Landau theory of crystallization and the capsid structures of small icosahedral viruses

Article *in* Physical review. B, Condensed matter · June 2008

DOI: 10.1103/PhysRevB.77.224109

2 authors, including:

Vladimir L Lorman
Université de Montpellier
121 PUBLICATIONS **1,140** CITATIONS

Landau theory of crystallization and the capsid structures of small icosahedral viruses

V. L. Lorman[1] and S. B. Rochal[1,2]

[1]*Laboratoire de Physique Theorique et Astroparticules, CNRS, Universite Montpellier 2, Place Eugene Bataillon, 34095 Montpellier, France*

[2]*Physical Department, South Federal University, 5 Zorge Street, 344090 Rostov-on-Don, Russia*

(Received 9 February 2008; revised manuscript received 27 April 2008; published 19 June 2008)

A new approach to the capsid structures of small viruses with spherical topology and icosahedral symmetry is proposed. It generalizes Landau theory of crystallization to describe icosahedral viral shells self-assembled from identical asymmetric proteins. An explicit method which predicts the positions of centers of mass for the proteins constituting the shell is discussed in detail. The method is based on irreducible density distribution function which generates the protein positions. The universal form of the density distribution function which contains no fitting parameter permits to classify the capsids structures of small viruses. The theory describes in a uniform way both the structures satisfying the well-known Caspar and Klug geometrical model for capsid construction and those violating it. A group theory analysis of the Caspar and Klug model and of the "quasiequivalence" principle for protein environments in viral capsids is given. The molecular basis of difference in protein environments and peculiarities in the assembly thermodynamics are also discussed.

DOI: 10.1103/PhysRevB.77.224109 PACS number(s): 64.70.dg, 64.70.Nd, 61.44.Br

I. INTRODUCTION

Viruses occupy the "gray area" between living and nonliving matter. On the one hand, like in living organisms, virus organization and functionning depend on the program encoded in their genome. On the other hand, virus structures and self-assembly process demonstrate properties typical for ordering in nonliving physical systems. In contrast with other forms of living matter, viruses cannot replicate themselves, they need a host cell and its biosynthetic machinery to reproduce new viral particles. Their genetic material is constituted by either desoxyribonucleic acid (DNA) or ribonucleic acid (RNA)[1] which can be single or double stranded, linear or circular, and shows no unique structure. The nucleic acid is vulnerable to degradation or biochemical agression before the infection into the cell can occur. Therefore, it is protected by the viral protein shell (capsid) which encloses the genome, preserves it and transmits the genetic material to an appropriate host cell. Though the final infective virus capsid formation involves biologically specific events, some steps of the self-assembly represent passive physical processes and show universal features. The viral capsid self-assembly does not need active local energy consumption like adenosine triphosphate (ATP) hydrolisis and the process can be reversible.[2,3] Moreover, in many cases the viral shells assembling does not even need genomes and proceeds *in vitro* in purified protein solutions.[1] The capsids of almost all viruses are made of many copies of identical subunits (one or few proteins). The positions and orientations of subunits display high level of spatial organization on a two-dimensional shell. Thus, capsids represent nanosystems well-suited to modern structural methods of study such as synchrotron-radiation diffraction or cryoelectron microscopy. The structural data obtained together with additional information coming from more specific biochemical experiments rise a whole number of questions concerning unconventional positional order of subunits in the shell, thermodynamics, and physical mechanisms of the self-assembly.

The problem of the capsid structure formation attracts the attention of physicists since fifty years. In their pioneer work, Crick and Watson[4] stated that spherical viruses should have the symmetry (but not necessarily the structure) of one of regular polyhedra with the faces formed by identical perfect polygons. Later in 1962, Caspar and Klug (CK) argued that spherical capsids adopt icosahedral point symmetry.[5] They have seen the physical reason why the Nature prefers this type of symmetry in the fact that the icosahedron has the largest volume-to-surface ratio among the regular polyhedra. Besides, CK obtained four prominent results:[5] (i) The capsid symmetry is lower than that of the regular icosahedron since the proteins are asymmetric. Identical asymmetric building blocks can compose the structures with rotational symmetry elements only, excluding inversion and mirror planes. (ii) The asymmetric proteins can be located only in regular (trivial) 60-fold positions of the rotational icosahedral point group, therefore the total number of proteins in a capsid is always equal to $60N$, where N is a positive integer number. (iii) CK concluded, for the first time, that "the self-assembly is a process akin to crystallization and is governed by the laws of statistical mechanics." (iv) They proposed a geometrical model for the viral capsid construction based on the properties of the almost regular mapping of the two-dimensional (2D) hexagonal structure on the icosahedron surface. Specific properties of the model impose the selection rules for the value of N (and, consequently, for the total number of proteins in the shell). Only the values which satisfy the relation $N=h^2+k^2+hk$, where h and k are nonnegative integers are allowed by the CK selection rules. All four points and their direct consequences resulted in the principle formulated by CK and put in the basis of modern virology. Though a big number of virus, capsid structures are in a good agreement with all the points of the CK scheme, there is a growing number of experimentally resolved structures which do not satisfy the CK selection rules nor their predictions about local proteins arrangement.[6] These facts show that point (iv) of the principle is not universal and needs to be generalized.

Recent progress of the x-ray and cryoelectron microscopy techniques[7] gave a new insight into the viral capsids problem. Due to the experimental advances, the capsid structures are now determined at high resoluton ~2.8 Å (they are the largest biomecular aggregates whose organization is known in such a detail). Refined structural data have stimulated a new wave of theoretical interest to the viral capsid structures and to the self-assembly process. From the theoretical point of view, the main effort was done in two directions (see Ref. 8). Along the first line, the mechanical properties of capsids and their relation to the capsid shape were investigated.[9,10] The possible buckling instability of the spherical capsid was studied in the frame of the nonlinear physics of thin elastic shells.[9] The results of this study explain why the relatively small viruses are always spherical while the larger ones have a more angular or faceted shape. In the continuous model developed, the shape is related to the dimensionless Fopple–von Kármán number which is, in turn, dependent on a square of the capsid shell radius. A big number of proteins in the shell of large viruses justifies the use of continuum elasticity approximation and makes the predictions of (Ref. 9) about the large-scale shape details universal. Along the second line, the mean-field studies of model systems were performed in order to approach the thermodynamics of the self-assembly process. The results in this direction were obtained using strong physical and geometrical simplifications. The protein-protein interaction was reduced to the interaction between protein oligomers (capsomers). The symmetry of the capsomers was considered to be isotropic and the shape of the capsomers was associated to a disk. The free energy of the viral structure has been approximated by that of a model system consisting of two types of disks located on the spherical surface.[11] The proposed pair potential of the disk interaction favors the icosahedral symmetry of the disk packing[11] provided an optimization of several model parameters. However, all recent theoretical works on the capsid structure do not take into account the asymmetry of capsid proteins and the restrictions on the capsid symmetry formulated in points (i) and (ii) of the CK principle. By contrast, the nonuniversal CK selection rules [point (iv)] are taken as an ingredient in all models.

