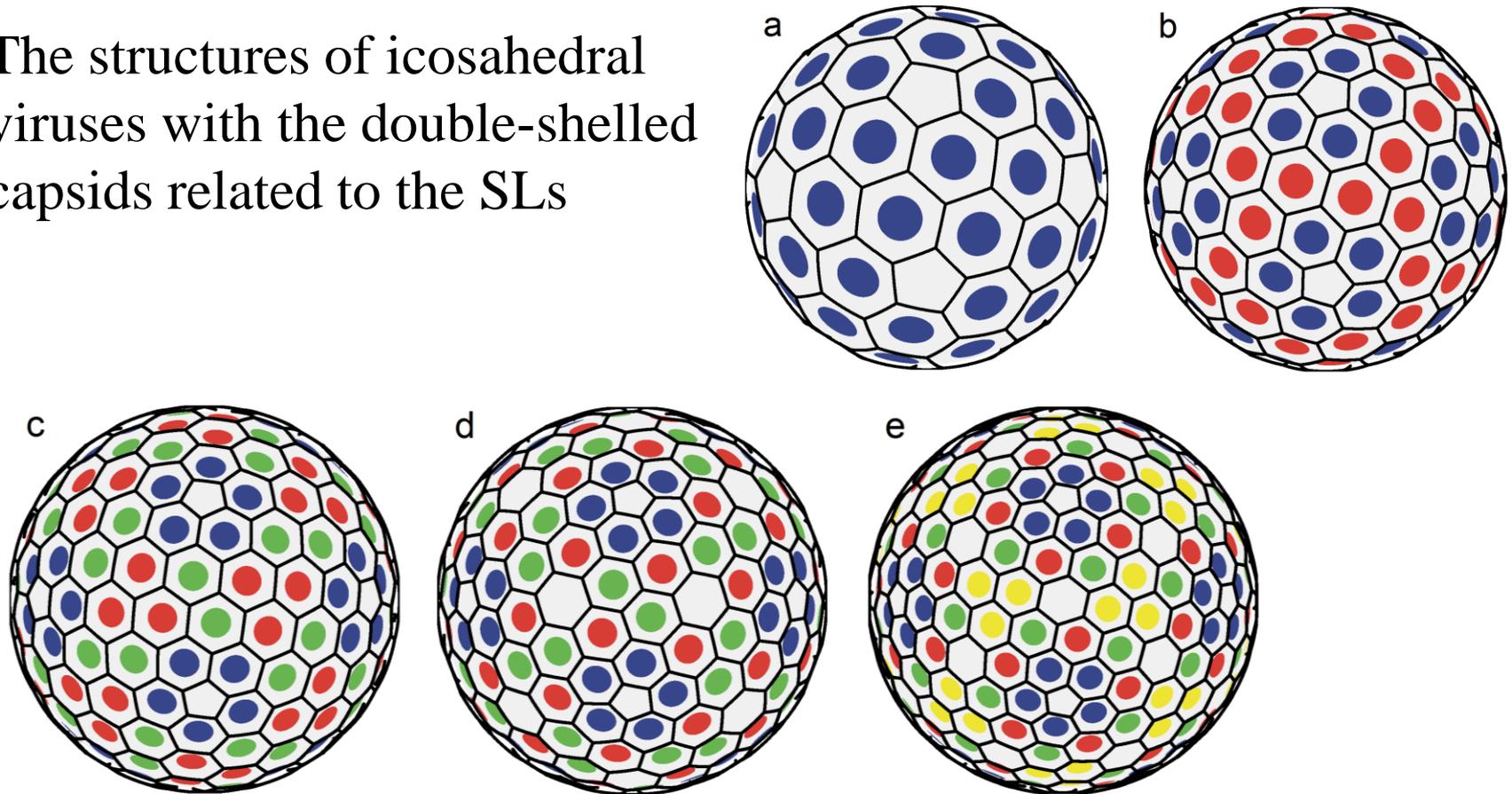


In the inner capsid shells the proteins occupy only the lattice nodes with the trivial symmetry (colored circles). Different positions are shown in different colors. The nodes with the non-trivial symmetry are excluded. These inner shell structures are similar to the experimentally observed single-shelled capsids.

