Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 24 (2005) 46–58
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Efficient determination of low-frequency normal modes of large protein structures by cluster-NMA

Adam D. Schuyler, Gregory S. Chirikjian *

Department of Mechanical Engineering, The Johns Hopkins University, 223 Latrobe Hall, 3400 North Charles Street, Baltimore, MD 21218, USA

Received 12 January 2005; received in revised form 18 March 2005; accepted 18 March 2005
Available online 28 June 2005

Abstract

The structure–function relationship is critical to understanding the biologically relevant functions of protein structures. Various experimental techniques and numerical modeling methods, normal mode analysis (NMA) in particular, have been employed to gain insight into this relationship. Experimental methods are often unable to provide all the desired information and comprehensive modeling techniques are often too computationally expensive. The authors build upon and optimize their cluster normal mode analysis (cNMA) tool, which uses embedded rigid-bodies and harmonic potentials to capture the biologically significant, low-frequency, oscillations of protein structures. cNMA represents atomic details with a scalable number of degrees-of-freedom, which can be chosen independent of structure size. This representation overcomes the otherwise quadratic order memory requirements and cubic order computational complexity associated with traditional all-atom NMA. cNMA is two orders of *magnitude* faster than traditional all-atom NMA when clustering by residue (very high resolution) and in the more traditional application using a fixed number of clusters, cNMA computationally scales as $\mathcal{O}(n)$, which is two orders of *complexity* faster than all-atom NMA. cNMA is presented and very large example structures with up to 10^6 atoms are analyzed on a notebook PC in the time scale of minutes/hours. The resulting mode shapes help identify biologically significant, conformational pathways.

Keywords: Protein mechanics; Normal mode analysis (NMA); Cluster-NMA (cNMA); Rigid-body motions; Elastic network; GroEL/GroES; Virus capsid

1. Introduction

X-ray crystallography yields atomic coordinates of protein structures in their (near) equilibrium conformations, which can be used to further understand the relationship between protein structure and function. Experimentally derived temperature factors [1] and Raman spectroscopy are able to identify regions of the structure that experience higher magnitudes of mobility. These techniques start to go beyond the static information of the crystal structure, but only provide magnitudes of motions, not directions. Regions with high mobility do not necessarily indicate regions containing conformational changes—rigid regions extending away from a hinge will experience maximal motion, but the region of interest is really the hinge. Other experimental methods, such as fluorescent resonance energy transfer and nuclear magnetic resonance can provide partial information about large-amplitude protein motions [2–4], but the details of the full dynamics are still incomplete.

Computational models have proved quite valuable in gaining insights into the structure–function relationship, particularly where the current experimental methods fall short in either capability or level of detail. Molecular dynamics provides a comprehensive picture, but is computationally prohibitive for all but the smallest structures due to the set of complex, coupled interactions, required to represent such high degree of freedom (DOF) structures. Consequently, the available computing power limits such simulations to very short time scales.

Classical normal mode analysis (NMA) techniques, based on very simple potential functions (see [5] for a simplified protein model, and [6] for energy function calculations along with a formulation of NMA equations) have proved to be viable alternatives, as demonstrated in [7]. The scalar Gaussian network model [8] and the directional

* Corresponding author. Tel.: +1 410 516 7127; fax: +1 410 516 7254.
E-mail address: gregc@jhu.edu (G.S. Chirikjian).

1093-3263/$ – see front matter
doi:10.1016/j.jmgm.2005.05.002

anisotropic network model [9] have been employed with great success in determining global motions. The Gaussian network model is computationally cheap, but only produces scalar magnitudes of displacement. The anisotropic network model provides directional information, but at a prohibitive computational cost for an all atom representation of all but relatively small structures.

The MBO(N)D method presented in Chun et al. [10] decomposes the structure into rigid and flexible bodies. NMA is performed on the flexible bodies and the corresponding modes are superimposed over the rigid motions of the body, based on various frequency matching criterion.

Computational limitations have been reduced by the usage of coarse-grained models [11]. In [12], the authors consider varying lengths of sequential units of residues and calculate corresponding radii of gyration to establish a relationship between an appropriate interaction cutoff distance and segment length. For varying levels of coarse graining, an anisotropic network model is used to compute atomic mean-square fluctuations, which compare favorably with crystallographic temperature factor data. In particular, as the model becomes coarser, the mean-square fluctuations smooth out. The authors conclude that the slowest mode shape is still adequately captured when the sequential segments contain up to 10 residues each.

A major reason for efficiently, and accurately, calculating slow modes follows from statistical mechanics, which dictates that the harmonic motions around the equilibrium conformation contribute proportional to the inverse of their frequencies (i.e. low-frequency motions contribute the most to the global motion of the structure [13,14]). It has been well observed that the biologically significant motions are characterized by low-frequency, large-amplitude motions [14–17], in which entire regions often move together. In contrast, the high-frequency, localized vibrations, may be more related to signal transmission and other internal processes [18].

The statistical dominance of slow modes along with their biological significance has lead to the development of a variety of efficient, low-frequency specific, NMA-based methods; many of which are discussed in Hinsen [15]. For example, a Fourier decomposition is used to represent the vector field of atomic displacements. The frequency parameter is constrained in order to capture the low-frequency modes. The author also discusses C_α coarse grained models which use simplified force fields. Such models have been shown to reproduce temperature factor data, thus supporting the concept of using simplified force fields to calculate slow, collective motions. Hinsen [15] also details a deformation analysis method which can be used to identify rigid regions.

The dominance of low-frequency motions as dictated by statistical mechanics and the successes of other NMA-based techniques are the motivating factors for the development of the rigid-body-based NMA technique, called cluster-NMA (cNMA), which was originally presented as an α-carbon, coarse grained method by Schuyler and Chirikjian [19]. In the present paper, the core concepts of cNMA are reviewed. A detailed comparison is made with the RTB method [20], which is also a cluster-based NMA method, but as the analysis shows, there are several fundamental characteristics that set the methods apart. Finally, the newly optimized cNMA method is applied at atomic resolution to a few very large structures (up to 10^6atoms) to illustrate how the large amplitude motions associated with biological function can be accurately and efficiently computed from the purely geometry-based model.

2. Review of cluster-NMA

The original presentation of cNMA is given by Schuyler and Chirikjian [19], and its primary purpose is to show the full derivation of cNMA, verify its accuracy through comparison with classical, all-atom NMA, and show the computational savings achieved on a set of sample structures. The core concepts and notation of cNMA are reviewed here[1] and the reader is referred to [19] for additional details and a more thorough literature review. In particular, the reader should consult Bahar et al. [8] for the formulation of the scalar Gaussian network model and Atilgan et al. [9] for the directional anisotropic network model. The theoretical successes, and computational limitations,[2] of these methods are the motivation for cNMA.

2.1. Coordinate system

Given a structure with n atoms it can be fully represented in Cartesian coordinates by

$$\boldsymbol{\sigma} = [\boldsymbol{\sigma}_1^{\mathrm{T}}, \ldots, \boldsymbol{\sigma}_n^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{3n} \tag{1}$$

where each atom is parameterized by its position in the global coordinate frame as

$$\boldsymbol{\sigma}_i = [x_i, y_i, z_i]^{\mathrm{T}} \in \mathbb{R}^3 \tag{2}$$

where the superscript T is used to indicate the transpose. Classical all-atom NMA is typically based on the displacement experienced by the generalized coordinate $\boldsymbol{\sigma}$. Other variables, such as torsion angles, can also be used, but they also require $\mathcal{O}(n)$ parameters.