In our previous Letter,[12] we proposed to apply the Landau theory of crystallization to explain and to classify the capsid structures of small viruses with spherical topology and icosahedral symmetry. We sketched the theory which describes in a uniform way both the structures satisfying the CK geometrical model for capsid construction and those violating it (e.g., L-A virus, Dengue virus, West Nile virus, Murine Polyoma virus, etc.). The aim of the present paper is to develop the density wave theory of small icosahedral capsid structures. The approach explicitly takes into account the protein asymmetry and satisfies points (i)–(iii) of the CK principle but it is free of nonuniversal CK selection rules. It includes an analysis of the CK model by means of the group theory methods and a discussion of the quasiequivalence notion which underlies the CK principle. An explicit method which predicts the positions of centers of mass for the proteins constituting viral capsid shell is presented in detail. The paper is organized as follows. In Sec. II, a detailed analysis of the CK geometrical model is given. Section III contains the main points of the theory of crystallization on the spherical surface and their comparison with the experimental viral capsid structures. Sec. IV is devoted to the discussion of several biologically relevant predictions of the theory and to its nonstructural features. Selection rules for the order parameters which drive the capsid crystallization transitions are derived in Appendix.

II. SYMMETRY ANALYSIS OF THE CASPAR AND KLUG GEOMETRICAL MODEL

Microscopic interactions involved in the packing of viral proteins into a shell are still poorly understood. It is clear that protein folding states and protein chemistry define particular anisotropies and docking preferences in the interactions. They result in specific geometrical arrangements and influence the capsid assembly phenomenon. The studies of these complex problems constitute a separate branch of science (at the interface between biochemistry and biophysics) and are out of the scope of the present work. Here we follow the approach adopted by CK (Ref. 5) which consists in the consideration of only the robust properties of shell elements and only the main consequences of protein structure and chemistry. Three basic properties used by CK in the study of viral capsid formation were: the lateral type of protein interaction, the intrinsic curvature, and especially the asymmetry of capsid proteins. Using, in addition, the fact that small viral genome can code only one (or few) type of proteins they formulated the main structural problem in the field of physical virology: how to construct a regular shell with the icosahedral symmetry formed by multiple copies of identical asymmetric proteins (considered as 2D units). They also stressed that the interaction of identical proteins should lead to identical local environments, including local orientational and positional order and local chemical bonding and proposed to construct the icosahedral shells possessing these properties. A whole number of physical and geometrical restrictions arises on this way and leads to the selection rules for the viral capsid structures.

Since the proteins are asymmetric, they can form the structures with rotational symmetry elements only, excluding inversion and mirror planes. For icosahedral viruses, it means that the ordered distribution of proteins in the capsid has the symmetry group I of the icosahedron rotations and not the I_h group of all icosahedron symmetries. Indeed, structures with mirror symmetries would impose to their constituents either high symmetry of individual proteins or the presence of two nonidentical proteins related by a mirror-plane operation (a racemic mixture of two distinct enantiomers of opposite chirality). Both conditions are incompatible with the asymmetry of capsid proteins.

Identical asymmetric proteins can be put in identical environments if they are located in the positions which form one regular (trivial) orbit of the rotational icosahedral point group I. The number of positions in a regular orbit of a discret symmetry group G is equal to the group order $|G|$ (total number of operations). Thus, for the I group, the orbit contains 60 different equivalent positions, which means that 60 would be the maximum possible number of proteins with

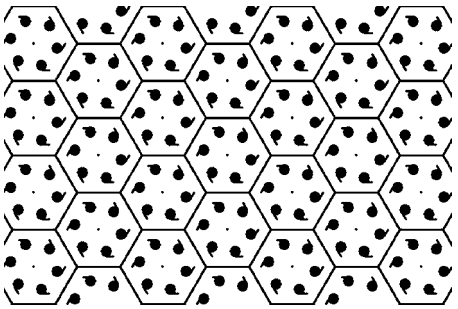

FIG. 1. Plane hexagonal structure with one sixfold regular position filled with asymmetric structural units (proteins). Proteins are presented by full circles with tails. Honeycomb lattice is added as a guide to the eye to visualize the symmetry axes (see in the text). A net of an icosahedron containing 20 equilateral triangles can be slitted as a finite piece of the structure.

identical local environments in an icosahedral shell. However, many viruses contain more than 60 identical proteins in the shell. In general, their number is equal to $60N$, where N is a positive integer. Their positions belong to different regular orbits and cannot be equivalent. The task of CK was to find a way to put identical proteins in different but nearly equivalent positions and to explain the origin of this "quasiequivalence." For that aim, CK proposed to use the properties of plane periodic structures formed by the same type of asymmetric particles. Indeed, the above restriction on the group orbit size is proper to point groups which describe the distributions with spherical or other nontrivial compact topology. By contrast, the translational symmetry of a 2D plane lattice makes infinite the number of positions belonging to the same regular orbit of the 2D crystal space group. Consequently, the number of asymmetric proteins which can be put in equivalent environments becomes infinite in a 2D plane crystal. CK were then looking for an almost regular mapping of the plane periodic structure on the icosahedron surface. Mapping splits one regular orbit of the plane structure into several different regular orbits of the icosahedron group but maintains some "traces" of their former "equivalence" in the plane structure, thus making the positions in different orbits "quasiequivalent." Symmetry analysis shows that the only type of crystalline order suitable for the almost regular mapping is the plane hexagonal structure with one sixfold regular position in the unit cell filled with asymmetric proteins (see Fig. 1). Proteins in this structure are presented by circles with tails. To visualize the symmetry axes, the structure is superimposed with the honeycomb lattice with sixfold, threefold, and twofold axes situated in each hexagon center, each hexagon vertex and in the middle of each edge, respectively. Note, however, that the honeycomb lattice symmetry is higher than that of the considered hexagonal structure and the lattice is given in Fig. 1 for practical reasons only.

A net of an icosahedron which consists of 20 equilateral triangles can be slitted as a finite piece of the considered plane hexagonal structure and then mapped (folded) onto the icosahedron surface (Fig. 2). The mapping is chosen in such a way that the vertices of regular triangular faces of the icosahedron coincide with the sixfold axes of the plane hexagonal structure. By cutting out 60° sector, one can transform the sixfold axis into the fivefold one and then join the sector edges on the icosahedron surface. Let us stress the remarkable symmetry property of this operation which has allowed CK to state that their folding is done "without destroying the bonding pattern in the lattice." Indeed, joining of the edges in this case represents a continuous matching operation because the edges of the cut out sector are equivalent by the symmetry of the hexagonal structure. This result has to be compared with the mapping of the plane trigonal (and not hexagonal) structure on the icosahedron surface which is, in principle, not forbidden by geometry. One could slit the triangular net of an icosahedron as a piece of a plane trigonal structure with threefold and not sixfold symmetry axes. But in this case, the edges of the 60° sector would not be equivalent by the threefold symmetry of the structure. Then, joining of the edges would be not matching but "stitching" operation and would lead to the "seam" formation. For the CK mapping of the plane hexagonal structure, the edge length of the icosahedron face is determined by the vector joining vertices of the net triangles. Vertices being situated in the nodes of the hexagonal lattice, the vector length square is given by $N=h^2+k^2+hk$, where h and k are non-negative integers. This number is evidently equal to the number of the lattice nodes contained in two net triangles. Figure 2 shows first two nontrivial icosahedron nets with $N=3$ and $N=4$, respectively. In the hexagonal coordinates (h,k), the first net is indexed as (1,1), the index of the second one is (2,0).

