Structure, Vol. 13, 373–380, March, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.02.002

Review

Usefulness and Limitations of Normal Mode Analysis in Modeling Dynamics of Biomolecular Complexes

Jianpeng Ma[1,2,3,*]
[1]Department of Biochemistry and Molecular Biology
Baylor College of Medicine
One Baylor Plaza
Houston, Texas 77030
[2]Department of Bioengineering
Rice University
Houston, Texas 77005

Various types of large-amplitude molecular deformation are ubiquitously involved in the functions of biological macromolecules, especially supramolecular complexes. They can be very effectively analyzed by normal mode analysis with well-established procedures. However, despite its enormous success in numerous applications, certain issues related to the applications of normal mode analysis require further discussion. In this review, the author addresses some common issues so as to raise the awareness of the usefulness and limitations of the method in the general community of structural biology.

Introduction

Normal mode analysis (NMA) is a powerful computational method for studying large-amplitude molecular deformational motions that are widely involved in biological functions of macromolecules (Brooks and Karplus, 1985; Levitt et al., 1985). Numerous examples have shown that functionally important transition pathways of biomolecules, especially those of supramolecular complexes, often follow the trajectories of one or a few low-frequency normal modes (Cui et al., 2004; Krebs et al., 2002; Li and Cui, 2002; Ma and Karplus, 1997; Ma and Karplus, 1998; Seno and Go, 1990a, 1990b; Thomas et al., 1996a, 1996b, 1999). An important conclusion that emerged from those studies is that protein structures have evolved in such a way that their intrinsic structural flexibility, as manifested in normal modes, facilitates the functionally important conformational variations, both statically and dynamically (Ma, 2004). Moreover, in recent years, several important new algorithms, especially the coarse-grained NMA (Atilgan et al., 2001; Bahar et al., 1997; Doruker et al., 2002; Haliloglu et al., 1997; Hinsen, 1998; Ming et al., 2002b, 2003a; Tama et al., 2002; Tirion, 1996), have enabled one to simulate deformational motions of supramolecular complexes at dramatically extended resolution and length scales (Beuron et al., 2003; Chacon et al., 2003; Keskin et al., 2002; Kong et al., 2003; Ming et al., 2002a, 2003b; Miyashita et al., 2003; Tama and Brooks, 2002; Tama et al., 2003; Wang et al., 2004; Xu et al., 2003). Those methods have also contributed to structural refinement against X-ray diffraction data (Kundu et al., 2002) and low- to intermediate-resolution structural data such as those from electron cryomicroscopy (cryo-EM) (Brink et al., 2004; Carazo, 2004; Tama et al., 2004) and fiber diffraction (Wu and Ma, 2004a, 2004b).

Despite its enormous success, concerns regarding the validity of NMA still widely exist. It is therefore the intention of this review to address some of those frequently asked questions. The discussion covers, but is not limited to, the following issues: Why can we use pure harmonic analyses, such as NMA, to analyze large-amplitude conformational dynamics that are presumably highly anharmonic? To what extent is NMA valid? When does it fail and what can we do when it fails? How do we address the timescale of protein motion in solvent by using normal modes calculated in vacuum? Can we use NMA to address the evolutionary directions of structural changes among homologous proteins of a given superfamily? How large can the amplitudes of protein motions be along harmonic modes? Can the harmonic modes be used to model activated nonequilibrium conformational transitions triggered by ligand binding? What are the energetics involved in motions along normal modes? How can one correctly identify biologically relevant mode(s)? Is a coarse-graining scheme necessary and useful for systems that have atomic coordinates available, e.g., the ribosome complexes?

In this article, the author offers comments on these focused issues, especially the ones raised from applications of NMA to large-amplitude conformational changes. A small portion of material in this article overlaps with a previous review (Ma, 2004), which emphasizes methodology and applications. The author sincerely asks forgiveness from those colleagues whose work is not explicitly referred to due to space limitation.