In cNMA, the structure is represented as N rigid bodies (clusters of atoms). The structure's conformation is thus defined by the position and orientation of the N embedded rigid bodies—or equivalently, the translational and rota-

[1] In Schuyler and Chirikjian [19], α-carbon traces are used to compare cNMA results with all-atom NMA. With the improved performance of cNMA, we now consider all atom representations and significantly larger structures.

[2] In Thomas et al. [21], a standard normal mode analysis in Cartesian coordinates has been performed on the 2757 residue ATCase using a Cray C98 supercomputer. The first 53 normal modes require 690 h of computation time.

tional displacement relative to the crystal structure. The translational displacement of cluster i is measured from its center of mass, $\boldsymbol{x}_i$, relative to its starting position and is given by

$$\boldsymbol{\chi}_i(t) = \boldsymbol{x}_i(t) - \boldsymbol{x}_i(0) \tag{3}$$

The t parameter is used to indicate the original crystal structure ($t = 0$) and motions away from this equilibrium state ($t > 0$). When it is understood that an arbitrary conformation is being discussed, t may not be explicitly written.

The rotational displacement around the cluster's center of mass is defined by the rotation matrix $R_i(t) \in \mathrm{SO}(3)$, where $\mathrm{SO}(3)$, is the special orthogonal group of rotations in three-dimensional space. The three-parameter axis-angle vector, $\boldsymbol{\gamma}_i(t)$, corresponding to the rotational displacement is defined by the relation

$$\begin{aligned} R_i(\boldsymbol{\gamma}_i(t)) &\triangleq \exp\left(J(\boldsymbol{\gamma}_i(t))\right) \\ &= \mathbb{I}_3 + \left(\frac{\sin\left(\|\boldsymbol{\gamma}_i(t)\|\right)}{\|\boldsymbol{\gamma}_i(t)\|}\right)[J(\boldsymbol{\gamma}_i(t))] \\ &\quad + \left(\frac{1-\cos\left(\|\boldsymbol{\gamma}_i(t)\|\right)}{\|\boldsymbol{\gamma}_i(t)\|^2}\right)[J(\boldsymbol{\gamma}_i(t))]^2 \end{aligned} \tag{4}$$

where $\|\boldsymbol{\gamma}_i(t)\|$ is the usual vector norm and the skew symmetric matrix function, $J : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$, is defined by

$$J([a,b,c]^{\mathrm{T}}) = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \tag{5}$$

once again, the t parameter may be omitted from $\boldsymbol{\gamma}_i(t)$ when arbitrary conformations are being discussed.

Each cluster's generalized coordinates are given by

$$\boldsymbol{\delta}_i = [\boldsymbol{\chi}_i^{\mathrm{T}}, \boldsymbol{\gamma}_i^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^6 \tag{6}$$

and the whole structure's generalized coordinates are the stacked vector

$$\boldsymbol{\delta} = [\boldsymbol{\delta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\delta}_N^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{6N} \tag{7}$$

From these definitions, we can express the Cartesian location of atom α in cluster i as

$$\boldsymbol{s}_{i,\alpha}(t) = \underbrace{[R(\boldsymbol{\gamma}_i(t))](\boldsymbol{s}_{i,\alpha}(0) - \boldsymbol{x}_i(0)) + \boldsymbol{x}_i(0)}_{\text{rotate about cluster's center of mass}} + \underbrace{\boldsymbol{\chi}_i(t)}_{\text{translate}} \tag{8}$$

The rotation is applied to the atom's position *relative* to its cluster's center of mass—hence the term $\boldsymbol{s}_{i,\alpha}(0) - \boldsymbol{x}_i(0)$, which is equivalent to translating the cluster's center of mass to the origin before applying the rotation. The corresponding cluster's center of mass is then translated back to its original position by adding back $\boldsymbol{x}_i(0)$. Finally, the entire cluster is translated by the displacement $\boldsymbol{\chi}_i(t)$.

2.2. cNMA interaction model

The protein structure is represented by a set of masses, one located at each atomic position as specified by the crystal coordinates. The atomic interactions are modelled by a set of harmonic potentials that connect all pairs of atoms within a cutoff distance of $r_c = 5$ Å. This system is typically used for all-atom NMA.

Many coarse-grained NMA techniques reduce the structure by representing groups of atoms by single point masses (i.e. NMA on the α-carbon trace). This effectively collapses all atoms of the group (and their interactions with other atoms) onto a single position. In contrast, cNMA represents groups of atoms as rigid-bodies, but maintains individual atom positions and interaction contributions by including translation *and* rotation parameters in the coordinate system.

Even though these analysis techniques are designed to capture the slow modes, the oversimplified structure representation of some models distorts the geometry of atomic interactions. In particular, forces acting on each atom of a cluster under cNMA will determine the net rotational motion of the cluster, whereas coarse-grained models which use point mass representations for collections of atoms will only be able to identify translation of the group of atoms, thus neglecting an entire component of the structure's motion.

In typical all-atom NMA, the network of atomic interactions can be determined with a very straight forward implementation by considering all pairs of atoms. This $\mathcal{O}(n^2)$ operation is insignificant because the $\mathcal{O}(n^3)$ eigenproblem dominates the overall complexity. However, in the case of cNMA, as will be shown, the computational complexity of the eigenproblem can be made independent of structure size (constant complexity). This leaves the contact calculation as a potential computation limiting step. In response, a new method for calculating atomic pair interactions is required—this material is presented in Appendix A.

It should be noted that, if desired, the masses of the model can be scaled according to each specific atom type and the potentials can be scaled according to the chemical properties of the atoms involved in each interaction. However, since low-frequency motions are of primary interest, unit masses and a uniform potential are sufficient to capture the structure's global characteristics.

2.3. Equation of motion

For an arbitrary small structure motion, $\boldsymbol{\delta}$, the associated kinetic and potential energies can be approximated very well as the quadratic expressions

$$T = \tfrac{1}{2}\dot{\boldsymbol{\delta}}^{\mathrm{T}} M \dot{\boldsymbol{\delta}} \tag{9}$$

and

$$V = \tfrac{1}{2}\boldsymbol{\delta}^{\mathrm{T}} K \boldsymbol{\delta} \tag{10}$$

The complete derivation and closed form expressions for the mass matrix, M, and stiffness matrix, K, are given in [19]. Lagrange's equation directly produces the equation of motion

$$M\ddot{\boldsymbol{\delta}} + K\boldsymbol{\delta} = \boldsymbol{0} \tag{11}$$

whose solutions contribute to the harmonic motions around the equilibrium conformation inversely proportional to their corresponding frequencies.

2.4. Clustering guidelines

The cNMA parameterization is used for two main reasons. First, when probing the structure–function relationship, low-frequency, large amplitude, motions are of the most interest. They exhibit large collective motions (i.e. parameterization by rigid clusters is sufficient to capture these motions). Second, the total number of degrees of freedom is drastically reduced, thus providing a significant enhancement of computational performance.

In [19], clustering was done evenly across a structure while keeping the ends of helices at the cluster interfaces. This construction was chosen because the favorable hydrogen bonding characteristic of helices is well conserved across conformational changes, especially under low-frequency motions. In the present context clustering is applied according to the structure size guidelines[3] shown in Table 1. These guidelines are intended to keep the total number of DOFs to no more than approximately 5000.

