Soft Matter

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

A minimal coarse-grained model for the low-frequency normal mode analysis of icosahedral viral capsids†

M. Martín-Bravo,[a] J. M. Gomez Llorente,[a] and J. Hernández-Rojas[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

The main goal of this work is the design of a coarse-grained theoretical model of minimal resolution for the study of the physical properties of icosahedral virus capsids within the linear-response regime. In this model the capsid is represented as an interacting many-body system whose composing elements are capsid subunits (capsomers), which are treated as three-dimensional rigid bodies. The total interaction potential energy is written as a sum of pairwise capsomer-capsomer interactions. Based on previous work [Gomez Llorente *et al.*, *Soft Matter*, 2014, 10, 3560], a minimal and complete anisotropic binary interaction that includes a full Hessian matrix of independent force constants is proposed. In this interaction model, capsomers have rotational symmetry around an axis of order $n > 2$. The full coarse-grained model is applied to analyse the low-frequency normal-mode spectrum of icosahedral $T = 1$ capsids. The model performance is evaluated by fitting its predicted spectrum to the full-atom results for the Satellite Tobacco Necrosis Virus (STNV) capsid [Dykeman and Sankey, *Phys. Rev. Lett.*, 2008, 100, 028101]. Two capsomer choices that are compatible with the capsid icosahedral symmetry are checked, namely pentamers ($n = 5$) and trimers ($n = 3$). Both subunit types provide fair fits, from which the magnitude of the coarse-grained force constants for a real virus is obtained. The model is able to uncover latent instabilities whose analysis is fully consistent with the current knowledge about the STNV capsid, which does not to self-assemble in the absence of RNA and is thermally unstable. The straightforward generalisability of the model beyond the linear regime and its completeness make it a promising tool to theoretically interpret many experimental data such as those provided by the atomic force microscopy or even to better understand processes far from equilibrium such as the capsid self-assembly.

1 Introduction

Viruses are the simplest biological systems that are complex enough at the molecular scale to exhibit intricate behaviours such as self-assembly and self-replication. These systems are characterised by the presence of a capsid, i.e. a protein shell that encloses and protects the genetic material (RNA or DNA), and plays an essential role in the delivery of this material in the host cell. This capsid has a well defined geometrical structure. It is formed by copies of a small number (just one in the simplest cases) of coat proteins.

One of the most remarkable features of viruses, which is relevant in their life cycle, is their capability to self-assemble forming regular capsids with well defined size and geometry. This process can be also observed in vitro without the genetic material [1–3]. Noteworthy, a given capsid architecture can be shared by many widely different viruses [4], which allows their collection as a particular family or class. All these features evince the existence of universal effective interactions acting between the basic building blocks of these biological systems [5].

There is extensive experimental and theoretical evidence of the existence of capsid subunits formed by a reduced number of coat proteins (protein fragments in some cases). Depending of this number and/or the compound geometry, one speaks of hexamers or hexons, pentamers or pentons, trimers and dimers. The generic name of capsomers is commonly used for these subunits. The particular choice of these subunits depends on the process being analysed. For instance, experimental results show that the kinetics of the self-assembly process of many viral capsids (in particular for the larger ones) is hierarchical and involve two time scales: first, proteins join into capsomers that later self-assemble into the final capsid structure [6–15]. Theoretical studies support this hierarchical mechanism [16]. In the final capsid structure, which may have required an additional maturation process after its self-assembly [17], capsid subunits can also be identified by alluding to specific criteria. One of these criteria is rigidity and identifies the subunits as those protein or protein-fragment compounds with high relative rigidity [18]. The theoretical implications of the exis-

[a]*Departamento de Física and IUdEA, Universidad de La Laguna, 38200 Tenerife, Spain.*
† Electronic Supplementary Information (ESI) available: Movies with the motion of some of the coarse-grained normal modes obtained for STNV capsid. See DOI: 10.1039/cXsm00000x/

tence of these kinetic or structural subunits would be consistent with the many physical properties shared by the viruses in a class.

The theoretical understanding of the more universal physical properties found in viruses requires to relinquish the details of the molecular scale that the computationally involved all-atom approaches [19–21] provide, and develop coarse-grained models that incorporate the relevant features [5,22–54]. Some of these models describe the capsid as a continuous medium in order to account for its mechanical and elastic properties [51–54]. Others use spring networks for this purpose [21,48–51,55]. On the other hand, most of the discrete models represent the capsid as a set of a reduced number of rigid subunits of one or at most two different classes, and propose a binary interaction potential energy between them [27,32–35,37,37–47,56–58]. These models provide potential energy landscapes [59] whose global minima reproduce the observed capsid structures. Moreover, theoretical analyses of this landscape and its connection with the kinetics of the self-assembly process can also be carried out. Besides, this work will show that an adequate model of this type can also be used to tackle the low-frequency normal-mode analysis of the capsid.

About half of all viruses known so far have capsids with icosahedral symmetry [60]. Since proteins are asymmetric units, capsids can not possess inversion centre and thus icosahedral capsids must have the symmetry of the I point group (if an inversion centre were possible one would obtain instead the I_h point group). For these viruses, the simple geometrical construction model introduced by Kaspar and Klug [61] is the basis of their structural classification. These authors build a geometrical model of the capsid by folding a 2D hexagonal lattice whose equivalent positions are given by a pair non-negative integers (h and k). The resulting icosahedral shell is made of 20 equilateral triangular faces whose area, expressed in units of the smallest icosahedron that can be built, is given by the triangulation number $T = h^2 + k^2 + kh$, with h and k integer numbers. Then the possible values for T are discredited, i.e. $T = 1,3,4,7,...$, and are used as structural classification of all but a few icosahedral viruses.

The simplest icosahedral viruses have $T = 1$. In this case, the capsid is formed by 60 repetitions of a single protein and the possible polymeric subunits, from those mentioned before, can be pentamers, trimers and dimers. These subunits have a rotation symmetry axis which is five-fold for pentamers (C_5 symmetry point group), threefold for trimers (C_3 group), and twofold for dimers (C_2 group). In the icosahedral $T = 1$ capsid, these rotation symmetry axes must coincide with the corresponding symmetry axes of the icosahedral symmetry group. Therefore, pentamers are located at the icosahedron vertices, trimers on the faces and dimers at the edges. Hence, if one identifies capsomers with polyhedron faces, then 20 equivalent trimers will form a regular icosahedron, while 12 equivalent pentamers will produce its dual polyhedron, i.e. a regular dodecahedron. These are the two convex regular polyhedra with icosahedral symmetry.

The elastic and mechanical properties of the viral capsids have been the focus of recent research [55,62,63]. In the last years with these properties in mind, proposals to attack the viruses have appeared in the scientific literature. This action would be carried out by means of mechanical procedures such as ultrasounds [64,65], or with electromagnetic radiation by using impulsive stimulated Raman dispersion [66–69].

Important structural and dynamical properties of the capsid dynamics that are involved in many processes like the previous ones are determined by the frequency spectrum and motions of the capsid normal modes. The normal mode analysis (NMA) provides a systematic methodology to investigate all these aspects. The use of NMA [70] as a tool to study the dynamics of proteins under equilibrium conditions and its relationship with structure started more than thirty years ago [71–73], and has been fruitfully extended, more recently, to many other biological systems such as viruses [20,21,48–51,74–77]. The most genuine dynamics is associated with the low-frequency modes, since these are related to the less rigid and possibly more labile motions of the capsid. Therefore, the use of coarse-grained interaction models would be totally justified to study this low-frequency spectrum.

The general procedure to accomplish such a study is the main goal of this work, which will be focused, for illustration, on the $T = 1$ capsid class. In order to attain this objective the coarse-grained model presented in a previous work [56], which describes the subunits as rigid bodies with axial symmetry, will be conveniently generalised. Additional terms will be required to account for the particular rotational symmetry of the capsomeric subunit and other necessary effects. In Section 2, the coarse-grained interaction energy model and the generalisation needed to account for the normal-mode dynamics of $T = 1$ capsids are introduced. Here, expressions for the force constants of the binary interaction as analytical functions of the model parameters are derived and the steps for the calculation of the capsid normal modes are given. Section 3 presents the normal-mode frequency spectra obtained by fitting the all-atom numerical data by Dykeman and Sankey for the Satellite Tobacco Necrosis Virus (STNV) [75,78]. The dependence of the normal-mode frequencies on the values of the force constants and on the type of subunits used (either trimers or pentamers) will be analysed. The real physical scale of the force constants will be provided. The model shall be shown to uncover latent structural instabilities fully consistent with the current knowledge about this capsid, which does not to self-assemble in the absence of RNA and is thermally unstable. Finally, Section 4 summarises the main conclusions of this work.

2 Model and methods

As has been mentioned in the Introduction, the coarse-grained interaction model chosen for this wok assumes rigid subunits and therefore six degrees of freedom are required to fix the position and orientation of each one of them. For $T = 1$ viruses, which are the object of this work, all subunits are identical and here two possible choices shall be made, namely trimers and pentamers.

