doi:10.1016/S0022-2836(02)01426-2 J. Mol. Biol. (2003) 326, 485–492

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Mega-Dalton Biomolecular Motion Captured from Electron Microscopy Reconstructions

Pablo Chacón, Florence Tama and Willy Wriggers*

Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 USA

*Corresponding author

The vibrational analysis of elastic models suggests that the essential motions of large biomolecular assemblies can be captured efficiently at an intermediate scale without requiring knowledge of the atomic structure. While prior work has established a theoretical foundation for this analysis, we demonstrate here on experimental electron microscopy maps that vibrational modes indeed describe functionally relevant movements of macromolecular machines. The clamp closure in bacterial RNA polymerase, the ratcheting of 30 S and 50 S subunits of the ribosome, and the dynamic flexibility of chaperonin CCT are extracted directly from single electron microscopy structures at 15–27 Å resolution. The striking agreement of the presented results with experimentally observed motions suggests that the motion of the large scale machinery in the cell is surprisingly independent of detailed atomic interactions and can be quite reasonably described as a motion of elastic bodies.

Keywords: normal mode analysis; flexing; domain motions; macromolecular assemblies; subcellular machines

Nearly all biomolecular assemblies of mega-Dalton molecular weight exhibit large-scale domain movements that drive the functional mechanisms of cellular machinery.[1,2] Examples are the transcription complexes, ribosome, and molecular chaperones. To fully understand the action of these subcellular machines one must know the dynamic states of all flexible segments, ideally, based on two or more conformations that are solved at sufficiently high resolution. The large size of macromolecular machines precludes the routine determination of atomic-resolution structures due to the considerable effort of crystallization and data collection. Structures of large assemblies are more readily visualized with cryo-electron microscopy (cryo-EM) that provides 3D reconstructions at intermediate levels of resolution.[3] Rapid advances in 3D imaging call for novel computational methods that estimate and visualize the intrinsic flexibility of large subcellular machines based on a single cryo-EM structure.

The computational prediction of such large-scale motions is a daunting task, even if the atomic structure is known. Traditional molecular simulation techniques such as molecular mechanics,[4] molecular dynamics,[5] or Monte Carlo[6] are able to describe structural fluctuations and transitions of biomolecules at the atomic level, but such calculations are very time-consuming and an adequate sampling of conformational space is feasible only for systems comprising a few thousand atoms.[7,8] Normal mode analysis (NMA) is an interesting alternative to molecular dynamics. Traditionally, NMA is applied to an atomic structure where the atomic interactions are described by a standard molecular dynamics force field.[9] The basic assumption (and limitation) of the vibrational analysis is that the potential energy of the system varies quadratically about a given minimum energy conformation. This idea is rooted in the observation that biomolecules behave, more than expected, as if the energy surface were parabolic, even though the potential contains many local minima.[10] NMA has been established in computational biology in the early 1980s[11,12] and as a computational tool has recently gained acceptance in structural biology, e.g. in the flexible refinement of actin filaments against X-ray fiber diffraction data[13] and as a model for diffuse X-ray scattering of protein crystals.[14] Thus, the harmonic approximation of biomolecular motion is in excellent agreement with many observations of structural polymorphism, where domains seemingly "move in relation to one another with only small expenditures of energy".[1]

Abbreviations used: cryo-EM, cryo-electron microscopy.