The computational complexity of assigning the clusters is, at worst, $\mathcal{O}(n)$ when clustering by residue, domain, chain, or unit—simply iterate though all atoms and assign cluster indices accordingly. If domains are pre-defined, then the cluster assignments can be done in constant time. In addition to assigning the cluster indices, each cluster's initial orientation and position must be calculated. By definition, the initial orientation is set to the identity matrix (rotational displacements are measured relative to this initial reference frame). Each cluster's initial position is set to its center of mass as measured in the global coordinate frame. Since every atom in the structure is in exactly one cluster, every atom will be considered exactly once during the center of mass calculations. Overall, the clustering assignments and initialization is an $\mathcal{O}(n)$ process.

The cNMA examples presented in the following sections show each of the three clustering methods mentioned in Table 1.

[3] As an alternative, cNMA resolution, as determined by cluster size, can be varied across the structure to achieve higher resolution in areas of particular functional interest and lower resolution over highly conserved regions. This strategy minimizes computational costs while maintaining the atomistic model. A method for varying resolution in the anisotropic network model is presented in Kurkcuoglu et al. [22]. The same theory and procedure holds for cNMA.

Table 1
Structure size cluster guidelines

Structure size	Clustering method
$\approx$100–750 residues	Cluster by residue
$\approx$ 750–5000 residues	Cluster by domain
$>$5000 residues	Cluster by chain/unit

3. Comparison with RTB method

The RTB method presented in Tama et al. [20], and implemented in CHARMM in Li and Cui [23], is an accepted coarse-grained, NMA technique. It uses an embedded rigid-body structure representation to circumvent the diagonalization of the Hessian matrix, which is the memory and computation limiting step of most NMA-based methods. By pre-computing the rigid-body modes of each block of consecutive residues, the Hessian is projected onto a reduced DOF subspace. The reduction of a structure's representation space is common to many NMA-based techniques, including cNMA. However, the way in which this is accomplished almost completely defines the method and has considerable impact on the remaining computations. To illustrate this point and highlight the key operational aspects of cNMA, a comparative analysis with the RTB method is presented.

Major distinctions between the RTB method and the cNMA method stem from the coordinate system choices. Tama et al. choose to use a full Cartesian atomic representation for the Hessian and then transform it into a reduced DOF subspace. In contrast, the present authors start by defining a rigid-body coordinate system and then formulate the *entire* cNMA procedure in the inherently reduced DOF representation (i.e. no transforms are necessary). Even though both options produce valid models, this key difference has many implications in the subsequent normal mode calculations and mode shape analyses.

First, the projection matrix used in the RTB method is a Cartesian representation of the clusters' rigid-body motions and its transform of the Hessian reshuffles the coordinate system. The resulting eigenvectors (mode shapes) of the reduced DOF Hessian must therefore be transformed back into all-atom Cartesian coordinates in order to "see" the modes. This extra transform is not computationally costly, but it forces an all atom representation when it is not desired. In contrast, the cNMA method maintains its coordinate system throughout the computations. Since the coordinate system is a natural reduced DOF representation, computational performance is enhanced and cooperative motions are easily identified directly from the cNMA mode shapes by comparing the clusters' translational and rotational components. Similar translational components indicate regions that move together and similar rotational components (i.e. the axis/angle vector describing the rotation) indicate common hinge axes.

Second, the coordinate system of cNMA is designed to describe rigid-body motions where atoms will experience

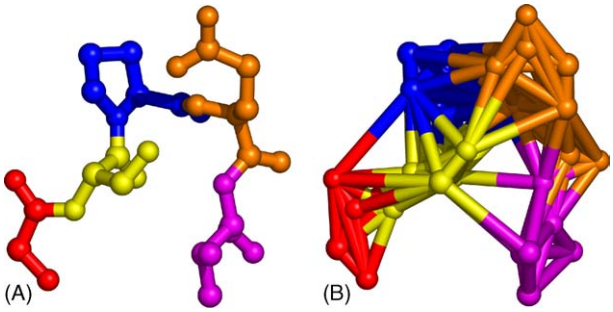

Fig. 1. Illustration of which pairwise atomic interactions are required by standard NMA (all sticks in B) and cNMA (only the sticks in B between different colored atoms).

(linear) translations and (curvilinear) rotations. Mode shapes can therefore be extrapolated under the cNMA method without distorting local geometries, whereas the RTB method will *not* maintain local geometries as Cartesian representations of rotations are amplified. It is best not to use significant extrapolation for pathway generation, but it does serve as a valuable visualization tool.

Third, cNMA is computationally faster than the RTB method. In Tama et al. [20], the RTB method is applied with one residue per block to the 99 residue HIV-1 protease

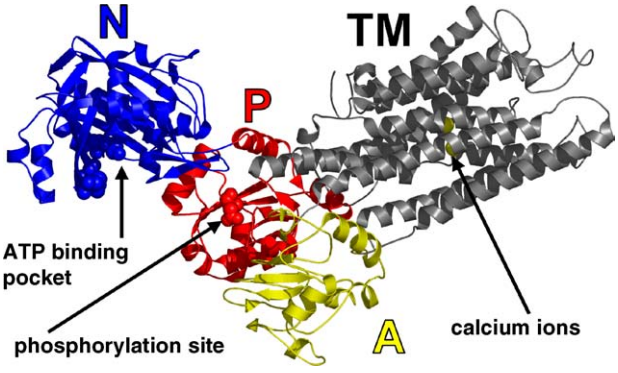

Fig. 2. The four clusters/domains are given by the sets of residues: $A = \{1-40, 126-245\}$, $N = \{360-604\}$, $P = \{330-359, 605-737\}$, and $TM = \{41-125, 246-329, 738-994\}$. The pair of calcium ions, phosphorylation site and ATP binding site are shown with space filling spheres and labelled accordingly.

Table 2
Descriptions of some of the lowest non-rigid-body modes of calcium ATPase with four domain cluster representation

Mode index	Characterization
7	N rotates about axis perpendicular to TMa and "closes" down on P
8	N rotates towards A about axis almost parallel to TMa
9	N and A rotate away from each other about axes parallel to TMa, fully exposing the face of P
10	A "closes" onto P by hinging around axis running through A/P interface
14	P rocks forward into the crease between A and N

Notational shorthands: Single letters are used to refer to domains, TMa = axis running down center of TM helix bundle.

Table 3
Computation times for cNMA (clustering by domain) on calcium ATPase

Calculation	Time (s)
Read in PDB file	59.9
Determine all contacts	29.5
Cluster representation	0.4
M = mass matrix	0.7
K = stiffness matrix	2.7
$Q = M^{-1/2}KM^{-1/2}$	0.3
Determine eigenpairs of Q	<0.1
Transform modes back into cluster coords	<0.1
cNMA total	4.1

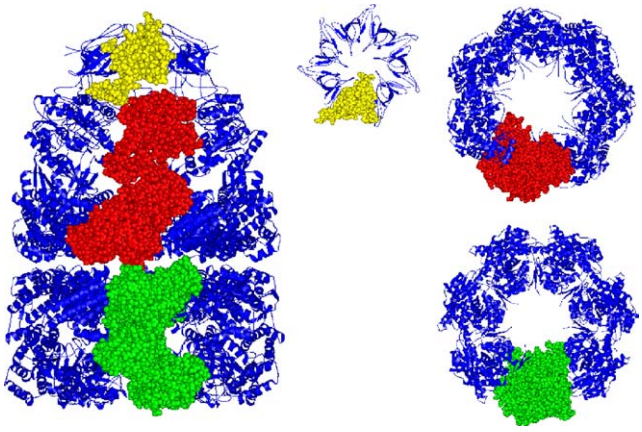

Fig. 3. The GroEL/GroES complex is shown on the left with a blue ribbon diagram. To illustrate how the seven chains on each level fit together, the following chains are shown with space filling spheres: chain P from GroES cap (yellow), chain A from upper GroEL ring (red), and chain N from lower GroEL ring (green). On the right are cross sectional slices of the cap and each seven chain ring.