2.1 The interaction energy.

The starting expression for the binary interaction between two subunits is the form presented in a previous work [56], which is an anisotropic interaction that includes the lowest order terms of a multipolar expansion, namely

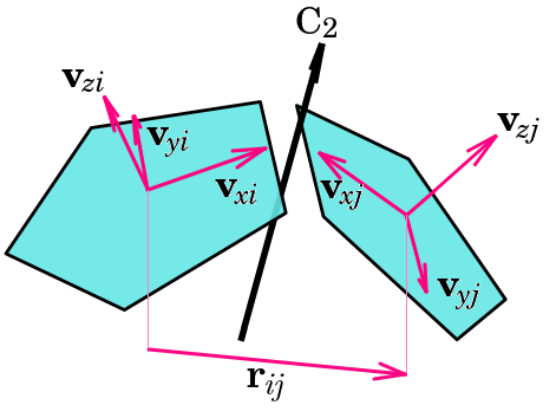

Fig. 1 Equilibrium configuration of a pair of pentamers around a twofold symmetry axis C_2.

$$V_{ij}^{(0)} = p_0 F_0(r_{ij}) + p_1 F_1(r_{ij}) \left[1 - \mathbf{v}_{zi}\cdot\mathbf{v}_{zj} + 2(\mathbf{v}_{zi}\cdot\mathbf{n}_{ij})(\mathbf{v}_{zj}\cdot\mathbf{n}_{ij})\right] + p_2 F_2(r_{ij}) \left[(\mathbf{v}_{zi}\cdot\mathbf{n}_{ij} + \cos\theta_{e_n})^2 + (\mathbf{v}_{zj}\cdot\mathbf{n}_{ij} - \cos\theta_{e_n})^2\right], \quad (1)$$

where $\mathbf{n}_{ij}$ is the unitary vector $\mathbf{n}_{ij} = \mathbf{r}_{ij}/r_{ij}$, $\mathbf{r}_{ij}$ being the intercapsomer position vector from capsomer i to capsomer j, and $\mathbf{v}_{zi}$ is the unitary vector giving the orientation of the z axis of a body-fixed frame on capsomer i. This axis is chosen to coincide with the subunit rotation symmetry axis. Here, p_0, p_1 and p_2 are parameters with dimensions of energy. The required form of the isotropic functions F_0, F_1 and F_2 and the role of the angular parameter θ_{e_n} will be given later.

This potential has been designed to ensure that each of the three terms in Eq. (1) fixes the equilibrium values of a subset of the six internal degrees of freedom of the two-body problem with a minimum number of terms with lowest possible order from the general multipolar expansion (this expansion is always possible for any kind of anisotropic interaction, not just electrostatic ones). The orientational degrees of freedom of both capsomers will be given momentarily taking as reference a laboratory frame with its Z axis pointing along $\mathbf{n}_{ij}$. Then, the body-fixed frame on capsomer i provides its orientation, which is determined by the corresponding three Euler angles θ_i, ϕ_i and χ_i. These angles shall be defined with the ZYZ convention[79], in which case θ_i and ϕ_i can be identified respectively with the polar and azimuthal angles of the spherical coordinates of the vector $\mathbf{v}_{zi}$ in the laboratory frame. Fig. 1 is an illustration of the equilibrium configuration of two capsomers.

In the first term of Eq. (1), $F_0(r_{ij})$ is dimensionless and fully isotropic and should be chosen to fix the equilibrium distance $r_{ij} = r_{e_n}$, where it shall take the value $F_0(r_{e_n}) = -1$. The parameter r_{e_n} determines the system length scale and thus can be used as length unit, i.e. $r_{e_n} = 1$ [the subscript n in a parameter denotes explicitly its dependence on the capsomer type (either $n = 3$ for trimers or $n = 5$ for pentamers)]. The other two terms have isotropic parts that have been chosen respectively to make their contributions vanish at the equilibrium orientational configuration, thus there $V_{ij}^{(0)} = -p_0$, this being the minimum value of this binary interaction energy if p_0 is positive and p_1 and p_2 are positive semi-definite. The value of p_0 can be then taken as the system natural energy unit in which case $p_0 = 1$. The second term includes the dipole-dipole anisotropic contributions and fixes the equilibrium configuration by the conditions $\phi_i - \phi_j = 0$ and $\theta_i + \theta_j = \pi$. The third term incorporates monopole-dipole and monopole-quadrupole interactions and fixes the equilibrium dihedral angle between the two subunits through the parameter θ_{e_n} and the conditions $\theta_j = \theta_{e_n}$ and $\theta_i = \pi - \theta_{e_n}$, which are fully compatible with the previous conditions. As will be shown later, the required value of the parameter θ_{e_n} depends on the type of capsomer. The potential energy $V_{ij}^{(0)}$ does not depend on the two remaining degrees of freedom, χ_i and χ_j, thus subunits in this model have axial symmetry. Both functions $F_1(r_{ij})$ and $F_2(r_{ij})$ in Eq. (1) are dimensionless isotropic functions which should be chosen to satisfy $F_1(r_{e_n}) = F_2(r_{e_n}) = 1$. It will soon become clear that these are the only conditions to impose to these functions within the linear response regime. Of course, beyond this regime the behaviour of the functions $F_0(r)$, $F_1(r)$ and $F_2(r)$ with r should fit additional physical requirements at the relevant length scale. For instance, V_{ij} should conveniently vanish at distances large with respect to the molecular scale.

The study of the capsid normal modes only requires the knowledge of the interaction form for small displacements around the capsid equilibrium configuration, and such form is determined by the corresponding symmetric Hessian matrix of force constants. As will be discussed later, global equilibrium is reached when each neighbouring pair of capsomers in the capsid is at its free equilibrium configuration. Therefore, under these circumstances, the capsid force-constant matrix will be determined by that of the binary interaction. One has as many diagonal force constants as the number of internal degrees of freedom, i.e. six for the binary interaction. As shown in Fig. 1, the dihedral equilibrium configuration for the two capsomers has a twofold symmetry axis perpendicular to the $\mathbf{n}_{ij}$ vector (this axis must coincide with one of the twofold symmetry axis of the icosahedral capsid). As a consequence, the diagonal force constants must be associated with modes that will be either symmetric or antisymmetric with respect to a rotation of π around that axis. One of these force constants, k_r, is connected to the small oscillation of the scaled coordinate r_{ij}/r_{e_n} around its equilibrium value 1 (stretching mode). This is a symmetric mode that is determined by F_0 in $V_{ij}^{(0)}$. The term proportional to p_1 in Eq. (1) determines the force constant, k_ϕ, associated with the displacement of the angular coordinate $\phi_j - \phi_i$, which shall be called torsion ϕ^+ since it is symmetric. One readily obtains for this force constant the expression $k_\phi = p_1 \sin^2\theta_{e_n}$. There are other two force constants related to the small oscillations of θ_i and θ_j (bending modes). The symmetric one corresponds to the displacement of the coordinate $\theta^+ = \theta_j - \theta_i$ while the antisymmetric one is associated with $\theta^- = \theta_i + \theta_j$. The corresponding force constants are $k_{\theta^+} = p_2 \sin^2\theta_{e_n}$ and $k_{\theta^-} = p_1 + p_2 \sin^2\theta_{e_n}$. Note that one has just two parameters to fix the three force constants k_ϕ, k_{θ^+} and k_{θ^-}. Since the interaction potential does not depend on χ_i and χ_j, the corresponding two force constants vanish.

In order to properly account for the normal modes, the most general binary interaction model must incorporate a non-vanishing symmetric Hessian matrix of independent force constants whose dimension coincides with the number of degrees of freedom. For this purpose, one firstly needs, on one hand, an extra parameter to allow for two independent k_θ force constants and, on the other, one has to break the axial symmetry of the capsomers in Eq. (1) by incorporating χ-dependent terms. Of course, these terms shall be different for trimers and pentamers by including the particular rotational symmetry of the subunit. The following additional terms are proposed:

$$V_{ij}^{(1)} = p_3 F_3(r_{ij})\Big\{\mathrm{Re}\big\{[-(\mathbf{v}_{xi}+i\mathbf{v}_{y_i})\cdot\mathbf{n}_{ij}]^n+[(\mathbf{v}_{xj}+i\mathbf{v}_{y_j})\cdot\mathbf{n}_{ij}]^n\big\} + [1-(\mathbf{v}_{zi}\cdot\mathbf{n}_{ij})^2]^{n/2}+[1-(\mathbf{v}_{zj}\cdot\mathbf{n}_{ij})^2]^{n/2}\Big\} + p_4 F_4(r_{ij})\left[(\mathbf{v}_{zi}\cdot\mathbf{n}_{ij}+\cos\theta_{e_n})(\mathbf{v}_{zj}\cdot\mathbf{n}_{ij}-\cos\theta_{e_n})\right] + p_5 F_5(r_{ij})\,\mathrm{Im}\big\{[-(\mathbf{v}_{xi}+i\mathbf{v}_{y_i})\cdot\mathbf{n}_{ij}]^n\big\}\,\mathrm{Im}\big\{[(\mathbf{v}_{xj}+i\mathbf{v}_{y_j})\cdot\mathbf{n}_{ij}]^n\big\}. \quad (2)$$

In this equation, The first and third terms introduce non-vanishing and non-degenerate χ force constants, respectively. Here, $\mathbf{v}_{xi}$ and $\mathbf{v}_{y_i}$ are unitary vectors giving (see Fig. 1), respectively, the orientation of the x and y axes of the body-fixed frame on capsomer i (and the same for j). The factor i in this expression denotes the imaginary number unit. Symbols Re and Im refer to real and imaginary parts respectively. The second term adds an extra parameter to fix the k_θ force constants. The new model parameters with dimension energy are p_3, p_4 and p_5. As before, F_3, F_4 and F_5 are general dimensionless functions satisfying the constraint $F_3(r_{e_n}) = F_4(r_{e_n}) = F_5(r_{e_n}) = 1$. Again, beyond the linear response regime other constraints may have to be imposed. The symbol n, either as a power or as a subscript, denotes the order of the rotational symmetry axis of the corresponding subunit, i.e. $n = 3$ for trimers and $n = 5$ for pentamers. None of terms in Eq. (2) alter the previous equilibrium values for θ_i, θ_j and $\phi_i - \phi_j$. Besides, the last two terms provide n equivalent equilibrium values for χ_i and χ_j, therefore breaking the axial symmetry of the capsomers, but keeping the twofold rotational symmetry of the equilibrium configuration. The design of $V_{ij}^{(1)}$ follows that of $V_{ij}^{(0)}$, and like the anisotropic terms proportional to F_1 and F_2 in Eq. (1), all the terms in Eq. (2) vanish at the equilibrium configuration.