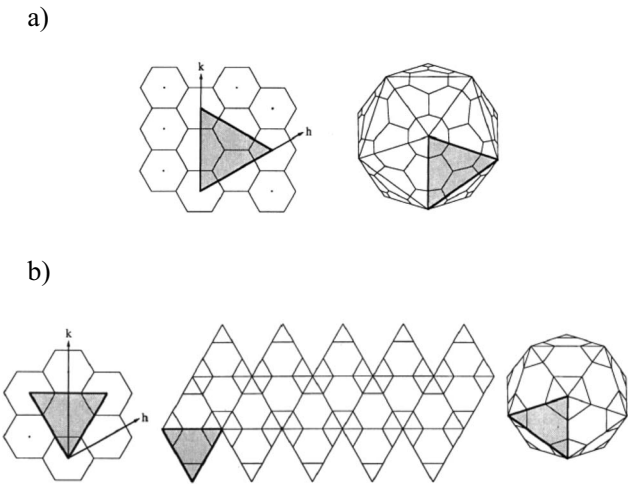

FIG. 2. Mapping of the plane hexagonal structure on the icosahedron surface. First two nontrivial icosahedron nets with $(h,k)=(1,1)$ and (a) $N=h^2+k^2+hk=3$ and (b) $(h,k)=(2,0)$ and $N=4$ are presented.

In what follows, we distinguish the consequences of the CK mapping for the hexagonal lattice nodes on the one hand, and for the protein positions on the other hand. This distinction is important for the understanding of the morphological unit (capsomer) concept introduced by CK. In the plane hexagonal structure (Fig. 1), all protein positions belong to one sixfold regular position in the unit cell. They can be conventionally divided in groups of six positions situated around the same lattice node. Corresponding six proteins can be then considered as a morphological unit, which is a hexamer in

the considered case. The CK geometrical model maps both the lattice nodes and the protein positions onto the icosahedron surface. After the mapping the number of positions around one "node" (the image of the lattice node mapping) is either six or five. Corresponding morphological units can be conventionally considered as hexamers or pentamers. All $60N$ proteins can be divided in 12 pentamers which surround the nodes situated in the fivefold axes and $10(N-1)$ hexamers situated around other nodes. Centers of pentamers and hexamers coincide with the nodes which form the image of the hexagonal Bravais lattice mapping. Total number of these nodes is given by $10N+2$.

The image of the hexagonal lattice mapping helps to illustrate the quasiequivalence notion. The mapping transforms the hexagonal lattice node surrounded by six equivalent positions into a "node" which looses its true sixfold symmetry. In the (1,1) structure (Fig. 2), the true local symmetry of a hexavalent node becomes threefold one, in the (2,0) structure (Fig. 2) this type of node has twofold symmetry, while in the (2,1) structure, the symmetry of the hexavalent node becomes trivial. However, the CK mapping preserves an approximate quasihexagonal symmetry of the structure. It maintains not only the approximately hexagonal shape of "hexamers" around each hexavalent node but also the approximate twofold and threefold local axes (mapping images of the true twofold and threefold axes of the hexagonal lattice) situated in general points of the icosahedron surface. In a whole number of viral capsids which satisfy to the CK geometrical model, the approximate twofold and threefold axes appear as nearly parallelograms and nearly-equilateral triangles, respectively. The proteins situated in the vertices of these parallelograms and triangles belong to different capsomers, thus illustrating the conventional character of the capsomer notion. Finally, we can resume that the CK construction transforms global symmetry axes of the hexagonal structure into the approximate local ones and the true structural equivalence of protein positions becomes the structural quasiequivalence.

Clear and simple rules of the CK principle constituted the basis for the x-ray and cryoelectron microscopy data interpretation. Note, however, that the CK geometrical model reduced possible capsid structures to those satisfying the selection rules for the number of different environements N and, consequently, for the total protein number $60N$. It imposed also the particular hexamer shape to protein positions situated around hexavalent nodes on the icosahedron surface thus reducing strongly possible types of local protein arrangements. Though a big number of virus capsid structures obtained at moderate resolution by x-ray crystallography[7] are in a good agreement with the CK geometrical model, there is a growing number of structures resolved at high resolution which do not satisfy the CK selection rule for N number nor their predictions about local proteins arrangement.[6] In several peculiar cases of capsids which do not satisfy the restriction imposed by the CK quasiequivalence principle, the geometrical model can be formally "repaired" by replacing hexamers in the nodes of the lattice mapping by dimers, trimers, or pentamers. This operation lacks evidently physical justification. In addition, it leads to the loss of local approximate symmetry axes and brakes the structural quasiequivalence of protein environments in the sense proposed by CK. Nevertheless, several capsids of this type and even more complex cases of viral capsids were observed experimentally by high-resolution crystallography methodes. These data show that identical proteins can occupy a small number of structurally not equivalent positions in shells with the icosahedral symmetry and that the notion of quasiequivalence needs to be revised. In Sec. III, we apply the Landau theory of crystallization to the problem of small capsid formation. Resulting approach to the icosahedral virus structure replaces purely geometrical quasiequivalence by physical equivalence based on the principles of statistical physics.

III. CRYSTALLIZATION ON THE SPHERICAL SURFACE

Basic ideas of the approach developed in this section have been introduced in Ref. 12. In a perfect accord with the hypothesis formulated in point (iii) of the CK principle (see Sec. I), we have associated the viral capsid formation with the crystallization process and proposed to describe the capsid self-assembly using a generalization of the Landau theory of crystallization. From the structural point of view, the resulting approach takes explicitly into account the protein asymmetry [point (ii) of the CK principle] but it is free of nonuniversal CK selection rules. It allows us to describe in a uniform way all experimentally observed small spherical viruses including those which cannot be obtained using the CK geometrical model. To generalize the structure description and to relate it to the thermodynamics of the self-assembly, we follow the main line of the Landau theory of structural phase transitions[13] and replace geometrical notions by statistical ones. Namely, the positions of identical proteins in the shell are associated in the frame of the theory with one statistical protein density distribution function. The assembly thermodynamics is described by the free energy which represents an invariant functional of the density function. Furthermore, we argue that similar to the case of atomic positions in simple atomic crystals the protein positions in the capsid shell can be located in the maxima of the corresponding density function.

Both the experimental data and the theoretical consideration[9] show that the shape of small viruses with the icosahedral symmetry is close to the spherical one. This fact gives the possibility to consider the crystallization on a spherical surface and to avoid the problems arising in the CK geometrical construction during the mapping of planar hexagonal structures upon the icosahedron surface. Like in the case of usual 3D crystal solidification[13] Landau theory of the assembly process gives simple and clear predictions in the vicinity of crystallization point. In this region, the probability density ρ of protein distribution in the capsid structure is presented as:

$$\rho = \rho_0 + \Delta\rho, \tag{1}$$

where ρ_0 is an isotropic density in the solution and $\Delta\rho$ corresponds to the density deviation induced by the ordering. The symmetry breaking during the crystallization is associated with one critical order parameter which spans an irre-

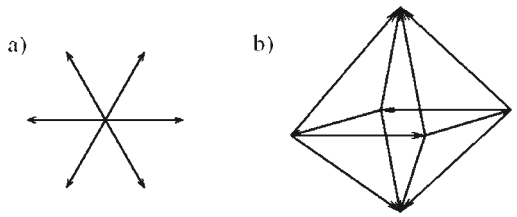

FIG. 3. Irreducible stars of wave vectors in the reciprocal space which characterize critical system of density waves for the structure crystallization. (a) Six waves which generate the plane hexagonale lattice; (b) twelve waves which generate the B.C.C. lattice.

ducible representation of the symmetry group of the disordered state. In addition, in the vicinity of crystallization point, the structure of the ordered state (defined by $\Delta\rho$) is determined by the critical order parameter only, the contribution of noncritical degrees of freedom being negligible in this region. For the crystallization process, the order parameter represents a critical system of density waves (CSDW) with the wave vectors of the *same length* and the transition free energy is an invariant function of the CSDW amplitudes.[13] The symmetry of crystals which condense from the isotropic state coincides exactly with that of the corresponding CSDW. In the vicinity of crystallization point, the atomic positions can be then associated with the positions of maxima of the CSDW. In the following consideration, we will call *physically equivalent* the positions which are generated by maxima of a *single irreducible* density function.