E-mail address of the corresponding author: wriggers@scripps.edu

0022-2836/03/$ - see front matter

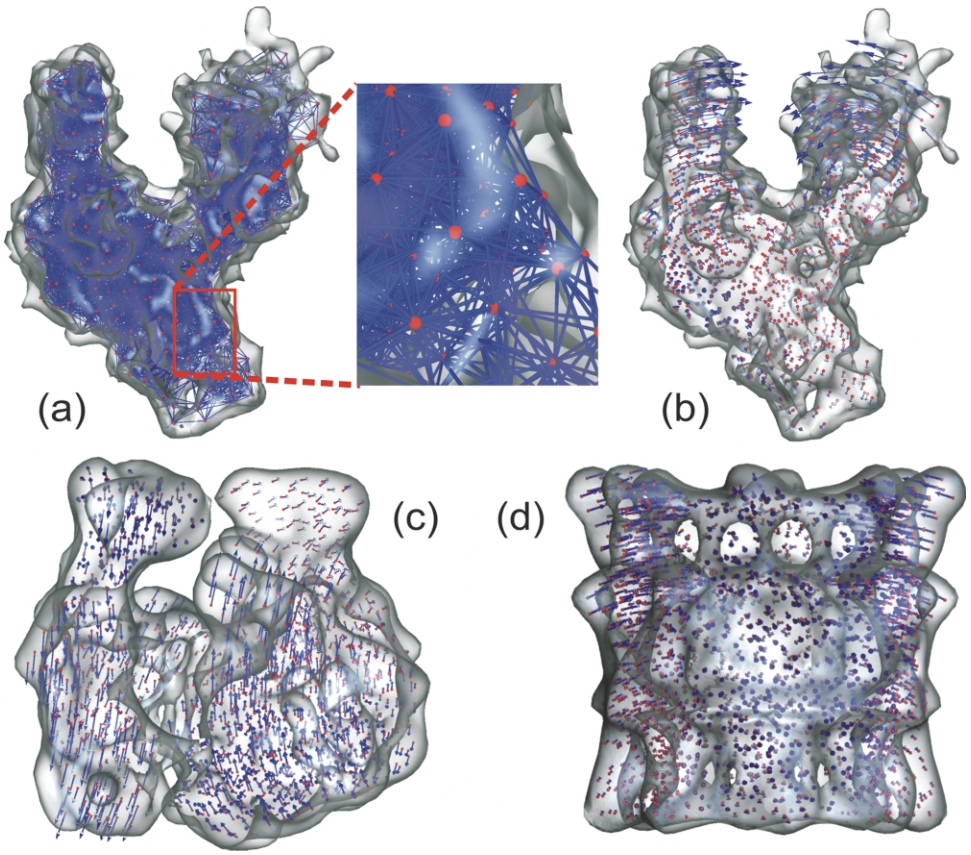

Figure 1. Lowest-frequency normal modes captured from electron microscopy maps (transparent isocontour surfaces). (a) The *E. coli* RNA polymerase density[27] is represented by the VQ method[23] using 993 landmarks (shown as read beads). The landmark connectivities (blue) were constructed using the Tirion convention[15] with a distance cutoff of 14 Å. (b) *E. coli* RNA polymerase[27] (resolution 15 Å). (c) *E. coli* ribosome[33] (resolution 25 Å, represented by 1481 landmarks, connectivity cutoff 18 Å). (d) Eukaryotic chaperonin CCT[36] (resolution 27 Å, represented by 1992 landmarks, connectivity cutoff 14 Å). The mode displacements in (b)–(d) have been rendered in the form of finite difference vectors (blue arrows) to indicate the intrinsic flexibility of the molecules (transparent isocontour surfaces). The three structures are not to scale. Figures 1–4 were created with Situs[37] and with the molecular graphics programs VMD[39] and Raster3D.[40] Movie animations of the 12 lowest-frequency modes for each molecule are presented at http://emotion.biomachina.org

Recent efforts by theorists aimed at extending the method to the large systems of interest. A first step in the reduction of the computational cost of NMA is the replacement of the atomic force field by a more simplified harmonic interaction potential of neighboring atoms. This approach, pioneered by Tirion,[15] showed that low-frequency modes depend more on the global character of the deformations than on the precise form of the interaction potential. At full atomic resolution the standard Cartesian method involves a diagonalization of a $3N \times 3N$ matrix, where N is the number of atoms. The memory requirements are prohibitive for large proteins or assemblies with more than 5000 atoms. A number of investigators proposed to reduce the amount of spatial detail in the model, while using the simplified harmonic interaction force field developed by Tirion. Excellent agreement with experimental temperature factors been obtained with a harmonic model including only the C_α positions[16] or with even sparser models.[17,18] These studies suggest that it is the global shape, and not the resolution or local detail of the force-field, that governs the low-frequency modes observed with NMA.

To apply NMA to low-resolution data, it was necessary to develop a bridging technology that allows one to create pseudo-atomic models from cryo-EM densities. Our prior work adapted information processing concepts, such as artificial neural networks and vector quantization (VQ)[19] to represent a 3D density map in real space with a finite number of so-called landmark points. Figure 1(a) presents a VQ of the 3D density of bacterial RNA polymerase. The VQ distributes the landmarks (red) over a 3D biological data set according to the map density, yielding a reduced representation of the data.[19] Proximity relationships (blue) among the landmarks can be constructed based on a nearest-neighbor distance cutoff (Figure 1(a)). This representation yields a deformable elastic network once the connections are modeled as Hookean springs (see below). A preliminary account of the theory and application to simulated low-resolution data has already been published: in Ref. 18 we have shown how NMA

can reproduce the experimentally observed (atomic-resolution) opening of the cleft in adenylate kinase in the reduced model at various levels of detail. The VQ routines made available by us[19,20] have been adapted for NMA independently and simultaneously also by Ma *et al*.[21,22]

While our earlier theoretical studies[18] established the robustness of the VQ-based elastic network model and the resulting NMA against changes in parameters, the feasibility of capturing functionally relevant motions remained to be demonstrated for experimental systems. We show in the following that the reduced description of the intrinsic flexibility can be applied directly to 3D image reconstructions from cryo-EM. We have chosen three "archetypal" macromolecular machines, bacterial RNA polymerase, bacterial ribosome, and chaperonin CCT, to demonstrate the validity of the approach.

