Normal mode analysis and applications in biological physics

TOPICAL REVIEW

Normal mode analysis and applications in biological physics

Eric C Dykeman and Otto F Sankey

Department of Physics, Center for Biological Physics, Arizona State University, Tempe,
AZ 85287-1504, USA

Abstract
Normal mode analysis has become a popular and often used theoretical tool in the study of
functional motions in enzymes, viruses, and large protein assemblies. The use of normal modes
in the study of these motions is often extremely fruitful since many of the functional motions of
large proteins can be described using just a few normal modes which are intimately related to
the overall structure of the protein. In this review, we present a broad overview of several
popular methods used in the study of normal modes in biological physics including continuum
elastic theory, the elastic network model, and a new all-atom method, recently developed, which
is capable of computing a subset of the low frequency vibrational modes exactly. After a review
of the various methods, we present several examples of applications of normal modes in the
study of functional motions, with an emphasis on viral capsids.

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction and background

One of the most important paradigms in biological physics
is that of the relationship between structure and function in
proteins, protein assemblies, or enzymes. This paradigm
is a simple but powerful one; that the structure of these
nanomachines ultimately determines the function of these
nanomachines. One can see similarities to the relation of the
geometrical and symmetry properties of a simple molecule,
such as water, to its overall macroscopic properties. Change
one of the atoms in the molecule and the electron structure and
overall 'shape' or geometry of the molecule changes, causing
the overall macroscopic properties to change too. Similarly,
change (or add/delete) one or several of the amino acids in
a functional protein and the overall function of this protein
changes—perhaps by inhibiting certain enzymatic reactions
because the mutation occurs in the active site of the protein
or by making the overall protein non-folding or 'shapeless'.

Proceeding further along, the function (such as an
enzymatic reaction) of the protein nanomachine often depends
on the motion or dynamical properties of its domains or a small
section of the protein, with the large motion dynamics typically
occurring at hinges, flaps, gates and otherwise floppy regions
of the protein. But how to theoretically study these types

of motions effectively in order to understand the connection between motion and function in these nanomachines? The traditional method has predominately been by molecular dynamics (MD) simulations of the system in a solvent [1–5]. These simulations produce great insights, and when the timescales of the dynamics is short enough, MD is a tool without equal. But there are other tools that also give insight into dynamical fluctuations that are at the heart of the structure–function paradigm. One of these tools is normal mode analysis, and this will be the subject of this review. Normal mode analysis seeks to find the natural concerted motion of the nanomachine by investigating the vibrational normal modes determined by its structure and the interactions within its various parts. Our review will be broad and describe several algorithms used to determine the normal modes and specific frequencies of each. The topic is an old one, so there are many twists of these basic algorithms that are useful, but we will attempt to stick to the main ideas. By way of application, we will use viral capsid assemblies as a recurring theme of a complex example.

The problems of investigating dynamical motions in biological proteins and assemblies using normal mode analysis is difficult. Normal mode analysis is a single tool and is often used in tandem with other tools such as traditional MD. In addition, normal mode analysis has several limitations. Often the dynamical motions in proteins involves large amplitude motions while the theory of normal modes is for small oscillations. Further, biomolecules are often highly damped due to solvent and exhibit either few or no oscillations. A review of some of the issues that limit the applicability of normal mode analysis are addressed by Ma [6].

We will discuss a variety of approaches in this review. An emerging new theme discussed in this review is that of all-atom determinations of normal modes. This area has not attracted much attention for large systems because it was an intractable problem. The all-atom phonon functional method described herein now makes these problems tractable [7–9]. A full all-atom picture removes some of the uncertainty of the approximations of other methods. A comparison will be made to coarse grained models to obtain evidence for validity or shortcoming of other models. Generally, the coarse grained models do well for the displacement patterns, although they do not predict frequencies well or not at all. Most applications in the literature do not focus on the frequencies as generally these are even more difficult to find and often what is interesting is the displacement pattern itself. But in some applications, partial oscillations are important and having information about the frequencies is useful. One such example is impulsive stimulated Raman scattering which 'pings' the system with a short pulse [10, 11]. The length of the pulse should be a quarter of a period of oscillation to be most effective. Estimates of frequencies are needed are to investigate such systems knowledgeably.

We now present an outline of this review paper. In section 2, basic principles of the harmonic approximation, the central theory of most normal mode analysis, will be given. After a description of the harmonic approximation, we discuss an often calculated quantity from normal modes

which is useful in describing the 'floppiness' or 'rigidity' of a protein, the thermal B factors or Debye–Waller factor. In section 3 we describe the mathematics of continuum elastic models. Examples will include the M13 bacteriophage and how the continuum theory can be advanced further to obtain estimates of relative Raman scattering profiles followed by a brief discussion of determining the overall macroscopic elastic properties of an icosahedral capsid from continuum models. Next, section 4 will discuss one of the most popular methods in coarse grained normal mode analysis, the elastic network model, and more briefly the rotation–translation block method. Some work on the coarse grained models of viral capsids and other enzymes are reviewed. In section 5, we switch to a statistical approach, principal component analysis. The foundations of principal component analysis will be related to normal mode analysis in the harmonic limit. An example of principal component analysis will be made from a molecular dynamics simulation. In addition, we will discuss an example of categorizing the static structures of the Ras kinase protein. Section 6 describes the all-atom phonon functional technique. This technique is quite new and determines a subset of the lowest frequency modes of a very large system, exactly, within an empirical force field model. By seeking only a subset of the lowest modes, the all-atom phonon functional method avoids diagonalizing large matrices. We compare the results of the all-atom phonon functional for the polio capsid to previously determined results using a restricted basis set of oscillation. The need for having reasonable estimates of capsids is useful for impulsive stimulated Raman scattering (ISRS) experiments which are a suggested mechanism for destruction of viral capsids using pulsed light [10]. We give an example of such a destruction for M13 viral capsids using an MD simulation with added forces from a laser. In section 7 we describe some of the ways in which the normal modes of a protein are used to search for a mode displacement subspace in order to better understand conformational changes of enzymes. One example is given in detail—that of open to closed transition of lactoferrin. Comparisons are made of the all-atom phonon functional and the elastic network model. Finally, the work concludes with mode displacement subspace of a quite complex system, a gram-negative bacterial ABC transporter. This membrane protein is an assembly of three distinct functional protein complexes-the periplasmic binding protein which functions as a 'cargo holder', a transmembrane protein with a functional gate, and an ATP–ADP driven 'engine'. An all-atom model is used to find cooperative motions throughout the quaternary structure of protein complexes.

2. Basic principles

We outline the basic physical and mathematical framework that is required to describe and compute normal modes of a molecule, protein complex, or virus. Normal mode analysis is appropriate for small excursions from equilibrium in which the restoring forces are linear. We first describe the point mass model in which the massive objects (atoms, residue/beads, or proteins) are discrete, and later we consider the continuum model. Thermal B-factors (the Debye–Waller factor) which

describe average squared displacements from equilibrium is given as an example of the use of normal modes to compute a qualitative quantity.

2.1. Harmonic approximation

The purpose of a normal mode analysis is to describe the principal collective motions of a group of N atoms that move in a potential energy of interaction V. Empirical potential energy functions extensively developed for classical molecular dynamics simulations are written as

$$V(\vec{r}_1, \vec{r}_2, \ldots) = \sum_{\text{bonds}} k_b (r - r_{\text{eq}})^2 + \sum_{\text{angles}} k_a (\theta - \theta_{\text{eq}})^2$$
$$+ \sum_{\text{dihedrals}} \frac{v_n}{2} \left[1 + \cos(n\phi - \gamma) \right]$$
$$+ \frac{1}{2} \sum_{ij} \left(\frac{q_i q_j}{\epsilon r} + \frac{A}{r^{12}} - \frac{B}{r^6} \right). \qquad (1)$$

The first and second terms in the empirical potential describe the energy due to bond stretching and the bending between two bonds as simple harmonic potentials with spring constants k_b and k_a.[1] The third term describes the potential energy due to rotation of a chemical group around a single bond (for example rotating the three hydrogens about the C–C bond in C_2H_6). The last terms are the electrostatic potential from Coulomb and van der Waals interactions. Force field models such as AMBER [12] and CHARMM [13] have extensively developed parameters for individual amino acids and nucleic acids.

Ordinarily, one solves for the motions of the atoms by integrating the second order differential equation

$$m_i \frac{\mathrm{d}^2 \vec{r}_i}{\mathrm{d}t^2} = -\vec{\nabla} V \qquad (2)$$

over time for each atom i. This is a general solution but is computationally demanding and requires deep analysis to understand the system's behavior. The harmonic approximation solves the problem of describing the motion but in the limit of small excursions from equilibrium. It treats the systems of atoms as a collection of *small* displacements about an equilibrium position R where the net force on all N atoms is zero. The solutions oscillate about equilibrium with characteristic angular frequencies ω_ν and displacement patterns $|\eta_\nu\rangle$. They are determined as eigenvalues and eigenvectors of a 'dynamical matrix'. To obtain these, the potential energy V can be expanded in a Taylor series about the equilibrium and truncated, i.e.,

$$V(\vec{r}_1, \vec{r}_2, \ldots) = V(\vec{r}_1, \vec{r}_2, \ldots)|_R + \sum_{i,\alpha} \frac{\partial V}{\partial r_{i\alpha}} \bigg|_R (r_{i\alpha} - R_{i\alpha})$$
$$+ \frac{1}{2} \sum_{i,\alpha} \sum_{j,\beta} \frac{\partial^2 V}{\partial r_{i\alpha} \partial r_{j\beta}} \bigg|_R (r_{i\alpha} - R_{i\alpha})(r_{j\beta} - R_{j\beta}), \qquad (3)$$

where α or β represent one of the directions x, y, or z while i and j represent one of the N atoms. Since atoms at equilibrium have zero net force, the second term in equation (3) is zero.

[1] A factor of $1/2$ could be included in front of these terms—historically it is not. This factor of $1/2$, whether it is there are not, affects the spring k values by a corresponding factor of two.

Furthermore, we are free to add any constant term to the potential energy without effecting the results. Neglecting the first term, equation (3) simplifies to

$$V(\vec{r}_1, \vec{r}_2, \ldots) = \frac{1}{2} \sum_{i,\alpha} \sum_{j,\beta} \frac{\partial^2 V}{\partial r_{i\alpha} \partial r_{j\beta}} \bigg|_R u_{i\alpha} u_{j\beta}, \qquad (4)$$

where $u_{i\alpha} = (r_{i\alpha} - R_{i\alpha})$ is the α component of the displacement of atom i from its equilibrium position $\vec{R}_i$. Equation (4) can be rewritten in compact form as

$$V(\vec{r}_1, \vec{r}_2, \ldots) = \tfrac{1}{2} \langle \vec{u} | \overset{\leftrightarrow}{\phi} | \vec{u} \rangle \qquad (5)$$

where $|\vec{u}\rangle$ is the column vector of length $3N$ which contains the atomic displacements $u_{i\alpha}$ where i is the atom index and α the Cartesian direction. The symmetric $3N \times 3N$ atomic force-constant matrix (or Hessian matrix) $\overset{\leftrightarrow}{\phi}$ has components

$$\phi_{i\alpha, j\beta} = \frac{\partial^2 V}{\partial r_{i\alpha} \partial r_{j\beta}} \bigg|_{r=R}. \qquad (6)$$

The equation of motion is easily constructed and solved exactly. The kinetic energy of the atoms,

$$T = \tfrac{1}{2} \sum_{i\alpha} m_i \dot{u}_{i\alpha}^2, \qquad (7)$$

along with equation (5) for the approximate potential energy produces the Lagrangian,

$$\mathcal{L} = T - V = \tfrac{1}{2} \sum_{i\alpha} m_i \dot{u}_{i\alpha}^2 - \tfrac{1}{2} \sum_{i\alpha, j\beta} u_{i\alpha} \phi_{i\alpha, j\beta} u_{j\beta}. \qquad (8)$$

The Lagrangian along with Euler's equations [14] yield the following set of $3N$ coupled equations of motion

$$m_i \ddot{u}_{i\alpha} = -\sum_{j\beta} \phi_{i\alpha, j\beta} u_{j\beta}. \qquad (9)$$

Harmonic solutions of equation (9) take the familiar form of

$$u_{i\alpha} = Q \eta_{i\alpha} \mathrm{e}^{-\mathrm{i}\omega t}, \qquad (10)$$

where $\eta_{i\alpha}$ are the components of the vector describing the displacement mode pattern of the masses, and Q describes the amplitude of the collective motion. Substitution into equation (9) gives

$$m_i \omega^2 \eta_{i\alpha} = \sum_{j\beta} \phi_{i\alpha, j\beta} \eta_{j\beta}, \qquad (11)$$

which in matrix form becomes

$$\overset{\leftrightarrow}{\phi} |\vec{\eta}_\nu\rangle = \omega_\nu^2 \overset{\leftrightarrow}{M} |\vec{\eta}_\nu\rangle \qquad (12)$$

where the $3N \times 3N$ matrix $\overset{\leftrightarrow}{M}$ is diagonal with elements $M_{i\alpha, j\beta} = m_i \delta_{ij} \delta_{\alpha\beta}$. The index ν labels the $3N$ possible solutions for the displacement eigenvectors $|\vec{\eta}_\nu\rangle$ and eigenfrequencies ω_ν.

Equation (12) is in the form of a generalized eigenvector equation with matrices $\overset{\leftrightarrow}{\phi}$ on the left and $\overset{\leftrightarrow}{M}$ on the right.

It is convenient to convert to 'mass weighted' coordinates, i.e., $e_{i\alpha} = \sqrt{m_i}\eta_{i\alpha}$ so that equation (12) can be written in the following simpler form of a standard Hermitian matrix eigenvalue problem,

$$\overset{\leftrightarrow}{D}\,|\hat{e}_\nu\rangle = \omega_\nu^2|\hat{e}_\nu\rangle \qquad (13)$$

where the matrix $\overset{\leftrightarrow}{D}$ is the $3N \times 3N$ dynamical matrix given by $\overset{\leftrightarrow}{D} = \overset{\leftrightarrow}{M}{}^{\frac{-1}{2}}\overset{\leftrightarrow}{\phi}\,\overset{\leftrightarrow}{M}{}^{\frac{-1}{2}}$, or more simply in component form, $D_{i\alpha,j\beta} = \phi_{i\alpha,j\beta}/\sqrt{m_i m_j}$. Diagonalizing the dynamical matrix yields a set of $3N$ orthonormal eigenvectors $|\hat{e}_\nu\rangle$ with corresponding eigenvalue ω_ν^2. The eigenvector $|\hat{e}_\nu\rangle$ describes the mass weighted collective displacement of the $3N$ atoms.

The connection between the displacement eigenvectors and the eigenvectors of the dynamical matrix and their orthonormality conditions can cause confusion, and so we summarize them here. The two set of vectors are related to each other by the transformation $|\vec{\eta}_\nu\rangle = \overset{\leftrightarrow}{M}{}^{\frac{-1}{2}}|\hat{e}_\nu\rangle$ and $|\hat{e}_\nu\rangle = \overset{\leftrightarrow}{M}{}^{\frac{1}{2}}|\vec{\eta}_\nu\rangle$. The eigenvectors $|\hat{e}_\nu\rangle$ of the dynamical matrix are normalized to unity and are orthonormal to each other; $\langle\hat{e}_\nu|\hat{e}_{\nu'}\rangle = \delta_{\nu\nu'}$. The displacement eigenvectors $|\vec{\eta}_\nu\rangle$ are normalized with the weighting of the mass matrix, $\langle\vec{\eta}_\nu|\overset{\leftrightarrow}{M}|\vec{\eta}_{\nu'}\rangle = \delta_{\nu\nu'}$.

The power of the normal modes is that a general displacement of atomic positions can be written as a linear combination of the displacement eigenvectors, $|\vec{u}\rangle = \sum_\nu Q_\nu(t)|\vec{\eta}_\nu\rangle$, where the $Q_\nu(t)$ are the amplitudes of each mode in the pattern and have harmonic ($\cos(\omega_\nu t)$ and $\sin(\omega_\nu t)$) time dependence. The amplitudes then become generalized coordinates and are useful for thermodynamic evaluations and for quantization of the system. Specifically, the kinetic and potential energies are easily represented in terms of the normal mode eigenfrequencies and mode amplitudes;

$$T = \tfrac{1}{2}\langle\dot{\vec{u}}|\overset{\leftrightarrow}{M}|\dot{\vec{u}}\rangle = \tfrac{1}{2}\sum_\nu \dot{Q}_\nu^2(t)$$

$$V = \tfrac{1}{2}\langle\vec{u}|\overset{\leftrightarrow}{\phi}|\vec{u}\rangle = \tfrac{1}{2}\sum_\nu \omega_\nu^2 Q_\nu^2(t). \qquad (14)$$

The Lagrangian in these generalized coordinates defines the generalized conjugate momenta, $P_\nu = \partial\mathcal{L}/\partial\dot{Q}_\nu = \dot{Q}_\nu$. The harmonic oscillator problem is one of the primary problems in statistical mechanics that has an exact solution. Normal mode analysis produces a system of $3N$ independent harmonic oscillator problems. In this regime, the classical partition function becomes a product of partition functions for each mode, $Z = \Pi_\nu Z_\nu$, where the partition function for each mode is

$$Z_\nu = \int\int e^{-\beta\left(\frac{1}{2}P_\nu^2 + \frac{1}{2}\omega_\nu^2 Q_\nu^2\right)}\,\mathrm{d}P_\nu\,\mathrm{d}Q_\nu. \qquad (15)$$

Thermodynamic averages of a quantity X are easily computed via,

$$\langle X(Q_\nu, P_\nu)\rangle = \frac{\int\int X(Q_\nu, P_\nu)e^{-\beta\left(\frac{1}{2}P_\nu^2 + \frac{1}{2}\omega_\nu^2 Q_\nu^2\right)}\,\mathrm{d}P_\nu\,\mathrm{d}Q_\nu}{Z_\nu}. \qquad (16)$$

The results are that averages over displacements satisfy $\langle Q_\nu\rangle = 0$, $\langle Q_\nu^2\rangle = k_\mathrm{b}T/\omega_\nu^2$, and $\langle Q_\nu Q_{\nu'}\rangle = 0\,(\nu \neq \nu')$. Averages over momenta satisfy similar relations $\langle P_\nu\rangle = 0$, $\langle P_\nu^2\rangle = k_\mathrm{b}T$, and $\langle P_\nu P_{\nu'}\rangle = 0\,(\nu \neq \nu')$.