(PDB = 1HHP, structure from Spinelli et al. [24]) and the 265 residue triglyceride lipase (PDB = 3TGL, structure from Brzozowski et al. [25]) on a "common HP workstation"[4] to obtain the slowest 50 modes of each structure in 90 and 1056 s, respectively. Application of cNMA[5] to HIV-1 protease took only 8.9 s to calculate 594 modes and cNMA on triglyceride lipase only took 101.5 s to compute 1590 modes. As an alternative, the slowest 50 modes are calculated with an iterative eigenvalue solver for each structure under cNMA in 6.0 and 33.1 s, respectively. The iterative method is less numerically stable than the full eigenvalue solver, which is why it is not used in the remainder of this paper, but it does allow for a more direct comparison with the published computation times of the RTB method. As these examples show, cNMA performs one to two orders of magnitude faster than the RTB method on these small to medium sized structures.

[4] No other platform data is given.

[5] All cNMA computations in this paper are performed on a notebook PC with a 1.6 GHz Pentium 4M processor and 512MB RAM running MATLAB 6.5.

Table 4
Computation times for cNMA (clustering by chain) on the GroEL/GroES complex

Calculation	Time (s)
Read in PDB file	168.1
Determine all contacts	414.7
Cluster representation	3.0
M =mass matrix	4.5
K = stiffness matrix	45.6
$Q = M^{-1/2}KM^{-1/2}$	0.7
Determine eigenpairs of Q	<0.1
Transform modes back into cluster coords	<0.1
cNMA total	53.8

Finally, the RTB method uses consecutive residues to define its clusters, whereas cNMA allows for any collection of atoms to define a cluster. The consecutive residue approach is simple in its application, but it has an inherent flaw. As clusters increase in size, they will tend to form elongated shapes, which are not nearly as rigid as more compact regions. Consequently, methods such as the RTB method must either limit cluster sizes to maintain an accurate model (as described in [12]) or risk locking up certain low-frequency motions. Further, the approach used in cNMA allows for varying cluster sizes to maintain high model resolution in regions of more structural/functional significance, while allowing bulk regions to remain coarse.

4. cNMA examples

4.1. Cluster by residue: a simple example

Fig. 1 shows clustering by residue on a 33 atom, five residue fragment with each residue shown in a different color. Image A shows all 33 covalent bonds in the structure and image B shows all 144 "sticks" between pairs of atoms within 5 Å. Under standard NMA, each stick in B, represents a contact and thus enters into the calculation of the stiffness matrix. When using cNMA (and clustering by residue) only the 91 contacts between atoms of *different* residues are needed. These are easily identified in B, as the sticks with two colors.

This simple example shows how cNMA, even when applied at a very high resolution, can reduce the complexity of the model. Standard NMA requires $33 \times 3 = 99$ DOFs, whereas cNMA only requires $5 \times 6 = 30$ DOFs. In the context of a cubic order eigenproblem, the reduction in DOFs by a factor of more than 3 corresponds to a computational performance factor of more than 35. In addition to the smaller matrices required in the cNMA equation of motion, the number of interactions has been slightly reduced, thus quickening the computation of the stiffness matrix. In larger structures, the cluster size can be set to maintain a constant number of DOFs. Once again, the total number of atomic interactions is also reduced, while preserving all atomic geometry. These fundamental concepts are common to all applications of cNMA.

4.2. Cluster by domain: calcium ATPase

The 994 residue calcium ATPase (PDB = 1EUL, structure from Toyoshima et al. [26]) pumps Ca^{2+}ions across the membrane of the sarcoplasmic reticulum against the concentration gradient in muscles during contraction [27]. This process is accomplished by a sequence of calcium ion bindings, ATP hydrolysis, and phosphorylation, which each cause rearrangements of the four domains [28,29]. The mechanical aspects of this process and the well-defined domains make cNMA (with clustering by domain) an ideal analysis tool.

Fig. 2 shows the positioning of the domains in the calcium ion bound conformation. The bundle of ten transmembrane α-helices is the largest domain (labelled as TM) and has a pair of Ca^{2+}-binding sites defined by pockets of negatively charged oxygen atoms. The first site is defined by the side-chain oxygen atoms of Asn^{768}, Glu^{771}, Thr^{799}, Asp^{800} and Glu^{908}. The second site is defined by the carbonyl oxygen atoms of Val^{304}, Ala^{305} and Ile^{307}, and the side-chain oxygen atoms of Asn^{796}, Asp^{800} and Glu^{309}. The other three domains

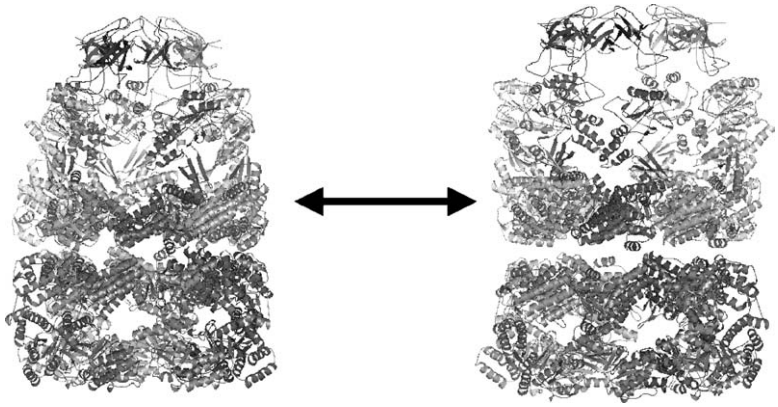

Fig. 4. Amplified visualization of mode 7. The lower ring rotates in the opposite direction as the upper ring and the GroES cap. There is also a slight radial expansion at the interface of the upper ring and the GroES cap.

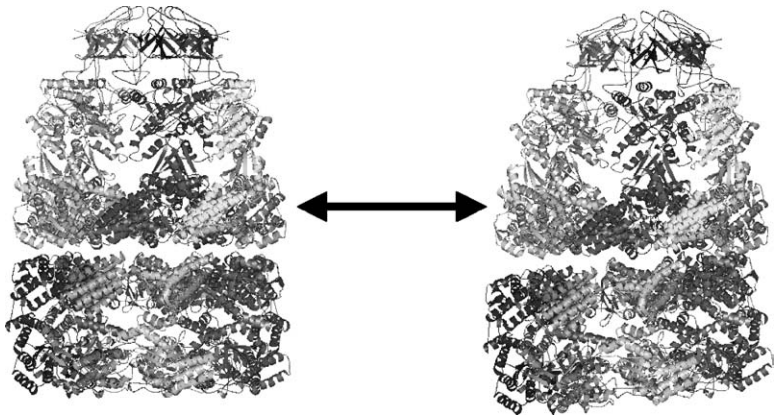

Fig. 5. Amplified visualization of mode 8. The lower ring rotates counter-clockwise about a radial axis perpendicular to the page and the upper ring and GroES cap experience a slight shear to the right in the plane of the page. Mode 9 is related to this motion by an approximate 90 ° rotation about the vertical axis.

(P, N and A) form the cytoplasmic headpiece, as referred to in Toyoshima et al. [26]. The phosphorylation site is in domain P at Asp^{351} and the ATP binding pocket is in domain N and is defined by Lys^{515} along with the supporting residues Phe^{487} and Lys^{492}. These critical regions are part of the calcium ion pump operation:

1. Calcium ions bind to sites within the helix bundle and cause conformational changes in TM.
2. This triggers the cytoplasmic headpiece to split open [26,30], thus allowing ATP hydrolysis on domain N and enzyme phosphorylation on domain P.
3. The phosphorylation causes conformational changes in domain TM and the calcium ions are transferred across the membrane.