The outer shells are shown as the honeycomb spherical "lattices" with the cells of the "lattice" occupied by the capsomers. Hexagonal cells are occupied by the hexamers while the pentamers are situated in the cells with the pentagonal shape. (a, c-e) Possible double-shelled structures predicted by the present approach and based on the following SLs: (a) <2,1>, (c) <3,2>; (d) <4,1>; and (e) <4,2>. (b) The structure with the indices <3,1> experimentally observed in the reoviridae and cystoviridae families.

# Experimental capsid structure of the cyanobacterial virus Syn5* and its slightly symmetrized model



a
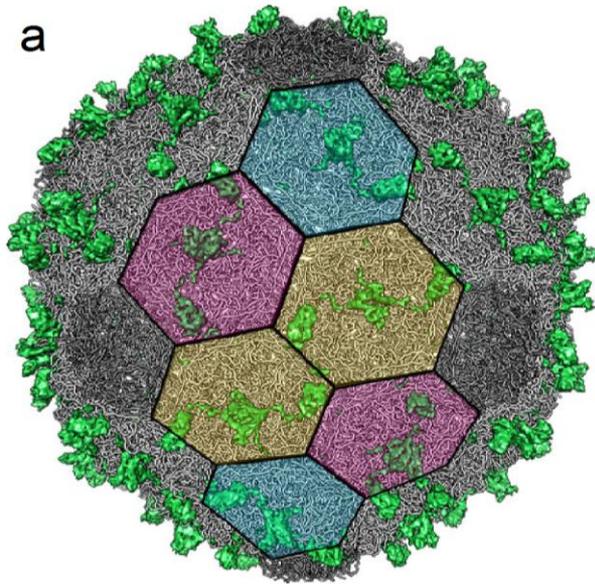
(a) The main capsid proteins in Syn5 and similar viruses are organized in the shell which corresponds to the original CK model with the indices <2,1> and the triangulation number $T_1=7$. The main protein capsid shell contains 60 hexamers. Knob-like proteins protruding from each hexamer are shown in green.



b
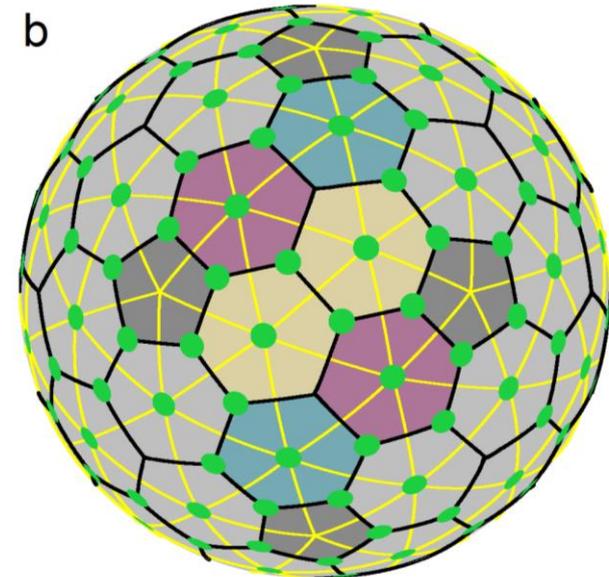
(b) Slightly symmetrized capsomers and positions of the protruding knob-like proteins. The edges of the SL with the indices <2,1> are given by yellow lines. Positions of the protruding knob-like proteins (green circles) form the SL with the indices <4,1> and the triangulation number $T_2=21$. The ratio $T_2/T_1=3$ corresponds to the simplest nontrivial commensurate relation between concentric icosahedral shells.