In summary, the potential energy function V which governs the motions of the atoms completely determines the *small* collective motions of the atoms that take place at a given local minimum. The challenge of applying normal mode analysis to a large protein or protein complex such as a virus is two fold; (i) to obtain a potential energy function which accurately describes the interactions between the atoms such that the resulting normal modes correctly predict the collective motions (plus frequencies) of the protein and (ii) to diagonalize the often large $3N \times 3N$ dynamical matrix. We will see in subsequent sections that the solution to these problems has required either approximations to the potential energy function and/or a reduction in the basis set in order to reduce the size of the dynamical matrix so that it is solvable. New methods (phonon functional method) have been recently developed which allow for the subset of low frequency modes of a large protein to be computed to atomic detail. This will be described in section 6.

In this review we will focus on developments in biological physics. However, normal mode analysis finds uses in many branches of engineering, physics, and chemistry. Polymer science is one area that made many novel developments and advances. For example, normal mode eigenvectors and eigenvalues can be averaged over trajectories in molecular dynamics to obtain a time average Hessian [15, 16]. Preconditioning iterative solutions of the normal mode problem have been advanced [18]. Spares matrix techniques based on a variant of the Arnoldi process in the software package ARPACK [17] which extend the popular Lanczos algorithm are often used [16].

2.2. Thermal B-factors from normal modes

Thermal B-factor (Debye–Waller factor) of atoms in a protein describe the average squared displacements of the atoms in the protein away from their equilibrium positions at a temperature T. They are a useful measure of a protein's flexibility or rigidity. Normal modes can be used to estimate the B-factors. Within a classical harmonic framework, the B-factor for atom 'i', B_i, is [19]

$$B_i = \frac{8\pi^2}{3}\langle U_i^2\rangle, \qquad (17)$$

where $U_i^2 = u_{ix}^2 + u_{iy}^2 + u_{iz}^2$, and $u_{i\beta}$ are the β components of the displacements away from the equilibrium position $\vec{R}_i$, (i.e. $\vec{r}_i = \vec{R}_i + \vec{u}_i$). The brackets in equation (17) denote an average computed at temperature T.

Computing the thermal B-factors requires the calculation of the average of the atoms' squared displacements from equilibrium, $\langle U_i^2\rangle$. The average squared displacement from equilibrium can be written in terms of classical harmonic normal modes, with the dominant contributions due to the *low* frequency modes. In a harmonic description of protein dynamics, the displacement $\vec{u}_i(t)$ of an atom i is

$$\vec{u}_i(t) = \sum Q_\nu(t)\vec{\eta}_i(\nu). \qquad (18)$$

Averaging the displacement squared, $\langle U_i^2 \rangle = \sum_{\mu\nu} \langle Q_\mu Q_\nu \rangle \vec{\eta}_i (\mu) \cdot \vec{\eta}_i (\nu)$ which simplifies using the thermodynamic averages following equation (15) to

$$\langle U_i^2 \rangle = \frac{k_b T}{m_i} \sum_\nu \frac{e_i^2(\nu)}{\omega_\nu^2}, \qquad (19)$$

where the vector $\vec{e}_i(\nu)$ is the mass weighted displacement vector $(\vec{\eta}_i(\nu) = \vec{e}_i(\nu)/\sqrt{m_i})$ for atom i in mode ν and is the normalized eigenvector of the dynamical matrix. The expression (19) for the thermal B-factors show that the contributions from each mode depends on $1/\omega_\nu^2$—the low frequency modes dominate. This of course assumes that the spectrum is somewhat uniformly distributed. One can obtain a reasonable approximation to the thermal B-factors by truncating the sum up to some maximum frequency.

As an example of a B-factor calculation for a protein we compare the experimental and theoretical B-factors for alpha lytic protease (PDB ID 2ALP) in figure 1. This problem was first investigated by Miller *et al* [20]. The B-factors that are displayed represent an averaged value for each amino acid residue. The average is computed by summing individual B-factors over the heavy atoms (no hydrogen) in each amino acid then dividing by the total number of heavy atoms. The theoretical B-factors are computed using two different regimes. In the top panel, the thick line indicates the theoretical B-factors computed with the all-atom phonon functional method (see section 6 below) for modes up to 20 cm^{-1} using the empirical AMBER potential energy model while the thin line indicates the experimental values. One sees qualitative agreement of the normal mode results and experiment at least in identifying regions with the largest motions (residues 82 is an exception). In the bottom panel, the curve with the high values shows the theoretical B-factors computed using the C_α elastic network model, ENM (to be described fully in section 4). For a proper comparison with the all-atom method, we scaled the frequencies computed from the elastic network model so that the frequency of the lowest elastic network mode was the same as the frequency for the lowest all-atom mode. The ENM also identifies the region with the largest motion, but exaggerates the displacements. This is likely due to the ENM's tendency to give a more dense spectrum of low frequency modes than the all-atom model—at least in the case when the frequency of the lowest ENM mode is fit to the all-atom model. This will be further discussed in section 7.

3. Continuum models

Continuum models of normal modes have been central to obtaining a fundamental understanding of the types of motions that a large protein or protein assembly can exhibit. Continuum models have mostly been applied to viral capsids and have been successfully used to describe some of the mechanical and physical processes that viral capsids exhibit. Such examples include the structural transitions involved in capsid maturation events [21] and the distribution of stress on viral shells in response to internal and external pressures [22]. Continuum
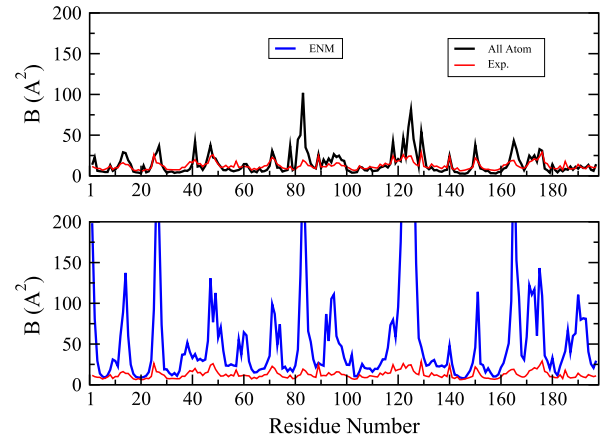

Figure 1. Experimental and theoretical B-factors for alpha lytic protease (PDB ID 2ALP). Top panel: theoretical B-factors computed from an all-atom normal mode calculation with AMBER force field (thick line) plotted with the experimental B-factors (thin line). Bottom panel: theoretical B-factors computed from a C_α ENM (curve with generally higher values) with experimental B-factors (generally lower values). The frequencies of all the modes in the ENM prediction were scaled such that the lowest mode had identical frequency with the all-atom results.

models have also been used to determine the stability of capsids of different shape (spherical, cylindrical, or conical shell capsids like HIV) based on the ratios of the Foppel von Karmen number and spontaneous curvature [23]. In all, continuum models have been useful for describing the vibrational modes and distortion of viral capsids and a variety of other physical phenomenon related to viral capsids.

In this section we provide a general overview of the theory of continuum models of vibrational modes. We then follow with a review of work that has been done on calculating the vibrational modes of tubular and spherical viral capsids and discuss how continuum models of vibrational modes have been used to predict Raman scattering profiles, Young's modulus, and the Poisson ratio of viral capsids. Continuum modeling can be found in other areas of science and engineering and plays an important role in structural engineering [24].

3.1. Continuum model theory

In continuum models of vibrational modes, one solves for the displacements per unit length, $\vec{u}(\vec{r}, t)$, as a function of position $\vec{r}$ and time t within the elastic material. The partial derivatives of the displacements per unit length $\vec{u}(\vec{r}, t)$, describe the shear and bulk stress at position $\vec{r}$ within the elastic material. Figure 2(a) illustrates the relations of the displacements per unit length and their derivatives to the stress tensor components $u_{ij} = \partial u_i / \partial x_j$. Assuming small deformations of the elastic material that obey Hook's law, one obtains the following equations of motion [25]

$$(\lambda + \mu)\vec{\nabla}(\vec{\nabla} \cdot \vec{u}) + \mu \nabla^2 \vec{u} = \rho \ddot{\vec{u}}. \qquad (20)$$

The parameters λ and μ are the Lamé coefficients [25] which describe the stiffness of the material while ρ is the density of the material. The Lamé coefficients are related to the Young's
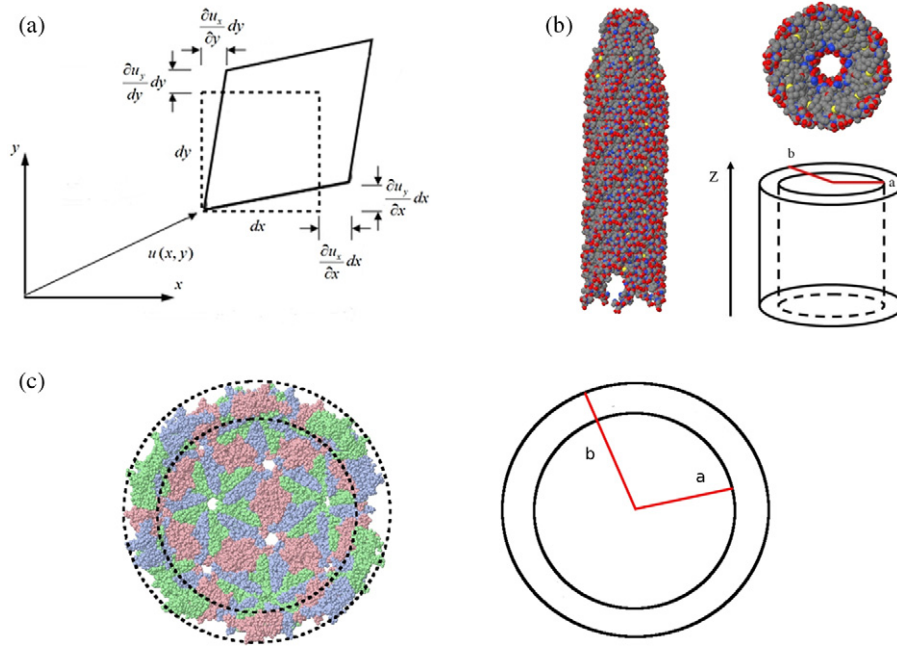
Figure 2. Continuum model of elastic vibrations. (a) Illustration of the displacement per unit length $\vec{u}(\vec{r}, t)$ in 2D for an infinitesimal square section (with area $dx \times dy$) of an elastic medium. The relative displacements between the corners of the infinitesimal square give the elements of the strain tensor ($u_{ij} = \partial u_i / \partial x_j$) for the elastic material. (b) Representation of the tubular M13 capsid as an isotropic elastic material with inner and outer radii of a and b, respectively. (c) Cross section showing the representation of an icosahedral virus as an isotropic elastic sphere with inner and outer radii of a and b, respectively.

modulus E and Poisson ratio σ of the material by the following relations:

$$\lambda + \mu = \frac{E}{2(1+\sigma)(1-2\sigma)}, \qquad \mu = \frac{E}{2(1+\sigma)}. \quad (21)$$

Solving equation (20) for $\vec{u}(\vec{r}, t)$ (subject to appropriate boundary conditions—see below) at every point within the material provides a complete description of the possible vibrational mode patterns that can be exhibited. The solution to equation (20) can be obtained numerically or, in cases where the material can be approximated as a tube or sphere, analytically [26–29, 31, 32]. When solving equation (20), it is usually useful to separate $\vec{u}(\vec{r}, t)$ into its longitudinal and transverse components $\vec{u}_l$ and $\vec{u}_t$ which have the following properties;

$$\vec{\nabla} \cdot \vec{u}_t = 0,$$
$$\vec{\nabla} \times \vec{u}_l = 0. \quad (22)$$

One can immediately see that if $\nabla \cdot \vec{u} = 0$ (i.e. $\vec{u} = \vec{u}_t$), then the first term in equation (20) is zero and the equation can be written as

$$\nabla^2 \vec{u}_t = \frac{1}{c_t^2} \frac{\partial^2 \vec{u}_t}{\partial t^2}, \quad (23)$$

while if $\nabla \times \vec{u} = 0$ (i.e. $\vec{u} = \vec{u}_l$), then $\nabla^2 \vec{u}_l = \vec{\nabla}(\vec{\nabla} \cdot \vec{u}_l)$ and equation (20) can be written instead as

$$\nabla^2 \vec{u}_l = \frac{1}{c_l^2} \frac{\partial^2 \vec{u}_l}{\partial t^2}. \quad (24)$$

The constants c_t and c_l are the transverse and longitudinal speeds of sound which are related to the Lamé coefficients by

the relations

$$c_t^2 = \frac{\mu}{\rho},$$
$$c_l^2 = \frac{\lambda + 2\mu}{\rho}. \quad (25)$$

Thus, the longitudinal and transverse portions of $\vec{u}$ each satisfy their own vector wave equation with a characteristic speed of sound c_t or c_l.

For tubular viruses such as M13, and tobacco mosaic virus (TMV), the capsid can be modeled as an infinite hollow tube with the tube axis along $\hat{z}$. Cylindrical coordinates r, θ, z are used. Figure 2(b) illustrates the geometry used for M13 with inner and outer radii a and b. In this case, one can solve equation (20) by writing the displacement per unit length in terms of gauge invariant scalar and vector potentials Φ and $\vec{H}$ [26],

$$\vec{u} = \vec{\nabla}\Phi + \vec{\nabla} \times \vec{H}. \quad (26)$$

With the appropriate choice of gauge ($\vec{\nabla} \cdot \vec{H} = 0$), substitution into equation (20) yields separate transverse and longitudinal wave equations for the vector and scalar potentials,

$$\nabla^2 \Phi = \frac{1}{c_l^2} \frac{\partial^2 \Phi}{\partial t^2}, \qquad \nabla^2 \vec{H} = \frac{1}{c_t^2} \frac{\partial^2 \vec{H}}{\partial t^2}. \quad (27)$$

The solutions for both equations take the form of a Bessel function times either cosine or sine of the polar angle θ [26],

$$\Phi = f(r)\cos(\theta)\exp(ikz - i\omega t),$$
$$H_r = h_r(r)\sin(\theta)\exp(ikz - i\omega t),$$
$$H_\theta = -h_r(r)\cos(\theta)\exp(ikz - i\omega t),$$
$$H_z = h_z(r)\sin(\theta)\exp(ikz - i\omega t),$$
(28)

where the radial functions $f(r)$, $h_r(r)$, and $h_z(r)$ are given by

$$f(r) = A_0 Z_n(\alpha r) + B_0 W_n(\alpha r),$$
$$h_r(r) = A_1 Z_{n+1}(\beta r) + B_1 W_{n+1}(\beta r),$$
$$h_z(r) = A_2 Z_n(\beta r) + B_2 W_n(\beta r).$$
(29)

The functions $Z_n(\alpha r)$ and $Z_n(\beta r)$ represent either regular or modified Bessel function of the first kind depending on the sign of $\alpha^2 = \omega^2/c_1^2 - k^2$ and $\beta^2 = \omega^2/c_t^2 - k^2$. Similarly $W_n(\alpha r)$ and $W_n(\beta r)$ are either regular or modified Bessel function of the second kind. An additional solution to equation (27) can be obtained by replacing $\cos(n\theta)$ for $\sin(n\theta)$ and $\sin(n\theta)$ for $-\cos(n\theta)$ in the solutions for the potentials Φ and $\vec{H}$ (equation (28)). This corresponds to a rotation of the displacement pattern by $\pi/2$ about the tube axis which results in an additional displacement pattern that is orthogonal to the first.

In the case of icosahedral viruses, the capsid can be modeled approximately as a spherical shell. Figure 2(c) illustrates the geometry for the spherical shell and its relation to an empty icosahedral virus which has inner and outer radii of a and b. In this case, the solution to equation (20) can be written as $\vec{u}(\vec{r}, t) = \vec{u}_0(\vec{r})e^{-i\omega t}$ with $\vec{u}_0(\vec{r})$ given by [27–29]

$$\vec{u}_0(\vec{r}) = A_0\vec{\nabla}\phi + A_1\vec{L}\psi + A_2\vec{\nabla} \times \vec{L}\psi$$
$$+ B_0\vec{\nabla}\phi' + B_1\vec{L}\psi' + B_2\vec{\nabla} \times \vec{L}\psi'$$
(30)

where the operator $\vec{L} = \vec{r} \times \vec{\nabla}$. The functions ϕ and ψ satisfy separate longitudinal and transverse wave equations respectively. The solutions can be written in terms of spherical harmonics and spherical Bessel functions

$$\phi = j_l(\omega r/c_l)Y_{lm}(\theta, \phi),$$
$$\phi' = n_l(\omega r/c_l)Y_{lm}(\theta, \phi),$$
$$\psi = j_l(\omega r/c_t)Y_{lm}(\theta, \phi),$$
$$\psi' = n_l(\omega r/c_t)Y_{lm}(\theta, \phi),$$
(31)

where j_l is a spherical Bessel function of the first kind and n_l is a spherical Bessel of the second kind.

The discrete frequencies and corresponding displacement patterns that are available to the spherical shell or hollow tube can be obtained by applying boundary conditions to the displacement patterns in either equation (29) or (30). For proteins with free boundaries, the boundary conditions are that the normal component of the stress tensor vanish at the surface; i.e. $\overleftrightarrow{\tau} \cdot \hat{n} = 0$. Requiring this to be true everywhere on the surface implies that the three components of the stress tensor $\tau_{r,i} = 0$ (with $i = r, \theta, z$ for cylindrical coordinates and $i = r, \theta, \phi$ for spherical coordinates) must separately vanish at the inner and outer radii a and b, respectively. Enforcement of these boundary conditions will determine the frequencies ω

and the six coefficients A_0, A_1, A_2, B_0, B_1, and B_2 that can be used in equation (29) for the tube or equation (30) for the sphere to determine the corresponding displacement patterns. The boundary conditions can be written in the form of a 6×6 matrix equation, $\overleftrightarrow{C} \cdot \vec{A} = 0$, where $\vec{A}$ is the column vector of coefficients, $\vec{A} = (A_0, A_1, A_2, B_0, B_1, B_2)$. The elements of the matrix will depend on ω, the inner and outer radii a and b, and the speeds of sound c_t and c_l. In the case of the hollow tube, the matrix elements also depend on the wavenumber k.