The sum of these two contributions, $V_{ij}^{(0)} + V_{ij}^{(1)}$, determines completely the six diagonal force constants, in the chosen coordinates, as linear functions of the six model parameters $p_i, i = 0,...5$. As already known, the stretching force constant k_r and the torsion force constant k_ϕ are determined by $V_{ij}^{(0)}$. The bending force constants k_{θ^+} and k_{θ^-} are determined by both $V_{ij}^{(0)}$ and $V_{ij}^{(1)}$. Finally, the torsional force constants k_{χ^+} and k_{χ^-} are fixed by $V_{ij}^{(1)}$. In order to fix k_r one needs the explicit functional form of $F_0(r)$. Actually, only its second derivative at the equilibrium distance $r = r_{e_n}$ is required. However, an expression shown in previous works to be very convenient for global optimisation and kinetical studies shall be used here, namely, the generalised Lennard-Jones potential in dimensionless form

$$F_0(r) = \left(\frac{r_{e_n}}{r}\right)^{2m} - 2\left(\frac{r_{e_n}}{r}\right)^{m}, \quad (3)$$

where the power parameter m should be chosen to fix the length range scale of the interaction. This form satisfies the required constraint $F_0(r_{e_n}) = -1$. Whence the stretching force constant $k_r = 2p_0m^2$ is obtained. For global optimisation tasks, a convenient and simple explicit form shall also be given to the other $F_i(r)$ functions, namely

$$F_i(r) = \left(\frac{r_{e_n}}{r}\right)^{m}; \quad i = 1,...5, \quad (4)$$

with the same m parameter. Note that they indeed satisfy the imposed constraint $F_i(r_{e_n}) = 1$. In Table 1, the expressions for the binary interaction diagonal force constants as functions of the model parameters together with the corresponding dimensionless coordinates and their equilibrium values are collected. There, s_n is the parameter $s_n = \sin\theta_{e_n}$.

Table 1 Dimensionless coordinates (first column) whose displacements from their equilibrium values (second column) provide the corresponding diagonal force constants (with dimensions of energy) as functions of the model parameters (third column, with $s_n = \sin\theta_{e_n}$). The $+,-$ signs give the symmetry (symmetric or antisymmetric respectively) of the mode with respect to a π rotation around the twofold symmetry axis of the equilibrium configuration.

Coordinates	Equilibrium	Force constants
$r^+ = r_{ij}/r_{e_n}$	1	$k_r = 2p_0m^2$
$\phi^+ = \phi_j - \phi_i$	0	$k_\phi = p_1 s_n^2$
$\theta^+ = \theta_j - \theta_i$	$2\theta_{e_n} - \pi$	$k_{\theta^+} = (p_2 - \frac{1}{2}p_4)\,s_n^2$
$\chi^+ = \chi_j + \chi_i$	$\pi \pm \frac{2\pi l}{n}$; $l \in Z$	$k_{\chi^+} = \frac{1}{2}n^2\,(p_3 + p_5 s_n^n)\,s_n^n$
$\theta^- = \theta_j + \theta_i$	π	$k_{\theta^-} = p_1 + (p_2 + \frac{1}{2}p_4)s_n^2$
$\chi^- = \chi_j - \chi_i$	$-\pi \pm \frac{2\pi l}{n}$; $l \in Z$	$k_{\chi^-} = \frac{1}{2}n^2\,(p_3 - p_5 s_n^n)\,s_n^n$

The form provided so far for the subunit binary interaction has vanishing non-diagonal force constants for the chosen coordinates. A complete general form requires additionally free values for these coupling elements. In order to achieve completeness the following terms are introduced:

$$V_{ij}^{(r)} = \frac{r_{ij}}{r_{e_n}} - 1,$$

$$V_{ij}^{(\phi)} = (\mathbf{v}_{zi}\times\mathbf{v}_{zj})\cdot\mathbf{n}_{ij},$$

$$V_{ij}^{(\theta^+)} = (\mathbf{v}_{zi}-\mathbf{v}_{zj})\cdot\mathbf{n}_{ij} + 2\cos\theta_{e_n},$$

$$V_{ij}^{(\theta^-)} = (\mathbf{v}_{zi}+\mathbf{v}_{zj})\cdot\mathbf{n}_{ij},$$

$$V_{ij}^{(\chi^+)} = \mathrm{Im}\big\{[-(\mathbf{v}_{xi}+i\mathbf{v}_{y_i})\cdot\mathbf{n}_{ij}]^n + [(\mathbf{v}_{xj}+i\mathbf{v}_{y_j})\cdot\mathbf{n}_{ij}]^n\big\},$$

$$V_{ij}^{(\chi^-)} = \mathrm{Im}\big\{[-(\mathbf{v}_{xi}+i\mathbf{v}_{y_i})\cdot\mathbf{n}_{ij}]^n - [(\mathbf{v}_{xj}+i\mathbf{v}_{y_j})\cdot\mathbf{n}_{ij}]^n\big\}.$$

Around the binary equilibrium configuration the leading contribution of each of these terms is linear in the corresponding internal coordinate. Therefore, the proposed additional contribution

to the subunit binary interaction is

$$V_{ij}^{(2)} = p_6 F_6(r_{ij}) V_{ij}^{(r)} V_{ij}^{(\phi)} + p_7 F_7(r_{ij}) V_{ij}^{(r)} V_{ij}^{(\theta^+)} + p_8 F_8(r_{ij}) V_{ij}^{(r)} V_{ij}^{(\chi^+)} + p_9 F_9(r_{ij}) V_{ij}^{(\phi)} V_{ij}^{(\theta^+)} + p_{10} F_{10}(r_{ij}) V_{ij}^{(\phi)} V_{ij}^{(\chi^+)} + p_{11} F_{11}(r_{ij}) V_{ij}^{(\theta^+)} V_{ij}^{(\chi^+)} + p_{12} F_{12}(r_{ij}) V_{ij}^{(\theta^-)} V_{ij}^{(\chi^-)}, \quad (5)$$

which, as required, only includes couplings between modes with the same symmetry. As before, $F_i, i = 6,..12$, are general dimensionless functions satisfying the constraint $F_i(r_{e_n}) = 1$, that have been given the explicit form in Eq. (4). Six additional parameters ($p_i, i = 6,..12$) are involved, each one fixing a corresponding non-diagonal force constant. Hence, the complete binary interaction is

$$V_{ij} = V_{ij}^{(0)} + V_{ij}^{(1)} + V_{ij}^{(2)}, \quad (6)$$

and the elements of the force-constant matrix (in energy units) as linear functions of the model parameters are presented in Tables 2 (for the symmetric modes) and 3 (for the antisymmetric modes).

Table 2 Force constant matrix elements (in energy units) for the binary interaction as functions of the model parameters for the symmetric internal coordinates (S. C.)

S. C.	r^+	ϕ^+	θ^+	χ^+
r^+	k_r	$k_{r\phi} = p_6 s_n^2$	$k_{r\theta^+} = p_7 s_n$	$k_{r\chi^+} = n p_8 s_n^n$
ϕ^+	$k_{r\phi}$	k_ϕ	$k_{\phi\theta^+} = p_9 s_n^3$	$k_{\phi\chi^+} = n p_{10} s_n^{n+2}$
θ^+	$k_{r\theta^+}$	$k_{\phi\theta^+}$	k_{θ^+}	$k_{\theta^+\chi^+} = n p_{11} s_n^{n+1}$
χ^+	$k_{r\chi^+}$	$k_{\phi\chi^+}$	$k_{\theta^+\chi^+}$	k_{χ^+}

Table 3 Force constant matrix elements (in energy units) for the binary interaction as functions of the model parameters for the antisymmetric internal coordinates (A. C.)

A. C.	θ^-	χ^-
θ^-	k_{θ^-}	$k_{\theta^-\chi^-} = n p_{12} s_n^{n+1}$
χ^-	$k_{\theta^-\chi^-}$	k_{χ^-}

Within the linear regime, the final expression for V_{ij} given in Eq. (6) is actually the most complete form that can be written with the minimum number of the lowest-order multipole-multipole terms, for the binary interaction between two rigid bodies satisfying the geometrical constraints imposed by the capsomer and capsid symmetries. Although not required for this work, the binary interaction model in Eq. (6) could also be used for hexamers by taking $n = 6$. Pentamers and hexamers are typical subunits appearing together in icosahedral virus capsids with larger triangulation numbers ($T \geq 3$). Then apart from the pentamer-pentamer and hexamer-hexamer interaction forms already given here, the pentamer-hexamer interaction would be also required to model these systems. This can be straightforwardly derived by taking the common terms and combining the different ones from the other two. Of course, parameters will be in general different for each of these three interaction forms.

From the binary potential, the total potential interaction energy for N identical capsomers (here $N = 60/n$) is readily obtained.

$$V = \sum_{i,j>i}^{N} V_{ij}. \quad (7)$$

The intercapsomer equilibrium distance ($r_{e_n} \sim 80$ Å for STNV) is much larger than the molecular interaction range (~ 1 Å). This requires m-parameter values $m \gtrsim 50$ in the isotropic term given in (3). Therefore and as the calculations presented in the next Section confirm, each subunit in the capsid equilibrium structure will interact only with its closest neighbours (contact interactions) and the capsid binding energy, E_b, will be the sum of these contact interactions. In each $T = 1$ geometrical structure that can be assembled either from pentamers (a regular dodecahedron, as shown in the Introduction) or from trimers (a regular icosahedron) ones has 30 equivalent contacts. If each of these contacts is at the equilibrium configuration of its binary interaction, then the $T = 1$ capsid binding energy will take the value $E_b = 30 p_0$ and if such equilibrium configuration corresponds to a stable stationary point of V_{ij} (i.e. a minimum), this value of E_b will be a global maximum (i.e. global minimum of V). When this happens, one generally speaks of a many-body equilibrium configuration with zero frustration[56]. As will be shown in the next Section, where we fit the model to the STNV capsid, it can happen that while V has a minimum at the capsid icosahedral configuration, the corresponding binary equilibrium configuration is not a minimum of V_{ij} but an unstable stationary point (a saddle point). In that case the capsid energy minimum $V = -30 p_0$ may not be the global minimum and other lower energy structures that break the original architecture may exist.