In simple atomic solids, e.g., in crystals of metals and especially in crystals of elements, maxima of the irreducible density function are located in high-symmetry positions. Structurally, it means that the corresponding identical atoms occupy the nodes of the crystal lattice only. The environments of these maxima are identical and the atomic positions are structurally equivalent. As it was first shown by Landau and discussed in detail by Alexander and McTague,[14] thermodynamics favors in the vicinity of crystallization point the body-centered cubic lattice in usual 3D crystals and plane hexagonal lattice in two dimensions. The CSDW in these two cases are characterized by the irreducible stars of wave vectors given in Figs. 3(a) and 3(b), respectively. In both structures, the maxima of the density function and, correspondingly, the positions of identical atoms have identical environments. However, physically equivalent positions are not always structurally equivalent. In general case, even in the vicinity of crystallization point, the maxima of irreducible density function have different environments but their number is small and they maintain "genetic" relations between them. It is the case in more complex periodic crystals but also in some aperiodic structures with long-range order.

The same principles are applied here to the assembly process on a sphere. The critical part $\Delta\rho_l$ of the density is determined by a CSDW with the *same wave number l*. The spherical harmonics Y_{lm} constituting CSDW on a sphere span one irreducible representation (IR) of the SO(3) symmetry group of the disordered state, thus $\Delta\rho_l$ is given by:

$$\Delta\rho_l(\theta,\phi) = \sum_{m=-l}^{m=l} A_{lm} Y_{lm}(\theta,\phi), \tag{2}$$

where l is the IR number, A_{lm} are the amplitudes of the spherical harmonics Y_{lm}, and θ and ϕ are the conventional angular variables of the spherical coordinate system.

According to points (i) and (ii) of the CK principle the ordered distribution of proteins in the viral capsid has the symmetry group I of the icosahedron rotations which does not contain spatial inversion nor mirror planes. This restriction is of major importance in the proposed theory. It selects the parity of the "active" IR's of the SO(3) symmetry group which induce the assembly of icosahedral shells of asymmetric proteins. Thus, the spherical harmonics Y_{lm} with *even* l numbers cannot form critical density [Eq. (2)] for viral capsids. The restriction also affects the free energy expansion of the assembly process taken in a standard for the crystallization theory form[13] $F=F_0+F_2+F_3+F_4+\ldots$ and containing invariant terms

$$F_2 = A(T,c) \sum_{m=-l}^{m=l} A_{l,m} A_{l,-m},$$

$$F_3 = B(T,c) \sum_{m_l,m_2,m_3} a_{m_l,m_2,m_3} A_{l,m_1} A_{l,m_2} A_{l,m_3} \times \delta(m_l + m_2 + m_3) \equiv 0,$$

$$F_4 = \sum_k C_k(T,c) \sum_{m_l,m_2,m_3,m_4} a^k_{m_l,m_2,m_3,m_4} A_{l,m_1} A_{l,m_2} A_{l,m_3} A_{l,m_4} \delta(m_l + m_2 + m_3 + m_4), \tag{3}$$

where a_i are weight coefficients of the SO(3) group (e.g., Clebsch-Gordan coefficients for the third-order term F_3), $\delta(0)=1$, $\delta(i\neq 0)=0$, $A(T,c)$, $B(T,c)$, and $C_k(T,c)$ are temperature- and composition-dependent coefficients of the Landau theory. For any *odd* wave number l, the third-order term F_3 is identically zero. This fact makes the thermodynamics of asymmetric proteins assembly different with respect to the thermodynamics of 3D icosahedral atomic clusters formation[15] in spite of several common points in formal description.

Next restriction on the choice of order parameters of the capsid formation comes from the fact that $\Delta\rho_l$ function with I symmetry can be constructed not for all but for particular odd l numbers only. The analysis based upon the theory of invariants (see Appendix) shows that any critical order parameter which drives the icosahedral assembly of asymmetric proteins has the wave number l satisfying the relation:

$$l = 15 + 6i + 10j, \tag{4}$$

where i and j are positive integers or zero. Equation (4) defines the list of l numbers for which the restriction of an IR of the SO(3) group on the icosahedral group I contains at least one totally symmetric representation. The sequence of the permitted values of the wave number l is given by: $l=(15,21,25,27,31,33,35\ldots)$. As we show below, this sequence determines possible capsid shell structures for small icosahedral viruses. Selection rule [Eq. (4)] gives the possi-

bility to obtain the explicit form of critical density [Eq. (2)]. Then the protein centers are associated with the positions of maxima of $\Delta\rho_l$ function (2). Thus, the density wave approach replaces nonuniversal geometrical model (iv) of the CK principle.

The explicit form of the critical density function $\Delta\rho_l(\theta,\phi)$ is given by the basis functions $f_l^i(\theta,\phi)$ $(i=1,2\ldots n_t)$ of all n_t totally symmetric representations of the icosahedral group I in the restriction of the active IR of the SO(3). The CSDW is a linear combination of these functions invariant with respect to the I group:

$$\Delta\rho_l(\theta,\phi)=\sum_{i=1}^{n_t} B_i f_l^i(\theta,\phi), \tag{5}$$

where B_i are arbitrary coefficients.

Their number n_t is equal to the number of integer nonnegative solutions (i,j) of Eq. (4) for a fixed permitted value of l. Another way to calculate n_t is to use the well-known relations of characters:[16]

$$n_t=(1/|G|)\sum_G \xi(\hat{g}) \tag{6}$$

where the sum runs over the elements $\hat{g}$ of the I group, $|G|=60$ is the I group order, and $\xi(\hat{g})$ is the character of the SO(3) group element which reads as:[16]

$$\xi(l,\alpha)=\frac{\sin[(l+1/2)\alpha]}{\sin(\alpha/2)},$$

where l is the IR number and the angle α is determined by the element $\hat{g}$. Then the explicit form of Eq. (6) becomes:

$$n_t(l)=\frac{1}{60}[2l+1+15\xi(l,\pi)+20\xi(l,2\pi/3)+12\xi(l,2\pi/5)+12\xi(l,4\pi/5)]. \tag{7}$$

For small icosahedral capsids, the practical construction of the protein density distribution is simplified because the CSDW [Eq. (5)] contains only one function $f_l(\theta,\phi)$. Indeed, according to Eq. (4) and/or Eq. (7) $n_t=1$ for all $l\leq 43$. In this simplest case $\Delta\rho_l(\theta,\phi)=Bf_l(\theta,\phi)$, where B is a single arbitrary coefficient. The positions of maxima of the density function do not depend on the value of B. They are generated by a single universal function $f_l(\theta,\phi)$ which has no fitting parameter. In the following consideration, the functions $f_l(\theta,\phi)$ possessing this properties are called *irreducible icosahedral density functions* and the structures generated by $f_l(\theta,\phi)$ are mentioned as *irreducible icosahedral structures*. The explicit form of the irreducible density function $f_l(\theta,\phi)$ for a given value of l is obtained by averaging of $Y_{lm}(\theta,\phi)$ harmonics over the I symmetry group.[17]

$$f_l(\theta,\phi)=\frac{1}{60}\sum_G Y_{l,m}[\hat{g}(\theta,\phi)]. \tag{8}$$

For any fixed value of m, procedure (8) gives either the same function $f_l(\theta,\phi)$ we are looking for, or zero. Functions are defined up to a constant complex multiplier. An appropriate choice of this multiplier makes functions (8) real.