* P. Gipson, M.L. Baker, D. Raytcheva, C. Haase-Pettingell, J. Piret, J.A. King, and W. Chiu, *Nature Comm.*, 2014, **5**, 4278

# Landau theory of crystalization: Irreducible density waves on a sphere
## V.L. Lorman and S/B. Rochal (2007)

$\rho = \rho_0 + \Delta\rho$
Density in the self-assembled state

2D spherical distribution of proteins reads:

$\Delta\rho = \sum_{l \in N} \sum_{|m| \leq l} \rho_{lm} Y^l_m(\Theta, \phi)$

*Main contribution to $\Delta\rho$ is caused by a critical density deviation* from its value $\rho_0$  In classical thery this deviation is irreducible.
*System of Waves on a Sphere with the fixed wave number l*

*Asymmetric Protein Units* have no Proper Symmetry. Because of the Asymmetry the final structure has *neither spatial inversion nor symmetry planes* elements => only *odd* spherical harmoniques in critical deviation
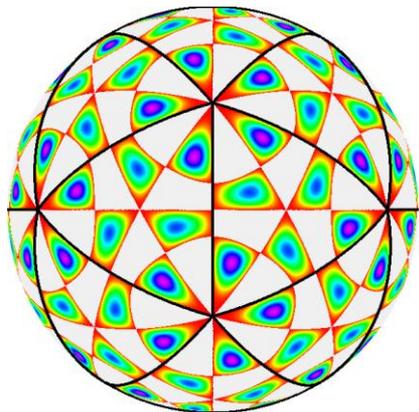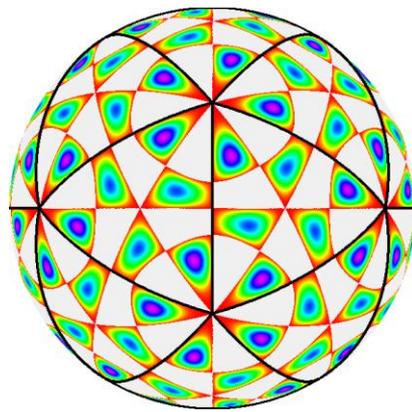


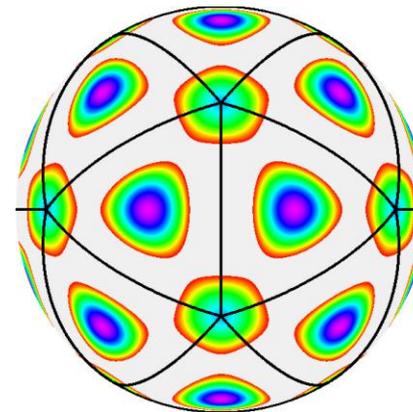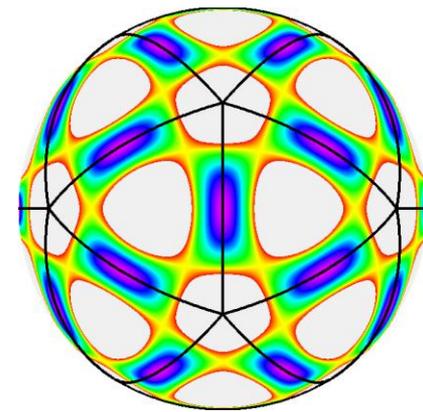$\Delta\rho > 0$        $\Delta\rho < 0$