In practice, one determines the discrete frequencies and their displacement patters by varying ω until the determinant of the matrix $\overleftrightarrow{C}$ vanishes. Once the determinant vanishes, the value of ω corresponds to a natural frequency of the tube or sphere. The matrix $\overleftrightarrow{C}$ can then be diagonalized for this value of ω. Any eigenvectors of the matrix $\overleftrightarrow{C}$ which have zero eigenvalue give a set of coefficients, $\vec{A} = (A_0, A_1, A_2, B_0, B_1, B_2)$, which can then be used in either equation (29) or (30) to determine the displacement patterns.

In summation, all that is required to determine a set of discrete frequencies and displacement patterns for any virus capsid in a continuum model are just four parameters: the inner and outer radii a and b, and the transverse and longitudinal speeds of sound c_t and c_l. Usually the transverse and longitudinal speeds of sound are not known *a priori*, and require a reasonable estimate to be used. Speeds of sounds of the lysozyme protein have been experimentally measured to be $c_t = 915$ and $c_l = 1817$ m s^{-1} [30]. These measurements provide a reasonable estimate of the speeds of sound in protein structures which can them be used to model the vibrational modes of a virus or large protein structure.

3.2. Applications of continuum models

Using the mathematical theory outlined above, one can construct the continuum vibrational modes for either spherical or tubular viruses. As was seen above, for the case where the shape of the virus can be modeled by a hollow spherical or cylindrical tube, the vibrational modes are completely determined analytically in the continuum framework and depend on just four parameters—the transverse and longitudinal speeds of sound c_t and c_l, and the inner and outer radii a and b of the spherical shell or hollow tube.

Balandin and Fonoberov performed some of the first predictions of the vibrational modes of tubular viral particles such as M13 and tobacco mosaic virus using continuum elastic theory [31]. Modeling M13 as a solid tube and TMV as a hollow tube, their work focused on prediction of the dispersion relations for M13 and TMV in mediums of water and air. Interestingly, their results on damping suggest that radial vibrations are in the overdamped to slightly underdamped regime with a quality factor of $\text{Re}(\omega)/\text{Im}(\omega) = 3.6$ [31].

Dykeman, Sankey, and Tsen also used the continuum elastic model to predict the vibrational modes of M13 [32]. However, in contrast to Balandin and Fonoberov's work, M13 was modeled instead as a hollow tube. Dispersion relations for $n = 0, 1, 2$ are shown in figure 3. In addition, the work also extended the popular bond polarizability model [33–36]
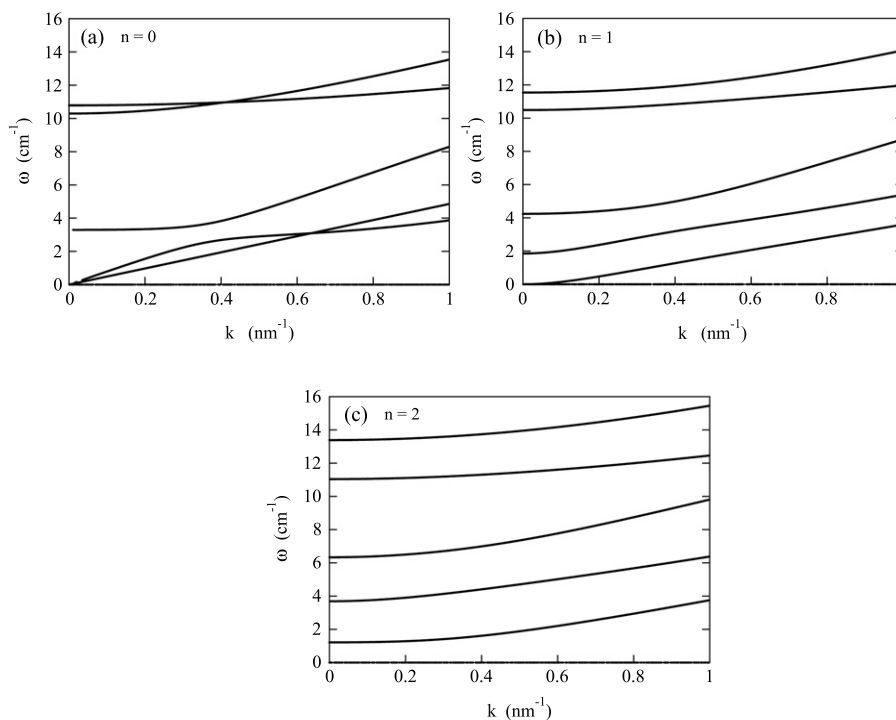
Figure 3. Dispersion relations for the M13 capsid. The wavenumber k is in nanometers and frequencies are in cm^{-1}. The dispersion curves are shown for (a) $n = 0$, (b) $n = 1$ and (c) $n = 2$.

used to predict Raman scattering profiles to an 'amorphous isotropic bond polarizability model' (AIBP) for use in a continuum elastic framework. This allowed the vibrational modes predicted from a continuum theory to be used to predict the corresponding Raman scattering cross section from each mode. A plot of the predicted Raman scattering profiles within the AIBP model can be seen in figure 4 for modes with $n = 0, 1, 2$. The AIBP predicts radial modes with $n = 1$ are Raman silent while radial modes with $n = 0, 2$ and axial modes with $n = 1$ contribute to the Raman cross section. The largest peak is from an $n = 1$ axial shearing mode at roughly 10 cm^{-1}.

More recently, Yang *et al* [29] predicted the vibrational modes of several icosahedral viruses and an icosahedral enzyme, lumazine synthase, using the generalizations of the continuum model formulas outlined above. By using vector spherical harmonics and their representations in the icosahedral group, they were able to classify normal modes computed in a continuum theory in terms of the irreducible representations of the icosahedron. Additionally, by fitting predictions for the mode frequencies from continuum elastic models to those from an anisotropic network model, Yang *et al* were able to obtain values of the ratio of the Young's modulus to the microscopic force constant E/γ and the Poisson ratio σ.

4. Elastic network models

Elastic network models (ENM) of normal modes are arguably one of the most popular methods today for computing the normal modes of a protein due to its simplicity and relative accuracy. It is especially useful for uncovering floppy mode displacement patterns. Their are many variations of ENM's,

but the central idea relies on reducing the number of degrees of freedom available to the protein. This reduces the size of the dynamical matrix allowing it to be diagonalized using standard methods. In this section we highlight the basic theory of elastic network models and discuss two popular methods for coarse graining the molecule. We then provide several examples of the application of such models to a wide variety of problems.

4.1. Elastic network model theory

The elastic network model was originally pioneered by Tirion [37] to model the large amplitude (or equivalently low frequency) vibrational modes of a molecular structure using a phenomenological model for the potential energy. The energy model consisted of a simple single parameter Hookean spring potential between all pairs of atoms that were separated by a distance less than a user specified cutoff distance r_c. In the following, we use the term 'atom' to refer to the objects connected by the spring which can be either individual atoms (hydrogen, nitrogen, carbon, etc) or a collection of atoms. Generally, the usefulness of the ENM is its ability to course grain the system so that each 'atom' represents a collection of atoms in an amino acid (often referred to as a bead), that move more-or-less rigidly together. An extra bonus of the ENM is that it determines the modes without the initial minimization step that is required in all-atom full force field methods. This allows raw x-ray coordinates of proteins to be used in the normal mode analysis.

The single parameter model is extremely simple, and surprisingly accurate in determining the low frequency displacement patterns of proteins. Since the contribution to the potential energy from each atom pair takes the form of
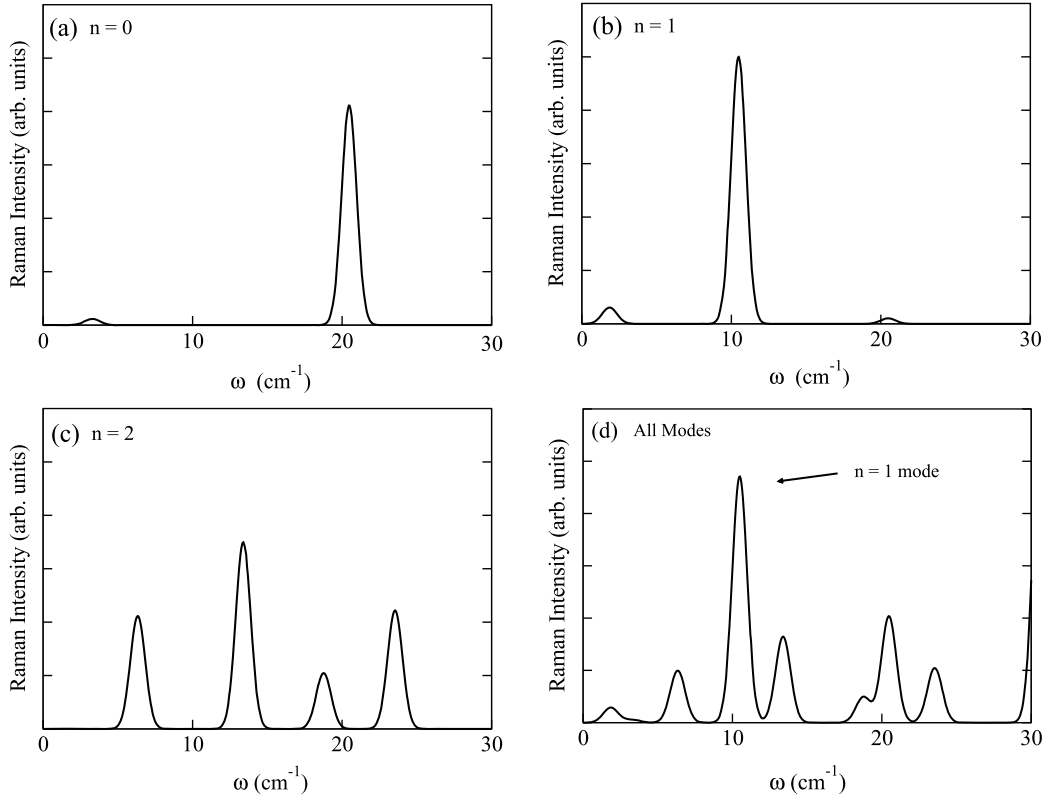
Figure 4. Raman scattering cross section predictions for the M13 virus capsid in the amorphous isotropic bond polarizability model (AIBP). Raman cross sections are illustrated for (a) $n = 0$, (b) $n = 1$ and (c) $n = 2$. (d) Combined Raman scattering intensity from $n = 0, 1,$ and 2.

$V = \frac{1}{2}k_e(d - d_e)^2$ where $d = |\vec{r}_i - \vec{r}_j|$ and d_e is the distance between the atoms at equilibrium, the dynamical matrix elements can be written in terms of the direction cosines of the bond between atom pairs as

$$D_{i\alpha,i\beta} = \frac{k_e}{m_i} \sum_j' \frac{r_{i\alpha} - r_{j\alpha}}{|\vec{r}_i - \vec{r}_j|} \frac{r_{i\beta} - r_{j\beta}}{|\vec{r}_i - \vec{r}_j|} \Delta_{ij}$$

$$D_{i\alpha,j\beta} = -\frac{k_e}{\sqrt{m_i m_j}} \frac{r_{i\alpha} - r_{j\alpha}}{|\vec{r}_i - \vec{r}_j|} \frac{r_{i\beta} - r_{j\beta}}{|\vec{r}_i - \vec{r}_j|} \Delta_{ij}. \tag{32}$$

The value k_e is the universal spring constant in the ENM, m_i is the mass of atom i, and the labels α and β denote one of the directions x, y, or z. The prime in the summation indicates that the sum is over all atoms j excluding the $j = i$ contribution. The term Δ_{ij} is a Heaviside step function used to restrict contributions to pairs that are separated by a distance less than a specified cutoff r_c i.e.

$$\Delta_{ij} = \begin{cases} 1 & \text{if } |\vec{r}_i - \vec{r}_j| < r_c \\ 0 & \text{otherwise.} \end{cases} \tag{33}$$

Note that in the potential energy used for the ENM above, motions which change the orientation of the bond direction do not change the potential energy. This potential energy model is the most popular and is often referred to as the 'anisotropic network model' [38]. Another choice for the potential energy between atom pairs i and j is $V = \frac{1}{2}\gamma(\vec{\Delta} \cdot \vec{\Delta})$ where $\vec{\Delta} = (\vec{r}_i - \vec{R}_i) - (\vec{r}_j - \vec{R}_j)$ and $\vec{R}_i$ is the equilibrium position of

atom i. When this potential energy is used in an ENM, the model is referred to as a 'Gaussian network model'. As one can see, motions which change either the bond orientation or its length contribute to a change in the potential energy in the Gaussian network model. However, a peculiarity of Gaussian network models (as noted in a comment by Thorpe [39]) is that the potential energy term in Gaussian network models do not enforce rotational invariance.

Over time, the ENM has come to be synonymous with a coarse graining of the molecule with the use of Tirion's single parameter potential. As discussed in section 2, the frequencies and displacement patterns of a molecule with N atoms is determined by diagonalizing the $3N \times 3N$ dynamical matrix. Many proteins and protein complexes, such as viruses, have hundreds of thousands or even millions of atoms, making a direct diagonalization of the dynamical matrix prohibitive on today's computers. Coarse graining reduces the number of degrees of freedom of the molecule by either neglecting the motions of certain atoms or by clumping sets of atoms together. We highlight the most popular methods in each category, specifically: (i) the C_α method and (ii) the rotation translational block method (RTB) [40].

In the C_α method, one only considers the motion of the C_α carbon atom in each amino acid of a protein. For a protein that has N_R amino acid residues, this results in a new coarse grained 'molecule' which has $N_R C_\alpha$ atoms with three degrees of freedom each. Since the average amino acid residue has approximately 20 atoms, this method of coarse graining can reduce the number of atoms in a molecule by roughly a factor
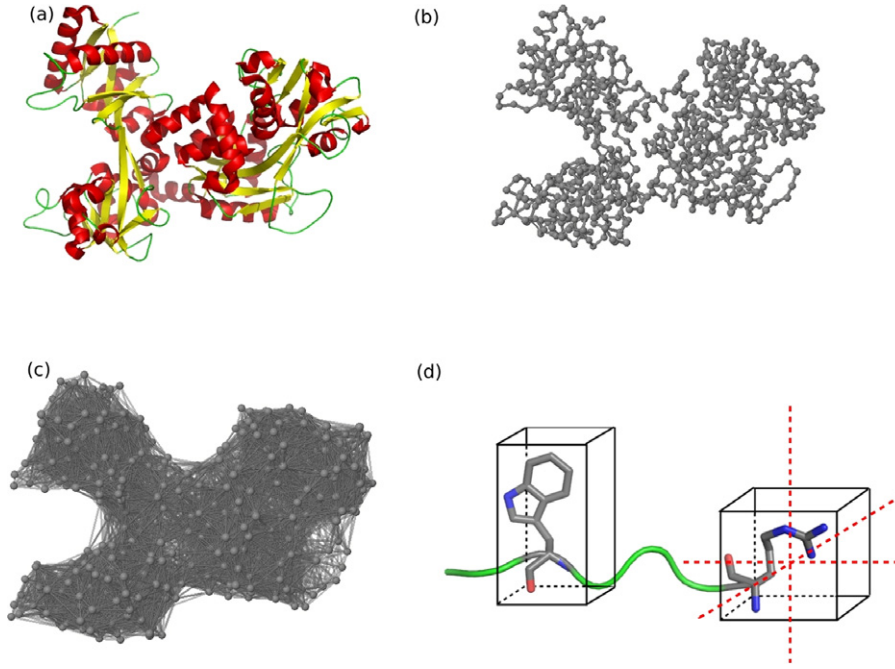
Figure 5. Elastic network representations of the lactoferrin protein (PDB ID 1LFH). (a) Ribbon diagram of the lactoferrin protein. Beta sheet and alpha helix structures are thick ribbons and coils respectively while uncoiled loops are thin lines. (b) Structure of the lactoferrin protein where only C_α atoms are shown. (c) Elastic network connections between the C_α atoms used to construct the Tiron potential in a C_α only elastic network model. Atoms within a cutoff of 8 Å are shown connected via a 'bond'. (d) Example of the rotational translation block coarse graining method where each amino acid constitutes a rigid 'block' that can undergo rotations and translations about three principal axes (shown in the dashed line).

of 20, or the dynamical matrix by a factor of 60^2. Figure 5 illustrates how a network of C_α carbon atoms is formed for the lactoferrin protein, shown in a ribbon diagram in figure 5(a). The backbone C_α atoms for the lactoferrin protein are shown in figure 5(b). The backbone C_α atoms are then connected by a spring with spring constant k_e if they are separated by a distance less than r_c. The springs (drawn as a C–C bond) between C_α atoms in figure 5(c) represent one of these identical elastic network springs. In the C_α method, a typical value for the cutoff is $r_c = 8$ Å.

In the RTB method, one 'clumps' a number of atoms together to create a rigid unit with six degrees of freedom, i.e. three rotations and three translations. An example is shown in figure 5(d) for the lactoferrin protein. Here each amino acid in the protein is a rigid object which can translate and rotate about each of the three principal axes represented by the dashed red line. Representing the motions of the amino acids in this way allows one to construct a 'reduced basis set' of vectors which can be used to form the displacement patterns for all atoms in the protein. A reduced basis set of vectors is simply any set of M orthogonal vectors $|\vec{v}_i\rangle$ with $i = [1, M]$ of length $3N$ where N is the number of atoms in the molecule and $M < 3N$. This is in contrast to all-atom methods which use the 'full basis set' i.e. the set of $3N$ orthogonal vectors that can be formed by considering the three translations for each atom in the molecule. Once one has the M orthogonal basis vectors, the dynamical matrix is reduced by applying the similarity transform

$$\overset{\leftrightarrow rtb}{D} = \overset{\leftrightarrow\dagger}{V}\overset{\leftrightarrow}{D}\overset{\leftrightarrow}{V} \tag{34}$$

where $\overset{\leftrightarrow}{V}$ is the $M \times 3N$ matrix formed using the 'reduced basis set' of M orthogonal basis vectors $|\vec{v}_i\rangle$. The rotational translational block dynamical matrix $\overset{\leftrightarrow rtb}{D}$ is then a smaller $M \times M$ matrix that can be diagonalized producing $\nu = [1, M]$ squared frequencies ($\omega_\nu^2 = \lambda_\nu$) with corresponding eigenvectors of length M; $\hat{e}^{rtb}(\nu)$. The frequencies ω_ν^2 directly approximate M of the $3N$ frequencies of the full dynamical matrix while the M components of its eigenvectors $\hat{e}^{rtb}(\nu)$ can be used construct M approximated normal modes $|\hat{e}_\nu^*\rangle$ using the reduced basis set;

$$|\hat{e}_\nu^*\rangle = \sum_{i=1}^{M} |\vec{v}_i\rangle\hat{e}_i^{rtb}(\nu). \tag{35}$$

This method of coarse graining is also adaptable to models other than the ENM which use a reduced basis set, such as the basis set of dihedral motions [41], but compute the dynamical matrix elements from the full atomic force field instead of Tiron's phenomenological force field constructed from a network of identical springs between C_α atoms. This type of model is also discussed in the all-atom methods section below. Within the ENM framework, the RTB method usually considers each amino acid as a rigid block, but in the case of viruses, entire proteins have been considered as rigid blocks or 'beads' [42].