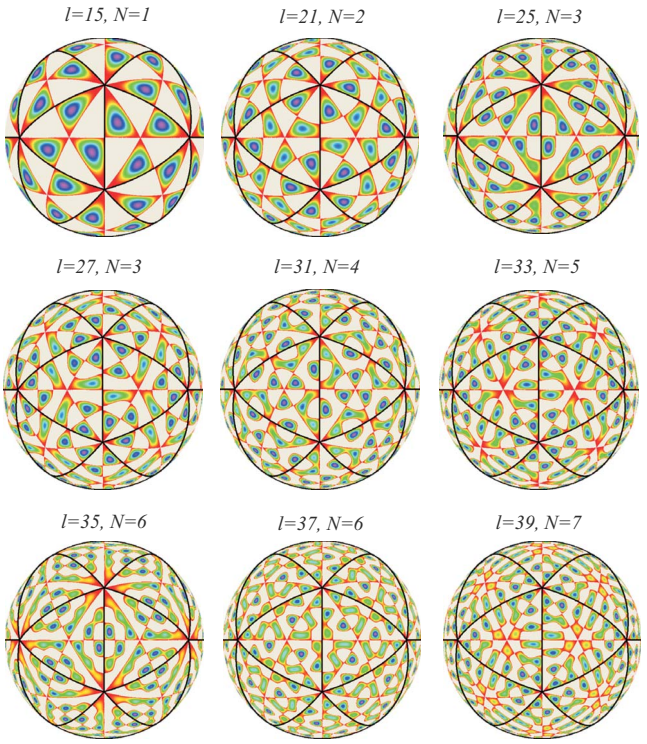

FIG. 4. (Color online) The first nine irreducible icosahedral density functions with the wave numbers $l=15$, 21, 25, 27, 31, 33, 35, 37, and 39, respectively. Corresponding numbers of different 60-fold positions of density maxima are $N=1$, 2, 3, 3, 4, 5, 6, 6, and 7.

Figure 4 resumes the irreducible density functions $f_l(\theta,\phi)$ permitted by selection rule (4) for the nine smallest icosahedral capsids. The value of $f_l(\theta,\phi)$ is represented using false color image: variation of colors from red to violet corresponds to the function growth. Note that all $f_l(\theta,\phi)$ functions are antisymmetric: they change their sign under the inversion of all coordinates or under the action of mirror planes of a regular icosahedron. Thus, for the sake of clariry, we present the positive part $f_l(\theta,\phi)>0$ only. The number of maxima of the density functions is equal to $60N$, where N is the number of different regular 60-fold positions of the I group. In the viral capsid, N corresponds to the number of different positions occupied by the proteins. Let us stress that in a sharp contrast with the CK geometrical model, the crystallization theory predicts the existence of capsids with all positive integer values of N and not only for $N=h^2+hk+k^2$. Functions $f_l(\theta,\phi)$ generate in a uniform way protein distributions which can be obtained by the CK mapping of the hexagonal lattice on an icosahedron and those which cannot. On the one hand, the distributions $(l=15,\ N=1)$, $(l=27,\ N=3)$, and $(l=31,\ N=4)$ in Fig. 4 give classical CK structures. The positions of protein centers in a big number of viral capsids are described by these density distribution functions. Figure 5 shows the correspondence between the maxima of f_{15}, f_{27}, and f_{31} and the protein arrangement in Satellite Tobacco Mosaic virus [Fig. 5(a)], Cowpea Chlorotic Mottle virus [Fig. 5(b)] and Sindbis virus [Fig. 5(c)], respectively. On the other hand, the distributions $(l=21,\ N=2)$, $(l=33,\ N=5)$, $(l=35,\ N=6)$, and $(l=37,\ N=6)$ in Fig. 4 do not satisfy the CK selection rules for N number. The distribution (l

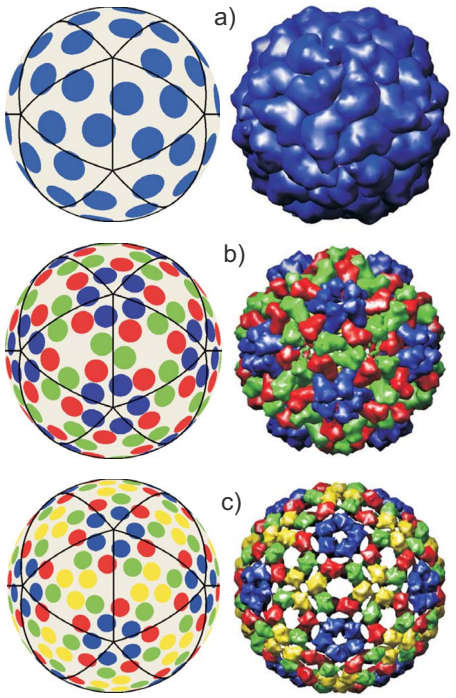

FIG. 5. (Color online) Comparison of the positions of protein centers predicted by our model (left panel) with the experimental viral structures[18] (right panel) for the capsids satisfying selection rules of the CK geometrical model. Capsids of (a) Satellite Tobacco Mosaic virus, (b) Cowpea Chlorotic Mottle Virus, and (c) Sindbis virus are presented. The corresponding density functions for $l=15$, $l=27$, and $l=31$, respectively, are shown in Fig. 4.

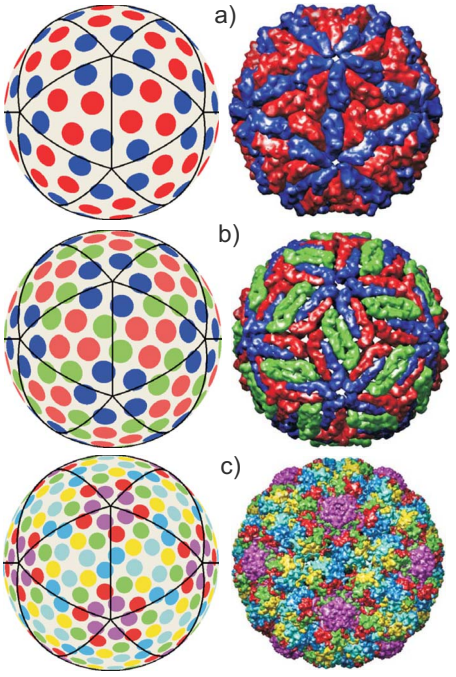

FIG. 6. (Color online) Comparison of the positions of protein centers predicted by our model (left panel) with the experimental viral structures[18] (right panel) for the capsids which cannot be explained by the CK geometrical model. Capsids of (a) L-A virus, (b) Dengue virus, (c) Murine Polyoma virus are presented. The corresponding density functions for $l=21$, $l=25$, and $l=37$, respectively, are shown in Fig. 4.

$=25$, $N=3$) shows no hexagonal arrangements of protein positions and cannot be obtained by the CK geometrical model, though the number of protein positions N satisfies the CK selection rules. In addition, the comparison of distributions ($l=25$, $N=3$) and ($l=27$, $N=3$) in Fig. 4 illustrates another striking result of the crystallization theory: there exist qualitatively different capsid structures (induced by $f_l(\theta,\phi)$ functions with different l) but with the same number N of protein positions.