In order to manage frustration, the relevant parameter in the model is θ_{e_n}, which is a geometrical parameter that fixes the dihedral angle between the two neighbouring capsomers. Of course, the design of a vanishing frustration requires that this dihedral angle coincide with the dihedral angle between neighbouring faces of the corresponding regular polyhedron, either a dodecahedron for pentamers or an icosahedron for trimers. Therefore, these target geometries are going to determine the value of θ_{e_n}, namely $\theta_{e_3} = \frac{1}{2}\arccos(\frac{\sqrt{5}}{3})$ for trimers and $\theta_{e_5} = \frac{1}{2}\arctan(2)$ for pentamers. These parameters will be fixed to their respective values in all the calculations performed in this work. If parameter m is chosen large enough to fit the real interaction length, the other model parameters, which have a direct effect on the force constants given in Table 1 and thus on the capsid elastic properties, hardly affect the capsid final geometrical structure. The parameter p_0 fixes the energy scale and r_{e_n} the length scale of the system. If the remaining parameters satisfy the previously established constraints and provide a stable binary equilibrium structure, the capsid icosahedral equilibrium geometry will remain unchanged. Furthermore, one can find that these regular structures are indeed global energy minima of the corresponding capsid potential energy surface, with the expected maximum binding energy $E_b = 30 p_0$. In order to find these global minima, the Basin-Hopping global optimisation technique[80,81] was employed. Illustrations of these structures are presented in Fig.2. Although the only geometrical constraint for the capsomers em-

ployed in this work is their rotational symmetry, in this figure a regular pentagon has been chosen for pentamers and a regular triangle for trimers. Note however that while the first two terms ($V_{ij}^{(0)}$ and $V_{ij}^{(1)}$) of the binary interaction induce in the capsid the symmetry of the group I_h, the third term $V_{ij}^{(2)}$ destroys the inversion centre and reduce the capsid symmetry to that of the group I, as happens in real capsids. Precisely, the terms in $V_{ij}^{(2)}$ responsible for this symmetry breaking are those proportional to p_6, p_8, p_9, p_{11} and p_{12}, which change sign under inversion.

2.2 The determination of the normal modes.

The methodology to apply this coarse-grained model to the study of the normal modes of icosahedral $T = 1$ capsids will be worked out in the remaining of this Section. For that purpose, the form of the many-body kinetic energy is required. This kinetic energy will have contributions, coming from the translational motion of the centre of mass of each of the capsomers and from their rotational motion as rigid bodies. In order to describe this last motion the reference frame fixed on each capsomer (xyz) with the three axes chosen along the inertia principal ones will be used. Then, the kinetic energy for N equivalent subunits is

$$K = \frac{1}{2}M_n \sum_{i=1}^{N} \dot{\mathbf{r}}_i{}^2 + \frac{1}{2}\sum_{i=1}^{N}(I_x w_{xi}^2 + I_y w_{yi}^2 + I_z w_{zi}^2), \tag{8}$$

where M_n is the capsomer mass, I_x, I_y and I_z its principal moments of inertia and w_{xi} w_{yi} and w_{zi} the components of the angular frequency on the axes of the capsomer-fixed frame. Since, as already established, the subunits are symmetric tops, one has $I_x = I_y$. Here M_n is taken as the system natural mass unit.

The normal-mode analysis requires the $6N \times 6N$ Hessian matrix, $\mathbf{F}$, of second derivatives of the interaction with respect to the $6N$ generalised coordinates, evaluated at the capsid equilibrium configuration. With the displacements of these coordinates from their equilibrium values one defines the $6N$-dimensional column vector $\mathbf{R}$. For each capsomer, this vector includes the three displacements of its centre of mass in a laboratory-fixed frame and the three angular displacements around its respective principal axes of inertia. By expanding the potential V around the equilibrium capsid configuration up to quadratic terms one obtains $V = V_e + \frac{1}{2}\tilde{\mathbf{R}}\mathbf{F}\mathbf{R}$, where $V_e = -30p_0$ and the tilde on a symbol means its transpose. In these displacement coordinates the quadratic form of the kinetic energy in Eq. (6) has an associated $6N \times 6N$ diagonal matrix $\mathbf{G}^{-1}$ with the capsomer masses and principal moments of inertia, i.e. $K = \dot{\tilde{\mathbf{R}}}\mathbf{G}^{-1}\dot{\mathbf{R}}$. The normal mode coordinates, $\mathbf{Q}$, are related to the displacement coordinates through the linear transformation $\mathbf{R} = \mathbf{L}\mathbf{Q}$. The matrix $\mathbf{L}$ is obtained by solving the secular problem given by the two matrix equations $\tilde{\mathbf{L}}\mathbf{F}\mathbf{L} = \mathbf{\Lambda}$ and $\tilde{\mathbf{L}}\mathbf{G}^{-1}\mathbf{L} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{\Lambda}$ is the diagonal matrix of the squares of the normal mode frequencies, i.e. $\Lambda_{ii} = \omega_i^2$. This procedure is known as Wilson GF-secular scheme [82]. This is a general scheme to compute the normal modes of an isolated many body system in terms of a complete set of displacement coordinates which is independent of the particular choice of such coordinates and therefore invariant under general (not just linear) coordinate transformations [83]. The changes in the capsomer orientation have been given as rotations around the three axes of the frame fixed in each capsomer, which have been chosen along the three principal axes of inertia. As shown in Eq. (8), in this particular coordinates the rotational kinetic energy is diagonal. Besides, these orientational displacements can be straightforwardly written in term of quaternions [84–86]. This nice property and the other advantages of quaternions [87,88] have motivated the choice of these coordinates to represent the orientational degrees of freedom in all the computations perfomed in this work.

The normal modes of this problem must have the symmetry of the corresponding irreducible representations of the capsid I point group. The number of modes of each symmetry species can be obtained by reducing the representation, Γ_n, whose basis is the set of $6N$ displacements $\mathbf{R}$. After eliminating the species corresponding to the three translations and three rotations of the whole capsid one obtains for pentamers ($n = 5$) and trimers ($n = 3$), respectively,

$$\Gamma_5 = 2A \oplus 4T_1 \oplus 2T_2 \oplus 4G \oplus 6H, \tag{9}$$

$$\Gamma_3 = 2A \oplus 4T_1 \oplus 6T_2 \oplus 8G \oplus 10H, \tag{10}$$

where the A symmetry species are not degenerate, the T_1 and T_2 species are triply degenerate, and the G and H species are, respectively, quadruply and quintuply degenerate. By using symmetry-adapted coordinates, the GF-secular problem can be decoupled in smaller ones, one for each symmetry species. The number of normal modes is $6N - 6$, hence this number is larger for trimers than for pentamers.

Both types of capsomers give two totally symmetric A modes. Their frequencies can be obtained analytically. The first component of these modes corresponds to the breathing of the capsid in which all capsomers oscillate radially in phase. The kinetic energy of this motion is then $K_{\text{breathe}} = \frac{1}{2}NM_n(\dot{\Delta R})^2$, where ΔR is the radial displacement. The other component of the A modes corresponds to the synchronised in-phase rotation of all capsomers around their symmetry axis. Its kinetic energy is $K_{\text{rot}} = \frac{1}{2}NI_z(\dot{\Delta\chi})^2$ where $\Delta\chi$ is the angular change. Hence, the total kinetic energy is $K_{\text{A}} = \frac{1}{2}NM_n(R\dot{\Delta r^+})^2 + \frac{1}{8}NI_z(\dot{\Delta\chi^+})^2$. The three force constants involved in these two motions are k_r, k_{χ^+} and $k_{r\chi^+}$ and the leading change in the potential energy is $V_{\text{A}} = \frac{1}{2}C\left[k_{\chi^+}(\Delta\chi^+)^2 + k_r(\Delta r^+)^2 + 2k_{r\chi^+}\Delta r^+\Delta\chi^+\right]$, where C is the number of contacts or bonds (i.e. $C = 30$) and R the radius of the sphere containing the capsomer centres of mass at equilibrium, i.e. the radius of the inscribed sphere of the corresponding equilibrium polyhedron. From these two equations the following two frequencies are obtained (it has been assumed here that $\omega_{\text{b}} < \omega_{\text{r}}$)

$$\omega_{\text{b}} = \sqrt{\frac{15(S-D)}{I_z N M_n}}, \tag{11}$$

$$\omega_{\text{r}} = \sqrt{\frac{15(S+D)}{I_z N M_n}}, \tag{12}$$

where

$$S = 4M_n k_{\chi^+} + I_z k_r / R^2$$

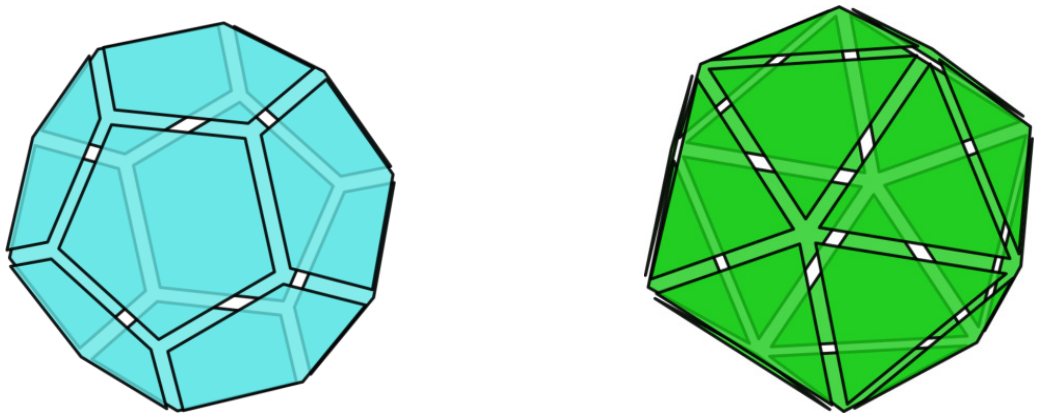

Fig. 2 Equilibrium configuration of $T = 1$ capsids for pentamers (left) and trimers (right).

and

$$D = \sqrt{(4M_n k_{\chi^+} - I_z k_r/R^2)^2 + 16 M_n I_z (k_{r\chi^+}/R)^2}$$

In these equations NM_n is the total capsid mass M_c. The values of R are $R = \frac{1}{2} r_e (1+\sqrt{5}) \sin(2\pi/3)$ for trimers and $R = r_e \sin(2\pi/5)$ for pentamers. When $k_{r\chi^+} = 0$, $\omega_b = \sqrt{30 k_r/(M_c R^2)}$ corresponds to the frequency of the breathing mode while $\omega_r = 2\sqrt{30 k_{\chi^+}/(N I_z)}$ gives the frequency of the rotation mode. All these are relevant equations since they directly relate three of the force constants of the model with the frequencies of the totally symmetric normal-modes.