We also mention a group theoretic approach for viral capsids by Peeters *et al* [43]. In this approach the coarse graining treats an entire protein as a single bead. However unlike Tama [42] in which the potential energy is computed

from the network of identical springs between all C_α atoms in the proteins, the beads are connected by single springs where the strength of the springs scales with the association energy between the proteins. This minimalist approach gives an overall picture of how tiling of the triangular faces of the icosahedron affect the global modes.

4.2. Applications of elastic network models

ENM's have been widely used in a large variety of biological problems ranging from the determination of functional protein motions [44–46] to the fitting of high resolution protein to a low resolution cryo-EM data in order to facilitate understanding of functional motions of protein complexes [47]. The success of the ENM in such a variety of situations and its overall simplicity have made it an important tool in theoretical biophysics.

Perhaps one of the most well known applications of the elastic network model has been in the study of the large motions of viral capsids [42, 48]. The large size of viral capsids have made ENM's a useful starting point in the study of the vibrational modes of capsids. Tama and Brooks were some of the first pioneers in this area. A comprehensive review on using the elastic network model to predict the vibrational modes of several viruses is presented in [42]. Figure 6 shows the two lowest non-degenerate modes for the hepatitis B virus that were calculated by Tama and Brooks [42] using an ENM.

In addition to viral capsids, the ENM has also been successfully used in the study of functional motions of other large protein structures such as the citrate synthase enzyme [44], and a mechanosensitive ion channel in *E. coli* MscL [45, 49]. The citrate synthase enzyme is an important protein that performs the first catalytic steps in the production of ATP from acetyl-CoA. The different functional motions that the enzyme undergoes in response to binding of different substrates is important for the regulation of binding acetyl-CoA and the subsequent production of citroyl-CoA. Similarly, the mechanosensitive ion channel MscL acts as a critical 'safety' valve for *E. coli* when osmotic pressure builds on the cell wall due to a change in environmental salt concentrations. Osmotic pressure on the cell wall induces a conformational change allowing the transmembrane pore to open and reestablish an equilibrium between the concentration of ions such as Na^+ inside and outside the cell wall. Both are prime examples of the work that has been done on studying the functional motions in proteins using the elastic network model.

Another example where the elastic network model has also been used successfully is in the prediction of flexible and rigid regions of proteins. This is usually done by calculating the B-factors of amino acids in the protein using the formulas discussed in section 2. Amino acids which have a large B-factor are undergoing large root mean squared fluctuations away from their equilibrium position and are thus predicted to be very flexible. Likewise, amino acids with a small B-factor values have small RMS deviations and are areas of the protein that are rigid. Bahar *et al* show that the B-factors for several proteins can be predicted using a simple ENM with the Gaussian network model potential [50]. As with all
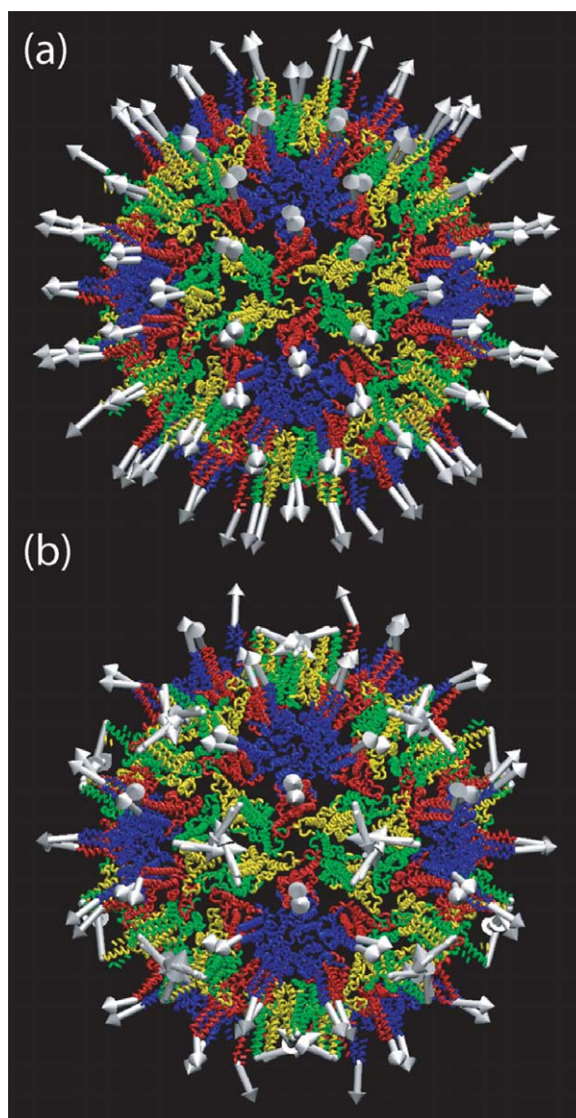

Figure 6. Diagram of the lowest non-degenerate normal modes of the hepatitis B virus (reproduced with permission from [42]. Copyright 2005, Elsevier.) (a) Lowest frequency non-degenerate mode and (b) second lowest frequency non-degenerate mode of the hepatitis B virus.

ENMs, the values for the frequencies produced by the ENM are arbitrary due to the free choice of the elastic network spring constant. Bahar *et al* show that reasonable agreement between theoretical and experimental B-factors can be achieved by adjusting the spring constant which ranges from 0.99 to 3.0 in their study.

One of the most recent applications of the elastic network model involve studying conformational changes in protein structures by flexibly fitting high resolution protein structures into low resolution cryo-electron microscopy images [47]. For example, the cowpea chlorotic mottle virus (CCMV) is a small $T = 3$ plant virus that undergoes a thermodynamically reversible swelling of its capsid from a 'closed' unswollen state to a 'open' swollen state which has 60 pores [48]. This motion is believed to be functionally important to the infection process by allowing the internal ssRNA to be released into the infected cell. The 'closed' state has been successfully crystallized and,

using standard x-ray diffraction, visualized to a high resolution of 2.8 Å [51]. However, the 'open' swollen state has only been able to be visualized using low resolution cryo-electron microscopy to a resolution of ∼28 Å. The flexible fitting procedure involves moving the proteins in the capsid of the high resolution 'closed' structure along the normal modes in small steps, until the predicted EM density of the structure closely fits the low resolution cryo-EM data. In this manner, one can both construct a best fit atomic model for the low resolution data, but also learn about which modes are important for the functional motion.

5. Principal component analysis—PCA

A major goal of normal mode analysis is to reduce the complexity of the full dynamics of a complex system, and to describe it in terms of a few generalized coordinates such as normal modes. The low frequency modes will exhibit larger excursions from equilibrium and these floppy modes play an especially important role. However, the usefulness of the normal modes is limited because of damping effects due to water and anharmonicity (which exists even in vacuum). There is an entirely different approach to the problem, and that is in terms of principal components (PCs). This alternate method is capable of reducing the dimensionality to a few components and is not restricted to harmonic systems—i.e. the system can be heavily damped or not oscillate at all. The method is called principal component analysis (PCA). PCA and is a fairly old technique [52] with a much wider range of applications than is described here. The method in effect computes second moments of a multivariate distribution and describes deviations from the average in terms of a set of principal components, which represent the collective motions that produce the largest deviations. The principle components are eigenvectors of a covariance matrix (described below) and the vectors with the largest eigenvalues are those producing the largest variance—thus the largest excursions from some average geometry.

Imagine we simultaneously both flip a coin and toss a six sided die. We add their scores (the coin score is $+1$ for heads and -1 for tails). A principal component analysis will produce the (obvious) result that the principle components of the variance are individually the die and the coin, and the component with the largest eigenvalue, which is the main source of the variance, corresponds to the die. This all appears unenlightening—but if instead one is dealing with thousands of variables, PCA can pick out which 'modes' are dominating the fluctuations. A mode then corresponds to a concerted motion of the masses producing large deviations away from the average structure.

The use of PCA is similar to the concept of 'essential dynamics' within molecular dynamics simulations [53]. Essential dynamics separates configuration space into two subspaces—the 'essential' subspace contains a few degrees of freedom in which anharmonic motion occurs that comprise most of the positional fluctuations, and the constrained region which in Gaussian distributed (normal mode like). Other ways of obtaining the 'essential' subspace containing the large positional fluctuations is through rigidity theory [54].

Thus, PCA is a technique to display the fundamental modes of motion contained in the complex dynamics of a biomolecule in equilibrium with its surroundings. Dynamical hinge like opening and closing modes can be obtained which are important for understanding the functionality of enzymes and biomolecules which operate as nanomachines. The method usually uses molecular dynamics (MD) simulations to explore the dynamics. The PCA is performed as a post-processing step of the trajectory data generated. The dynamics occurs in water where normal mode analysis has limited applicability, so PCA offers a very powerful method for developing a conceptual understanding of important floppy regions of the system responsible for large dynamic distortions and transitions between conformers in realistic settings.

The core quantity in PCA is the covariance matrix,

$$C_{\alpha,\beta} = \langle (r_\alpha - \langle r_\alpha \rangle)(r_\beta - \langle r_\beta \rangle) \rangle \qquad (36)$$

where r_α is one of the $3N$ atomic coordinates of the system. The matrix $\overleftrightarrow{C}$ is $3N \times 3N$, is symmetric (hence Hermitian) and thus has real eigenvalues and orthonormal eigenvectors. There are two types of averages typically performed. The first type is over an MD trajectory (or several trajectories). The second type is to perform the average over several static structures such as different conformations of the same protein or very similar proteins. The different conformations in the latter type are due perhaps to ligand binding and the structures come from x-ray or other structural measurements.

We first consider dynamic systems and relate PCA to harmonic analysis. But remember that PCA does not require harmonic systems. Both PCA and the thermodynamics of normal modes describe fluctuations, and it is instructive to see the relationship between the two. We finally describe a use of PCA from MD on ubiquitin and then describe a use to categorize static structures in the Ras protein.

We consider the case of an MD simulation of a *harmonic* system. Both PCA and NMA must produce the same description. Let us see how this occurs. The system oscillates around equilibrium and so $r_\alpha - \langle r_\alpha \rangle$ reduces to u_α, the harmonic distortion from equilibrium. The displacement correlation matrix $C_{\alpha\beta}$ is formed from the particles' trajectory,

$$C_{\alpha\beta} = \langle u_\alpha u_\beta \rangle. \qquad (37)$$

The average is over the long-time trajectory of the MD simulation which is simulated at constant T using a thermal bath. (The displacements $\vec{u}$ have been corrected by removing the center of mass motion and rigid body rotations.) The matrix $C_{\alpha\beta}$ is the central quantity evaluated. But what information does it contain about dynamical modes?

A clear interpretation of $C_{\alpha\beta}$ is obtained when we expand the displacements into normal modes. The displacement vector of length $3N$, $|\vec{u}\rangle$, when expanded in terms of the normal modes results in the correlation matrix

$$C_{\alpha\beta} = \sum_\nu \langle Q_\nu^2 \rangle (\eta_\nu)_\alpha (\eta_\nu)_\beta. \qquad (38)$$

Here $(\eta_\nu)_\alpha$ is the αth component of the displacement vector $|\vec{\eta}_\nu\rangle$ and $\langle Q_\nu Q_{\nu'} \rangle = 0$ for $\nu \neq \nu'$. Recalling that the

squared deviations $\langle Q_\nu^2 \rangle$ in thermal equilibrium are $k_{\rm b}T/\omega_\nu^2$ (see equation (16)), the correlation matrix can be written in a compact notation as

$$\langle u_\alpha u_\beta \rangle = k_{\rm b}T \langle \alpha | \overset{\leftrightarrow}{\lambda} | \beta \rangle. \tag{39}$$

Here $|\alpha\rangle$ is the unit vector $(0, 0, \ldots, 0, 1, 0, \ldots 0)$ where the 'one' is at the site corresponding to the index α. The $\overset{\leftrightarrow}{\lambda}$ matrix is

$$\overset{\leftrightarrow}{\lambda} = \sum_\nu{}' |\vec{\eta}_\nu\rangle \frac{1}{\omega_\nu^2} \langle \vec{\eta}_\nu|. \tag{40}$$

We have added a prime to the sum to indicate that the zero frequency modes (rigid body translation and rotation) are omitted from the mode sum. Zero frequency modes of course make the dynamical matrix singular with no inverse. But all other modes (in this derivation) are included.

This analysis gives us the connection we seek, and that is that the correlation matrix is proportional to the sum over all modes weighted by the inverse frequency squared (*sans* the zero frequency modes) with proportionality constant $k_{\rm b}T$,

$$\overset{\leftrightarrow}{C} = k_{\rm b}T \sum_\nu{}' |\vec{\eta}_\nu\rangle \frac{1}{\omega_\nu^2} \langle \vec{\eta}_\nu|. \tag{41}$$

This central relationship adds a foundation for meaning and understanding of the correlation matrix. The floppy modes with small ω now contribute to large eigenvalues of the correlation matrix. Thus they stand out prominently in the analysis.

We take as a toy model to illustrate the connection between the covariance and dynamical matrices two equal masses m connected by a spring k which are allowed to move in one dimension. There is one non-zero frequency mode, $\omega_0^2 = 2k/m$, with eigenvector $|\hat{e}_0\rangle = (1, -1)/\sqrt{2}$ and displacement vector $|\vec{\eta}_0\rangle = \frac{1}{\sqrt{2m}}(1, -1)$. The displacement vector (without the zero frequency mode corresponding to the center of mass translation) is $|\vec{u}\rangle = \frac{Q(t)}{\sqrt{2m}}(1, -1)$. The correlation matrix $\overset{\leftrightarrow}{C}$ has components $C_{11} = \langle u_1^2 \rangle = Q(t)^2/(2m)$. Similarly, the other matrix elements are $C_{22} = \langle u_2^2 \rangle = Q(t)^2/(2m)$ and $C_{12} = C_{21} = \langle u_1 u_2 \rangle = -Q(t)^2/(2m)$. Recalling that the average $\langle Q^2 \rangle$ is $\frac{k_{\rm b}T}{\omega_0^2}$ (see equation (16)) the $\overset{\leftrightarrow}{C}$ matrix is

$$\overset{\leftrightarrow}{C} = \frac{k_{\rm b}T}{2m\omega_0^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \tag{42}$$

This covariance matrix has eigenvalues $\frac{k_{\rm b}T}{m\omega_0^2}$ and 0. The zero eigenvalue reflects the stationary center of mass. The eigenvector for the dynamical portion is $(1, -1)/\sqrt{2}$. Thus the correlation matrix has eigenvalues proportional to the inverse squared of the dynamical matrix and the eigenvectors (displacement patterns) are identical to the dynamical matrix.

One of the early uses of PCA from molecular dynamics was by Hayward *et al* [55]. They followed in MD the structure of bovine pancreatic trypsin inhibitor (BPTI) protein in vacuum and determined the covariance matrix $C_{\alpha\beta}$. An effective frequency is defined for each of the PCs and the frequency is given by $\omega_{\rm Eff}^2(i) = k_{\rm b}T/\lambda_i$ where λ_i are the eigenvalue for the i'th principal component. Analysis of the trajectory in terms of PCs shows that the first principal component describes an apparent barrier crossing phenomenon, while the trajectory projected onto higher PCs show Gaussian-like distributions similar to a quasi-harmonic theory. The key here is that only a small number of effective PCs are needed to describe the motion. This theme arises frequently in this area (either PCA or normal modes) and that is the fact, or hope, that a few generalized coordinates are useful to describe the complex dynamic behavior.

More recently Ramanathan *et al* [56] have compared PCA (in a form known as the quasi-harmonic analysis (QHA)) and normal mode analysis for ubiquitin from MD using eight different starting structures. Their goal is to model the slow conformational motions from an ensemble of runs totaling 0.5 μS of MD simulations and compare them to a ensemble of 116 structures from NMR refinements. The quasi-harmonic modes are similar to PCA but are determined from the similar matrix

$$F_{\alpha\beta} = m_\alpha^{1/2} m_\beta^{1/2} \langle (r_\alpha - \langle r_\alpha \rangle)(r_\beta - \langle r_\beta \rangle) \rangle \tag{43}$$

instead of $C_{\alpha\beta}$, making it similar to the mass weighted eigenvectors of the dynamical matrix. The compact form of $\overset{\leftrightarrow}{F}$ for a harmonic system is $\overset{\leftrightarrow}{F} = k_{\rm b}T \sum_\nu{}' |\hat{e}_\nu\rangle \frac{1}{\omega_\nu^2} \langle \hat{e}_\nu|$ which shows a one-to-one connection with eigenvalue and eigenvectors of the dynamical matrix. Eight different starting x-ray structures of ubiquitin were used corresponding to PDB ID codes of 1UBQ, 1P3Q, 1S1Q, 1TBE, 1Y1W, 2D3G, 2FCQ, and 2G45. The entire ensemble was used to find the QHA modes. They find that ten QHA modes contribute 78% of all the dynamical properties. Their modes describe the fluctuations of the structures in the NMR ensemble of structures.

Figure 7 shows the projection of the QHA modes of the MD ensemble and the NMR ensembles. Most of the NMR structures fall into a subclass of all the MD trajectories, but some MD trajectories sample regions of QHA-space not sampled by the NMR ensemble. Their comparison of the QHA to BNMA highlights advantages of QHA. These advantages center on the QHA averaging over a much larger configurational space and is better able to describe the transitions that occur in an enzyme.