Calcium ATPase has 7673 atoms and thus requires 23,019 DOFs for all-atom NMA, whereas the low-frequency analysis of cNMA with four domain clusters requires only 24 DOFs (the Ca^{2+} ions are not included in the analysis). A summary of some of the cNMA non-rigid-body modes is given in Table 2. Of particular interest are modes 9 and 10, which show possible pathways for the opening and closing of the cytoplasmic headpiece.

The computation times are shown in Table 3. The times for reading in the PDB file and calculating the atomic contacts are not included in the total cNMA computation time listed because they must be done for *every* atomic resolution normal mode analysis method. These two ''set-up'' operations are $\mathcal{O}(n)$, and do not compare to the $\mathcal{O}(n^3)$ computational cost of the typical all-atom eigenpair

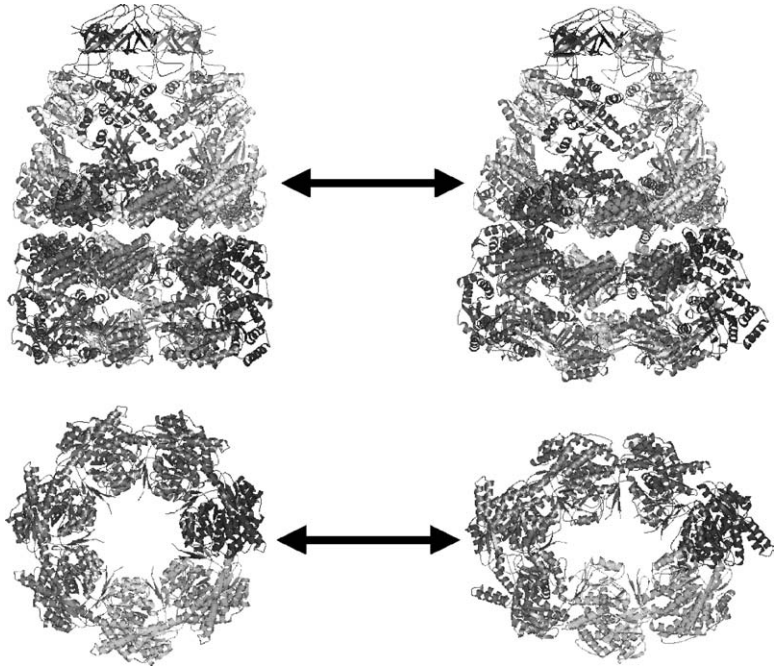

Fig. 6. Amplified visualization of mode 10. The lower ring stretches along a cross sectional direction and contracts in the perpendicular cross sectional direction. Mode 11 is related to this motion by an approximate 51 ° rotation about the vertical axis. This corresponds to a rotation of the seven chain ring by 1 position.

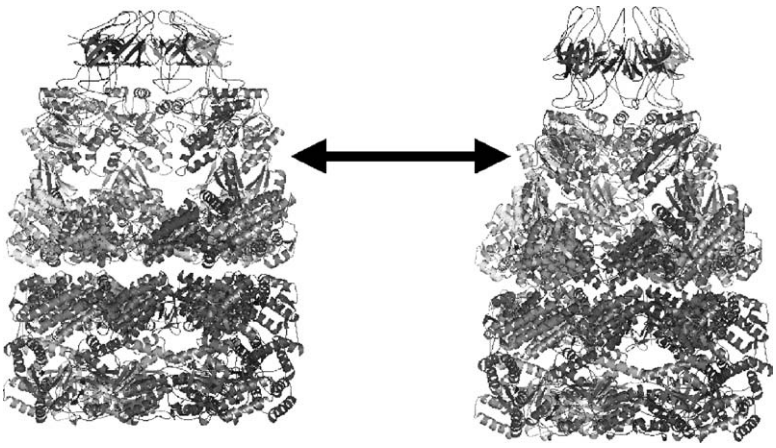

Fig. 7. Amplified visualization of mode 12. Symmetric radial breathing of the interface between the GroES cap and upper ring. The upper ring slightly twists counter-clockwise while the cap and lower ring twist together slightly clockwise.

calculation. However, as Table 3, and the examples to follow show, the cNMA calculations actually take *less* time than these traditional set-up costs—this represents a massive savings in computational complexity.

4.3. Cluster by chain: GroEL/GroES complex

GroEL is a protein folding/unfolding chaperonin complex composed of two rings, each with seven repeated chains. These rings stack about a central axis to form a tube, which can be capped by the seven chain GroES [31], see Fig. 3, for a ribbon representation and cross-sectional views of the GroEL/GroES complex (PDB = 1AON, structure from Xu et al. [32]). Most of the functionally significant sites are located on the interior channel walls and at the ends of the cylinder [33]. The central channel acts as an environment which helps facilitate proper folding [34]. It is suggested that some incorrectly folded structures are recognized by their exterior hydrophobic regions [32]. The central channel of the chaperonin can bind these regions and manipulate the target structure out of the incorrect fold before discharging the structure into the surrounding environment to naturally complete its fold [31].

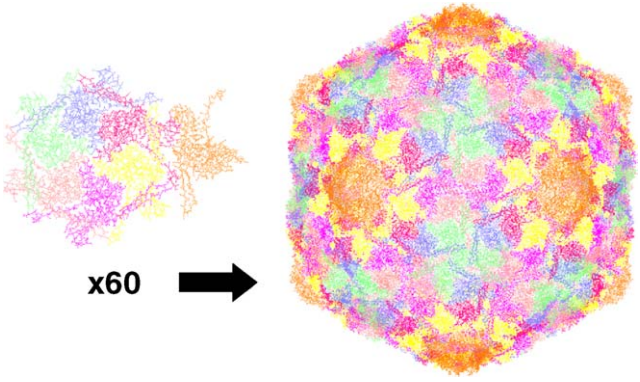

Fig. 8. Single seven chain unit (left) that when reproduced 60 times by symmetric transforms gives the complete virus capsid (right).

Table 5
Computation times for cNMA (clustering by unit and by chain) on the HK97 virus capsid

Calculation	Time (s)	
	By chain	By unit
Read in PDB file	9665.5	
Determine all contacts	85350.0	
Cluster representation	35.4	125.7
M = mass matrix	68.6	68.9
K = stiffness matrix	28253.5	14422.8
$Q = M^{-1/2}KM^{-1/2}$	154.6	3.5
Determine eigenpairs of Q	265.9	0.9
Transform modes back into cluster coords	5.1	0.1
cNMA total	28783.1	14621.9

Note: The crystal structure in the PDB file is only given for a single seven chain unit, the other 59 units are obtained by homogeneous transforms. The time given in the table for reading in the structure is 60 times the time required to read in the single seven chain unit.

This 8015 residue complex has 58,667 atoms—well beyond a personal computer's capacity for all-atom NMA. In Keskin et al. [35], the authors use an α-carbon trace to make the computations feasible. The authors state, the first 40 modes, a mere 0.1% of the 24,045 element mode set, contribute $\sim 75\%$ of the overall dynamics. The desired number of low-frequency modes can be determined one at a time (i.e. it is not necessary to compute all 24,045 modes only to use the first 10), but the computations still require the manipulation of matrices and vectors of dimension 24,045.