Constructing the Elastic Net

VQ is a data clustering technique that has been used since the 1950s for digital signal compression and for speech and image processing. The method allows one to approximate the density distribution $\rho(\mathbf{r})$ of 3D data signals, $\mathbf{r} \in \Re^3$, using a finite number of landmark points $\mathbf{w}_i \in \Re^3$, $i = 1, \ldots, N$. Several algorithms exist that solve the VQ problem by systematic updating of the $\mathbf{w}_i$ until the landmarks approximate the density $\rho(\mathbf{r})$ according to a statistical optimization criterion. Here, we implemented the Growing Cell Structures method by Fritzke.[23] Convergence assures that the set $\{\mathbf{w}_i\}$ represents the underlying biological data in a well-defined and reproducible fashion, as described earlier.[19] A number N of landmarks on the order of 1000–2000 was found to provide a level of detail suitable for NMA, but the resulting model and the normal modes are quite robust under changes in complexity.[18]

The elastic net was constructed by assigning Hookean springs between pairs of adjacent landmarks. The pairwise Hookean potential between adjacent landmarks is:

$$U_{ij} = \frac{c}{2}(\|\mathbf{w}_{ij}\| - \|\mathbf{w}_{ij}^0\|)^2 = \frac{c}{2}\left(\frac{\mathbf{w}_{ij}^0 \cdot \Delta\mathbf{w}_{ij}}{\|\mathbf{w}_{ij}^0\|}\right)^2 + O((\mathbf{w}_{ij}^0)^2) \qquad (1)$$

where $\mathbf{w}_{ij} \equiv \mathbf{w}_i - \mathbf{w}_j$, $\Delta\mathbf{w}_{ij} \equiv \mathbf{w}_{ij} - \mathbf{w}_{ij}^0$, and the zero superscript indicates the initial configuration. The strength of the potential c is an empirical constant for the system that can be adjusted such that the normal mode amplitudes match those from atomic detail NMA or to ensure consistency with experimentally observed properties.

The potential energy within the system is then given by:

$$U = \sum_{i<j} U_{ij} D_{ij} \qquad (2)$$

where the connections D_{ij} are assigned within a certain distance cutoff (Figure 1). The particular choice of cutoff depends on the level of detail.[18] As a rule of thumb, a cutoff distance should be chosen that encompasses the first peak of the pair-distance distribution function. It is straightforward to compute the $3N \times 3N$ Hessian matrix of second derivatives by expanding the U_{ij} to second order about $\mathbf{w}_{ij}^0$ (equation (1)). Since each landmark is assumed to have unit mass, the normal modes are the eigenvectors of the Hessian.[9] The modes in this work were obtained by direct diagonalization of the Hessian, unless noted otherwise. After generating a sparse estimation of the motion at the landmarks, displacements were extended to the full space by interpolation with the 3D thin-plate spline method.[24]

Vibrational Analysis

The vibrational analysis returns $3N - 6$ eigenmodes of landmark displacements ordered by ascending frequencies. These modes form an orthonormal basis of displacements, i.e. any conformational change can be expressed as a linear combination of the modes. What makes this analysis useful is that the essential global motions of large-scale biomolecules have been found to concentrate among the low-frequency modes. The first twelve modes contain about 70 % of the total fluctuations of a biomolecule.[25] Thus, NMA can be used as a low-pass filter in frequency space to reduce the dimensionality of the motions and to separate the essential (global, low-frequency) from the non-essential (local, high-frequency) motions. In particular, information on a significant conformational change of a biomolecule is often found in a single, low-frequency mode of its open form.[26] Therefore, we focus in this paper on the lowest-frequency modes, whereas higher-frequency modes are depicted at a web site†.

The intrinsic flexibility of three quintessential macromolecular assemblies was investigated with NMA based on available intermediate (15–27 Å) cryo-EM resolution data (Figure 1(b)–(d)). The clamp opening and closing motion of the *E. coli* RNA polymerase (Figure 1(b)) and the ratcheting motion of the ribosome (Figure 1(c)) were already characterized experimentally at atomic and intermediate resolution,[27–29] and the functional mechanisms of these molecular machines continue to be of significant interest. In addition, we chose eukaryotic chaperonin CCT (Figure 1(d)), whose interaction with partially folded substrates is believed to involve significant flexibility.[30] Figure