The high-resolution x-ray and cryomicroscopy data show the existence of a whole series of viral capsids which violate the CK geometrical model but correspond to the distributions generated by density functions $f_l(\theta,\phi)$. Figure 6(a) illustrates the correspondence between the positions of maxima of $f_{21}(\theta,\phi)$ and the structure[18] of L-A virus with $N=2$; Fig. 6(b) relates the maxima of $f_{25}(\theta,\phi)$ to the structure[18] of Dengue virus with $N=3$ and in Fig. 6(c) the maxima of $f_{37}(\theta,\phi)$ are compared with the protein distribution[18] of Murine Polyoma virus with $N=6$.

IV. DISCUSSION

Note that in the structures presented in Sec. III, the protein environments violate the CK geometrical model and are not quasiequivalent in the CK sense. However, they do not violate the physical equivalence of protein positions induced by a single irreducible density function. Furthermore, they lead at the molecular scale to the constraints on the protein quaternary structure similar to those imposed by the CK quasiequivalence principle. Indeed, the CK geometrical construction[5] ensures approximate structural equivalence of proteins from different orbits of the icosahedron symmetry group. Nevertheless, the number of different environments remains equal to the number of the group orbits. In other words, the asymmetric identical building blocks should be "deformable" and in different local environments should find themselves in different states. The latter property is intrinsic not only to the quasieqivalent CK structures but also for the capsid structures with physically equivalent positions shown in Fig. 6. Physical equivalence assumes that there exists a small number of different environments and that the states of a protein in different positions are related by a physical mechanism.

Physical chemistry basis of molecular "deformation" which allows the proteins to occupy different positions has been extensively studied.[19,20] It was shown that in small viruses, capsid proteins have special segments called "molecular switches" which provide the protein with the necessary structural properties. Two main mechanisms of "switching" were evidenced: (1) molecular switch can be ordered in some positions and disordered in similar, but not symmetry equivalent, positions; (2) the switch can have a small number of descrete conformational states which are realized in slightly different protein positions. Both mechanisms are quite general and are not related to the specific properties of the CK geometrical model. Witout any change in reasoning, these mechanisms constitute physical chemistry basis for the struc-

tures obtained in the framework of the density wave theory.

It is interesting to note that for a fixed number of proteins in a shell, the icosahedaral symmetry minimizes the number of different conformational states for switching domains in a protein molecule. Indeed, with respect to all other point groups which could, in principle, be suitable for a self-assembly of asymmetric structural units, the icosahedral group I has the maximal number $|G|=60$ of positions in one regular orbit (to be compared with $|G|=24$ and $|G|=12$ for the octahedral and tetrahedral rotational groups, respectively). Thus, for this symmetry, the relative number of proteins in structurally equivalent positions, and, consequently in the same conformational state, is maximal with respect to all other symmetry groups, and the number of different states is minimal.

Let us also briefly discuss the particular features of the assembly thermodynamics. Due to the absence of the cubic term in free energy [Eq. (3)], the icosahedral capsid assembly can be second-order phase transition. Thermodynamic processes of this type have two advantages for the assembly optimization: they need no latent heat to be involved in; and they take place without nucleation process. Several experimental facts in favor of this possibility can be found for a number of small viruses:[21] at equilibrium, either intact virus shell or free proteins are dominant species while assembly intermediates (capsid germs) are found in trace concentration. Possibly, collective character of second-order phase transition can help virus to reduce the probability of assembly errors which are more probable for strongly discontinuous processes. Of course, high nonlinearity of the free energy can make any phase transition discontinuous even in the case when the symmetry allows it to be continuous.[22–24] Thus, no clear-cut conclusion concerning the assembly kinetics can be done on the basis of symmetry considerations only. By contrast, structural features of the density wave theory are practically independent on the order of phase transition. The original Landau theory of crystallization[13] was introduced to describe discontinuous transformations of liquids into classical crystals. Even more, the degree of the transition discontinuity was directly related[13,14] to the choice of thermodynamically favorable phase in the vicinity of crystallization point. There exist only one class of phase transitions for which the Landau theory cannot give simple structural conclusions: these are reconstructive strongly discontinuous transformations. In the available experimental data on equilibrium thermodynamics of the viral capsid assembly, there is no indication in favor of this type of process.

Note also that the absence of the cubic term in the free-energy expansion and possible continuous character of the transformation makes the thermodynamics of asymmetric proteins assembly different with respect to the thermodynamics of 3D icosahedral atomic clusters formation.[15] The orientational symmetry breaking during the cluster formation is described by the free energy related to even spherical harmonics, in contrast to the expansion in odd harmonics in the case of the protein assembly. In the capsid shells, the asymmetry of proteins forbids the spatial inversion operation in the protein density distribution. This asymmetry has its incidence on the free-energy expression, while the isotropic symmetry of atoms in the clasters induces no additional selection rule. Corresponding free energy of the cluster formation contains third-order term and thus, in no case can be continuous. Discontinuous character of the transition at equilibrium favors nucleation process in kinetics.

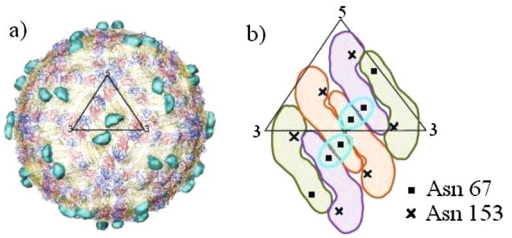

FIG. 7. (Color online) Carbohydrate recognition domain (CRD) of a dendritic cell ancillary receptor interacting with the surface of Dengue virus (DENV).[25](a) Surface-shaded representation of the DENV-CRD complex at 25 Å. Three different capsid protein environnements ($N=3$) are colored in different colors. The bound CRD molecules are shaded in cyan. An icosahedral asymmetric unit (fundamental domain of the icosahedral I group) is outline in black. (b) Schematic of the DENV icosahedral asymmetric unit showing positions of glycosylation sites Asn 67 and Asn 153. Asn 67 sites of all three identical proteins constituting the asymmetric unit are biologically identical. However, CRD (in cyan) binds to two Asn 67 sites leaving the third Asn 67 residue vacant. Comparison with the density function f_{25} (see Fig. 4, $l=25$, $N=3$) corresponding to the protein density distribution in DENV, shows that two Asn 67 binding sites are situated in the deepest minima of f_{25}. Their positions are related, in addition, by the twofold-symmetry axis. By contrast, the vacant Asn 67 residue is situated in a shallow minimum and its position in the global capsid structure is not equivalent to two other binding sites.