The multi-chain construction makes the complex well suited for cNMA. The 21 chains are modelled using 126 DOFs. The cNMA computation times are shown in Table 4, and visualizations[6] of the four lowest frequency families of

[6] *Note*: All mode shape visualizations in this paper are amplified to achieve RMS displacements on the order of 10 Å. These extrapolations go beyond the small-motion assumptions of the model and are only intended to ease visualization of mode *direction*.

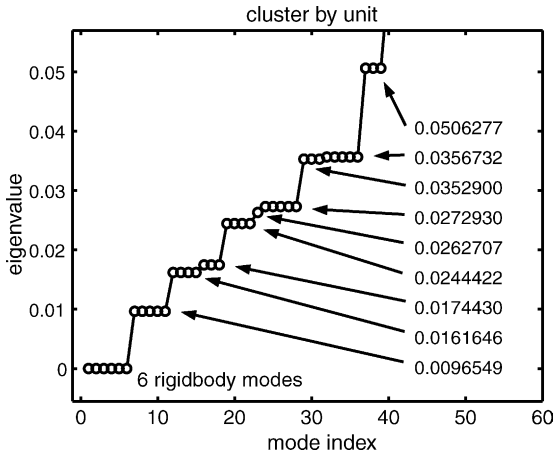

Fig. 9. Eigenvalue spectrum for cNMA on virus capsid clustered by chain.

mode shapes are shown in Figs. 4–7. Families 1 and 2 exhibit contortions of the central cavity which may be critical in binding incorrectly folded proteins to the hydrophobic sites on the inner channel walls and physically manipulating them out of the incorrect fold. Family 3 shows orthogonal stretching/contracting of the lower ring opening. Family 4 shows radially symmetric expansion of the upper ring/GroES cap interface coupled with extension of the cap away from the main structure along the central axis.

4.4. Cluster by chain: virus capsid HK97

Fig. 8 shows the icosahedral virus capsid HK97 (PDB = 1FH6, structure from Helgstrand et al. [36]), which is created from 60 repetitions of a seven chain unit [37]. This structure is chosen for analysis because its size (903,420 atoms) will test the computational performance of cNMA and its symmetry will allow us to compare the motions of cNMA with the motions calculated by other symmetry constrained methods [38]. Without the symmetry it would be nearly impossible to attempt all-atom NMA on a personal computer.

In this section we consider two different clustering approaches: one cluster per seven chain unit and one cluster per chain. Clustering by unit requires 60 clusters using 360 DOFs and clustering by chain requires 420 clusters using 2520 DOFs. Calculation times for both of these clustering schemes are shown in Table 5. In comparison, all-atom NMA on this 903,420 atom structure would require 2,710,260 DOFs!

As a consequence of the symmetry we obtain families of motions, each with a common frequency of oscillation. Fig. 9 shows a plot of the eigenvalue (squared frequency) spectrum for cNMA when clustering by unit. The plateau regions correspond to modes of the same ‘‘family’’ of symmetry related motions. A representative of each of the first four families is shown in Figs. 10–13 .

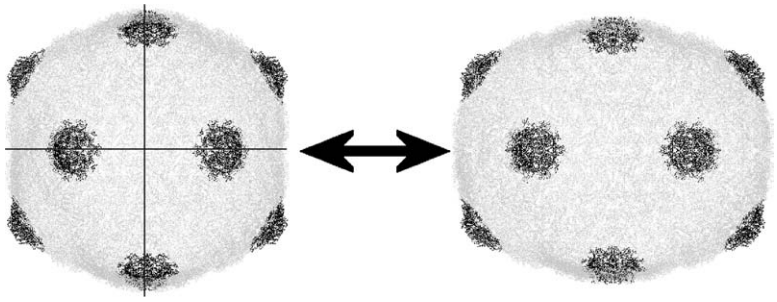

Fig. 10. Breathing motion along two perpendicular axes, each running from the center of the capsid though opposite edges (cluster by unit modes: 7, 8, 10 ; cluster by chain modes: 7, 8).

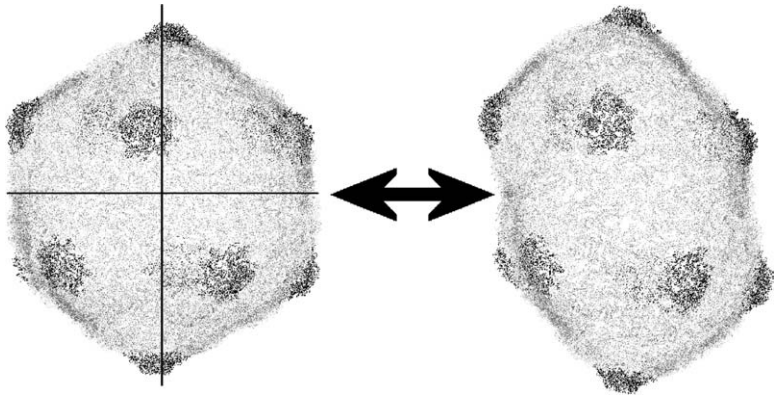

Fig. 11. Breathing motion along two perpendicular axes from the center of the capsid, one running though opposite edges and the other axis through opposite vertices (cluster by unit modes: 9, 11 ; cluster by chain modes: 9–11).

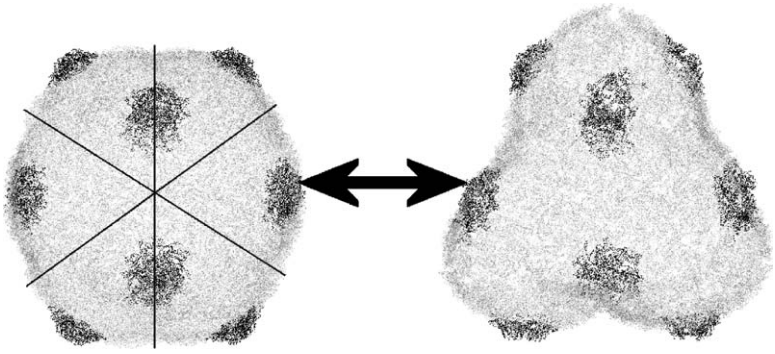

Fig. 12. Breathing motion along three evenly spaced axes around the ‘‘equator’’, each running from the center of the capsid though opposite edges (cluster by unit modes: 12–15 ; cluster by chain modes: 15, 16).

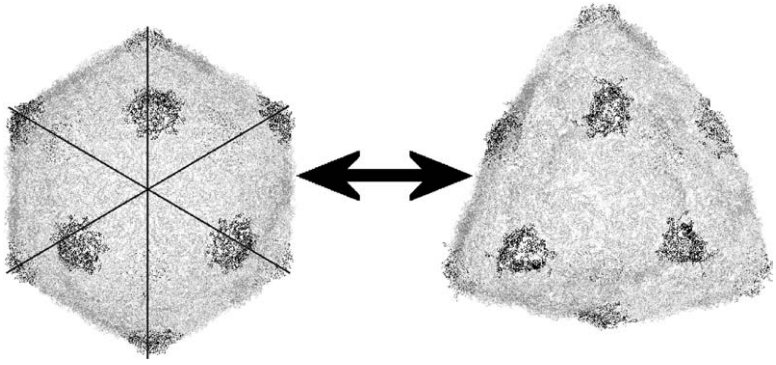

Fig. 13. Breathing motion along three evenly spaced axes around the ‘‘equator’’, each running from the center of the capsid though opposite vertices (cluster by unit modes: 16; cluster by chain modes: 12–14).