† http://emotion.biomachina.org

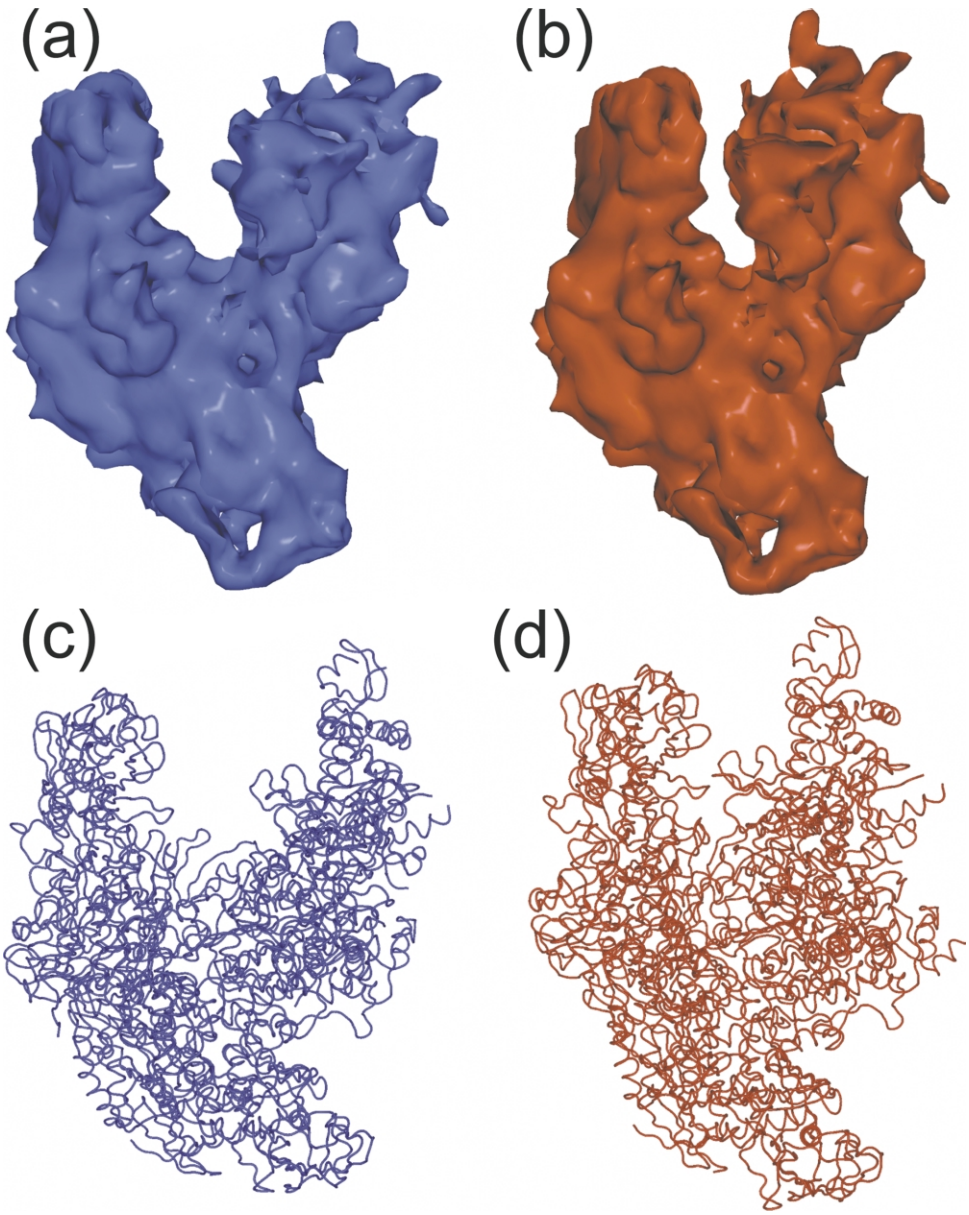

Figure 2. Clamp closure in RNA polymerase. (a) The 15 Å resolution map of *E. coli* RNA polymerase.[27,38] (b) Closing motion of the *E. coli* map predicted by the lowest-frequency normal mode Figure 1(b). (c) The flexed structure of *T. aquaticus* RNA polymerase, i.e. the atomic structure in (d) fitted to the map in (a).[27,38] (d) The original *T. aquaticus* RNA polymerase crystal structure.[28] Note that the structures in panels (a) and (d) represent original experimental data, whereas the models in panels (b) and (c) were created by NMA and flexible fitting, respectively. The amplitude of the motion in (b) was based on the observed displacements (c) *versus* (d).

1(b)–(d) gives a general overview of the lowest-frequency displacements observed in the analysis that will be described in more detail in the following.

E. coli RNA Polymerase

RNA is synthesized in the cell by DNA-dependent RNA polymerase (RNAP), a complex macromolecular assembly.[27,28] The structure of bacterial RNAP in an open conformation has been determined to a resolution of 15 Å by cryo-EM (Figure 2(a)), and to atomic resolution in a closed conformation by X-ray analysis[28] (Figure 2(d)). A flexible fitting of the atomic structure to the cryo-EM map revealed a prominent swinging motion of the so-called "clamp" domain by up to 25 Å[27] (Figure 2(c)). The differences between the crystal and the cryo-EM isoforms can be attributed to crystal packing effects and reveal a closing of the RNAP jaws relative to the cryo-EM data. Large conformational changes of the clamp domain have also been observed in crystallographic studies of yeast RNAP II.[31]

We systematically evaluated the overlap of low-frequency normal modes of *E. coli* RNAP at 15 Å resolution (Figure 1(b)) with the experimentally derived conformational change (Figure 2(c) and