Static Harmonic Deformation

One of the most fundamental questions regarding NMA is, given the ruggedness of the potential surface and severe solvent damping, are the biomolecular conformational transitions really harmonic? This question is usually raised for systems that undergo dynamic transitions. However, harmonic deformation in biomolecules can also be found to occur in a static fashion. For most systems, both types of deformation exit, but their extent varies drastically from system to system. To illustrate the underlying harmonicity involved in biomolecular conformational transitions, let's begin with two less familiar, but extremely informative, examples on the static deformation.

In certain cases, structural components in a complex can deform harmonically along their own intrinsic normal modes (which can be determined when the structural elements are in isolation), and then remain statically in their deformed conformation locked by surrounding interactions in the complex. Those structural components can vary widely, ranging from subunits to secondary structural elements. An excellent example (Emberly et al., 2003) was seen in a principal component analysis of the flexibility of α helices in all coiled-coil proteins in the SCOP database (Murzin et al., 1995).

*Correspondence: jpma@bcm.tmc.edu
[3]Lab address: http://sigler.bioch.bcm.tmc.edu/MaLab

In that study, it was demonstrated that, although every α helix in coiled-coil proteins is in a permanently deformed state, the low-frequency principal modes (bending and twisting modes) of helices calculated from the ensemble of all coiled-coil proteins are in extremely good agreement with dynamic normal modes calculated from an undeformed helix. Thus, one can envisage that, in these coiled-coil proteins, each of the statically deformed helices in the coiled-coil proteins is captured in a frozen harmonic wave.

The second example was from a study of how structural cores modify their shape across homologous proteins during evolution (Leo-Macias et al., 2005). For a set of 35 well-populated protein families, multiple structural alignment was first used to superimpose the structures and extract conserved cores. Then, principal component analysis was employed to extract the main deformational modes, which were regarded as the evolutionary directions of deformation, from the three-dimensional superimposition. Meanwhile, NMA was used to analyze the mechanical core plasticity of these protein families. It was found that evolutionary deformation spans a low-dimensional space (~4–5 dimensions on average), and there is significant correlation between these principal deformational modes and the ~20 lowest-frequency normal modes accessible to a particular topology. This remarkable result suggests that the structural response of a protein topology upon sequence change during evolution collectively takes place along the directions of a small set of combined low-frequency normal modes.

The above two examples are excellent demonstrations that static harmonic deformation of structural components can be used to build complex structures. In many proteins, structural deformation can also occur in a dynamic and reversible fashion as triggered by, for example, ligand binding and dissociation. That is the subject of the next section.

Dynamic Harmonic Deformation

From numerous studies of NMA, functionally important deformation of biomolecules is usually large-amplitude, low-frequency motions. For dynamic systems, it is known from experiments, such as neutron scattering (Balog et al., 2004; Cusack et al., 1988; Reat et al., 2000; Smith et al., 1990; Smith, 1991), that large-amplitude, low-frequency motions of biomolecules are highly *anharmonic* because of the rugged energy landscape and severe solvent damping (Kottalam and Case, 1990; Lamm and Szabo, 1986). The question is how one can use *harmonic* normal modes to describe those large-amplitude, *anharmonic*, biomolecular motions?

In many systems, few people dispute the fact that, on a very short timescale, or at the onset of a conformational transition, biomolecular deformation is of a small amplitude, for which NMA is a good approximation. An example is the results of transient phase grating spectroscopy used as a probe of ligand photolysis in heme protein (Deak et al., 1998). However, the question is about the ultimate amplitudes of harmonic motions, i.e., how far can biomolecules be stretched along normal modes in a biologically correct way since large-amplitude motions are more likely to be anharmonic? In other words, there is a discrepancy between the amplitude of thermal fluctuations of harmonic oscillators described by normal modes and the amplitude of biomolecular deformation that is experimentally observed. From the principles of statistical mechanics (McQuarrie, 1976), at thermal equilibrium, the square of the average amplitude of thermal fluctuation of any harmonic oscillator is k_BT/k, where k_B is the Boltzmann constant, T is the temperature, and k is the force constant for the harmonic oscillator. For a typical low-frequency normal mode calculated based on a realistic molecular mechanics force field, e.g., CHARMM force field (MacKerell et al., 1998), k_BT/k gives a very small number at 300 K (usually less than 1 Å^2) (Brooks et al., 1988). However, in reality, a ligand-induced conformational change, which can be effectively approximated by one or a few normal modes, can have a much larger amplitude than the equilibrium thermal fluctuation. This discrepancy is due to the fact that a ligand-induced conformational transition is usually an activated process, not a spontaneous thermal fluctuation. The binding of the ligand brings in additional energy that can stretch the structure along the normal mode much further. Please note that the term "ligand" is used in a very broad sense here and can be, for example, enzyme substrates, protein interacting partners, and structural components within a protein.