PCA need not be applied to dynamic fluctuations, but can be used to explore a mapping of many different conformers or mutants of a family of proteins. There has been recent work by Gorfe *et al* [57] exploring the variety of structures associated with Ras kinase proteins. Ras proteins are switches that affect cell growth and are important in cancer. Ras proteins bind GTP/GDP and their binding affects the shape of the protein allowing signaling process to occur within the cell. Gorfe *et al* studied a set of over 40 different x-ray structures containing different conformers. Specifically they find that the variance of a specific domain, called the catalytic domain, can be described at the 57% level by just two principal components. Continuing further, 88% of the variance is contained in eight dimensions. Using just 2 PCs, a separation of Ras was clearly visible concerning the state of the bound ligand—those with
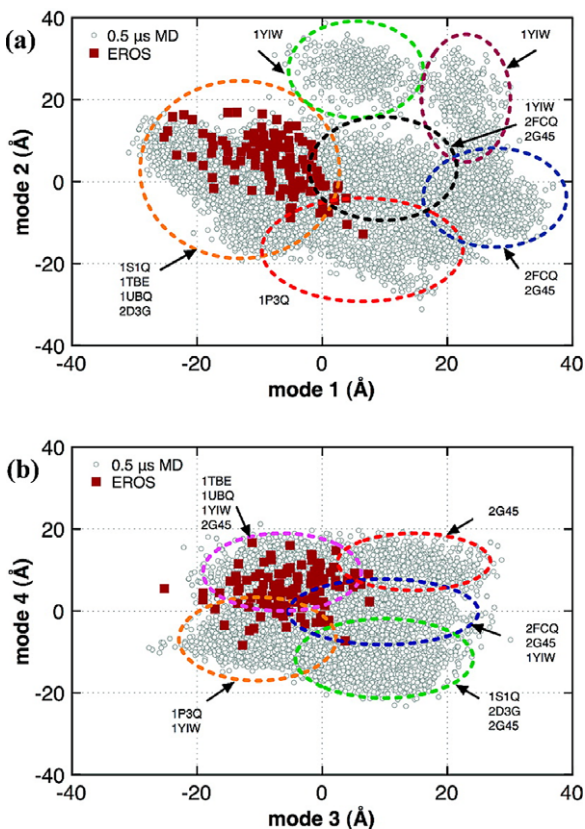
Figure 7. Quasi-harmonic approximation (QHA) of modes for ubiquitin from [56]. Plotted are (a) mode 2 versus mode 1 and (b) mode 3 versus mode 4. These projections highlight the similarity of the structures explored in various MD trajectories or for the similarities of structures in the ensemble of 116 NMR structures (red squares). Each of the MD runs started in from one of the eight different structures (PDB ID codes 1UBQ, 1P3Q, 1S1Q, 1TBE, 1Y1W, 2D3G, 2FCQ, 2G45), and the circled sections indicate regions primarily sampled by an MD trajectory from one of the specific eight starting structures.

GTP generally resided in one region within this subspace while those with GDP resided in another. Such a conformer plot [58] is extremely useful in highlighting similarities and dissimilarities in structure and their relationships.

6. All-atom phonon functional method

We have discussed several approximate methods to determine the vibrational modes of large complex biomolecules. There has been a recent development, called the all-atom phonon functional (AAPF), that allows all-atom calculations of the vibrational modes to be computed from an empirical force field. As there are many empirical force fields, this has not been the traditional barrier to solving this problem. The traditional barrier has been the huge size of the dynamical matrix. Even the elastic network model has difficulty with large complexes of proteins such as viral capsids. The phonon functional is a general technique that can be applied to any force field model (of finite range) and it need not be all atom. It applies equally well to a generalized Born all-atom force field model or even to a simplified coarse grained ENM. What is the catch? The

phonon functional only gives a subset of vibrational modes—the lowest frequency ones where the number M is chosen by the user. The computational expense goes up approximately linearly with the number of modes M chosen, at least where M is not too large. A typical number of modes chosen is about 100. These low frequency subset of modes usually are those involved in the 'essential dynamics' of the system or the 'mode displacement subspace' for conformer transitions. In this section we highlight the basic theory behind the phonon functional and its deep ties to electronic structure calculations.

6.1. The phonon functional method

Essentially the phonon functional method is a special matrix diagonalization technique for large matrices that are especially prone to being ill-conditioned. The goal of the phonon functional method is to (i) determine the eigenfrequencies ω_ν^2 and (ii) the eigenvectors $|\hat{e}_\nu\rangle$ (or equivalently the displacement vectors $|\vec{\eta}_\nu\rangle$) exactly (within an empirical energy model), (iii) work primarily with vectors and avoid storage of large matrices, and (iv) do this for a user chosen number of M modes starting with the lowest. We focus here on proteins and show applications to viral capsids; the technique can be applied in any type of normal mode analysis which can be described as 'atoms' (including RTB or C $_\alpha$ models) with any potential energy function.

The phonon functional method is a technique borrowed from electronic structure theory. We will initially frame the problem within this context since it is highly motivated from this viewpoint. The reader can skip to equation (44) to forgo this. The basic bottleneck steps of the electronic structure and vibrational analysis are the same—determine eigenvalues (electron energies or vibrational frequencies) and eigenfunctions (wavefunctions or displacement patterns).

Electronic structure theory, requires wavefunction eigenvectors of a Hamiltonian, H, with a few (perhaps a hundred) occupied states and a very many (perhaps many thousands) of unoccupied excited levels. In density functional methods, the wavefunction for the ground state is written as a product of single particle wavefunctions, with each electron in a different state to satisfy the exclusion principle. The occupied states determine the total energy and are the main focus of attention. The occupied states also determine the charge density of the material. Often only a few excited states are also considered since the affect of low lying excitations determine the doping properties, electron or hole transport, or the optical properties of the material. The difficult quantity to evaluate is G, the sum of single particle energies often called the band-structure energy. It is $G = \sum_{i_{\text{occupied}}} \lambda_i$, where the sum is over occupied (lowest lying) electron orbital energies satisfying $H\psi_i = \lambda_i\psi_i$. Thus G, as defined, requires only the eigenvalues. Generally, however, to get eigenvalues requires computing the eigenvectors. An energy functional that determines G is $G = \text{Tr}(Hf(H,\mu))$ where f is the Fermi function, $f(H,\mu) = \frac{1}{e^{\beta(H-\mu)}+1}$. The trace is over any complete orthonormal basis. The property of the trace makes the functional G invariant to the set of orthonormal basis states used to construct the matrix H. The Fermi function acts as a filter to pick out the

subspace of eigenvectors that are occupied. Thus (assuming low temperature) $f(\lambda, \mu)$ is 1 (occupied) or 0 (unoccupied) depending on whether λ is below or above the Fermi level μ.

This form of G is not useful in practice because the Fermi function acting on a state ψ is difficult to compute. Instead the functional is first changed to $\text{Min}(\text{Tr}_M H)$ which is to minimize the trace over the number M of occupied states (which depends on the number of electrons). The well know problem with this form is that the exclusion principle is not enforced—the minimum is where all electrons drop down to one single orbital (e.g. all electrons of oxygen are in the 1s state). To enforce the exclusion principle the M states must be made orthonormal. This is awkward, at best, since the functional itself is ignorant of these issues and minimization cannot continue in a straightforward fashion.

We now come to the problem at hand—and that is finding the lowest M frequency states of the dynamical matrix which is analogous to the occupied subspace in electronic structure. There are many solutions offered from electronic structure theory and we choose that of Ordejon–Drabold–Martin–Grumbach (ODMG) [59]. The functional is simple in form but it's action is surprisingly powerful. We define a phonon functional G_p as,

$$G_\text{p} = \text{Min}\{\text{Tr}_M(D^s + D^s(1 - S))\}. \tag{44}$$

Here D^s is the dynamical matrix shifted by its largest eigenvalue $(\overset{\leftrightarrow}{D} - \omega_\text{max}^2 \overset{\leftrightarrow}{I})$ which results in a matrix with a completely negative eigenvalue spectrum, and S is the matrix of overlaps of the vectors in the subspace of M basis states. The meaning of the overlap and the reason for shifting the dynamical matrix will soon be made clear. The subscript M on Tr_M is to remind us that the minimization problem is not over the enormously large basis of the full atom system (perhaps with millions of modes), but over the M modes in the basis set (say 100). Thus the matrices D^s and S are small ($M \times M$) and trivially diagonalized. The difficulty is obtaining these matrices and varying the subspace so as to minimize G_p. However, once G_p is minimized, the resulting subspace of M vectors that is produced from the minimization is that of the lowest M frequency modes—exactly!

So what are the matrices D^s and S? We start off by taking a wild guess and form M random vectors of length $3N$ to form a subspace of M dimensions. The M basis vectors we denote as $|u_i\rangle$ where the i is a label for each of the M vectors. These vectors need not be normalized or orthogonal. The vectors must be linearly independent, but choosing them at random from a much larger subspace offers essentially zero chance of them not being linearly independent. We form the small $M \times M$ subspace dynamical matrix D^s and overlap S as

$$D_{ij}^s = \langle u_i | \hat{D}^s | u_j \rangle \qquad S_{ij} = \langle u_i | u_j \rangle. \tag{45}$$

Note that the full dynamical 'matrix' $\hat{D}^s$ is not represented as a matrix—rather it is an operator. This is very similar to standard practice in electronic structure where the Hamiltonian is not represented as a matrix, but rather is an operator that changes a vector into a resultant vector, i.e. $\hat{H}|\psi\rangle = |\text{result}\rangle$. Here, the operator $\hat{D}^s$ acts on a vector $\hat{D}^s|u_j\rangle$ and that resultant vector

is what must be computed. This operation of the dynamical matrix on a vector proceeds in nearly exactly the same fashion as the computation of the force vector in an MD simulation. In an MD simulation, the derivatives $\partial V/\partial x_i$ are generated for each term in each sum in the empirical potential energy model one at a time (see equation (1)) and are added into a work vector of length $3N$. At the end of the loop over all energy terms, the work vector contains the correct force vector. The procedure for the dynamical matrix operating on a vector proceeds in a similar fashion. The second order derivatives $\partial^2 V/\partial x_i \partial x_j$ are generated analytically for each energy term. For the bond energy terms in equation (1), there are 36 such terms for each bond between two atoms in the sum. After computing these derivatives, they are multiplied by the 36 appropriate components of the vector being operated on and added into a work vector of length $3N$. As with the computation of the force, after the loop over all energy terms is complete, the work vector contains the result of the operation of the dynamical matrix on the vector.

As an example, we illustrate how the functional works with the most trivial model—a space of just one dimension. The dynamical matrix D^s is 1×1 with eigenvalue $-\lambda_0$ (recall that the dynamical matrix is modified so as to have a completely negative definite spectrum). The trial vector $|u_1\rangle$ must be proportional to the eigenvector since the space is one-dimensional in this example. The only degree of freedom is simply the norm of the vector. Having chosen a random trial vector, its squared norm is $\langle u_1 | u_1 \rangle = S$. Using $\text{Tr}(D^s) = -\lambda_0 S$ and $\text{Tr}(D^s(1 - S)) = -\lambda_0 S + \lambda_0 S^2$, the energy functional G_p becomes (after completing the square)

$$G_\text{p} = -\lambda_0 + \lambda_0(1 - S)^2 = \text{Term}_1 + \text{Term}_2. \tag{46}$$

Figure 8 plots Term_1, Term_2, and the full functional G_p. The minimum of G_p occurs at $S = 1$ (i.e. $|u_1\rangle$ is the normalized eigenvector $|\hat{e}_1\rangle = 1$) and the value of G_p is $-\lambda_0$, the sum (here just one contribution) of the eigenvalues. This simple example illustrates how the functional works—it produces a kind of energy landscape in displacement vector space whose minimum is the sum of the lowest eigenvectors and the minimization can proceed by varying M states without regard to normalization, or more importantly but not evident in this 1D example, orthogonality. After G_p is minimized, the M vectors found span the subspace of the M lowest frequency modes. A final diagonalization of the (small) $M \times M$ matrix D^s in that subspace yields the linear combination of vectors that are true eigenvectors of the full $3N \times 3N$ dynamical matrix and the eigenvalues are the eigenvalues of the full dynamical matrix.

We next address the iteration procedure that takes us from the 'wild guess' initial set of vectors to a better and better set of vectors until iteration converges the functional to the exact result (up to a tolerance of precision). The concept is to adjust the basis vectors $|u_i\rangle$ defining the 'occupied' subspace of M vectors until G_p is minimized. Optimization problems have many solutions, and we choose a conjugate gradient (CG) method [60]. A computationally expensive part of the CG algorithm is to minimize the function along a chosen line, and this generally requires many function evaluations.
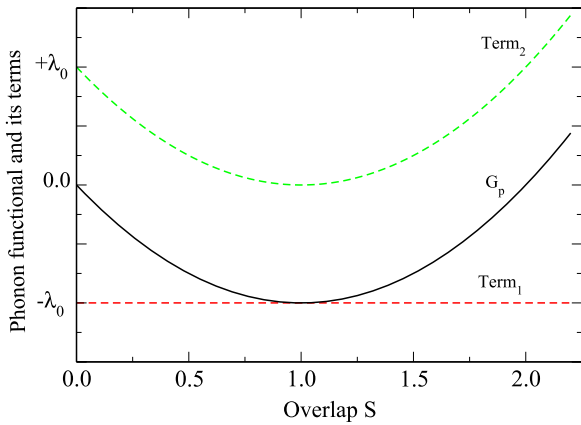
Figure 8. A plot of the phonon functional versus the overlap of the sought after state for a one-dimensional model in which the (shifted) dynamical matrix D^s has just one eigenvalue $-\lambda_0$. The trail eigenvector is a multiple of the single eigenvector, but is not normalized and has squared norm of S. The functional is written as a sum of two terms, each plotted individually. The full functional G_p, which is the sum of the two terms, has a single minimum at eigenvalue $-\lambda_0$ with normalized eigenvector.
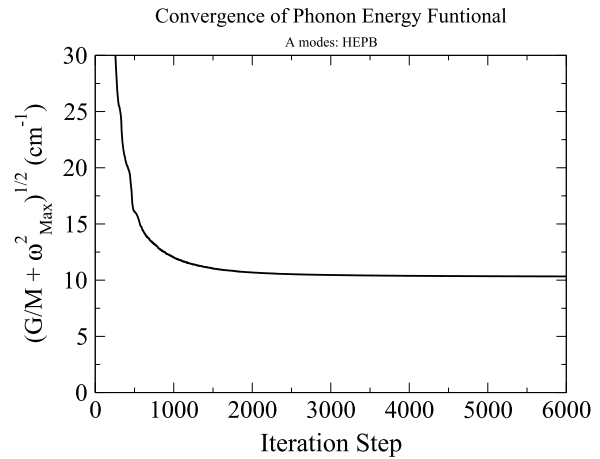
Figure 9. The convergence of the phonon functional G_p for the lowest $M = 100$ modes of A symmetry for the capsid of hepatitis B. The shift $-\omega_{max}^2$ is removed. The plotted quantity, when converged is equal to the RMS average of the first 100 modes. The convergence is smooth and the precision is far better that 0.1 cm^{-1}. The minimization is performed by a conjugate gradient procedure of the random vectors.

Fortunately for the G_p functional, the minimization along the line is analytic, requiring only a cubic equation to be solved. Details concerning the computation of the gradients and the construction of the cubic equation can be found in [8] and [9]. Since a cubic equation may have more than one real root, the root with the deepest minimum is chosen.

Each step of the minimization procedure of G_p will require M dynamical matrix operations on a vector to be computed— one necessary for the update of each $|u_i\rangle$. Although the time to minimize the phonon functional is likely to have a complicated dependence on M and the number of atoms N, some estimates can be made. Since the dynamical matrix operating on each vector can be computed in order N steps, a very simple approximation is that the time for minimization of the phonon functional scales roughly as order MN.

The convergence of G_p is illustrated in figure 9. Taken as an example is the $T = 4$ viral capsid of hepatitis B (HEPB) and the specific PDB coordinates are those of 1QGT [9]. Details of how such a calculation proceeds will be discussed later. For now we need to know that the capsid has symmetry of the icosahedron. The capsid can be represented (geometrically) as a truncated icosahedron (i.e. C$_{60}$) with each of the 60 vertexes containing four identical proteins. The entire capsid has 1.6 million modes, and symmetry (group theory) is used reduce the calculation. The value of M used is 100. In figure 9 we show the A mode convergence of G_p in the subspace corresponding to modes from the A irreducible representation of the icosahedron (point group I). These modes are non-degenerate and are the easiest to compute. Figure 9 shows what is equivalent to $\sqrt{(\text{Tr}(G_p)/M)}$, which is the square root of the average frequency squared. The extra term ω_{max}^2 is a technical detail (described more below) resulting from a shift of the frequencies to insure that G_p has a single minimum. The important points of the figure are that (i) the convergence is monotonic, (ii) that around 1000 iterations the result is

reasonably converged, and (iii) at several thousand iterations, convergence is far less than a fraction of a cm^{-1}. It should be recalled that the lowest frequency is less than 1 cm^{-1} while the highest is greater that 3000 cm^{-1} making the ratio of high to low eigenvalues larger than 10^{+7}. The ill-conditioned nature of the dynamical matrix apparently causes no difficulty in the phonon functional method.

We conclude this section with some quirks of the method which one must be aware of, and some comments about computational effort. The first is that the spectrum has to be shifted for the functional to have a minimum and the second is that the converged vectors $|u_i\rangle$ are not yet eigenvectors. First the shift quirk Ordejon *et al* [59] have shown that the spectrum of eigenvalues must be negative definite for the functional to converge. The vibrational problem with eigenvalues ω^2 is the antithesis of this—all eigenvalues are positive. To shift the spectrum, we replace $\overset{\leftrightarrow}{D}$ by $\overset{\leftrightarrow}{D}{}^s = \overset{\leftrightarrow}{D} - \omega_{max}^2 \overset{\leftrightarrow}{I}$ where ω_{max} is the largest eigenvalue in the spectrum. It usually is a local vibration of hydrogen. Fortunately, it is easy to compute as the application of $\hat{D}$ to any random vector a few times (a Lanczos type method) rapidly converges a vector to the upper extreme of the spectrum. We add a small cushion to this estimate to ensure we are safely above this maximum frequency. Secondly, the converged set of vectors $|u_i\rangle$ are not eigenvectors, but span the subspace of M lowest eigenvectors of $\overset{\leftrightarrow}{D}$. Since we are minimizing a *Trace* which is invariant under orthogonal transformations, each vector is a linear combination of eigenvectors within this low frequency 'occupied' space. The final eigenvectors and eigenvalues are however trivial to determine. Simply solve the generalized $M \times M$ eigenvalue problem $\overset{\leftrightarrow}{D}{}^s \vec{\psi} = \lambda \overset{\leftrightarrow}{S} \vec{\psi}$. Here $\overset{\leftrightarrow}{S}$, upon exact convergence, will be the unit matrix. It is included in the diagonalization here to more accurately allow for convergence of the phonon functional to within a finite tolerance. The eigenvalues λ_ν

are ω_ν^2 and the vector $\vec{\psi}$ of length M determines the linear combination of large vectors $|u_i\rangle$ of length $3N$ that makes an eigenvector of the dynamical matrix.