Table 6 shows the absolute values of taking the dot product of each of the first 12, non-rigid modes when clustering by unit with each of the first 10, non-rigid modes when clustering by chain. It is important to keep in mind that due to the rotational symmetry of the capsid, the mode ordering is arbitrary within each motion family. In order to compare the motions generated by each clustering scheme, one must consider sub-blocks of the decomposition matrix as defined by the motions of each family. The first 10 non-rigid modes, when clustering by chain, include motions up to the third plateau. To get a full decomposition we must consider all cluster-by-chain modes up to, and including, the third plateau. This requires 12 cluster-by-unit modes—hence the uneven number of rows and columns in Table 6.

The very high underlined values in each column of Table 6, show a correspondence between each cluster-by-

Table 6
Decomposition of Cartesian representation of modes when clustering by unit over Cartesian representation when clustering by chain

By unit	By chain									
	7	8	9	10	11	12	13	14	15	16
7	0.7776	0.6127	0.0005	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
8	0.6127	0.7776	0.0003	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0003	0.0185	0.9471	0.2875	0.0000	0.0000	0.0000	0.0000	0.0000
10	0.0000	0.0001	0.0715	0.2881	0.9444	0.0000	0.0000	0.0000	0.0000	0.0000
11	0.0006	0.0000	0.9872	0.0031	0.0738	0.0000	0.0000	0.0000	0.0000	0.0000
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8294	0.4792
13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1059	0.2738
14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5221	0.8168
15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0059	0.6239	0.7643	0.0000	0.0000
17	0.0000	0.0000	0.0000	0.0000	0.0000	0.4481	0.6793	0.5579	0.0000	0.0000
18	0.0000	0.0000	0.0000	0.0000	0.0000	0.8790	0.3504	0.2793	0.0000	0.0000
	0.9899	0.9899	0.9899	0.9899	0.9899	0.9867	0.9867	0.9867	0.9857	0.9857

The largest value in each column is underlined and the vector 2-norm of each 12 element column is written at the bottom.

chain mode and some cluster by unit mode. The vector norms under each column, all approaching a value of 1, show that each of the first 10, non-rigid, cluster-by-chain modes is almost entirely captured by the set of the first 12, non-rigid, cluster-by-unit modes. These numerical results precisely show that the low-frequency mode shapes are relatively insensitive to the chosen cluster size.

5. Conclusion

Classical all-atom NMA is dominated by the $\mathcal{O}(n^3)$ computation of the eigenproblem and the $\mathcal{O}(n^2)$ memory requirements [15]. cNMA overcomes these obstacles by using an $\mathcal{O}(n)$ method for representing arbitrarily large protein structures in a reduced DOF coordinate system that is designed to capture the low-frequency, biologically significant motions. Even for the very high resolution case of clustering by residue, cNMA exhibits two orders of magnitude worth of computational savings over classical NMA, as seen in Fig. 14. The computational performance can also be evaluated by observing that the series of computations associated with cNMA for each of the example structures takes about an order of magnitude less time than it takes to simply read in the PDB files and determine contacts. In other words, in the very high resolution (cluster-by-residue) application of cNMA, the computational complexity still scales as $\mathcal{O}(n^3)$, but significant computational performance improvements are still evident.

In the more common application of cNMA, one would fix the number of desired modes (i.e. the lowest 50 will be more than sufficient for most applications) and either use uniform cluster sizes across the entire structure or finely cluster the regions of interest, leaving a few larger clusters for the remaining bulk of the structure. In either case, the dimensions of the equation of motion are no longer a function of the overall structure size and consequently, the cubic scaling of the eigenproblem is reduced to a constant. Further, by using the $\mathcal{O}(n)$ approach for calculating atomic interactions, as discussed in Appendix A, the entire cNMA method operates at a computational complexity of $\mathcal{O}(n)$. Not only is this two orders of complexity faster than traditional all atom NMA, but also the compact cluster representation enables cNMA to escape the overwhelming memory requirements that often cripple other methods.

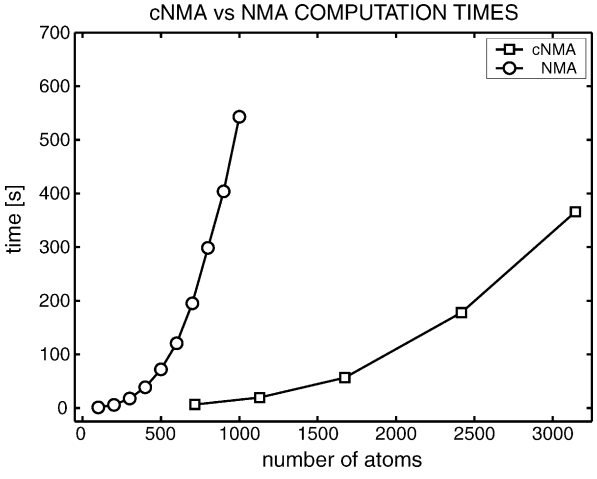

Fig. 14. Comparison of computation times between all-atom NMA (circles) and cNMA (squares) as a function of structure size. Even in this high resolution application of cNMA with clusters defined by residue, the computational savings factor approaches two orders of magnitude.

Appendix A. Interaction algorithms

Constructing the network of harmonic potentials by considering all possible pairs of atoms is very inefficient because an expensive distance calculation[7] is made between *all* pairs of atoms, when in fact, there is a limited number of contacts for each atom. The two alternative methods, referred to as box and voxel, are both based on discretizing Cartesian space and then using simple coordinate comparisons to reduce the number of distance calculations to a local neighborhood of candidates around each atom.

In the box method, there is a single iteration through the n atoms. For atom i, we calculate the relative x coordinate position of each of the $n-i$ atoms with larger indices (considering only larger indices eliminates "double counting"). From this data we keep only the atoms that are within the cutoff distance of atom i in the x coordinate dimension. This set is further reduced by making similar comparisons in the y and then the z coordinate dimensions—effectively trimming Cartesian space down to a box of edge length $2r_c$ centered at atom i. Since the sphere of radius r_c, centered at atom i, fits entirely in this box, all atoms within the cutoff distance of atom i (with index greater than i), are contained by the box. The Cartesian distance is computed *only* for this subset. The box method requires $\mathcal{O}(n^2)$ comparisons and $\mathcal{O}(n)$ vector norms.

The voxel method partitions Cartesian space into voxels (volume elements) with edge length r_c. Each voxel is identified by 3 integer coordinates. Each atom's voxel coordinates are computed and a lookup table is created by storing the voxel coordinates with each atom index *and* by storing each atom's index with the voxel coordinate. In a second iteration through the atoms, we consider atom i and all atoms within the $3 \times 3 \times 3$ voxel region centered around the voxel of atom i. The lookup table provides the atom indices contained in each neighboring voxel. Finally, the vector norm is used to calculate the distance to each candidate atom. This method requires $\mathcal{O}(n)$ computations to create the lookup table and $\mathcal{O}(n)$ vector norm calculations.

[7] Calculating the distance between two points in Cartesian space using the vector norm requires three multiplications, two additions, and one square root.

At first glance the voxel method appears to be the best choice for all cases. However, there is a subtlety which gives the computational advantage to the box method for structures below a threshold size. In the box method, a box of edge length $2r_c$ is created specifically for each atom. The volume of space with candidate atoms is thus $8r_c^3$. The voxel method uses a single partitioning for all computations (thus saving on ''setup'' costs), but as a consequence, 27 voxels of edge length r_c must be considered for each atom—a candidate volume of $27r_c^3$. Even though both methods require $\mathcal{O}(n)$ vector norm calculations, the box method operates on candidate volumes which are more than three times smaller than the corresponding volumes used by the voxel method. For structures with fewer than 12,500 atoms, the vector norm calculations on the additional volume of the voxel method are more computationally significant than the $\mathcal{O}(n^2)$ comparisons required by the box method. Conversely, for structures with more than 12,500 atoms, the voxel method performs more efficiently.