The coarse-grained model introduced in this work takes rigid subunits as the interacting bodies in the capsid. These capsomers, however, are not completely rigid but have internal motions, which, for an isolated capsomer can be expressed in terms of its own set of normal modes. These modes may involve more or less collective motions or even very localised ones. Since the chosen capsomers have an n-fold rotation symmetry axis (either $n = 3$ or $n = 5$), their normal modes must have the symmetries of the irreducible representations of the corresponding rotation symmetry point group (either C_3 or C_5). Both groups have an irreducible non-degenerate and totally symmetric representation A. Besides, while the group C_3 has a doubly degenerate additional representation E, the group C_5 has two of this kind, E_1 and E_2. If one takes the same particular mode in each capsomer, the set so formed defines a reducible representation of the icosahedral point group I. By reducing this representation one finds the symmetry of the corresponding capsid normal-modes in which the internal capsomer normal-mode participates. This reduction process is summarised in Table 4. The rigid-capsomer approximation would require the frequencies of these internal modes to be much larger than the frequencies of the modes derived with the rigid model. But situations may occur in which one of these internal modes corresponds to a very localised and floppy motion with a low frequency that could appear in the low frequency region of the coarse-grained modes. For the $T = 1$ capsids considered in this work, such a mode shall appear in each of its 60 coat proteins (the capsid asymmetric unit). The set of these 60 equivalent local modes defines a reducible representation for the icosahedral symmetry group, whose reduction provides the symmetry species of the corresponding capsid normal modes. These species are given in the last row of Table 4. If the local mode of this kind in a protein is far enough from the atoms in neighbouring proteins, its coupling to the other equivalent modes and even to other motions will be negligible, and its frequency will be observed practically unchanged in all the corresponding capsid normal modes belonging to the symmetry species just given. If coupling between the local modes can not be neglected, changes in the frequencies may be observed. Furthermore, the local low-frequency mode may also couple to the coarse-grained modes and thus contaminate them and alter their frequencies. These effects will be discussed for the STNV capsid in the next Section.

Table 4 Reduction of the reducible representations (R. R.) made with equivalent copies in all the capsid subunits (either 20 or 12) of each symmetry species of the capsomer symmetry group, C_n (either $n = 3$ or $n = 5$), into the irreducible representations (I. R.) of the icosahedral symmetry group I of the capsid. The last row corresponds to the case of a representation formed with the 60 equivalent repetitions of a mode in the asymmetric unit (C_1 symmetry group). Note that the sums of the I I. R. in each of the three groups (C_3, C_5 and C_1) coincide

Capsomer group	R. R.	I I. R.
C_3	$20A$	$A \oplus T_1 \oplus T_2 \oplus 2G \oplus H$
	$20E$	$2T_1 \oplus 2T_2 \oplus 2G \oplus 4H$
C_5	$12A$	$A \oplus T_1 \oplus T_2 \oplus H$
	$12E_1$	$2T_1 \oplus 2G \oplus 2H$
	$12E_2$	$2T_2 \oplus 2G \oplus 2H$
C_1	$60A$	$A \oplus 3T_1 \oplus 3T_2 \oplus 4G \oplus 5H$

To conclude this Section, an illustration of the effect of the coupling between intracapsomer and intercapsomer modes will be presented. Suppose one has a capsid intermode (i.e. coarse-grained) Q_1 of frequency ω_1, and an intramode Q_2 of frequency ω_2, both belonging to the same symmetry species A, and a potential coupling between them of strength λ. The energy (kinetic + potential) of these two modes in the harmonic approximation is then

$$E = \frac{1}{2}(\dot{Q_1}^2 + \dot{Q_2}^2) + \frac{1}{2}(\omega_1^2 Q_1^2 + \omega_2^2 Q_2^2) + \lambda Q_1 Q_2. \qquad (13)$$

The expected situation is that $\omega_2 >> \omega_1$ and in this case the new

frequencies derived from Eq. (13) are $\omega_1' \sim \omega_1 - \lambda^2/(2\omega_1\omega_2^2 - 2\omega_1^3) \leq \omega_1$ and $\omega_2' \sim \omega_2 + \lambda^2/(2\omega_2^3 - 2\omega_2\omega_1^2) \geq \omega_2$, thus while the intermode low frequency decreases the higher intramode one increases. Besides, some mixing will be observed in the new modes. Of course, this effect disappears here and in general in the limit of a very large intracapsomer mode frequency ω_2 (rigid capsomer limit). Contrarily, the effect becomes more pronounced when the two frequencies get closer and/or the coupling λ increases. When more modes are coupled (these must all belong to the same symmetry species), the situation is more complex but there is always a decrease in frequency for the lower frequency modes and an increase for the higher frequency ones.

3 Results and discussion

In this Section the performance of the designed coarse-grained model in the description of the low-frequency normal modes of an icosahedral $T = 1$ capsid will be analysed. The first step in this analysis is the determination of the model force constants and thus the parameters values required to describe real capsids. One direct way to perform this task would be, for instance, from a fit of the model binary interaction to an atomistic force-field. Here a different and more profitable approach will be carried out by fitting the predictions of the coarse-grained model to a known low-frequency normal-mode spectrum. The data obtained by Dykeman and Sankey from an atomistic approach for the Satellite Tobacco Necrosis Virus (STNV) capsid[75,78] will be used for that purpose. These authors provide a table with the five lowest frequencies for the normal modes in each of the symmetry species of the icosahedral I group (A, T_1, T_2, G and H). The frequencies of these modes are collected in Table 5.

Table 5 Frequencies (cm^{-1}) of the normal modes of the STNV capsid for each of the irreducible representations of the symmetry group I, as provided by Dykeman and Sankey[78]

Symmetry	ω_1	ω_2	ω_3	ω_4	ω_5
A	2.41	4.73	5.29	6.08	7.10
T_1	2.89	3.57	3.60	4.70	4.89
T_2	2.02	3.42	3.43	3.84	4.30
G	2.38	2.60	3.55	3.95	4.14
H	1.95	2.39	2.86	3.57	3.82

Among these normal modes those authors[75,78] identify the first clear intracapsomer collective mode, the so called "puckering" A mode about the five-fold symmetry axes, at a frequency of 6.08 cm^{-1}. Therefore, this value sets, on one hand, an upper bound on the frequencies of the normal modes that can be reproduced with the coarse-grained model and, on the other, the onset of intracapsomer mode-frequencies. A second intracapsomer A mode of very floppy and local character (involving around 140 atoms per protein) was identified at the relatively low frequency of 4.73 cm^{-1}. As discussed in the previous Section such a mode would imply the existence of similarly localised modes for the other symmetry species in the same frequency range, namely three T_1 modes, another three T_2, four G modes and five H (see Table 5). The two T_1 modes with the highest frequencies (4.70 and 4.89 cm^{-1}) in Table 5 lie precisely within such a range. However, the number of atoms participating in these modes is found to be[75] around five times larger than in the related A mode, which would indicate possible coupling with the intercapsomer modes that may appear at these low frequencies. The highest frequency mode with A symmetry in Table 5 at 7.10 cm^{-1} has also a very local character.

Of the remaining two A modes, the lowest frequency one (2.41 cm^{-1}) corresponds predominantly to the collective breathing mode of the capsid (radial expansion-contraction). In this mode the three proteins closer to any threefold rotation axis of the capsid appear to have a dominant in-phase oscillatory displacement along the direction of this axis[75,78], which could be fairly well described as a quasi-rigid translation of the corresponding trimer. The last A mode at 5.29 cm^{-1} (a collective rotation mode) corresponds dominantly to an in-phase oscillatory rotation of the five proteins closer to any of the five-fold rotation axis around it[75,78]. Seen from above a threefold rotation axis this motion appears to be equally well an oscillatory rotation around this axis of its three closer proteins. Definitely, these two collective A modes are in one-to-one correspondence with the two A modes (breathing and rotating modes respectively) provided by the coarse-grained model, whose frequencies, as analytical functions of model parameters, were given in equations (11) and (12). However, there are other modes that should be reproduced and for that a least-squares fit of the normal mode frequencies predicted by the coarse-grained model to a subset of those collected in Table 5 has been implemented by minimising the cost function

$$\mathscr{C} = \frac{1}{2}\sqrt{\frac{1}{f}\sum_{i=1}^{f}\left(\frac{\omega_{ci}^2 - \omega_{ti}^2}{\omega_{ti}^2}\right)^2}, \qquad (14)$$

where f is the number of target frequencies ω_{ti} from Table 5 and ω_{ci} the corresponding coarse-grained predictions. The cost $\mathscr{C}$ is a measure of the average frequency relative error of the fit. Different sets of target frequencies have been considered. In all cases, the two A local-character modes (at 4.73 and 7.10 cm^{-1}), the A "puckering" mode (at 6.08 cm^{-1}) and the two T_1 modes (at 4.70 and 4.89 cm^{-1}) that are suspected of being contaminated by the floppy local motion seen in the 4.73 cm^{-1} A mode, have been discarded.