Finally a few words about computational effort. One expects several thousand iterative minimization steps ($N_{\text{Iter}} \sim 1000$) and each step will require the operation of M dynamical matrix $\hat{D}$ operations on a vector. Each operation is very similar to an MD step as discussed above (instead of computing a first derivative it computes a second derivative). Roughly, the scaling with number of atoms and size of subspace gives a computational scaling with M and N that goes as order (MN). Put in terms of traditional MD simulations, the computer time is similar to $N_{\text{Iter}} \times M$ MD steps. A rough estimate of the computational effort uses $N_{\text{Iter}} = 1000$ and $M = 100$, to yield a computational effort similar to 10^5 MD steps, which is approximately equivalent to a 100 ps MD simulation.

6.2. Tubular virus M13—vibrations, impulsive stimulated Raman scattering, and inactivation of virions

Ultra-short laser pulses of visible or near infrared light have been shown to inactivate viruses [10, 61, 62]. By tuning the length of the laser pulse (100 fs to several ps) and the intensity of the laser, conditions can be found that do not damage other organisms such as cells or bacteria. This discovery has profound implications in the prevention and spread of infection, and especially has implications for the cleansing of blood within blood-banks or in undeveloped regions of the world.

The physical mechanism responsible for the inactivation of viruses is not fully understood. One intriguing mechanism suggested is impulsive stimulated Raman scattering (ISRS). ISRS scatters light through the Raman coupling of the light to the polarizability of the electrons. However, the polarizability is dependent on the configuration of the atoms within the structure. This dependence produces an indirect coupling of the electric field of the laser with the dynamical and structural properties of the material. The electric field of the laser oscillates far to quickly to be in resonance with the low frequency oscillations of the capsid. Rather, the effect couples distortions quadratically to the electric field. The electric field squared is proportional to the intensity, and the intensity is modulated into pulses. The length of the laser pulse can be made to be compatible with the timescale of an oscillation. This produces a 'kick' on the capsid. The kick should last for about $\frac{1}{4}$ of a vibrational period to produce the largest effect. If the kicks are large enough, the virus can, theoretically, break apart. This interesting application does not require full oscillations as in a standard resonance conditions, such as the shattering of a wine glass through sound. If the system is undamped, completion of just $\sim\frac{1}{4}$ of an oscillation will likely be enough to produce damage to the capsid.

In the ISRS technique, (pioneered by Nelson *et al* [63, 64]), an electric field from the laser polarizes the proteins through their polarizability $\overset{\leftrightarrow}{\alpha}$, producing a time dependent dipole moment on the proteins in the capsid; $\vec{p}(t) = \overset{\leftrightarrow}{\alpha} \cdot \vec{E}(t)$. The energy of interaction, $U(t)$, is then simply $U(t) = -\frac{1}{2}\vec{E}\cdot \overset{\leftrightarrow}{\alpha} \cdot \vec{E}$. The electric field we are considering is applied

as a Gaussian pulse of duration τ_L with the frequency of the laser being ω_L. The applied electric field from the laser then has the form $\vec{E}(t) = \vec{E}_0 e^{-t^2/2\tau_L^2} \cos(\omega_L t)$, which produces an energy of interaction

$$U(t) = -\frac{1}{2}\vec{E}_0\cdot \overset{\leftrightarrow}{\alpha} \cdot \vec{E}_0 e^{-t^2/\tau_L^2} \cos^2(\omega_L t). \qquad (47)$$

The polarizability of the capsid depends on its structure, so it implicitly depends on the displacements of a normal mode pattern if it exhibits a distortion. Raman scattering depends on the polarizability derivatives and for a displacement pattern $|\vec{\eta}_\nu\rangle$ of amplitude Q_ν there exists a driving force on this mode due to the external field given by

$$\begin{aligned} F_{Q_\nu} &= -\frac{\partial U(t)}{\partial Q_\nu} \\ &= \frac{1}{2}\vec{E}_0 \cdot \frac{\partial \overset{\leftrightarrow}{\alpha}}{\partial Q_\nu} \cdot \vec{E}_0 e^{-t^2/\tau_L^2} \cos^2(\omega_L t). \end{aligned} \qquad (48)$$

Evaluating such derivatives is somewhat involved and details can be found in [11] and [65]. The equation of motion for the mode becomes that of a driven harmonic oscillator with a time dependent Gaussian force, $\ddot{Q}_\nu + \omega_\nu^2 Q_\nu = F_{Q_\nu}$. The solution to this can be determined from the harmonic oscillator Green's function and is approximately $Q_\nu(t) = Q_\nu^{\max} \sin(\omega_\nu t)$, where the maximum amplitude is given by

$$Q_\nu^{\max} = \sqrt{\pi}\left(\frac{\tau_L}{2\omega_\nu}\right) F_0 e^{-\omega_\nu^2 \tau_L^2/4}. \qquad (49)$$

This is a very interesting result and exhibits a kind of 'resonance' due to the kick. The prefactor τ_L is large for long pulses and the $1/\omega_\nu$ prefactor is large for low frequencies. However, the major dependence on the size of the amplitude of the mode is controlled by the exponential Gaussian factor that includes both τ_L and ω_ν. It is maximum when both $\omega_\nu \tau_L$ is small. If we imagine a specific mode ν, then the largest amplitude excitation will occur when the laser pulse time is $\tau_L = \sqrt{2}/\omega_\nu$. This gives the result that one should 'tune the laser' to $\tau_L = 0.225 T_\nu$ where T_ν is the period of oscillation for mode ν. The interpretation is that the laser must 'kick' the oscillator for about 1/4th of a cycle to obtain the largest amplitude of oscillation. Unlike a resonance with a driving frequency, the ISRS mechanism requires less than a single oscillation to excite the mode. If the excitation is large enough, past some (difficult to determine) threshold, the kick can produce damage, and inactivate the virus.

Simulations of the ISRS effect of the M13 virus have been performed in previous studies [11, 65]. In these simulations, a low frequency symmetric breathing mode was found to damage the capsid. Figure 10 shows the trajectory of the M13 capsid from a MD simulation which had an impulsive force from ISRS. The pulse width of the light used in the simulation was $\tau_L = 1$ ps and had a maximum force on the capsid at $t = 0$. Figure 10(a) shows the capsid at $t = -5$ ps, or $5\tau_L$ prior to the maximum force of the ISRS pulse. Figure 10(b) shows the capsid at $t = 0$, when the maximum force is being exerted on the capsid. As can be seen in the figure, very little has happened to the capsid at the peak of the pulse. Clear damage to the capsid is seen at $t = 5$ ps (figure 10(c)) and $t = 10$ ps
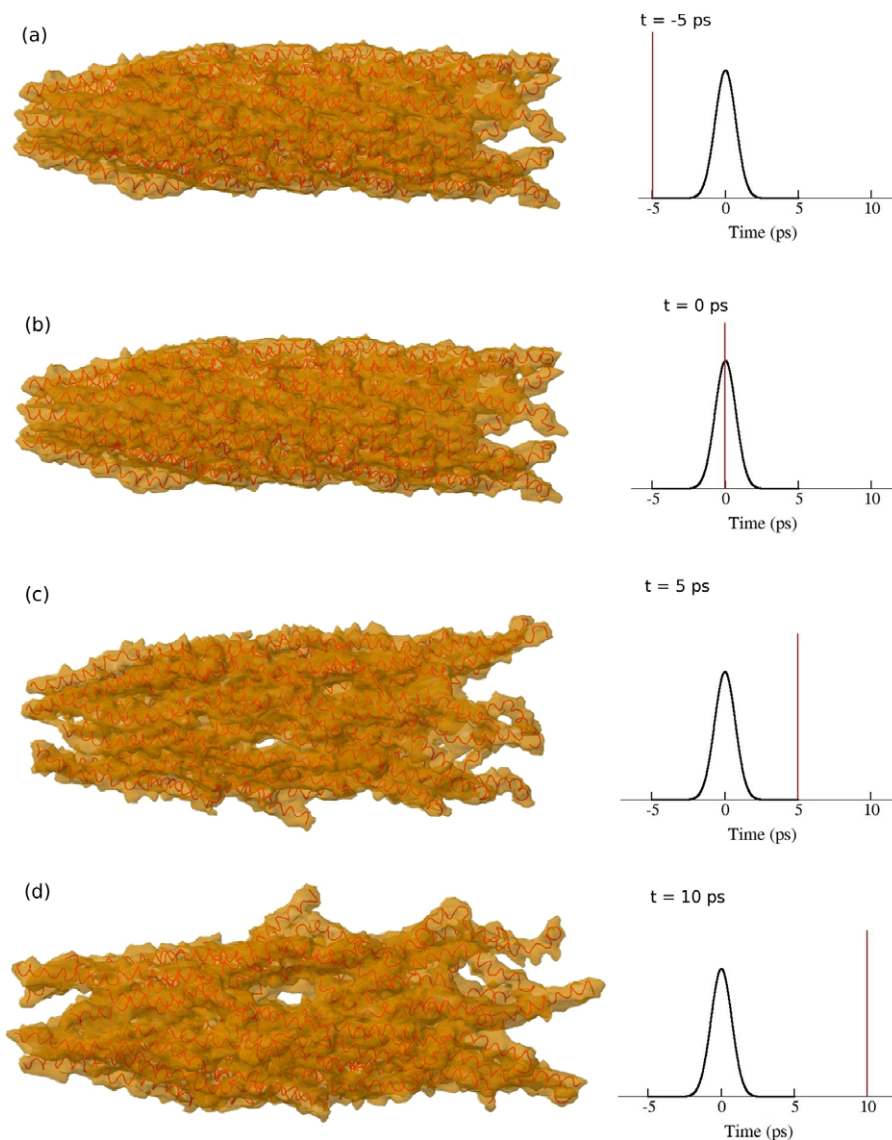
Figure 10. Trajectory snapshots of a molecular dynamics simulation of the M13 bacteriophage with an additional force from a pulsed (ISRS) light source. The magnitude of the ISRS force on each atom is proportional to the average intensity of the light pulse $\bar{I}(t) = I_0 \exp(-t^2/\tau_L^2)$, where I_0 is the maximum intensity of the light, and τ_L is the duration of the light pulse. We used $I_0 = 1.5$ PW/cm^2 and $\tau_L = 1$ ps for this simulation. A qualitative plot of the force on the capsid from the light pulse during the MD simulation is shown on the right-hand side of the figure with the red line indicating the current force at the given MD simulation time. (a) The capsid at $t = -5$ ps. At this time, there is negligible force from the light. (b) The capsid at $t = 0$ ps. At this time the force on the capsid is maximal as indicated by the red line. Note the capsid has almost no response to the light at this point. (c) Snapshot of the capsid at $t = 5$ ps showing large structural distortions. At this time, the light pulse is nearly gone and the force on the capsid is negligible. (d) Snapshot of the capsid at $t = 10$ ps showing the structural distortions remaining and permanent damage to the capsid.

(figure 10(d)), after the light pulse has stopped interacting with the capsid. For a pulse width of $\tau_L = 1$ ps, the most damage is produced to modes which have a period of $\sim$4 ps-consistent with the approximate period of the breathing mode ($\sim$6 ps).

It should be noted that these simulations produce breakup at intensities much greater than that of experiment. Experimental intensities for several viruses are in the range of approximately 100 MW cm^{-2} to 100 GW cm^{-2}. Notice the wide range of intensities necessary. Our simulation required an intensity of order 1 PW cm^{-2}, which is orders of magnitude larger than experiment. This brings up several questions. Of course the first question is whether ISRS is the correct mechanism for inactivation of viruses. The inactivation shows

a threshold like behavior, which is what one may expect with ISRS. But other mechanisms, such as multi-photon absorption, could also be expected to produce threshold behavior. Further experimental and theoretical work will be necessary to reveal the science that is occurring during inactivation of viruses with ultra-short laser pulses.

6.3. Vibrations of icosahedral viral capsids

The all-atom phonon functional allows previously intractable vibrational problems to be solved. An example of this is the determination of the mode displacement patterns and frequencies for icosahedral viral protein capsid particles. A

viral capsid consists of a spherical-like assembly of many individual proteins, often identical, arranged in a spherical cage-like assembly. The minimum number of proteins on the cage is 60. The purpose of the icosahedral cage is to protect the viral genome which resides inside.

Dykeman *et al* [9] have used the phonon functional method and applied it to the human virus capsids of polio and hepatitis B, and to the plant virus cowpea chlorotic mottle virus (CCMV). To our knowledge these are the only normal mode calculations using a full atomistic model will all degrees of freedom. We briefly review the results here for the polio virus, as we can compare its result to those with the work of Vlijmen [41] who used an atomistic potential energy model but restricted the basis set by allowing only dihedral distortions.

The polio virus (PV) is a small RNA virus that is a severe pathogen in humans. It has the potential to be eliminated as was successfully accomplished by smallpox, although this work is ongoing and there are severe obstacles. An x-ray structure of the virus is taken from Grant *et al* [66] (PDB ID 1VBD) and PV is a $T = 3$ virus (180 proteins in the capsid). The number of atoms in the icosahedron is nearly 800 000 and there are nearly $2.4M$ modes. An implicit solvent mode is used for the empirical force field between atoms of the capsid [67, 68]. The x-ray structure is energy minimized to a zero force structure from which the normal mode calculation proceeds. Group theory is used to simplify the problem and modes of A, T_1, T_2, G and H symmetry of the icosahedral group I. There is far more programming effort in the development of the group theory code than in the simpler phonon functional code.

Figure 11 shows a stick spectrum of the vibrational modes of various symmetries and compares them to the dihedral distortion only model of Vlijmen *et al*. The figure shows the first 50 modes for each representation and just the lowest three modes from Vlijmen are shown for comparison. The frequencies are very similar, but fine agreement between the two calculations is not there. Part of this difference is attributable to the different force fields used. Also shown in figure 11 for the A modes are the participation percentages W_ν for a few low frequency modes. The participation number gives an estimate of the total number of atoms participating in the distortion of a mode. The participation number is obtained from a kind of information entropy $W_\nu = e^{S_\nu}$ for each mode. Here, S_ν is the (information) entropy of each mode. The more entropic a mode is the more spread out the distortion is. The information entropy is given by $S_\nu = -\sum_\alpha p_\alpha(\nu) \ln(p_\alpha(\nu))$. The probabilities $p_\alpha(\nu)$ are the squared component of the normalized relative displacements for each atom/direction $\alpha = [1, 3N]$, $p_\alpha(\nu) = |\eta_\alpha(\nu)|^2$, where $\eta_\alpha(\nu) \propto M^{-\frac{1}{2}} e_\alpha(\nu)$. The quantity W_ν is a kind of number of accessible sites determined by the entropy. The number W_ν is made into a percentage by $W_\nu(\text{percent}) = 100 \times W_\nu/N_1$ where N_1 is the number of atoms in one of the 60 icosahedral 'cells' of the capsid ($N_1 = 13,074$ for polio virus). Most of the low frequency modes have high participation numbers indicating a concerted motion of large groups of atoms. These modes will be described well by simpler models like the ENM. However, some modes have low participation number and are usually random coils or
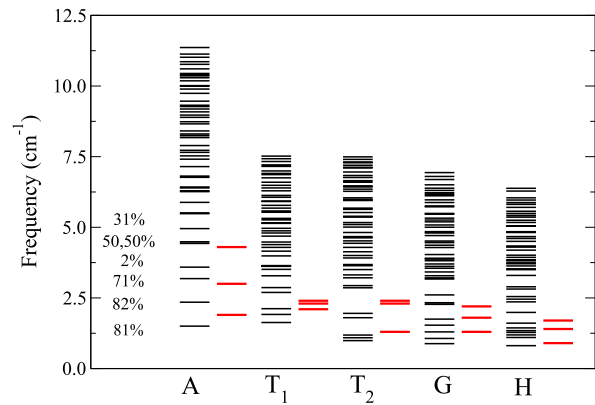

Figure 11. A stick spectrum of the vibrational modes of the polio virus of their group theoretic symmetries. The black lines are from the all-atom phonon functional and the red lines are a few of the modes from [41] which allowed only dihedral distortions. The percentages on the left refer to the per cent of atoms participating in the first few A symmetric modes determined by the all-atom phonon functional.

other unrestricted regions directed in toward the interior of the capsid. For ssRNA viruses such as Polio, these random coils often interact with the genome inside—an effect not included in this potential energy model since structural information on how the genome interacts with these random coils is often unavailable.

7. Normal modes and conformational changes

Many proteins, such as enzymes, can toggle between two conformers and these conformers represent the active and the inactive enzyme. Describing the conformational change as they change functionality, often due to binding of a ligand, is an important yet not satisfactorily solved problem. The simplest example of such a transition is a 'pac-man' transition between open- and closed states. A binding ligand changes the barrier between the initial and final states so that a transition can occur and the free-energy barrier can be overcome. Likewise, binding with inhibitors may change the conformation and inhibit binding with effectors shutting the functionality down. Other, more complex examples such as the ABC transporter describe below, have several moving parts that re-organize cooperatively so that the whole assembly can provide a function.