A good example of ligand binding-induced low-frequency motions is the hinge-bending motion of lysozyme. It is well known that such a motion can be effectively described by a single lowest-frequency normal mode (Brooks and Karplus, 1985; Bruccoleri et al., 1986; Levitt et al., 1985). In the open conformation of lysozyme, without the ligand, the structure can be bent, along the single bending mode with a magnitude of displacement much larger than spontaneous thermal fluctuation, to a state similar to the ligand bound closed state, but with an elevated energy. However, the conformational transition can stably take place *only* upon ligand binding, which also locks the system in the new state until the ligand dissociates from it. Another visually striking example is found in Leu/Ile/Val binding protein (LivJ) (Trakhanov et al., 2005). The protein has a typical bilobate structure with a ligand binding site located in a cleft between the two domains. The structure is in an open state in the absence of the ligand and closes upon ligand binding via a large-amplitude, hinged, rigid-body domain movement. As shown in Figure 1, the experimentally observed domain closure can be very well approximated by a single lowest-frequency hinge-bending mode. Other intriguing examples include a comparative study of motor proteins (Zheng and Doniach, 2003), which found that, for myosin, one or two lowest-frequency modes were sufficient for outlining the conformational changes; while, in kinesin, multiple modes were needed. The results suggested potentially different mechanisms for these motor proteins despite their strong evolutionary ties and structural similarities. Lastly, we present an example of the application of coarse-grained NMA to a gigantic supramolecular complex of the tE2 core of pyruvate dehydrogenase complex (PDC) based on low-resolution cryo-EM density maps (Kong et al., 2003). Figure 2 shows the motional pattern revealed in the symmetric mode

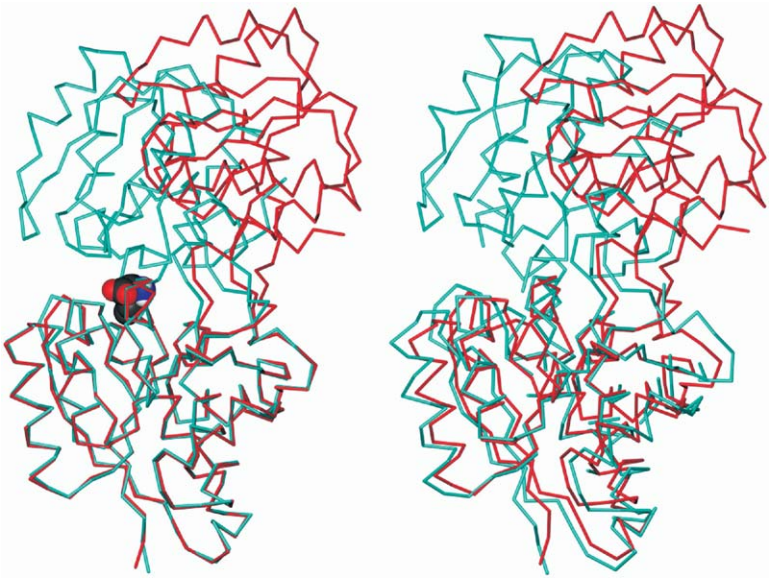

Figure 1. Harmonic Modal Displacement for LivJ Protein

See Trakhanov et al. (2005). On the left-hand side, the experimentally determined open (red) and closed (light blue) structures are superimposed on each other based on the N-terminal globular domain to show the domain closure upon ligand binding in the cleft between two domains (the ligand isoleucine is shown in a space-filling model). On the right-hand side, the experimentally determined open structure (red) is superimposed on a conformation (light blue) generated by walking along the eigenvector of the lowest-frequency mode (a classic hinge-bending mode) calculated on the open structure. This conformation qualitatively resembles the experimentally determined closed structure.