$l = 15 + 6i + 10j$ ;   i, j $\in$ N  $l=21$
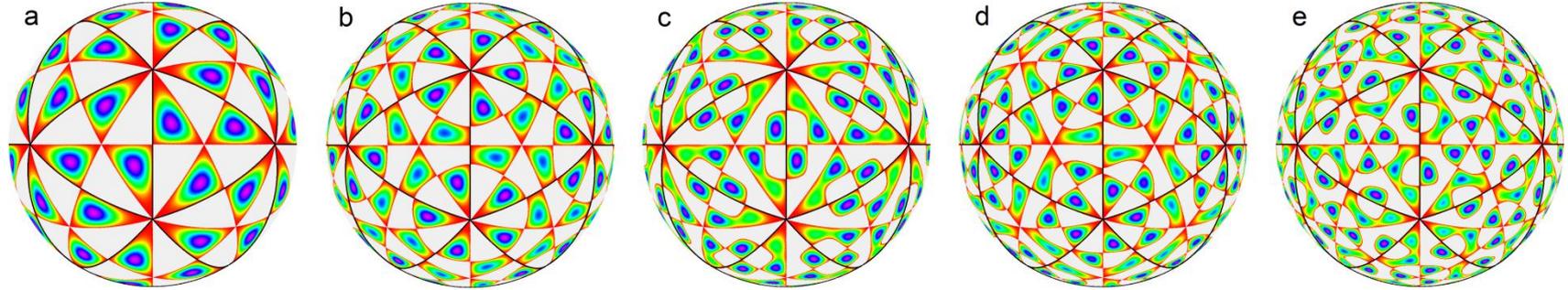
$\Delta\rho > 0$        $\Delta\rho < 0$

$l = 6i + 10j$ ;   i, j $\in$ N  $l=10$

# Chiral sferical lattices and iIrreducible even icosahedral density functions
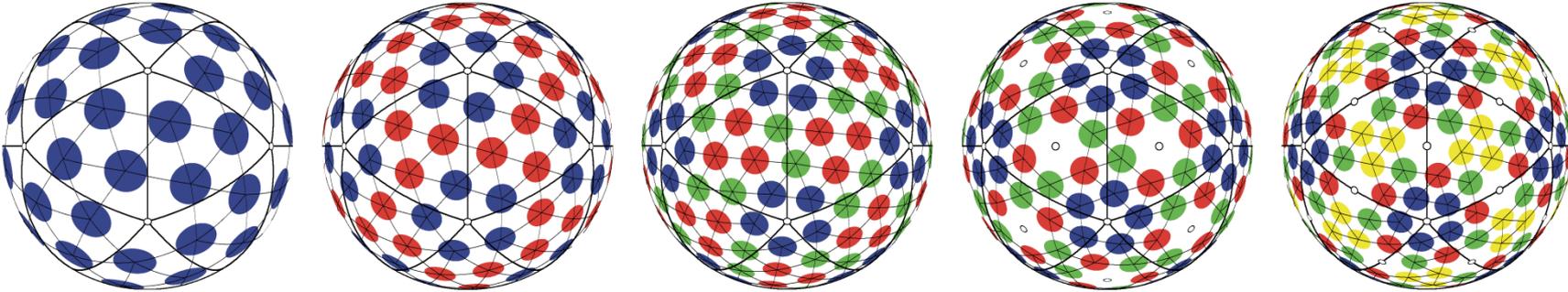


Density functions  small icosahedral viruses

a) $l = 15$; T = 1 (Caspar-Klug structure)          ===== SL <2,1>
b) $l = 21$; T = 2 (**non Caspar-Klug structure**) ===== SL <3,1>
c) $l = 25$; T = 3 (**non Caspar-Klug structure**) ===== SL <3,2>
d) l = 27; T = 3 (Caspar-Klug structure)          ===== SL <4,1>
e) l = 31; T = 4 (Caspar-Klug structure)          ===== SL <4,2>