The question then arises 'Can normal mode displacement patterns be used to describe the cooperative motion within a proteins, or protein assembly?' The answer appears to be 'sometimes'. We first will describe a simple open–closed system, lactoferrin, where we use the all-atom phonon functional (AAPF) and find that the answer is 'yes', in that just two modes describe 63% of the transition and six modes describes it at the 81% level. We then compare this with a determination of the same quantity using the simpler ENM and find that the ENM produces qualitatively similar results, but requires more modes to describe the motion than the AAPF. We follow this up with a discussion of the work of Petrone *et al* [69] who use an ENM and find examples where even 20 modes do not describe the transition to better than 50%.
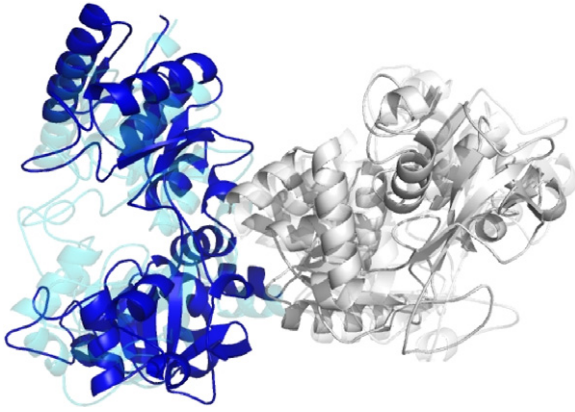
Figure 12. Superimposed open and closed structures of lactoferrin. On the left, the open (dark) and closed (light) regions of the enzyme are evident. The regions on the right are the similar areas of the two conformers.
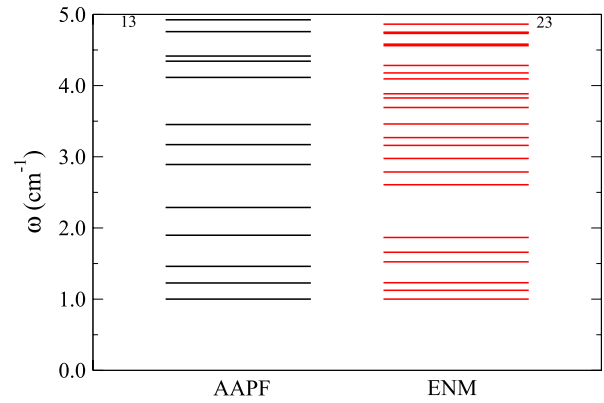

Figure 13. A stick spectrum of the frequencies of the open structure of lactoferrin computed by the all-atom phonon functional and the elastic network model. The ENM has its spring adjusted to exactly fit the frequency of the lowest AAPF mode. Notice that the density of modes in the ENM (23 modes below 5 cm^{-1}) is higher than that in the AAPF model (13 below 5 cm^{-1}).

Lactoferrin belongs to the class of transferrin proteins that shuttle around iron and control iron levels in animals. It is found in milk and leucocytes. The binding and unbinding with iron complexes produces a large-scale conformational change within the protein. Optimization (energy minimization) was first performed on the original x-ray PDB [70] structures 1LFH (open) and 1LFG (closed).

An implicit solvent generalized Born model was used in the force field to determine the optimized structure and the second derivatives required for the dynamical matrix. The two optimized solutions for the open and closed states were then aligned. The center of mass of the two were made coincident, and a grid of 10^6 Euler angle rotations were made to minimize the RMS deviation between the two structures. Figure 12 shows the trace of the backbone for each of the two aligned structures. The functionally important parts of the protein are dark for the open state and light for the closed state. The remaining domains of the open and closed lactoferrin proteins that are not involved in the functionally important motion are on the right. As can be seen in the figure, the major difference between the two is the region on the left which shows a transition between open and closed structures that occurs via a moving of the two flaps.

The harmonic vibrational modes of the open structure were then determined using the phonon functional method with the goal to determine if a single floppy low frequency mode or several low frequency modes can describe the transition from the open to the closed conformer. The number of degrees of freedom is $3N$, where N, the number of atoms is 10 509 including hydrogen. The component displacements from the open to the closed are $\delta d_\alpha = r_\alpha^o - r_\alpha^c$ where the composite label α denotes a specific atom *and* Cartesian direction, while the 'o' and 'c' superscript correspond to open and closed configurations. For each normal mode ν, we define a projection p_ν and a projection angle θ_ν as

$$p_\nu = \cos^2 \theta_\nu = \frac{|\langle \vec{\eta}_\nu | \overleftrightarrow{M} | \vec{\delta d} \rangle|^2}{|\langle \vec{\delta d} | \overleftrightarrow{M} | \vec{\delta d} \rangle|^2}. \qquad (50)$$

It is easy to show that the sum, $\Sigma_\nu \cos^2 \theta_\nu$, over all 31 527 modes, is 1. The projection p_ν is the 'amount' of normal mode displacement contained in $|\vec{\delta d}\rangle$ and is one when summed over all modes.

Figure 13 shows the results of the normal mode analysis. Figure 13 is a stick spectrum of the frequencies of lowest frequency modes computed by both the all-atom phonon functional (AAPF) method, and by the ENM model. The ENM model used only the C_α atoms of the protein backbone and a spring with the same spring constant connected all neighbors within 10 Å. The ENM model spring was fit to reproduce the lowest mode near 1 cm^{-1} in the AAPF calculation. One feature worth noting is that the three lowest modes in both methods are fairly close in frequency. A second feature is that the density of levels (i.e. the number of levels per cm^{-1}) is higher in the ENM than in the AAPF, being roughly twice as high.

Figure 14 plots the projection computed for each mode ν and a partial sum of all projections up to mode ν. This is done for both the AAPF and ENM methods. (The six zero frequency modes are not shown.) The results are interesting and somewhat surprising. First, by adding up around seven modes, both models show that the conformational change can be recovered by over 0.8 (80%). The quantitative correspondence between the two methods is very good. Next when we try to describe the transition with just a couple of modes, the AAPF result is more willing. In the AAPF method, one mode ($\nu = 2$) projects around 40% and adding in just the first two modes yields 64%. Thus, the projection is highly localized into a couple of modes. The ENM model tends to spread out its projection and is more delocalized. It takes about five modes to achieve this 64% level. Finally the projection saturates around 7–10 modes in both cases. Adding dozens of modes additional modes in either case does little to increase the summed projection.

In this lactoferrin system, there exists a mode displacement subspace (MDS) that generates the displacement pattern connecting to the two conformers. These modes are similar to the essential dynamics excursions determined by principal
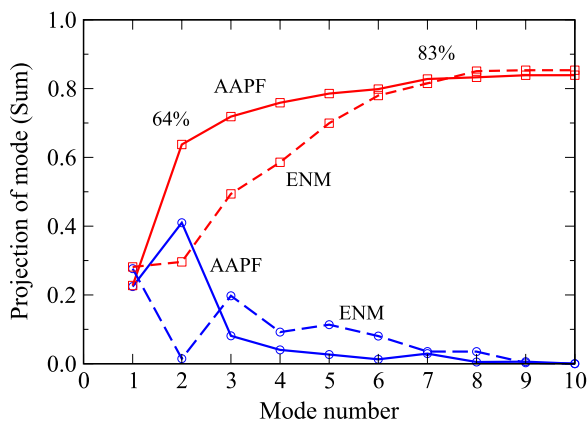
Figure 14. The overlap of the open to closed transition by modes of the open structure. The projection p_v (equation (50)) are the lower two curves and the partial sum (Σp_v, the upper two curves) of the open lactoferrin on the displacement vectors of the transition to the closed conformer. The solid lines are the all-atom phonon functional and the dashed are the elastic network model. The AAPF converges more rapidly than the ENM, yet converges to roughly the same sum of over 80% projection with ~7 modes. The AAPF yields a 64% projection with just two modes.

Figure 15. Illustration of the three domains of a gram-negative bacterial ABC transporter. The three domains are the periplasmic binding protein (PBP), the transmembrane domain (TMD), and the nucleotide binding domain (NBD). The PBP, TMD and NBD are stacked from top to bottom respectively. The transmembrane domain penetrates the inner lipid membrane which separates the periplasm from the cytoplasm of the bacterium. Binding of ATP to the nucleotide binding domain facilitates conformational changes which allow transport of substrates (direction indicated by arrow) from the periplasm to the cytoplasm.

components that a protein takes as it thermally samples configuration space. The existence of an MDS can be a useful aid in probing the pathways in the complex energy landscape. This conclusion in line with that of Tama *et al* [71] who find that even one mode is useful in the description of an open to closed conformational change.

A comprehensive study of the connection between normal modes and conformational changes has been performed by Petrone *et al* [69]. Their model of the normal modes was an elastic network model, a far simpler model than AAPF and especially useful for a rapid analysis. In the few test cases (myosin, calmodulin, NtrC, and hemoglobin) at least 20 modes were found to be necessary to achieve a description of even 50% of the conformational change. There was a very large variation from one test case to the next suggesting that sometimes a normal mode analysis offers a small subspace to describe the MDS, while in other cases the MDS is too large that a NMA is not very useful.

7.1. Example of functional motions in a complex system—normal modes of an ABC transporter

As an example of application of normal mode analysis in the study of complex biological proteins, we demonstrate the use of the phonon functional method to compute the low frequency modes of an ATP-binding cassette (ABC) transporter with over 20 000 atoms. ABC transporters are a family of membrane proteins that are involved in the active transport of diverse substrates across cellular membranes [73]. The clinical relevance of ABC transporters cannot be understated. Mutations of the ABC transporter complex are responsible for disorders in humans such as cystic fibrosis and Tangier disease [74, 75]. In bacteria and cancerous cells, over expression of multi-drug ABC exporters confers drug resistance, making antibiotic and chemotherapy
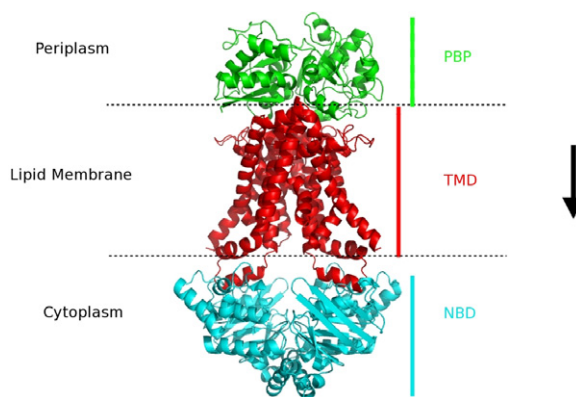
treatments ineffective [76]. The normal mode analysis of the ABC transporter illustrated here shows how several distinct proteins can come together and exhibit a concerted functional motion. A review of NMA of membrane proteins can be found in [72].

The general ABC transporter (figure 15) consists of a two subunit nucleotide binding domain (NBD), which drives the transport process by binding and hydrolyzing ATP, and a two subunit transmembrane domain (TMD) that forms a translocation pathway through the membrane. The entrance to the translocation pathway is blocked by a gating region which opens as a result of the conformational changes that occur in the NBD from ATP binding and hydrolysis. In gram-negative bacteria ABC importers, the substrate is delivered to the translocation pathway by a periplasmic binding protein (PBP) which binds to the substrate in the periplasm and is essential for transport [77]. The three components consisting of the PBP, TMD, and NBD must follow a concerted motion in order to achieve transport of the substrate. The function of the assembly is to deliver the cargo trapped in the PBP through the TMD which has a gated channel. The NBD is the 'engine' that drives the delivery.

Our example illustrates the conformational changes of a *complete* molybdate ABC transporter of the archaeal bacteria *Archaeoglobus fulgidus* using the fully atomistic phonon functional method to compute the normal modes from the x-ray crystal structure [78]. Although NMA is not capable of predicting the complete pathway that each atom will take as the gating region opens, it can provide information on how each of the three domains move in concert, at least for small displacements. Thus, one can think of the NMA of the ABC transporter as analogous to examining how the connecting rod and piston of a locomotive work cooperatively to rotate the wheels by making small displacements of the locomotive forwards and backwards.
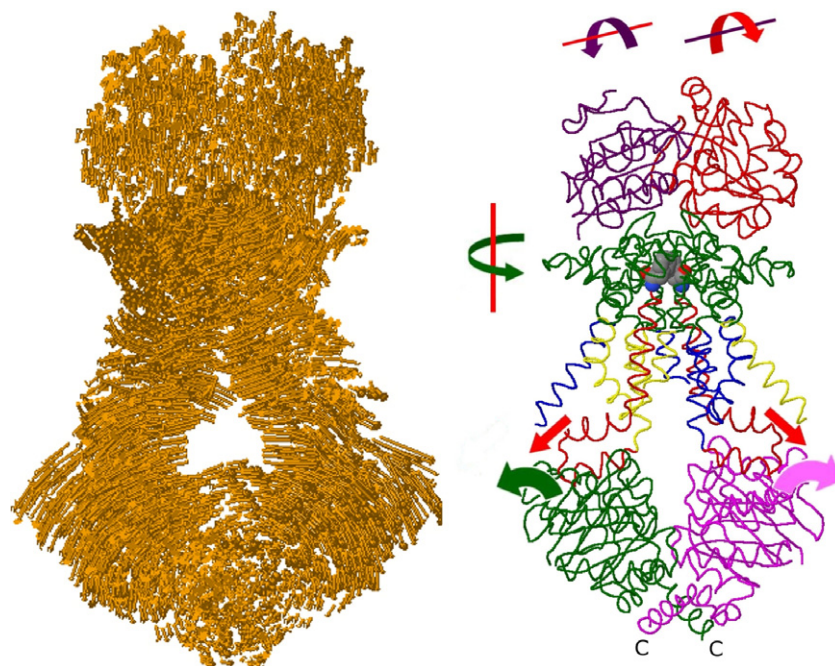
Figure 16. Low frequency 'corkscrew' mode of the *Archaeoglobus fulgidus* molybdate ABC transporter with frequency of 2.62 cm^{-1}. The left-hand figure shows all carbon atoms in the structure with arrows indicating the direction of the atomic displacements. The right-hand figure gives an overall schematic of the coupled motions between domains. The individual motions for various sections of the ABC transporter are illustrated with arrows.

The *A. fulgidus* molybdate ABC transporter (figure 15) was prepared for NMA by first protonating the x-ray structure [78] (PDB code 2ONK). The bound tungstate and two ATP molecules that bind to the NBD were not included in the simulation. It is unlikely that this will appreciably affect the global low frequency modes of the entire complex. The lipid membrane was also excluded from the calculation. The final structure consisted of the NBD, TMD and the PBP and had a total of 20 667 atoms. The structure was then energetically minimized using the AMBER 94 force field [12] with the generalized Born (GB) implicit solvent model [67, 68] to a root mean squared (RMS) force gradient less than 10^{-4} eV Å^{-1}. A cutoff of 10 Å was used for electrostatic interactions. The resulting RMS deviation of the optimized structure from the original x-ray structure was 1.75 Å.

Using the optimized structure and the AMBER 94 force field with the GB implicit solvent model, the lowest $M = 100$ normal modes of the ABC transporter were found using the phonon functional method described in section 6. A low frequency mode at 2.62 cm^{-1} (78.6 GHZ) is of particular interest and is illustrated in figure 16. This mode shows how the three different domains (PBP, TMD, and NBD) operate in a concerted motion and can best be described as a 'corkscrew' like motion. An examination of the mode reveals that the lower portion of the NBD has a rigid region near the C-terminal hinge while the upper portion of the NBD (at the NBD/TMD interface) couples to two alpha helices (coupling helices, CH) of the TMD. As the NBD opens/closes, the TMD/PBP interface twists. The twisting of the TMD/PBP interface causes the two lobes of the PBP to move in a shearing like motion that

is consistent with theoretical studies of other PBP [79]. The PBP lobe shearing also has similarities to the ligand to unligand hinge-twist transition found in a crystallographic study of a maltodextrin binding protein [80]. Thus, the shearing of the lobes could be related to motions that will result in an opening of the PBP binding cleft and the subsequent release of the bound substrate. Overall, the normal mode describes a combined motion of the three domains that is connected and highly concerted. The three components (NBD, TMD, PBP) operate together like a machine, similar to the gears in a motor.

Although the normal mode analysis of the ABC transporter cannot predict the exact pathway that the three domains of the ABC transporter undergoes during the transport cycle, it does provide clues about how the three domains operate in concert, and what the important interactions between domains are. Although it would be very interesting to follow the normal mode to see how (or if) the PBP binding cleft would open, mode following in a computational setting is extremely difficult and is plagued by many problems. However, in figure 17 we offer a hypothesis how this 'corkscrew' motion could lead to an opening of the PBP binding cleft. One can see in figure 17(b) a hypothetical illustration of how the twisting like motion at the TMD/PBP interface combined with the shearing motion in the lobes of the PBP simultaneously forces the substrate binding cleft and gating region open.