(mode 31) from coarse-grained NMA. At this resolution, the symmetric expansion and contraction of the dodecahedral complex in this mode is nearly identical to what was observed in cryo-EM measurements (Zhou et al., 2001).

A major resource of anharmonicity is the energy barrier for conformational transitions. It seems a rather common perception that the energy of the total system, protein plus ligand, is higher in the partially engaged transition state, a picture coming from the traditional mechanism of bimolecular association reaction. Therefore, the energy barrier for conformational transitions is high. Normal modes, however, as harmonic vibrational modes around a local energy minimum, cannot describe the system crossing over any energy barrier. In reality, for many systems, including the examples mentioned above and especially in large ATP-driven molecular motors such as ATP synthase F_1-ATPase (Cui et al., 2004; Gao et al., 2003; Karplus and Gao, 2004; Ma et al., 2002; Yang et al., 2003), the heights of barriers encountered during transitions can be significantly lowered by the continuing engagement of favorable protein-ligand interactions in the process of binding, i.e., the energy landscape is tipped over toward the ligand bound state. This makes the actual modal displacement of harmonic motions much larger than those of simple vibrations around the structural minimum in the ligand-free state. A very common feature of conformational transitions in those proteins is en bloc domain movements around some well-evolved structural hinges for maximal efficiency of binding-induced conformational transitions. For those systems, NMA is a particu-

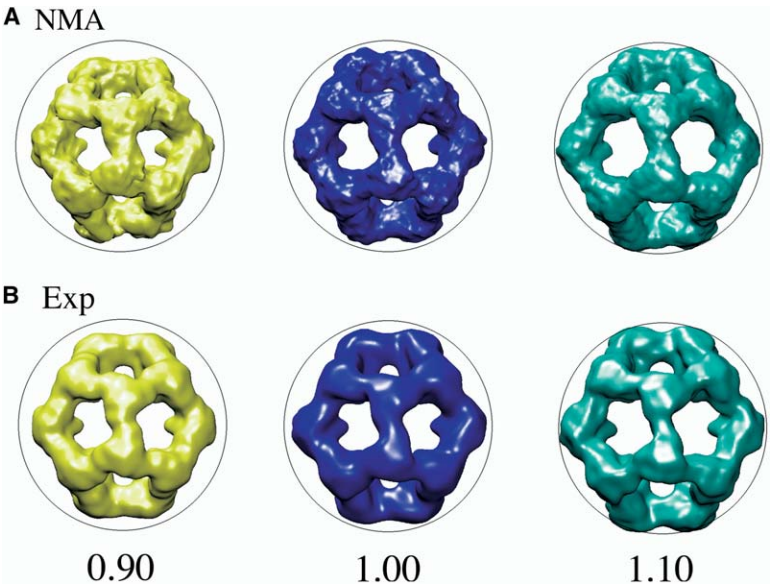

Figure 2. Motional Patterns of the Symmetric Breathing Mode in the tE2 Component of Pyruvate Dehydrogenase Complex Revealed by Coarse-Grained NMA

(A and B) See Kong et al. (2003) and Ming et al. (2002a). The structures representing the 0.90 and 1.10 size groups in (A) are the two end point structures in the symmetric harmonic mode revealed by NMA. The structures in (B) are generated by three-dimensional reconstruction from experimental data (Zhou et al., 2001). The symmetric breathing motion revealed in the symmetric mode is nearly identical to the experimental observation. The figure was modified from Figure 1 in Kong et al. (2003).

larly effective tool for analyzing large conformational changes. There are other systems such as p21^{ras} (Ma and Karplus, 1997), however, in which NMA is better suited for studying the initial stages of the transitions when only localized transitions are involved.

It is noteworthy that the observation that the amplitudes of conformational dynamics can be larger than the equilibrium thermal fluctuations of pure harmonic oscillators is also supported by the results of normal mode-based structural refinement in X-ray crystallography (Diamond, 1990; Kidera et al., 1992; Kundu et al., 2002; Suhre and Sanejouand, 2004), cryo-EM (Brink et al., 2004; Tama et al., 2004), and fiber diffraction (Tirion et al., 1995; Wu and Ma, 2004b), in which the eigenvectors are used to designate the directions of motions, and the amplitudes are optimized against experimental data.