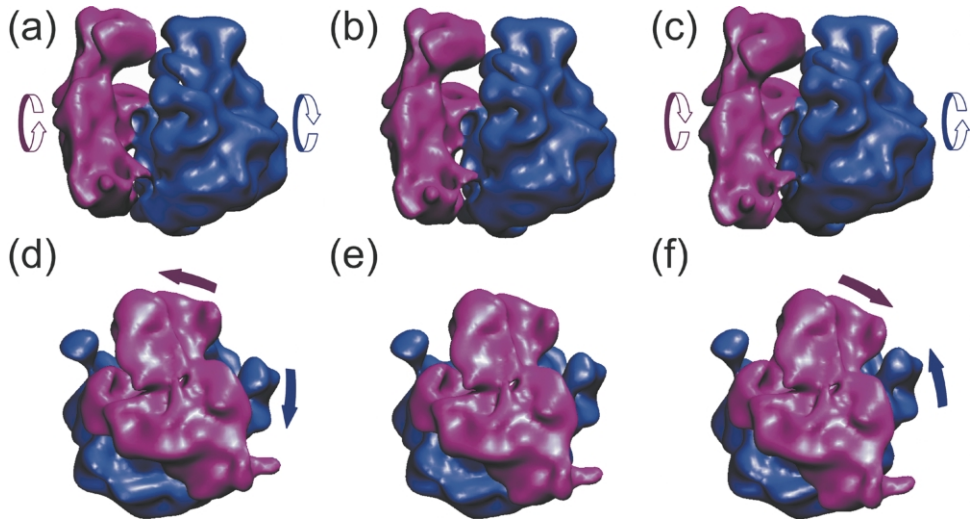

Figure 3. Ratcheting motion of the 25 Å electron microscopy map of the *E. coli* ribosome[33] captured by the lowest-frequency normal mode. The 30 S subunit is shown in purple, the 50 S subunit in blue. Panels (a)–(c) show the front view of the ribosome, panels (d)–(f) the side view. Panels (a) and (d) correspond to the maximum positive deformation, panels (b) and (c) to the original map (minimum of the elastic energy), and panels (c) and (f) to the maximum negative deformation. The amplitude of the motion has been amplified by a factor 2 relative to the observed EF-G-dependent ratcheting[29] to better visualize the mode.

(d)). To this end, the sparsely sampled displacements were interpolated to the atomic positions with the 3D thin-plate spline method.[24] For the overlap analysis we employed the RTB diagonalization method[32] of the Hessian matrix. The overlap is a measure of similarity of the direction of the observed displacements with the direction of the modes,[26] and has a maximum value of one for identical directions.

In our analysis, only the lowest frequency mode (Figure 1(b)) exhibited significant overlap with the experimental change (overlap value 0.71), while the overlap values of higher frequency modes were below 0.24. This similarity demonstrates that the lowest-frequency mode derived from the cryo-EM map is a reasonable approximation of the clamp-closing motion observed at atomic resolution. In Figure 2(b) we have used the predominant mode to visualize the clamp closure directly on the cryo-EM map to demonstrate the similarity with the atomic structure in panel (d).

E. coli Ribosome

The ribosome is a complex molecular machine that synthesizes proteins by translating genetic information encoded in messenger RNA into sequences of amino acids. *E. coli* ribosome is composed of two unequal subunits. The smaller, 30 S subunit binds messenger RNA, and the larger, 50 S subunit catalyzes the formation of the peptide bonds. Binding of elongation factor G (EF-G) and GTP hydrolysis promote the translocation process, and lead to ratchet-like rotations of the 30 S subunit relative to the 50 S subunit.[29] The ratchet-like motion has been proposed as a key mechanical step that opens the messenger RNA channel following binding of GTP to EF-G.[29] Here, we demonstrate that this mechanical motion is intrinsic to the shape of the ribosome itself, and that the information can be inferred from the classic 25 Å resolution cryo-EM map.[33]

Figure 3 shows that the motion deduced from comparisons of cryo-EM maps of functional complexes[29] is reproduced in the lowest-frequency normal mode using the reduced representation of the ribosome map as an elastic network of landmarks. The large conformational rearrangements can be attributed to a mechanical coupling with the binding sites of EF-G, the stalk base of the 50 S subunit, and the head and shoulder of the 30 S subunit. The density bridge between the 30 S and 50 S subunit (B2 in the nomenclature of Ref. 34), later resolved into four bridges,[35] is observed to act as flexible pivot for the motion. The overall motion of mode 1 in our study is practically identical to the motion of mode 3 in a separate NMA study of the atomic structure of the ribosome. The low-resolution map has merged four connecting RNA bridges into a single one, but still there appears to be little difference between the pivoting behavior of a four-component connection and the behavior of a single rod obtained by merging them.