To get a feeling for how this motion impacts the whole transport process, we illustrate a hypothetical transport cycle of the ABC transporter in figure 18. This transport cycle makes use of the information gained from the normal mode analysis. As can be seen in the figure, the transport cycle consists
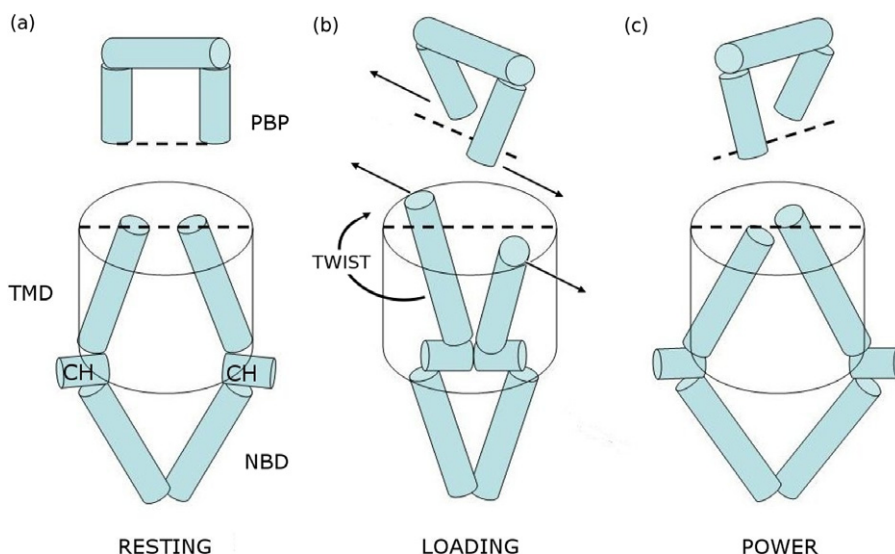
Figure 17. Proposed conformational changes that occur as the ABC transporter moves between the resting state with a closed translocation pathway and the transition state with an open translocation pathway. The proposed conformational changes are based on the normal modes. (a) The resting state. (b) Illustration of the 'loading stroke', which moves the ABC transporter from the resting state to the transition state. The twist at the PBP/TMD interface causes the binding cleft in the PBP to separate. (c) Illustration of the 'power stroke', which moves the ABC transporter back to the resting state.
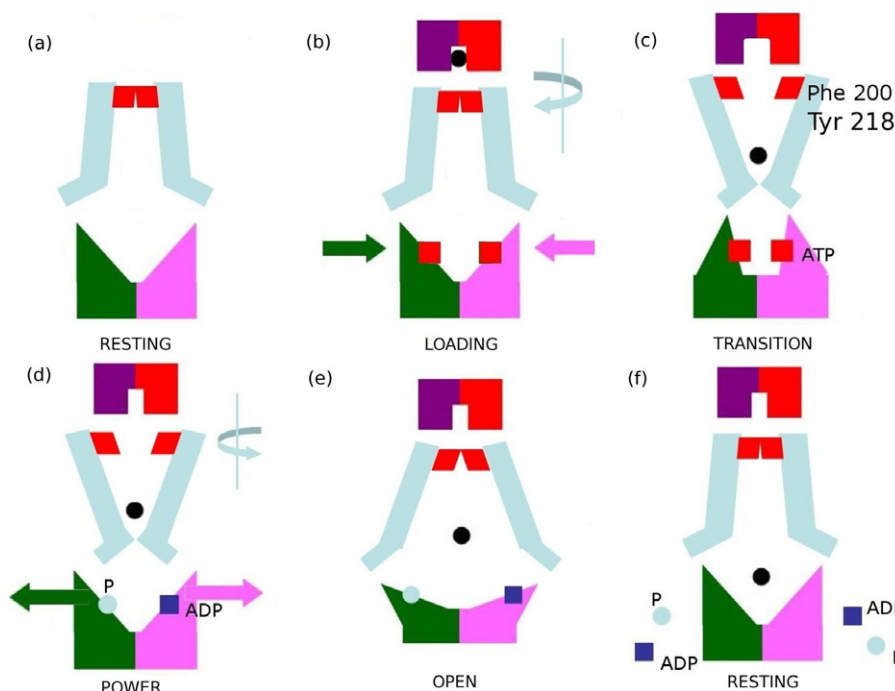

Figure 18. Diagram of the hypothetical transport process for the molybdate ABC transporter inferred from the all-atom normal modes. The two Phe 200/Tyr 218 residues making up the gating region are indicated by blocks. The substrate is indicated by a black circle. (a) Resting state with no PBP present. (b) The PBP and two ATP molecules bind to the NBD causing the NBD to close. This begins the loading stroke. (c) The transition state after completion of the loading stroke. The substrate becomes trapped in the TMD after being released from the PBP. (d) One or two of the ATP molecules are hydrolyzed forcing the NBD open. This begins the power stroke. (e) Open state after completion of the power stroke. The substrate is now free to enter the cytoplasm. (f) The transport process is complete and the ABC transporter returns to a resting state.

of five stages, resting, loading, transition, power, and open. The resting stage represents the state of the ABC complex prior to interaction with the PBP. After a PBP carrying the substrate makes contact with the TMD, binding of ATP to the NBD facilitates the loading stage. It is the loading stage in which we believe the normal mode computed in the AAPF

method may describe the beginning stages of the functional motion that takes place. The result of this functional motion is the transition state where the PBP binding cleft and gating region (consisting of PHE 200 and TYR 218) are forced open, allowing transport of the substrate through the channel. Finally in the power stage, the NBD 'engine' catalyzes the ATP–ADP reaction and results in a conformational transition to the open stage, which resets the ABC transporter to the resting stage.

8. Conclusions

As we have shown in this review, there are a variety of methods that can be used to calculate the normal modes of proteins, viruses, or large protein assemblies to predict their functional motions. Each of the methods have their strengths and weaknesses. Continuum elastic theories can provide simple insights into the basic dependence of functional motions and structural transitions in terms of just a few quantities such as Young's modulus or the Foppel von Karmen number. Elastic network and other coarse grained models are powerful tools that can simplify a complex protein into a simple network of connections which can give a reasonable description of its dynamical properties. And finally, recent developments of tractable all-atom methods provide quantitative estimates of frequencies and more precision in describing displacement patterns. In the case of lactoferrin discussed in this review, the increased precision in the displacement patterns were able to describe the open to closed transition with the first two modes better then coarse grained methods.

Normal mode analysis is just beginning to scratch the surface into the connections of protein structure to protein function and their are many unanswered questions that remain. New novel applications of normal modes, such as in the study of impulsive stimulated Raman scattering of viral capsids, or the use of normal modes in the construction of a transition pathway between two protein conformers, continue to provide new challenges and new areas of application for normal modes. Indeed, the wide range of successes of normal mode analysis in predicting functional motions and other properties have made it a tool that will be continued to be used and explored in many biophysics applications. Indeed the future for normal modes in biological physics appears bright.

References

[1] Allen M P and Tildesley D J 1987 *Computer Simulation of Liquids* (Oxford: Clarendon)
[2] Rahman A 1964 Correlations in the motion of atoms in liquid argon *Phys. Rev.* A **136** 405–11
[3] Verlet L 1967 Computer experiments on classical fluids. i. Thermodynamical properties of Lennard-Jones molecules *Phys. Rev.* **159** 98–103
[4] Field M J 1999 *A Practical Introduction to the Simulation of Molecular Systems* (Cambridge: Cambridge University Press)
[5] Leach A R 2001 *Molecular Modeling: Principles and Applications* 2nd edn (Englewood Cliffs, NJ: Prentice Hall)
[6] Ma J 2005 Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes *Structure* **13** 373–80

[7] Dykeman E C and Sankey O F 2008 Low frequency mechanical modes of viral capsids: an atomistic approach *Phys. Rev. Lett.* **100** 028101
[8] Dykeman E C and Sankey O F 2009 Theory of the low frequency mechanical modes and Raman spectra of the M13 bacteriophage capsid with atomic detail *J. Phys.: Condens. Matter* **21** 035116
[9] Dykeman E C and Sankey O F 2010 Atomistic modeling of the low-frequency mechanical modes and Raman spectra of icosahedral virus capsids *Phys. Rev.* E **81** 021918
[10] Tsen K-T, Tsen S-W, Chang T-C W C L, Hung C-F and Kiang J 2007 Inactivation of viruses by coherent excitations with a low power visible femtosecond laser *Virol. J.* **4** 50
[11] Dykeman E C, Benson D, Tsen K-T and Sankey O F 2009 Simulations of impulsive laser scattering of biological protein assemblies: application to M13 bacteriophage *Phys. Rev.* E **80** 041909
[12] Cornell W D, Cieplak P, Bayly C I, Gould I R, Merz K M, Ferguson D M Jr, Spellmeyer D C, Fox T, Caldwell J W and Kollman P A 1995 A second generation force field for the simulation of proteins, nucleic acids, and organic molecules *J. Am. Chem. Soc.* **117** 5179–97
[13] Brooks M R, Bruccoleri R E, Olafson B D, States D J, Swaminathan S and Karplus M 1983 CHARMM: a program for macromolecular energy minimization, and dynamics calculations *J. Comput. Chem.* **4** 187–217
[14] Goldstein H 1980 *Classical Mechanics* (London: Addison-Wesley)
[15] Fukui K, Sumpter B G, Noid D W, Yang C and Tuzun Fukui R E 2001 Analysis of eigenvalues and eigenvectors of polymer particles: random normal modes *Comput. Theor. Polym. Sci.* **11** 191–6
[16] Hathorn B C, Sumpter B G, Noid D W, Tuzun R E and Yang C 2002 Vibrational normal modes of polymer nanoparticle dimers using the time-averaged normal coordinate analysis method *J. Phys. Chem.* A **106** 9174–89
[17] Lehoucq R B, Gray S K, Zhang D H and Light J C 1998 Vibrational eigenstates of four-atom molecules, a parallel strategy employing the implicitily restarted Lanczos method *Comput. Phys. Commun.* **109** 15–26
[18] Yang C, Peyton B W, Noid D W, Sumpter B G and Tuzun R E 2001 Large-scale normal coordinate analysis for molecular structures *Siam J. Sci. Comput.* **23** 563–82
[19] Kuriyan J and Weis W I 1991 Rigid protein motion as a model for crystallographic temperature factors *Proc. Natl Acad. Sci. USA* **88** 2773–7
[20] Miller D W and Agard D A 1999 Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease *J. Mol. Biol.* **286** 267–78
[21] Gurin T and Bruinsma R 2007 Theory of conformational transitions of viral shells *Phys. Rev.* E **76** 061911
[22] Zandi R and Reguera D 2005 Mechanical properties of viral capsids *Phys. Rev.* E **72** 021917
[23] Nguyen T T, Bruinsma R F and Gelbart W M 2006 Continuum theory of retroviral capsids *Phys. Rev. Lett.* **96** 078102
[24] Sohlberg K, Sumpter B G, Tuzun R E and Noid D W 1998 Continuum methods of mechanics as a simplified approach to structural engineering of nanostructures *Nanotechnology* **9** 30–6
[25] Landau L D and Lifshitz E M 2001 *Theory of Elasticity* (New York: Pergammon)
[26] Graff K F 1991 *Wave Motion in Elastic Solids* (New York: Ohio State University Press)
[27] Lamb H 1882 *Proc. Lond. Math. Soc.* **13** 187
[28] Talati M and Jha P K 2006 Acoustic phonon quantization and low-frequency Raman spectra of spherical viruses *Phys. Rev.* E **73** 011901

[29] Yang Z, Bahar I and Widom M 2009 Vibrational dynamics of icosahedrally symmetric biomolecular assemblies compared with predictions based on continuum elasticity *Biophys. J.* **96** 4438–48

[30] Tachibana M, Kojima K, Ikuyama R, Kobayashi Y and Ataka M 2000 Sound velocity and dynamic elastic constants of lysozyme single crystals *Chem. Phys. Lett.* **332** 259–64

[31] Balandin A A and Fonoberov V A 2005 Vibrational modes of nano-template viruses *J. Biomed. Nanotechnol.* **1** 90–5

[32] Dykeman E C and Sankey O F 2007 Raman intensity and spectra predictions for cylindrical viruses *Phys. Rev.* E **76** 011906

[33] Snoke D W and Cardona M 1993 A bond polarizability model for the C60 Raman spectrum *Solid State Commun.* **87** 121–6

[34] Go S, Bilz H and Cardona M 1975 Bond charge, bond polarizability, and phonon spectra in semiconductors *Phys. Rev. Lett.* **34** 580–3

[35] Guha S, Menndez J, Page J B and Adams G B 1996 Empirical bond polarizability model for fullerenes *Phys. Rev.* B **53** 13106–14

[36] Bermejo D, Montero S, Cardona M and Muramatsu A 1982 Transferability of the bond polarizabilities: from saturated hydrocarbons to diamond *Solid State Commun.* **42** 153–5

[37] Tirion M M 1996 Large amplitude elastic motions in proteins from a single parameter, atomic analysis *Phys. Rev. Lett.* **77** 1905–8

[38] Bahar I, Lezon T R, Yang L-W and Eyal E 2010 Global dynamics of proteins: bridging between structure and function *Annu. Rev. Biophys.* **39** 2342

[39] Thorpe M F 2007 Comment on elastic network models and proteins *Phys. Biol.* **4** 60–3

[40] Tama F, Gadea F X, Marques O and Sanejouand Y H 2000 Building block approach for determining low-frequency normal modes of macromolecules *Proteins* **41** 1–7

[41] van Vlijmen H W T and Karplus M 2001 Normal mode analysis of large systems with icosahedral symmetry: application to (Dialanine)$_{60}$ in full and reduced basis set implementations *J. Chem. Phys.* **115** 691–8

[42] Tama F and Brooks C L III 2005 Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis *J. Mol. Biol.* **345** 299–314

[43] Peeters K and Taormina A 2009 Group theory and icosahedral virus capsid vibrations: a top down approach *J. Theor. Biol.* **256** 607–24

[44] Marques O and Sanejouand Y-H 1995 Hinge-bending motion in citrate synthase arising from normal mode calculations *Proteins* **23** 557

[45] Valadie H, Lacapcre J J, Sanejouand Y-H and Etchebest C 2003 Dynamical properties of the MscL of *Escherichia coli*: a normal mode analysis *J. Mol. Biol.* **332** 657

[46] Tama F and Brooks C L III 2006 Symmetry, form, and shape: guiding principals for robustness in macromolecular machines *Annu. Rev. Biophys. Biomol. Struct.* **35** 115

[47] Tama F, Miyashita O and Brooks C L III 2004 Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM *J. Struct. Biol.* **147** 315–26

[48] Tama F and Brooks C L III 2002 The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus *J. Mol. Biol.* **318** 733

[49] Sukharev S I, Blount P, Martinac B and Kung C 1997 Mechanosensitive channels of *Escherichia Coli*: the MscL gene, protein, and activities *Annu. Rev. Physiol.* **59** 633

[50] Bahar I, Atilgan A R and Erman B 1997 Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential *Fold. Des.* **2** 173–81

[51] Speir J E, Johnson J A, Munshi S, Wang G, Timothy S and Baker J E 1995 Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by x-ray crystolography and cryo-electron microscopy *Structure* **3** 63

[52] Jackson E J M 1991 *A Users Guide to Principal Components* (London: Wiley)

[53] Amadel A, Linssen A and Berendsen H J C 1993 Essential dynamics of proteins *J. Comput. Chem.* **17** 412–25

[54] Hespenheide B M, Jacobs D J and Thorpe M F 2004 Structural rigidity in the capsid assembly of cowpea chlorotic mottle virus *J. Phys.: Condens. Matter* **16** S5055–64

[55] Hayward S, Kitao A and Go N 1994 Harmonic and anharmonic aspects in the dynamics of BPTI: a normal mode analysis and principal component analysis *Protein Sci.* **3** 936–43

[56] Ramanathan A and Agarwal P K 2009 Computational identification of slow conformational fluctuations in proteins *J. Phys. Chem.* B **113** 16669–80

[57] Gorfe A A, Grant B J and McCammon J A 2008 Mapping the nucleotide and isoform-dependent structural and dynamical features of Ras proteins *Structure* **16** 885–96

[58] Grant B J, Rodrigues A P C, ElSawy K M, McCammon J A and Caves L S D 2006 Bio3d: an R package for the comparative analysis of protein structures *Bioinformatics* **22** 2695–6

[59] Ordejon P, Drabold D A, Martin R M and Grumbach M P 1995 Linear system-size scaling methods for electronic-structure calculations *Phys. Rev.* B **51** 1456–76

[60] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1986 *Numerical Recipies* (Cambridge: Cambridge University Press) p 303

[61] Tsen K-T, Tsen S-W, Sankey O F and Kiang J 2007 Selective inactivation of micro-organisms with near-infrared femtosecond laser *J. Phys.: Condens. Matter* **19** 472201

[62] Tsen K-T, Tsen S-W D, Hung C-F, Wu T-C and Kiang J G 2008 Selective inactivation of human immunodeficiency virus with subpicosecond near-infrared laser pulses *J. Phys.: Condens. Matter* **20** 252205

[63] Yan T-X, Gambel E B and Nelson K A 1985 Impulsive stimulated scattering: general importance in femtosecond laser pulse interactions with matter, and spectroscopic applications *J. Chem. Phys.* **83** 5391–9

[64] De Silvestri S, Fugimoto J G, Ippen E P, Gamble E B Jr, Williams L R and Nelson K A 1985 Femtosecond time-resolved measurements of optic phonon dephasing by impulsive stimulated Raman scattering in α-perylene crystal from 20 to 300 K *Chem. Phys. Lett.* **116** 146–52

[65] Dykeman E C and Sankey O F 2009 Vibrational energy funneling in viruses simulations of impulsive stimulated Raman scattering in M13 bacteriophage *J. Phys.: Condens. Matter* **21** 505102

[66] Grant R A, Hiremath C N, Filman D J, Syed R, Andries K and Hogle J M 1994 Structures of poliovirus complexes with anti-viral drugs: implications for viral stability and drug design *Curr. Biol.* **4** 784–97

[67] Tsui V and Case D A 2001 Theory and applications of the generalized Born solvation model in macromolecular simulations *Biopolymers* **56** 275–91

[68] Bashford D and Case D A 2000 Generalized Born models of macromolecular solvation effects *Annu. Rev. Phys. Chem.* **51** 129–52

[69] Petrone P and Pande V S 2006 Can conformational change be described by only a few normal modes? *Biophys. J.* **90** 1583–93

[70] Norris G E, Anderson B F and Baker E N 1991 Molecular replacement solution of the structure of apolactoferrin, a

protein displaying large-scale conformational change *Acta Crystallogr.* B **47** 998

[71] Tama F and Sanejouand Y H 2001 Conformational change of proteins using normal mode calculations *Protein Eng.* **14** 1–6

[72] Bahar I, Lezon T R, Bakan A and Shrivastava I H 2010 Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins *Chem. Rev.* **110** 1463–97

[73] Holland I B, Cole S P C, Kuchler K and Higgins C F 2003 *ABC Proteins: from Bacteria to Man* (London: Academic)

[74] Lewis H A 2005 Impact of the ΔF508 mutation in first nucleotide-binding domain of human cystic fibrosis transmembrane conductance regulator on domain folding and structure *J. Biol. Chem.* **280** 1346

[75] Kolovou G D *et al* 2006 Tangier disease four decades of research: a reflection of the importance of HDL *Curr. Med. Chem.* **13** 771

[76] Ueda K, Cardarelli C, Gottesman M M and Pastan I 1987 Expression of a full-length cDNA for the human 'MDR1' gene confers resistance to colchicine, doxorubicin, and vinblastine *Proc. Natl Acad. Sci.* **84** 3004

[77] Nikaido H and Hall J A 1988 Overview of bacterial ABC transporters *Methods Enzymol.* **292** 3

[78] Hollenstein K, Frei D C and Locher K P 2007 Structure of an ABC transporter in complex with its binding protein *Nature* **446** 213

[79] Kandt C, Xu Z and Tieleman D P 2006 Opening and closing motions in the periplasmic vitamin B_{12} binding protein BtuF *Biochemistry* **45** 13284

[80] Sharff A J, Rodeseth L E, Spurling J E and Quiocho F A 1992 Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis *Biochemistry* **31** 10657