We would like to stress that irreducible icosahedral density function $f_l(\theta,\phi)$, which is the main object of the present work that contains much more physical information than simple positions of proteins centers. The full density distribution generated by $f_l(\theta,\phi)$ is very useful for understanding of biologically important properties such as virus infectivity. Recent advances in virology[25] have shown that infectivity promoted by interaction of cell receptors with virus surface depends not only on biospecific binding properties but also on the capsid proteins distribution. Along this line, the relation can be established between the minima of $f_l(\theta,\phi)$ and binding sites on the capsid surface. One-to-one correspondence of the deepest minima of f_{25} (Fig. 4 $l=25$) and the binding sites for the carbohydrate recognition domains of the dendritic cell receptors on the Dengue virus surface (Fig. 7)[25] can be taken as an illustration of the relation. The detailed correspondence together with a series of other predictions concerning the family of Dengue virus, West Nile virus, and Yellow Fever virus will constitute a separate (more biologically oriented) publication. Constructive character of the proposed density wave approach and the explicit form of the density functions can, in principle, help to understand assembly time evolution. The major obstacle on this way is the fact that the time evolution consists of several successive steps which are not the same for different virus families. A big number of viruses undergo on the first step a procapsid formation with the following maturation process. The latter process involves specific biochemical features which cannot be

modeled by a simple structural or thermodynamical model. Among them we can distinguish large conformational changes of proteins[26,27] during time evolution, protein cleavage,[28] new covalent bond formation,[29] elastic instability of the global shape[9] etc. In the present work, we focused mainly on the first step of the complex assembly process which can be understood in the frame of physical theories. Further studies are needed to clarify its successive steps.

APPENDIX: IRREDUCIBLE ICOSAHEDRAL FUNCTIONS ON THE SPHERICAL SURFACE

Following analysis in terms of the theory of invariants allows to establish the selection rules [Eq. (4)]. They limit the wavenumber l of the order parameter driving the icosahedral virus shell crystallization. The analysis starts with general properties of an arbitrary scalar density function defined on a spherical surface and invariant with respect to the icosahedron rotational symmetry group I (which does not contain spatial inversion nor symmetry planes). We fix three Goldstone degrees of freedom related to the orientation of the icosahedral point symmetry group in space. In the present article, we choose the orientation, for which the x, y, and z directions of the Cartesian coordinate system are parallel to the twofold axes of the icosahedron symmetry (another reasonable set was used in Ref. 10, where one of the icosahedron fivefold axes was chosen parallel to the z direction). Nevertheless, the results obtained are independent on the coordinate system orientation.

Any scalar function invariant with respect to the group I can be expanded in formal polynomial series of x, y, and z. All the terms in the series are the monomials of four following basis invariants (full set of generators for the ring of invariant polynomials: they constitute so-called integrity basis[30,31] of the I symmetry group):

$$J_0 = x^2 + y^2 + z^2, \quad J_1 = \prod_{i=1}^{6} (\mathbf{n}_i \mathbf{r}),$$
$$J_2 = \prod_{i=1}^{10} (\mathbf{p}_i \mathbf{r}), \quad J_3 = \prod_{i=1}^{15} (\mathbf{q}_i \mathbf{r}), \tag{9}$$

where $\mathbf{r}=\langle x,y,z\rangle$ is the radius vector and $\mathbf{n}_i$, $\mathbf{p}_i$, and $\mathbf{q}_i$ are the vectors parallel to the icosahedron rotational axes. Namely, to the six fivefold axes, to the ten threefold axes, and to the fifteen twofold axes, respectively. Let us note that the properties of scalar functions invariant with respect to the I group differ greatly from those of functions invariant with respect to the full icosahedral group I_h (which includes all symmetry elements of a regular icosahedron). The point symmetry group I_h is generated by reflections. Therefore, the number of invariants in the integrity basis of this group is equal to the dimension of its vectorial representation, $\dim V(I_h)=3$ and the basis invariants are functionally independent.[30,31] The corresponding integrity basis includes the invariants J_0, J_1, and J_2, but not the J_3 one. Consequently, the expansion in formal series of a function invariant with respect to the I_h group has rather simple form. The situation is different for the functions invariant with respect to the I group. Basis invariants given by Eq. (9) are functionally dependent and form a sizigia,[30,31] an algebraic relation of 30th degree. Indeed, the square of J_3 can be presented as a polynomial function of J_0, J_1, and J_2. This implies several peculiarities in the formal expansion in series, namely the terms containing J_3^2 in any degree can be expressed in function of J_0, J_1, and J_2. Thus, all J_3-containing terms in any series invariant with respect to the group I are linear in J_3. On the second step, we will follow the incidence of the I_h or I symmetry on the irreducible functions defined on a spherical surface.

Let us start our consideration with a more simple case of irreducible functions invariant with respect to the I_h symmetry group. By definition, any irreducible function $F_l(\theta,\phi)$ of this type is a linear combination of spherical harmonics Y_{lm} with the same l number. The I_h symmetry implies that the l number is even. Corresponding harmonics and, consequently, the density function are *homogeneous* functions of the lth degree in $\mathbf{r}=\langle x,y,z\rangle$:

$$F_l(x,y,z) = \sum_{2k+6i+10j=l} A_{k,i,j} J_0^k J_1^i J_2^j. \tag{10}$$

Here, $A_{k,i,j}$ are numerical coefficients, k, i, and j are the degrees of the J_0, J_1, and J_2 invariants, respectively. Homogeneity of function (10) implies that $l=2k+6i+10j$, where k, i, and j are non-negative integers.

On the unit sphere surface ($r=1$), the invariant J_0 becomes constant: $J_0=1$. Consequently, the form of the density function changes. Let us introduce a radial unit vector $\mathbf{e}_r = \mathbf{r}/r = \langle \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\rangle$ depending on angular coordinates θ et ϕ in a standard way. The irreducible scalar function $F_l(\theta,\phi)=F_l(\mathbf{e}_r)$ invariant with respect to the I_h group is then presented on the spherical surface in an inhomogeneous with respect to $\mathbf{e}_r$ components form:

$$F_l(\theta,\phi) = A_{0,0} + A_{1,0}J_1 + A_{0,1}J_2 + A_{2,0}J_1^2 + A_{1,1}J_1J_2 + \ldots \sum_{6i+10j=l} A_{i,j} J_1^i J_2^j. \tag{11}$$

Here, $A_{s,t}$ are numerical coefficients multiplying the terms of $(6s+10t)$-th degree in $\mathbf{e}_r$ components, s and t are the degrees of the J_1 and J_2 invariants, respectively. Terms in Eq. (10) for which $k=0$ remain terms of lth degree in Eq. (11), while those containing J_0 (i.e., those with $k\neq 0$) in Eq. (10) contribute in Eq. (11) to the terms of degree lower than l. Thus, for irreducible function [Eq. (11)] $l=6i+10j$, where i and j are non-negative integers.

The proof of Eq. (4) is based on the same considerations but applied to the irreducible functions invariant with respect to the I symmetry. By taking into account the sizigia between invariants [Eq. (9)], any irreducible scalar function $G_l(x,y,z)$ *antisymmetric* under the coordinates inversion and invariant with respect to the rotational group I can be presented as:

$$G_l(x,y,z) = \sum_{2k+6i+10j+15=l} J_3 (A_{k,i,j} J_0^k J_1^i J_2^j). \tag{12}$$

k, i, and j being the degrees of the J_0, J_1, and J_2 invariants, respectively. Due to homogeneity of function (12), its wave number l verifies the relation $l=2k+6i+10j+15$ where k, i,

and j are non-negative integers. Then, on the spherical surface the irreducible scalar function $G_l(\theta,\phi)=G_l(\mathbf{e}_r)$ has the form:

$$G_l(\theta,\phi)=J_3\left(A_{0,0}+A_{1,0}J_1+A_{0,1}J_2+A_{2,0}J_1^2+A_{1,1}J_1J_2 + \ldots \sum_{6i+10j+15=l} A_{i,j}J_1^iJ_2^j\right) \quad (13)$$

Similar to the relation between the terms in functions (11) and (10) considered above, the terms of l-th degree in Eq. (13) correspond to the terms with $k=0$ in Eq. (12). This gives, finally, that the wave number l for irreducible density functions [Eq. (13)] verifies the relation $l=15+6i+10j$, where i and j are non-negative integers. In Sec. III, it was shown that the CSDW driving the icosahedral assembly of asymmetric proteins [see Eq. (5)] is a linear combination of functions similar to function (13). Then, any critical order parameter of the assembly satisfies Eq. (4).