Eukaryotic Chaperonin CCT

The chaperonins are large cylinder-shaped protein assemblies divided into two subfamilies: group I including the GroEL-GroES system and group II including the thermosome and CCT.[30] We focus in this work on CCT which shares a common architecture with other group II members. The large apical and equatorial domains are connected to the smaller intermediate domain by hinges. The

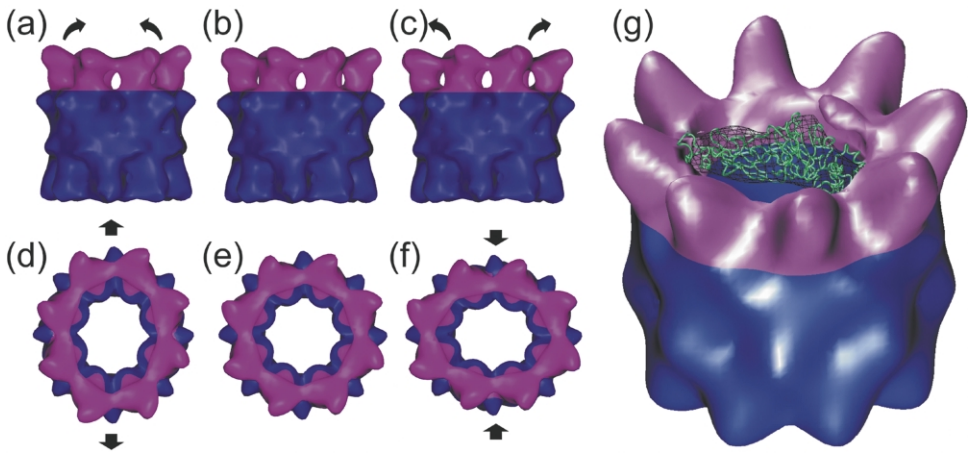

Figure 4. Breathing motion of the apical domain of eukaryotic chaperonin CCT. The tips of the apical domain are shown in purple, the remainder of the assembly in blue. Panels (a)–(c) show the side view of the chaperonin in the apo form,[36] panels (d)–(f) show the corresponding top view. Panels (a) and (d) correspond to the maximum positive deformation, panels (b) and (c) to the original 27 Å resolution map, panels (c) and (f) to the maximum negative deformation. The last panel (g) shows in comparison actin-bound CCT[36] (purple, blue), the flexibly fitted structure of actin (green, see text), and the corresponding actin density (black wire mesh), as viewed from above the apical domain.

equatorial domain contains the binding site for ATP, whose hydrolysis is necessary for the chaperonin working cycle, while the apical domain is involved in substrate binding. Both actin and tubulin in open conformations have been visualized with cryo-EM bound to the apical domain of CCT.[36] Upon binding of the substrate the helical extensions of the apical domain exhibit considerable conformational variability, which was hypothesized to represent various states in the ATPase cycle that leads to an encapsulation of the folding substrate.[30]

The lowest-frequency mode of the apo-CCT electron microscopy map suggests an increased flexibility of the apical domain, whereas other parts of the CCT structure remain relatively rigid (Figures 1(d) and 4(a)–(f)). Although the precise local functional motion of the apical domains tips is difficult to detect with the global NMA, this mode suggests that these regions involved in substrate binding are the maximally deformable regions of the CCT (Figure 1(d)), in agreement with the known structual variability of the chaperonin.[36]

To create a model of unfolded, open actin bound to CCT chaperonin (Figure 4(g)) the closed atomic structure of actin was moved towards the open EM density by forcing eight corresponding landmarks in both data sets to coincide. This was done in a molecular dynamics refinement of the atomic structure with the Situs docking package[37] where a quantity equivalent to the landmark discrepancy forms a global penalty that is imposed by distance constraints[38] while preserving the moved structure at the local level. The detailed protocol of the actin flexing is described elsewhere.

The contacts of the tips of the apical domains with the flexed actin molecule in its unfolded state are shown in Figure 4(g). As it can be seen, the contacts of the actin density bridge are localized at the maximally deformable regions, suggesting a strong mechanical coupling of actin folding and the ATP-dependent flexibility of CCT's apical domain. Based on these observations one could hypothesize that the contact provides an allosteric mechanism by which the movement of the apical domain in group II chaperonins is coupled to the release of the nucleotide and subsequent expulsion of the substrate.

Discussion

The motions of large-scale biomolecular assemblies are very complex. A reliable modeling of the intrinsic flexibility of a given structure of mega-Dalton molecular weight is therefore also very complicated and out of reach for classical force-field based simulation methods. Only by imposing a number of often quite stringent approximations can this problem be reduced in complexity so that it can be investigated in some detail. As this work has demonstrated, the crucial question is how to approximate and what physical properties to neglect while maintaining a close resemblance with experimentally observed biomolecular motions. It is useful to extend classical NMA, a well-established technique in structural biology,[13,14] to the large systems if interest by reducing the level of detail in the elastic network representation of the biophysical data. This approach has the added advantage that low-resolution density maps from cryo-EM can be animated that would otherwise yield only static information.

What are the general features of the elastic network approximation? The most remarkable observation is that in the three systems studied

the lowest-frequency modes clearly corresponded to functionally relevant motions of the molecular assemblies. The intrinsic flexibility of the three molecular assemblies was independent of detailed atomic interactions and can be ascribed to the general shape of the molecules. Conformational changes were similar to the motion of elastic bodies.