# Free energy expansion near the isotropic phase :
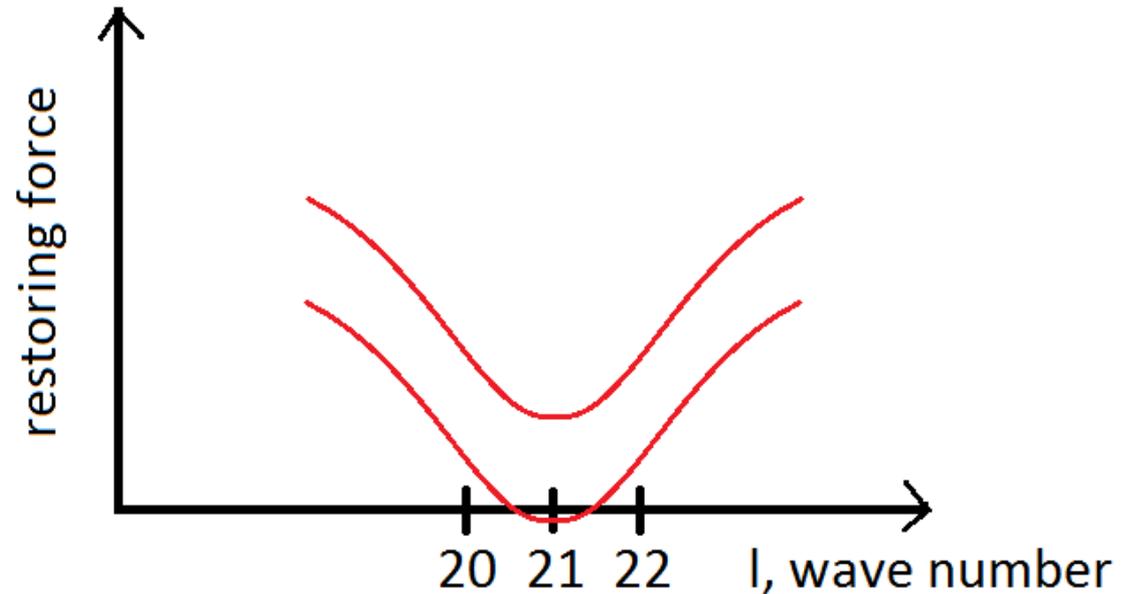
$$F_0 + F_2 + F_3 + F_4 + \ldots$$

$$F_2 = A(T, c) \sum_m a_m \, \rho_{lm} \, \rho_{l(-m)}$$

$$F_3 = B(T, c) \sum_{m1, m2, m3} a_{m1, m2, m3} \, \rho_{lm1} \, \rho_{lm2} \, \rho_{lm3} \, \delta(m_1 + m_2 + m_3) \equiv 0$$

$$F_4 = \sum_k C_k(T, c) \sum_{m1, m2, m3, m4} a_{m1, m2, m3, m4} \, \rho_{lm1} \, \rho_{lm2} \, \rho_{lm3} \, \rho_{lm4} \, \delta(m_1 + m_2 + m_3 + m_4)$$

# Taking into account of nearest even irreducible icosahedral functions

- Odd and even density functions can couple. This coupling should be more strong between functions with closest l values.



- So such odd ($l+1$, $l-1$) functions yield the second order contribution to $\Delta \rho$ in addition to the primary contribution of the even Ir function. Effective third order invariant appears [Robijn Bruinsma] and the crystallization becomes the (weak) first order phase transition.

# Conlusion

-We have modified the CK geometrical model, which is the basic paradigm in structural virology. Our approach gives rational physical interpretation for a variety of the experimentally obtained small viral capsid structures including anomalous ones. In our theory the CK projection scheme is preserved but the position are filled with proteins only after the order is transferred onto the icosahedron surface.

-The "parent" hexagonal lattice is the common origin of both the "anomalous" and conventional capsid structures. Even for small capsids described within the original CK approach, the modified model points out the additional hidden symmetry in the capsid structure.

-The developed approach clarifies the peculiarities in structural organization of double-shelled capsids. We have demonstrated the commensurability between the inner and outer capsid shells of these composite concentric nanoassemblies. Our approach also explains the location of the protruding knob-like proteins in some marine viruses.

-. The main results of the proposed geometrical approach are in a good agreement with the conclusions obtained previously in the frame of the thermodynamic Landau crystallization theory.

Rochal, S. B., Konevtsova, O. V., Myasnikova, A. E., & Lorman, V. L. (2016). *Nanoscale*, *8*(38), 16976-16988.

# Thank you for your attention