Table 6 Values of the fixed model parameters for trimers ($n = 3$) and pentamers ($n = 5$)

n	m	θ_{e_n}	r_{e_n} [10^{-9} m]	M_n [10^{-22} Kg]
3	50	$\frac{1}{2}\arccos(\frac{\sqrt{5}}{3})$	6.400	1.089
5	50	$\frac{1}{2}\arctan(2)$	8.587	1.815

If the capsomer mass M_n is fixed as well as the geometry of the capsid by fixing the geometrical model parameters n (to either 3 or 5) and θ_{e_n} (to the corresponding required value given in Section 2), the normal-mode frequencies obtained by solving the GF secular problem shall depend on 15 variables, namely the two different moments of inertia (I_z and $I_x = I_y$) and the 13 force constants of the binary interaction model as given in tables 1-3. The equilibrium distance r_e has also been fixed so as to reproduce the real size of the virus. Finally a fixed integer value has been given to the power parameter m, which determines the interac-

tion length range. The values assigned to these fixed parameters are collected in Table 6. Hence, the number of force constants coincides with the number of free interaction-model parameters p_i, $i = 0, ...12$, and both sets are related by the linear transformation given in tables 1-3; thus one can readily obtain the values of one set from those of the other.

The optimal values of the 15 free parameters will be obtained by minimising $\mathscr{C}$. For this purpose the Basin-Hopping global optimisation technique[80] was also employed. Here results shall be presented for the largest subset of objective frequencies taken from Table 5 which contains the two A-breathing modes at 2.41 cm^{-1} and 5.29 cm^{-1}, the first three T_1 modes (at 2.89, 3.57 and 3.60 cm^{-1}), all five T_2 modes for trimers, and the first two of these for pentamers (for this type of capsomers the model predicts just two T_2 modes), all five G modes for trimers and the first four of these for pentamers (the model predicts just four G modes in this case), and all five H modes. Hence, one has $f = 20$ target frequencies for trimers and $f = 16$ for pentamers.

The optimal parameter sets obtained for this sample are given in Table 7, together with the corresponding values of the cost function. An interesting result is that since the proposed complete binary interaction model leads to a capsid equilibrium configuration with the symmetry of the icosahedral group I, there are indeed two equivalent enantiomeric global minima of the cost function which are transformed into each other by inversion and provide the same normal-mode frequency spectrum. These two global minima share the values of all the force constants except for $k_{r\phi}$, $k_{r\chi^+}$, $k_{\phi\theta^+}$, $k_{\theta^+\chi^+}$, and $k_{\theta^-\chi^-}$, which change sign. In Table 7, just one of these two sets is given. It turns out from the scaling properties of the normal-mode problem that the optimal force constants given in Table 7 (in units of force/length) are indeed independent of the length scale of the problem, r_{e_n}, and only change if the mass scale, given by the capsomer mass M_n, change; once these two scales are fixed, the energy scale, p_0, is determined by the fit and depends on the value given to the interaction-range parameter m.

Finding the global minimum of the cost function was quite fast and only required a few basin-hopping steps (< 10) when starting from a random parameter set. Among these steps, just one competing minimum with slightly higher cost and no essentially different force constants was sampled. These features provide strong significance to the derived optimal model parameters. Fig. 3 compares the frequency spectrum predicted by the optimal coarse-grained models with the target data from Table 5.

The first observation from these results is that the frequencies of the two totally symmetric modes (A modes) agree, within the computational precision, with the analytical predictions given in equations (11) and (12), which were derived assuming contact interactions with zero frustration. Therefore the results confirm the hypotheses made in the design of the interaction model. Of course, the expected capsid binding energy $E_b = 30p_0$ was always obtained.

Both models (trimers and pentamers) are able to reproduce the target frequencies within a 5% error. However while the trimer-model is able to account for more modes as intercapsomer modes, the pentamer-model would assign the observed modes outside its predicted spectrum to intracapsomer modes. One may argue that the rigidity of the subunits is a rough approximation. Indeed the atomistic results by Dykeman and Sankey[78] reveal some degree of mixing between rigid and nonrigid motions. Despite this behaviour, the work presented here demonstrates that coarse-grained models that adequately freeze many of the degrees of freedom of the problem can fit very accurately its low frequency normal-mode spectrum. Hence, for the STNV capsid there are not significant difference between the two fits to make a choice between trimers and pentamers as the optimal subunits.

In general, the normal modes involve the excitation of all internal degrees of freedom of the capsomer pair and thus their frequencies are going to depend on all the force constants and moments of inertia when one changes them in the neighbourhood of their optimal values. There is of course the exception of the A modes which, as seen in equations (11) and (12), only involve the coordinates χ^+ and r and the corresponding force constants k_r, k_{χ^+} and $k_{r\chi^+}$, and the moment of inertia I_z. The other exceptions are found only for the pentamer capsid. By reducing the representations Γ_r and Γ_t whose basis are, respectively, the set of the 12 rotation vectors and 12 displacement vectors along each pentamer body-fixed z axis one obtains the decomposition $\Gamma_r = \Gamma_t = A \oplus T_1 \oplus T_2 \oplus H$. The pentamer capsid has just 2 A modes and 2 T_2 modes. Therefore the T_2 modes share the known property of the A modes, namely, all these modes are linear combinations of just the 24 previous coordinates. As a consequence of this, the T_2 modes can not excite ϕ^+ and θ^+ and do not depend, therefore, on the force constants in which any of these coordinates appears as a subscript, neither on the I_x moment of inertia. Another implication of the previous decompositions is that the pentamer-capsid G modes can not excite any of the 24 previous rotations and translations, hence the frequency of this modes can not depend on I_z. For the trimer capsid the corresponding 20-dimensional representations Γ_r and Γ_t reduce as $\Gamma_r = \Gamma_t = A \oplus T_1 \oplus T_2 \oplus 2G \oplus H$ and symmetry does not impose restrictions on the parameter dependence of the modes with the exception of the two A modes. All these predictions are confirmed by the calculations. Movies with the motion of some of the coarse-grained normal modes obtained for the STNV capsid are provided in the Electronic Supplementary Information.

With the fixed parameters given in Table 6 and the optimal values of the force-constants given in Table 7 one can readily calculate the six normal-mode frequencies of an isolated pair of capsomers at the equilibrium configuration of its binary interaction. The values obtained for trimers and pentamers are collected in table 8. This table presents a remarkable result: one of the mode-frequencies for the pentamer pair has an imaginary value ($1.21i$ cm^{-1}), which means that the equilibrium configuration corresponds to a stationary point which is unstable along this mode coordinate (a transition state). The trimer pair, on the other hand, provides a stable fixed point for the corresponding equilibrium configuration. This is a robust result that also appears from the fit of the simpler binary interaction form $V_{ij} = V_{ij}^{(0)} + V_{ij}^{(1)}$ to the same target frequencies, and even from the fit of $V_{ij} = V_{ij}^{(0)}$ to a lower-frequency target set. In the following, arguments will be

Table 7 Minimum values of the cost functions $\mathscr{C}$ and the corresponding optimal values of the force constants and moments of inertia for trimers ($n=3$) and pentamers ($n=5$). Force-constant units have been converted to N/m by dividing them by $r_{e_n}^2$. In each case, an equivalent enantiomeric set of force constant values exists with only a change in the signs of the five force constants $k_{r\phi}$, $k_{r\chi^+}$, $k_{\phi\theta^+}$, $k_{\theta^+\chi^+}$, and $k_{\theta^-\chi^-}$

n	$\mathscr{C}$	k_r	k_ϕ	k_{θ^+}	k_{χ^+}	$k_{r\phi}$	$k_{r\theta^+}$	$k_{r\chi^+}$	$k_{\phi\theta^+}$	$k_{\phi\chi^+}$	$k_{\theta^+\chi^+}$	k_{θ^-}	k_{χ^-}	$k_{\theta^-\chi^-}$	$I_x=I_y$	I_z
		[N/m]													$[M_n r_{e_n}^2]$	
3	0.053	39.72	6.874	27.54	15.67	3.568	-24.53	10.99	-9.213	9.508	14.29	2.543	9.770	0.671	0.295	0.904
5	0.047	19.86	0.335	0.389	3.115	1.414	-1.747	3.535	-0.153	1.162	0.180	0.657	0.997	0.254	0.057	0.196

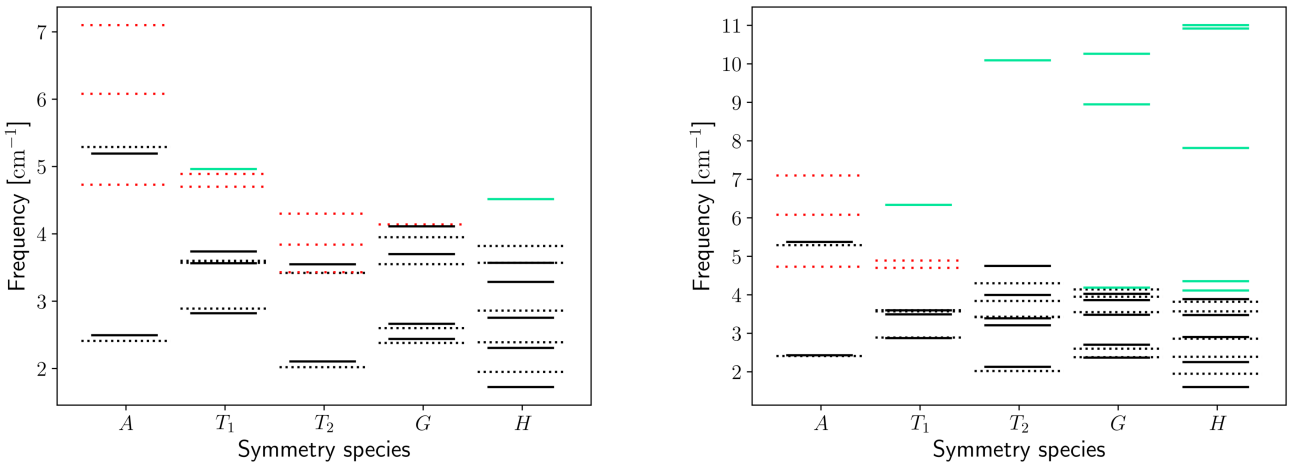

Fig. 3 Capsid frequencies separated by symmetry species. Dotted lines are used for the target frequencies from table 5 and solid lines give the model results. Black colour is used for the frequencies included in the cost function $\mathscr{C}$ and red for the frequencies excluded from the target set. Turquoise is used for the additional frequencies predicted by the coarse-grained model.