Two additional remarks should be done on the irreducible icosahedral functions on the spherical surface. First, remark concerns the number of parameters in these functions. Since function (13) is irreducible (and not arbitrary) the numerical coefficients in Eq. (13) are linearly dependent. Using also its orthogonality properties (the irreducible function $G_l(\theta,\phi)$ should be orthogonal to all spherical harmonics Y_{nm} with $n<l$) one can show that function (13) has independent coefficients in lth degree (with respect to $\mathbf{e}_r$ components) only. It is also easy to see that the number of different terms of lth degree in Eq. (13) is equal to the number of non-negative integer solutions (i,j) for the equation $l=15+6i+10j$.

As a second remark, note that the J_3 invariant plays an exceptional role for the antisymmetric irreducible functions. Vanishing of J_3 leads to the vanishing of all the functions of this type [see Eq. (13)]. From Eq. (9), it follows that J_3 turns zero when $\mathbf{q}_i\mathbf{r}=0$, i.e., at the symmetry planes of the icosahedral group. It means that the antisymmetric irreducible functions change their sign not only under the coordinates inversion, but also under reflections in the icosahedral symmetry planes. The intersections of these planes with a spherical surface form the well-known Möbius tessellation of a sphere.[32] Consequently, the J_3 invariant, $\mathbf{q}_i\mathbf{r}=0$ planes and the Möbius tessellation are of primary importance for the protein density distributions in viral capsids. By contrast, the planes $\mathbf{n}_i\mathbf{r}=0$ and $\mathbf{p}_i\mathbf{r}=0$ perpendicular to the fivefold and threefold icosahedral symmetry axes do not change qualitatively the antisymmetric irreducible function (13). Thus, their role for capsid structures is not important.

[1] S. J. Flint, L. W. Enquist, V. R. Racaniello, and A. M. Skalka, *Principles of Virology: Molecular Biology, Pathogenesis, and Control* (ASM, Washington, 2000).

[2] H. Fraenkel-Conrat and R. C. Williams, Proc. Natl. Acad. Sci. U.S.A. **41**, 690 (1955).

[3] J. B. Bancroft, *Advances of Virus Research* (Academic, New York, 1970), Vol. 16, p. 99.

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

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

[6] H. Naitow, J. Tang, M. Canady, R. B. Wickner, and J. E. Johnson, Nat. Struct. Biol. **9**, 725 (2002); I. Rayment, T. S. Baker, D. L. D. Caspar, and W. T. Murakami, Nature (London) **295**, 110 (1982); R. C. Liddington, Y. Yan, J. Moulai, R. Sahli, T. L. Benjamin, and S. C. Harrison, *ibid.* **354**, 278 (1991); R. Kuhn, W. Zhang, M. Rossmann, S. Pletnev, J. Corver, E. Lenches, C. Jones, S. Mukhopadhyay, P. Chipman, and E. Strauss, Cell **108**, 717 (2002); S. Mukhopadhyay, Bong-Suk Kim, P. R. Chipman, M. G. Rossmann, and R. J. Kuhn, Science **302**, 248 (2003).

[7] T. S. Baker, N. H. Olson, and S. D. Fuller, Microbiol. Mol. Biol. Rev. **63**, 862 (1999), and references therein.

[8] A. Zlotnick, Proc. Natl. Acad. Sci. U.S.A. **101**, 15549 (2004).

[9] J. Lidmar, L. Mirny, and D. R. Nelson, Phys. Rev. E **68**, 051910 (2003).

[10] T. T. Nguyen, R. F. Bruinsma, and W. M. Gelbart, Phys. Rev. Lett. **96**, 078102 (2006).

[11] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Phys. Rev. Lett. **90**, 248101 (2003).

[12] V. L. Lorman and S. B. Rochal, Phys. Rev. Lett. **98**, 185502 (2007).

[13] L. D. Landau, Phys. Z. Sowjetunion **11**, 26 (1937); **11**, 545 (1937).

[14] S. Alexander and J. McTague, Phys. Rev. Lett. **41**, 702 (1978).

[15] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).

[16] J. P. Elliot and P. G. Dawber, *Symmetry in Physics* (Macmillan, London, 1979).

[17] A computer program composed to calculate the irreducible icosahedral structures can be found at http://www.lpta.univ-montp2.fr/article.php3?id_article=132

[18] Experimental structures are reproduced using the UCSF Chimera package. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, J. Comput. Chem. **25**, 1605 (2004).

[19] S. C. Harrison, A. J. Olson, C. E. Schutt, F. K. Winkler, and G. Bricogne, Nature (London) **276**, 368 (1978); C. Abad-Zapatero, S. Abdel-Meguid, J. E. Johnson, and Andrew G. W. Leslie, I. Rayment, M. G. Rossmann, D. Suck and T. Tsukihara, *ibid.* **286**, 33 (1980); A. J. Fisher and J. E. Johnson, *ibid.* **361**, 176 (1993); J. A. Speir, S. Munshi, G. Wang, T. S. Baker, and J. E. Johnson, Structure (London) **3**, 63 (1995); C. Qu, L. Liljas, N. Opalka, C. Brugidou, M. Yeager, R. Beachy, C. Fauquet, J. Johnson, and T. Lin, *ibid.* **8**, 1095 (2000).

[20] J. E. Johnson and J. A. Speir, J. Mol. Biol. **269**, 665 (1997), and references therein.

[21] A. Zlotnick, J. Mol. Biol. **241**, 59 (1994).

[22] A. F. Devonshire Philos. Mag. **40**, 1040 (1949); Adv. Phys. **3**, 85 (1954).

[23] Yu. A. Izyumov and V. N. Syromyatnikov, *Phase Transitions and Crystal Symmetry* (Kluwer, Amsterdam, 1990).

[24] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 2000).
[25] E. Pokidysheva, Y. Zhang, A. Battisti, C. Bator-Kelly, P. Chipman, C. Xiao, G. Gregorio, W. Hendrickson, R. Kuhn, and M. Rossmann, Cell **124**, 485 (2006).
[26] M. A. Canady, M. Tihova, T. N. Hanzlik, J. E. Johnson, and M. Yeager, J. Mol. Biol. **299**, 573 (2000).
[27] W. Jiang, Z. Li, Z. Zhang, M. L. Baker, P. E. Prevelige, and W. Chiu, Nat. Struct. Biol. **10**, 131 (2003).
[28] D. J. Taylor, N. K. Krishna, M. A. Canady, A. Schneemann, and J. E. Johnson, J. Virol. **76**, 9972 (2002).
[29] Ch. Helgstrand, W. R. Wikoff, R. L. Duda, R. W. Hendrix, J. E. Johnson, and L. Liljas, J. Mol. Biol. **334**, 885 (2003).
[30] T. A. Springer, *Invariant Theory*, Lecture Notes in Mathematics Vol. 585 (Springer, Berlin, 1977).
[31] A. J. M. Spencer, *Continuum Physics*, edited by C. Eringen (Academic, New York 1971).
[32] H. S. M. Coxeter, *Regular Polytopes*, 3rd ed. (Dover, New York, 1973).