The computational advantage shifts to the voxel method as structure size increases, thus the computational complexity of calculating atomic contacts scales as $\mathcal{O}(n)$. The threshold value has been used to determine the optimal algorithm for the computations associated with each sample structure presented in this paper.

References

[1] S. Kundu, J.S. Melton, D.C. Sorensen, G.N. Phillips Jr., Dynamics of proteins in crystals: comparison of experiment with simple models, Biophys. J. 83 (2002) 723–732.

[2] T. Ha, Single-molecule fluorescence resonance energy transfer, Methods 25 (2001) 78–86.

[3] A.G. Palmer III, Probing molecular motion by NMR, Curr. Opin. Struct. Biol. 7 (1997) 732–737.

[4] S. Doniach, P. Eastman, Protein dynamics simulations from nanoseconds to microseconds, Curr. Opin. Struct. Biol. 9 (1999) 157–163.

[5] M. Levitt, A. Warshel, Computer simulation of protein folding, Nature 253 (5494) (1975) 694–698.

[6] S. Lifson, A. Warshel, Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and *n*-alkane molecules, J. Chem. Phys. 49 (11) (1968) 5116–5129.

[7] D. ben Avraham, M.M. Tirion, Normal modes analyses of macromolecules, Physica A 249 (1998) 415–423.

[8] I. Bahar, A.R. Atilgan, B. Erman, Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, Folding Design 2 (3) (1997) 173–181.

[9] A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, Biophys. J. 80 (2001) 505–515.

[10] H.M. Chun, C.E. Padilla, D.N. Chin, M. Watanabe, V.I. Karlov, H.E. Alper, K. Soosaar, K.B. Blair, O.M. Becker, L.S.D. Caves, R. Nagle, D.N. Haney, B.L. Farmer, MBO(N)D: a multibody method for long-time molecular dynamics simulations, J. Comput. Chem. 21 (3) (2000) 159–184.

[11] P. Doruker, R.L. Jernigan, I. Bahar, Dynamics of large proteins through hierarchical levels of coarse-grained structures, J. Comput. Chem. 23 (1) (2002) 119–127.

[12] P. Doruker, R.L. Jernigan, I. Navizet, R. Hernandez, Important fluctuation dynamics of large protein structures are preserved upon coarse-grained renormalization, Int. J. Quantum Chem. 90 (2002) 822–837.

[13] T. Haliloglu, I. Bahar, B. Erman, Gaussian dynamics of folded proteins, Phys. Rev. Lett. 79 (16) (1997) 3090–3093.

[14] I. Bahar, A.R. Atilgan, M.C. Demirel, B. Erman, Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability, Phys. Rev. Lett. 80 (12) (1998) 2733–2736.

[15] K. Hinsen, Analysis of domain motions by approximate normal mode calculations, Proteins: Struct. Funct. Genet. 33 (1998) 417–429.

[16] K. Hinsen, Domain motions in proteins, J. Mol. Liq. 84 (2000) 53–63.

[17] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins: Struct. Funct. Genet. 17 (4) (1993) 412–425.

[18] O. Keskin, S.R. Durell, I. Bahar, R.L. Jernigan, D.G. Covell, Relating molecular flexibility to function: a case study of tubulin, Biophys. J. 83 (2002) 663–680.

[19] A.D. Schuyler, G.S. Chirikjian, Normal mode analysis of proteins: a comparison of rigid cluster modes with C_α coarse graining, J. Mol. Graphics Modell. 22 (2004) 183–193.

[20] F. Tama, F.X. Gadea, O. Marques, Y.-H. Sanejouand, Building-block approach for determining low-frequency normal modes of macromolecules, Proteins: Struct. Funct. Genet. 41 (2000) 1–7.

[21] A. Thomas, M.J. Field, L. Mouawad, D. Perahia, Analysis of the low frequency normal modes of the T-state of aspartate transcarbamylase, J. Mol. Biol. 257 (1996) 1070–1087.

[22] O. Kurkcuoglu, R.L. Jernigan, P. Doruker, Mixed levels of coarse-graining of large proteins using elastic network model succeeds in extracting the slowest motions, Polymer 45 (2004) 649–657.

[23] G. Li, Q. Cui, A coarse-grained normal mode approach for macromolecules: an efficient implementation and application to Ca^{2+}-ATPase, Biophys. J. 83 (2002) 2457–2474.

[24] S. Spinelli, Q.Z. Liu, P.M. Alzari, P.H. Hirel, R.J. Poljak, The 3-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU, Biochimie 73 (11) (1991) 1391–1396.

[25] A.M. Brzozowski, Z.S. Derewenda, E.J. Dodson, G.G. Dodson, J.P. Turkenburg, Structure and molecular model refinement of rhizomucor miehei triacylglyceride lipase: a case-study of the use of simulated annealing in partial model refinement, Acta Crystallogr., Sect. B 48 (3) (1992) 307–319.

[26] C. Toyoshima, M. Nakasako, H. Nomura, H. Ogawa, Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution, Nature 405 (2000) 647–655.

[27] N. Reuter, K. Hinsen, J.-J. Lacapre, Transconformations of the SERCA1 Ca-ATPase: a normal mode study, Biophys. J. 85 (2003) 2186–2197.

[28] A. Barth, F. von Germar, W. Kreutz, W. Mäntele, Time-resolved infrared spectroscopy of the Ca^{2+}-ATPase: the enzyme at work, J. Biol. Chem. 271 (48) (1996) 30637–30646.

[29] A. Barth, C. Zscherp, Substrate binding and enzyme function investigated by infrared spectroscopy, FEBS Lett. 477 (2000) 151–156.

[30] H. Ogawa, D.L. Stokes, H. Sasabe, C. Toyoshima, Structure of the Ca^{2+} pump of sarcoplasmic reticulum: a view along the lipid bilayer at 9-Å resolution, Biophys. J. 75 (1998) 41–52.

[31] W.A. Fenton, A.L. Horwich, GroEL-mediated protein folding, Protein Sci. 6 (1997) 743–760.

[32] Z. Xu, A.L. Horwich, P.B. Sigler, The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex, Nature 388 (1997) 741–750.

[33] K. Braig, Z. Otwinowski, R. Hegde, D.C. Boisvert, A. Joachimiak, A.L. Horwich, P.B. Sigler, The crystal-structure of the bacterial chaperonin GroEL at 2.8Å, Nature 371 (1994) 578–586.

[34] J. Ma, M. Karplus, The allosteric mechanism of the chaperonin GroEL: a dynamic analysis, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 8502–8507.

[35] O. Keskin, I. Bahar, D. Flatow, D.G.. Covell, R.L.. Jernigan, Molecular mechanisms of chaperonin groel–groes function, Biochemistry 41 (2002) 491–501.

[36] C. Helgstrand, W.R. Wikoff, R.L. Duda, R.W. Hendrix, J.E. Johnson, L. Liljas, The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44Å resolution, J. Mol. Biol. 334 (2003) 885–899.

[37] W.R. Wikoff, L. Liljas, R.L. Duda, H. Tsuruta, R.W. Hendrix, J.E. Johnson, Topologically linked protein rings in the bacteriophage HK97 capsid, Science 289 (2000) 2129–2133.

[38] M.K. Kim, R.L. Jernigan, G.S. Chirikjian, An elastic network model of HK97 capsid maturation, J. Struct. Biol. 143 (2003) 107–117.