Table 8 Optimal values of the normal-mode frequencies for the binary interaction for trimers and pentamers in cm^{-1}.

n	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
3	0.76	2.73	2.84	3.22	3.98	8.98
5	$1.21i$	0.99	1.65	2.31	2.43	3.32

presented to show that these last results are completely consistent with the current knowledge on the structural properties of the STNV. In first place, the fit performed here use the target frequencies obtained by Dykeman and Sankey for the STNV empty capsid using an atomistic approach that started from a structure that was determined by local energy minimisation of the Protein-Data-Bank 2BUK structure in a classical force-field[75,78]. The real values of the capsid frequencies confirm the stability, at least locally, of such a structure. Obviously, the optimal parameters obtained from the fit of our coarse-grained model to these data equally lead in both cases, trimer and pentamers, to locally stable capsid configurations. However, while the configuration for trimers corresponds to a global minimum of the capsid potential energy surface V, that for pentamers is just a local minimum and there are other structures with significantly lower energies. The existence of these more stable configurations for the pentamer capsid is a consequence of the inferred instability of the corresponding binary equilibrium configuration along its unstable mode, where lower energy structures will exist for the pentamer pair. The unstable pair configurations are locally stabilised in the capsid but beyond this stable point in the capsid configuration space one shall find the lower energy asymmetric structures in which some of the capsomer pairs have abandoned their unstable structures. The icosahedral capsid structure might then be unstable against thermal fluctuations. Although our optimised model would be strictly accurate within a small enough neighbourhood around the chosen equilibrium configurations of both the capsomer pair and the full capsid, and extending this accuracy beyond this region would require a more careful choice of the $F_i(r)$ functions, it can still provide some insight on the structural changes that the pentamer-pentamer instability will induce. Namely, by moving in the direction of the unstable mode of the binary interaction we find stable energy minima in which the C_2 symmetry is broken and the intercapsomer distance is reduced. From the corresponding eigenvector one concludes that the pathway of the unstable mode corresponds to a coordinate that involves excitation of all the symmetric internal coordinates r^+, ϕ^+ , θ^+, and χ^+ in similar magnitudes, but with χ^+ excitation having opposite direction to the other three. It is therefore a symmetric mode, i.e. it does not change after a π rotation around the original twofold axis of the transition state. Therefore, the two opposite directions of the unstable mode are unrelated by symmetry and thus lead to symmetry unrelated stable configurations, as can be checked numerically (they even have different energies). These new equilibrium structures are also observed in the deeper energy minima that now exist in the capsid.

There is experimental and theoretical work that indicates that

the empty STNV capsid is unstable at room temperatures[19,89–93]. Arkhipov et al.[19] use a coarse-graining molecular dynamics method with a resolution of ∼200 atoms to show that this empty capsid collapses to an asymmetric and significantly squeezed structure. These results are therefore fully consistent with our previous findings. But why is this instability absent in the coarse-grained trimer model? The experimental work by Ford et al.[91] can help to answer this question. The work by these authors focuses on the role of RNA-capsid binding interactions in the assembly of STNV for which the inability of its coat proteins to self-assemble in the absence of RNA was observed. Their results imply the relevance of these RNA-capsid bonds in stabilising the capsid structure by overcoming electrostatic repulsion barriers between coat-protein sites around the threefold symmetry axes and in making trimers the relevant capsid subunits when RNA lies inside. In other words, the instability of the empty capsid is located around its threefold symmetry axes and the bonds appearing there with the RNA introduce the required trimer rigidity to stabilise the capsid. Our rigid-trimer coarse-grained model produces in the capsid the same effect and removes the unstable modes. In contrast, these are still present in the model with pentamers, since the three subunits around each threefold symmetry axis can, in this case, move independently. The effects of those repulsion barriers between the coat proteins must be present in the empty-capsid force-field used by Dykeman and Sankey and therefore in the corresponding low-frequency normal-mode spectrum. As demonstrated here, a fit of an appropriate coarse-grained model to this spectrum can uncover the interaction features behind the spectral features.

4 Conclusions

A minimal coarse-grained model for the interaction energy of viral capsids as a sum of pairwise interactions between properly chosen equivalent subunits (trimers and pentamers for $T = 1$ capsids) has been designed. The model treats these subunits as rigid bodies and writes for them the most complete form, within the linear response regime, of their binary interaction as a sum of the minimum number of anisotropic terms chosen from those of the general multipolar expansion with lower orders and satisfying the geometrical constraints imposed by the high symmetry of the capsid equilibrium structure.

The model has been employed in this work to analyse the normal-mode frequency spectrum of the capsid. For this particular purpose and once the two geometrical parameters of the model (θ_{e_n} and r_{e_n}) are fixed, the interaction energy depends on 13 parameters linearly related to the 13 independent elements of the Hessian force-constant matrix for the binary-interaction fixed point. Therefore within the rigid-subunit approximation the model is complete to tackle the analysis of the small capsid oscillations. The system kinetic energy introduces three additional parameters. One of them, the capsomer mass, fixes the mass scale of the problem, and the other two correspond to the different capsomer moments of inertia $I_x = I_y$ and I_z. A fit of the predicted normal mode frequency spectrum to the atomistic data obtained by Dykeman and Sankey[75,78] for STNV empty capsid, a T_1 capsid, has been performed, and the values of the 15 free model parameters (13 force-constant matrix elements and 2 moments of inertia) have been determined for both trimers and pentamers as fundamental subunits.

Fair fits with $\sim 0.5\%$ error have been obtained in both cases, with no significant difference to favour the choice of a capsomer type over the other. However, the normal mode spectrum of the binary system (two isolated capsomers) obtained with the fitted parameters has revealed in the case of pentamers an imaginary frequency mode which makes the capsomer-pair equilibrium configuration unstable along this mode coordinate (a transition state). The theoretically and experimentally known properties the STNV capsid, which does not self-assemble in vitro without the RNA and is unstable at room temperature, would justify the unstable mode found for pentamers and the inability of the trimer capsid to detect the instability.

Together with its simplicity, the straightforward generalisability of the model beyond the linear regime or to include other capsomers like hexamers make it a promising tool to theoretically interpret many experimental data such as those provided by the atomic force microscopy or even to better understand processes far from equilibrium such as the capsid self-assembly.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work has been funded by the Spanish Ministerio de Economía y Competitividad through Grant FIS2016-79596-P (MINECO/FEDER, UE). Manuel Martín Bravo thanks the Spanish Ministerio de Economía y Competitividad (MINECO) for the predoctoral grant BES-2017-081104.

Notes and references

1 K. Adolph and P. Butler, *J. Mol. Biol.*, 1974, **88**, 327–341.
2 L. Lavelle, M. Gingery, M. Phillips, W. M. Gelbart, C. M. Knobler, R. D. Cadena-Nava, J. R. Vega-Acosta, L. A. Pinedo-Torres and J. Ruiz-Garcia, *J. Phys. Chem. B*, 2009, **113**, 3813–3819.
3 S. N. Kanesashi, K. I. Ishizu, M. A. Kawano, S. I. Han, S. Tomita, H. Watanabe, K. Kataoka and H. Handa, *J. Gen. Virol.*, 2003, **84**, 1899–1905.
4 T. S. Baker, N. H. Olson and S. D. Fuller, *Microbiol. Mol. Biol. Rev.*, 1999, **63**, 862–922.
5 R. Zandi, D. Reguera, R. F. Bruinsma, M. Gelbart and J. Rudnick, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 15556–15560.
6 J. E. Johnson and J. A. Speir, *J. Mol. Biol.*, 1997, **269**, 665–675.
7 S. Tonegawa and M. Hayashi, *J. Mol. Biol.*, 1970, **48**, 219–242.
8 D. M. Salunke, D. L. Caspar and R. L. Garcea, *Cell*, 1986, **46**, 895–904.
9 S. Flasinski, A. Dzianott, J. A. Speir, J. E. Johnson and J. J. Bujarski, *J. Virol.*, 1997, **71**, 2500–2504.
10 A. Zlotnick, R. Aldrich, J. M. Johnson, P. Ceres and M. J. Young, *Virology*, 2000, **277**, 450–456.
11 D. Willits, X. Zhao, N. Olson, T. S. Baker, A. Zlotnick, J. E.