Does NMA adequately describe barrier-crossing transitions between multiple states? It is clear that the elastic bending motions that can be simulated by the simple harmonic analysis involve very low energy barriers. However, more complex transitions across high energy barriers that involve a non-harmonic description of the energy landscape are beyond reach of the method. Our local predicitons can give a reasonable lead how the molecule behaves when deviating from the known conformation, but for large deformations the harmonic approximation will be less reliable. Nevertheless, the proposed physics-based model provides a reasonable predictive insight into conformational changes of large biomolecular assemblies.

Why would one expect a similarity between the elasto-mechanical model and experimentally observed conformational changes? Our results are consistent with a recent survey of domain motions in 20 proteins in open and closed forms[26] that suggests that much of the information of the conformational change is carried by a single low-resolution mode of the open form of a biopolymer.[26] Although one cannot predict *a priori* which mode is the relevant one, an experimentally observed change was found to be well characterized by one of the three lowest-frequency modes[26] in most proteins. The difference in the frequency order of the ribosomal ratchet modes from cryo-EM (1, this study) and atomic structure (3) are therefore expected and within the limits suggested by the survey. The remarkable result is that the ratchet motion can be captured already at 25 Å resolution.

Does the ranking of the modes reflect their functional relevance? The frequency-dependent ranking is a good indicator of the "relevance likelihood", but ultimately not stringent enough to predict whether such motions could be observed experimentally in cryo-EM or X-ray crystallographic structures. One cannot guarantee *a priori* which particular low-frequency mode (or a linear combination of more modes) is relevant.

Given these limitations, what are meaningful applications of NMA in structural biology? There are situations where NMA can be employed directly as a dimensionality reduction filter of the motion space, e.g. in the refinement of structures against low-resolution biophysical data,[13] or in cryo-EM image processing where normal modes can be added as additional degrees of freedom to capture structural polymorphism in the averaging and classification phase. As we have seen in the case of the chaperonin CCT, NMA may help predict functional motions directly from a single cryo-EM reconstruction in the absence of multiple structures or structures at atomic resolution. In practical applications an unequivocal prediction requires additional standards for parametrization, i.e. a screening against complementary experimental data to select the relevant mode and amplitude. This is where a public forum such as a web site of computed modes can provide a valuable service, because expert users can rate the modes based on agreement with experimental observations (e.g. polymorphism in electron micrographs, FRET or NMR spectroscopy data, diffuse X-ray scattering) and post their findings to the public. We are currently building such a site (E. Metwally, P.C., and W.W., unpublished results†). Visitors can explore the animated normal modes shown for (currently) 13 EM structures (including those discussed in this work) and are welcome to post their insights and comments.

Acknowledgements

This work was supported by NIH grants 1R01GM62968 and P41RR12255 and by the LJIS Interdisciplinary Training Program/Burroughs Wellcome Fund. We thank J. M. Valpuesta for kindly providing the EM maps of the apo and actin-bound CCT structures, and S. Darst for providing the *E. coli* RNA polymerase reconstruction.

References

1. Gerstein, M., Lesk, A. M. & Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry*, **33**, 6739–6749.
2. Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, **92**, 291–294.
3. Frank, J. (1996). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*, Academic Press, San Diego.
4. Packer, M. J. & Hunter, C. A. (1998). Sequence dependent DNA structure: the role of the sugar phosphate backbone. *J. Mol. Biol.* **280**, 407–420.
5. McCammon, J. A. & Harvey, S. C. (1987). *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge.
6. Leach, A. R. (1996). *Molecular Modelling: Principles and Applications*, Addison Wesley Longman, Essex, UK.
7. Clarage, J. B., Romo, T., Andrews, B. K., Pettitt, B. M. & Phillips, G. N. (1995). A sampling problem in molecular dynamics simulations of macromolecules. *Proc. Natl Acad. Sci. USA*, **92**, 3288–3292.
8. Balsera, M. A., Wriggers, W., Oono, Y. & Schulten, K. (1996). Principal component analysis and long time protein dynamics. *J. Phys. Chem.* **100**, 2567–2572.
9. Case, D. A. (1997). Normal mode analysis of biomolecular dynamics. In *Computer Simulation of Biomolecular Systems* (van Gunsteren, W. F., Weiner, P. K. & Wilkinson, A. J., eds), vol. 3, pp. 284–301, Kluwer Academic, Dordrecht

† http://emotion.biomachina.org

10. Horiuchi, T. & Go, N. (1991). Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: human lysozyme. *Proteins: Struct. Funct. Genet.* **10**, 106–116.
11. Tasumi, M., Takeuchi, H., Ataka, S., Dwivedi, A. M. & Krimm, S. (1982). Normal vibrations of proteins: glucagon. *Biopolymers*, **21**, 711–714.
12. Noguti, T. & Go, N. (1982). Collective variable description of small-amplitude conformational fluctuations in a globular protein. *Nature*, **296**, 776–778.
13. Tirion, M. M., ben-Avraham, D., Lorenz, M. & Holmes, K. C. (1995). Normal modes as refinement parameters for the F-actin model. *Biophys. J.* **68**, 5–12.
14. Faure, P., Micu, A., Perahia, D., Doucet, J., Smith, J. C. & Benoit, J. P. (1994). Correlated intramolecular motions and diffuse X-ray scattering in lysozyme. *Struct. Biol.* **1**, 124–128.
15. Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter atomic analysis. *Phys. Rev. Letters*, **77**, 1905–1908.
16. Bahar, I., Atilgan, A. R. & Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* **2**, 173–181.
17. Doruker, P., Jernigan, R. L. & Bahar, I. (2002). Dynamics of large proteins through hierarchical levels of coarse-grained structures. *J. Comp. Chem.* **23**, 119–227.
18. Tama, F., Wriggers, W. & Brooks, C. L. (2002). Exploring global distortions of biological macromolecules and assemblies from low-resolution structural information and elastic network theory. *J. Mol. Biol.* **321**, 297–305.
19. Wriggers, W., Milligan, R. A., Schulten, K. & McCammon, J. A. (1998). Self-organizing neural networks bridge the biomolecular resolution gap. *J. Mol. Biol.* **284**, 1247–1254.
20. Wriggers, W., Milligan, R. A. & McCammon, J. A. (1999). Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.* **125**, 185–195.
21. Ming, D., Kong, Y., Lambert, M. A., Huang, Z. & Ma, J. (2002). How to describe protein motion without amino acid sequence and atomic coordinates. *Proc. Natl Acad. Sci. USA*, **99**, 8620–8625.
22. Ming, D., Kong, Y., Wakil, S. J., Brink, J. & Ma, J. (2002). Domain movements in human fatty acid synthase by quantized elastic deformation model. *Proc. Natl Acad. Sci. USA*, **99**, 7895–7899.
23. Fritzke, B. (1994). Growing cell structures—a self-organizing network for unsupervised and supervised learning. *Neural Networks*, **7**, 1441–1460.
24. Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data*, Cambridge University Press, Cambridge, UK.
25. Tirion, M. M. & ben-Avraham, D. (1993). Normal mode analysis of G-actin. *J. Mol. Biol.* **230**, 186–195.
26. Tama, F. & Sanejouand, Y.-H. (2001). Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **14**, 1–6.
27. Darst, S. A., Opalka, N., Chacón, P., Polyakov, A., Richter, C., Zhang, G. & Wriggers, W. (2002). Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **99**, 4296–4301.
28. Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. & Darst, S. A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell*, **98**, 811–824.
29. Frank, J. & Agrawal, R. K. (2000). A ratchet-like intersubunit reorganization of the ribosome during translocation. *Nature*, **406**, 318–322.
30. Zhang, X., Beuron, F. & Freemont, P. S. (2002). Machinery of protein folding and unfolding. *Curr. Opin. Struct. Biol.* **12**, 231–238.
31. Cramer, P., Bushnell, D. A. & Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Angstrom resolution. *Science*, **292**, 1863–1876.
32. Tama, F., Gadea, F. X., Marques, O. & Sanejouand, Y.-H. (2000). Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins: Struct. Funct. Genet.* **41**, 1–7.
33. Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A. *et al.* (1995). A model of protein synthesis based on cryo-microscopy of the *E. coli* ribosome. *Nature*, **376**, 441–444.
34. Frank, J., Verschoor, A., Li, Y., Zhu, J., Lata, R. K., Radermacher, M. *et al.* (1995). A model of the translational apparatus based on a three-dimensional reconstruction of the *Escherichia coli* ribosome. *Biochem. Cell Biol.* **73**, 757–765.
35. Gabashvili, I. S., Agrawal, R. K., Spahn, C. M. T., Grassucci, R. A., Svergun, D. I., Frank, J. & Penczek, P. (2000). Solution structure of the *E. Coli* 70 S ribosome at 11.5 Å resolution. *Cell*, **100**, 537–549.
36. Llorca, O., Martin-Benito, J., Ritco-Vosonvici, M., Grantham,, J., Hynes, G. M., Willison, K. R. *et al.* (2000). Eukaryotic chaperonin CCT stabilizes actin and tubulin folding intermediates in open quasi-native conformations. *EMBO J.* **19**, 5971–5979.
37. Wriggers, W. & Birmanns, S. (2001). Using Situs for flexible and rigid-body fitting of multi-resolution single molecule data. *J. Struct. Biol.* **133**, 193–202.
38. Wriggers, W. & Chacón, P. (2001). Modeling tricks and fitting techniques for multi-resolution structures. *Structure*, **9**, 779–788.
39. Humphrey, W. F., Dalke, A. & Schulten, K. (1996). VMD—visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.
40. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.

Edited by W. Baumeister

(Received 6 September 2002; received in revised form 23 November 2002; accepted 27 November 2002)