Johnson, T. Douglas and M. J. Young, *Virology*, 2003, **306**, 280–288.
12 S. J. Hanslip, N. R. Zaccai, A. P. J. Middelberg and R. J. Falconer, *Biotechnol. Progr.*, 2006, **22**, 554–560.
13 A. Oppenheim, O. B. nun Shaul, S. Mukherjee and M. Abd-El-Latif, *Comput. Math. Methods Med.*, 2008, **9**, 265–276.
14 T. Dokland, *Structure*, 2000, **8**, 157–162.
15 L. Cardarelli, K. L. Maxwell and A. R. Davidson, *Proc. Nat. Acad. Sci. USA*, 2011, **108**, 10168–10173.
16 J. E. Baschek, H. C. R. Klein and U. S. Schwarz, *BMC Biophysics*, 2012, **5**, 22–40.
17 W. R. Wikoff, J. F. Conway, J. Tang, K. K. Lee, L. Gan, N. Cheng, R. L. Duda, R. W. Hendrix, A. C. Steven and J. E. Johnson, *J. Struct. Biol.*, 2006, **153**, 300–306.
18 G. Polles, G. Indelicato, R. Potestio, P. Cermelli, R. Twarock and C. Micheletti, *PLoS Comput. Biol.*, 2013, **9**, 1–13.
19 A. Arkhipov, P. L. Freddolino and K. Schulten, *Structure*, 2006, **14**, 1767–1777.
20 E. C. Dykeman and O. F. Sankey, *Phys. Rev. E*, 2010, **81**, 021918.
21 Y.-C. Hsieh, F. Poitevin, M. Delarue and P. Koehl, *Front. Mol. Biosci.*, 2016, **3**, 85.
22 H. D. Nguyen, V. S. Reddy and C. L. Brooks, III, *Nano Lett.*, 2007, **7**, 338–344.
23 A. Luque, R. Zandi and D. Reguera, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 5323–5328.
24 R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick and R. Zandi, *Phys. Rev. Lett.*, 2003, **90**, 248101.
25 R. Zandi and D. Reguera, *Phys. Rev. E*, 2005, **72**, 021917.
26 A. Luque, D. Reguera, A. Morozov, J. Rudnick and R. Bruinsma, *J. Chem. Phys.*, 2012, **136**, 184507.
27 T. Chen, Z. Zhang and S. C. Glotzer, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 717–722.
28 T. Chen and S. C. Glotzer, *Phys. Rev. E*, 2007, **75**, 051504.
29 S. Li, P. Roy, A. Travesset and R. Zandi, *Proc. Nat. Acad. Sci. USA*, 2018, **115**, 10971–10976.
30 B. Berger, P. Shor, L. Tucker-Kellogg and J. King, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 7732–7736.
31 R. Schwartz, P. W. Shor, P. E. Prevelige and B. Berger, *Biophys. J.*, 1998, **75**, 2626–6636.
32 D. J. Wales, *Phys. Biol.*, 2005, **2**, S86–S93.
33 D. C. Rapaport, *Phys. Rev. E*, 2004, **70**, 051905.
34 H. D. Nguyen and C. L. Brooks, *Nano Lett.*, 2008, **8**, 4574–4581.
35 S. N. Fejer, T. R. James, J. Hernández-Rojas and D. J. Wales, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2098–2104.
36 M. F. Hagan and D. Chandler, *Biophys. J.*, 2006, **91**, 42–54.
37 K. V. Workum and J. F. Douglas, *Phys. Rev. E*, 2006, **73**, 031502.
38 O. M. Elrad and M. F. Hagan, *Nano Lett.*, 2008, **8**, 3850–3857.
39 S. N. Fejer, D. Chakrabarti and D. Wales, *ACS Nano*, 2010, **4**, 219–228.
40 A. Ahadi, J. Colomo and A. Evilevitch, *J. Phys. Chem. B*, 2009, **113**, 3370–3378.
41 A. Arkhipov, W. H. Roos, G. J. L. Wuite and K. Schulten, *Biophys. J.*, 2009, **97**, 2061–2069.
42 A. Zhmurov, K. Rybnikov, Y. Kholodov and V. Barsegov, *J. Phys. Chem. B*, 2011, **115**, 5278–5288.
43 O. Kononova, J. Snijder, M. Brasch, J. Cornelissen, R. I. Dima, K. a. Marx, G. J. L. Wuite, W. H. Roos and V. Barsegov, *Biophys. J.*, 2013, **105**, 1893–1903.
44 M. Cieplak and M. O. Robbins, *J. Chem. Phys.*, 2010, **132**, 015101.
45 M. Cieplak and M. O. Robbins, *PLoS One*, 2013, **8**, 1–15.
46 M. Zink and H. Grubmüller, *Biophys. J.*, 2009, **96**, 1350–1363.
47 K. J. Boyd, P. Bansal, J. Feng and E. R. May, *Front. Bioeng. Biotechnol.*, 2015, **3**, 1–8.
48 K. Peeters and A. Taormina, *Comput. Math. Method. M.*, 2008, **9**, 211–220.
49 B. H. Lee, S. Jo, M. ki Choi, M. H. Kim, J. B. Choi and M. K. Kim, *Comput. Biol. Chem.*, 2018, **72**, 53–61.
50 M. Widom, J. Lidmar and D. R. Nelson, *Phys. Rev. E*, 2007, **76**, 031911.
51 Z. Yang, I. Bahar and M. Widom, *Biophys. J.*, 2009, **96**, 4438–4448.
52 A. Levandovsky and R. Zandi, *Phys. Rev. Lett.*, 2009, **102**, 198102.
53 A. Lošdorfer Božič, A. Šiber and R. Podgornik, *J. Biol. Phys.*, 2013, **39**, 215–228.
54 C. Tiwari, V. Sharma, P. K. Jha and A. Pratap, *J. Biomol. Struct. Dyn.*, 2019, **0**, 1–8.
55 E. R. May and C. L. Brooks, *J. Phys. Chem. B*, 2012, **116**, 8604–8609.
56 J. M. Gomez Llorente, J. Hernández-Rojas and J. Bretón, *Soft Matter*, 2014, **10**, 3560–3569.
57 M. Aznar and D. Reguera, *J. Phys. Chem. B*, 2016, **120**, 6147–6159.
58 D. Reguera, J. Hernández-Rojas and J. M. Gomez Llorente, *Soft Matter*, 2019, **15**, 7166–7172.
59 D. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, 2004.
60 A. Zlotnick, *J. Mol. Recognit.*, 2005, **18**, 479–490.
61 D. L. D. Caspar and A. Klug, *Cold Spring Harbor Symp. Quant. Biol.*, 1962, **27**, 1–24.
62 C. Carrasco, A. Luque, M. Hernando-PÃľrez, R. Miranda, J. Carrascosa, P. Serena, M. de Ridder, A. Raman, J. GÃşmez-Herrero, I. Schaap, D. Reguera and P. de Pablo, *Biophys. J.*, 2011, **100**, 1100–1108.
63 M. Aznar, S. Roca-Bonet and D. Reguera, *J. Phys.: Condens. Matter*, 2018, **30**, 264001.
64 M. Babincová, P. Sourivong and P. Babinec, *Med. Hypotheses*, 2000, **55**, 450–451.
65 M. Zupanc, Ž. Pandur, T. S. Perdih, D. Stopar, M. Petkovšek and M. Dular, *Ultrason Sonochem*, 2019, **57**, 147–165.
66 Y. Yan, E. B. Gamble and K. A. Nelson, *J. Chem. Phys.*, 1985, **83**, 5391–5399.
67 K. T. Tsen, S.-W. D. Tsen, C.-L. Chang, C.-F. Hung, T.-C. Wu

and J. G. Kiang, *J. Phys.: Condens. Matter*, 2007, **19**, 322102.
68 K. T. Tsen, S.-W. D. Tsen, O. F. Sankey and J. G. Kiang, *J. Phys.: Condens. Matter*, 2007, **19**, 472201.
69 K. T. Tsen, S.-W. D. Tsen, C.-F. Hung, T.-C. Wu and J. G. Kiang, *J. Phys.: Condens. Matter*, 2008, **20**, 252205.
70 Q. Cui and I. Bahar, *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*, Chapman and Hall/CRC, New York, 1st edn, 2005.
71 B. Brooks and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.*, 1985, **82**, 4995–4999.
72 N. Go, T. Noguti and T. Nishikawa, *Proc. Natl. Acad. Sci. U.S.A.*, 1983, **80**, 3696–3700.
73 M. Levitt, C. Sander and P. S. Stern, *J. Mol. Biol.*, 1985, **181**, 423–447.
74 H. W. van Vlijmen and M. Karplus, *J. Mol. Biol.*, 2005, **350**, 528–542.
75 E. C. Dykeman and O. F. Sankey, *Phys. Rev. Lett.*, 2008, **100**, 028101.
76 P. Koehl, *Comput. Biol. Chem.*, 2018, **14**, 3903–3919.
77 G. Song, *J. Mol. Graph.*, 2019, **87**, 30–40.
78 E. C. Dykeman, *PhD thesis*, Arizona State University, 2008.
79 L. C. Biedenharn, J. D. Louck and P. A. Carruthers, *Angular Momentum in Quantum Physics*, Addison-Wesley Publishing Company, 1st edn, 1981, vol. 8.
80 D. J. Wales and J. P. K. Doye, *J. Phys. Chem. A*, 1997, **101**, 5111–5116.
81 Z. Li and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 6611–6615.
82 E. B. Wilson, J. Decius and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*, Dover Publications, Inc., 1980.
83 V. RÃijhle, H. Kusumaatmaja, D. Chakrabarti and D. J. Wales, *J. Chem. Theory Comput.*, 2013, **9**, 4026–4034.
84 D. J. Evans, *Mol. Phys.*, 1977, **34**, 317–325.
85 D. J. Evans and S. Murad, *Mol. Phys.*, 1977, **34**, 327–331.
86 K. Refson and G. S. Pawley, *Mol. Phys.*, 1987, **61**, 669–692.
87 D. C. Rapaport, *The art of molecular dynamics simulation*, Cambridge University Press, 1995.
88 M. Arribas, A. Elipe and M. Latasa, *Celestial Mechanics and Dynamical Astronomy*, 2006, **96**, 239–251.
89 N. Ban, S. B. Larson and A. McPherson, *Virology*, 1995, **214**, 571–583.
90 D. H. Bunka, S. W. Lane, C. L. Lane, E. C. Dykeman, R. J. Ford, A. M. Barker, R. Twarock, S. E. Phillips and P. G. Stockley, *J. Mol. Biol.*, 2011, **413**, 51–65.
91 R. J. Ford, A. M. Barker, S. E. Bakker, R. H. Coutts, N. A. Ranson, S. E. Phillips, A. R. Pearson and P. G. Stockley, *J. Mol. Biol.*, 2013, **425**, 1050–1064.
92 N. Patel, E. C. Dykeman, R. H. A. Coutts, G. P. Lomonossoff, D. J. Rowlands, S. E. V. Phillips, N. Ranson, R. Twarock, R. Tuma and P. G. Stockley, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 2227–2232.
93 N. Patel, E. Wroblewski, G. Leonov, S. E. V. Phillips, R. Tuma, R. Twarock and P. G. Stockley, *Proc. Natl. Acad. Sci. U.S.A.*, 2017, **114**, 12255–12260.

A minimal coarse-grained model unveils relevant structural properties of icosahedral viral capsids when fitted to reproduce their low-frequency normal-mode spectrum.