One must keep in mind that the anharmonicity due to small energy barriers such as those from local "sticky" side chain interactions along the major conformational pathways (which also serve as the final "catch" at the end of ligand binding) are omitted from the picture revealed by NMA. Usually, the low-frequency modes are used only to describe the overall directions, or trends so to speak, of the conformational changes and do not reflect the side chain motions that are much higher in frequency.

In summary, the essential picture for ligand-induced low-frequency motions that follow normal modes is that, before ligand binding, the macromolecule oscillates around the ligand-free equilibrium state, but the energy does increase when the molecule moves away from the equilibrium conformation by stretching along harmonic normal modes (Figure 3A). In other words, the ligand bound state is inherently unstable (even if it is accessible) without a ligand (the so-called preexisting equilibrium model [Goh et al., 2004; Kern and Zuiderweg, 2003; Kumar et al., 2000]). The actual transition takes place when a ligand binds, along with the gain of binding energy that shifts the equilibrium toward the ligand bound state (Figure 3C). The energy barrier along the binding is significantly lowered because the energy landscape is gradually tipped over toward the final bound state via the continuous energetically favorable engagement of the ligand with the protein (Figure 3B).

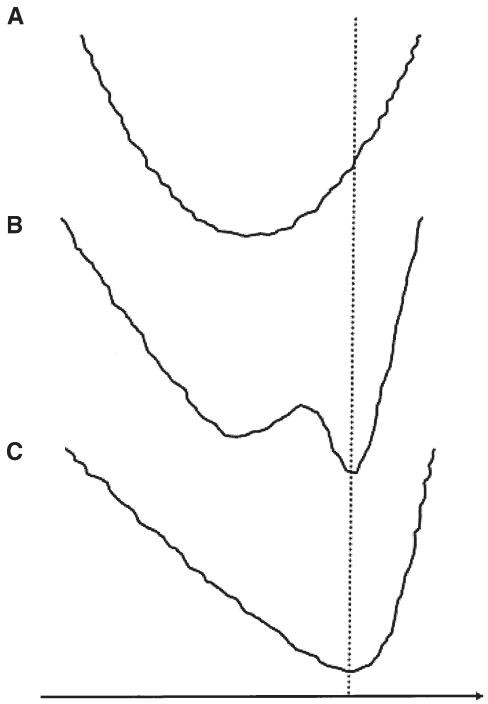

Figure 3. The Schematic Illustration of Changes in the Energy Landscape during Ligand Binding to a Protein

(A–C) The ligand-free structure is in (A), the conventional view of an energy barrier during binding is in (B), and a more accurate way of describing the energy landscape tilting toward the final ligand bound state is in (C). The horizontal axis gives the progression of the reaction coordinate, and the vertical dashed line indicates the ligand bound conformation for the protein.

Flexibility of Building Block and the Entire Complex

From the studies of supramolecular complexes, it has been shown that, for certain systems such as molecular chaperonin GroEL (Xu and Sigler, 1998), there is a very intricate relationship between the flexibility of individual building blocks (the subunits) and that of the entire complex. In the case of GroEL, each individual subunit carries an ATP binding site. It receives ATP and changes conformation closely following a few low-frequency normal modes (Ma and Karplus, 1998; Ma et al., 2000). In fact, more than 80% of the huge conformational changes in GroEL can be approximated by a combination of the two lowest-frequency modes that account for downward motions of the intermediate domain and upward motions and a dramatic twist of the apical domain (Ma and Karplus, 1998). However, the very same types of motion of the subunits were not obvious in the modes calculated for the entire GroEL complex in its closed state (Keskin et al., 2002). This is because the subunits in the GroEL complex in its closed state are interlocked by seven intersubunit salt bridges, or latches, between Arg197 and Glu386 (Braig et al., 1994) (see schematic illustration in Figure 4). Those salt bridges are completely broken upon the large conformational changes of GroEL triggered by the binding of ATP and co-chaperonin GroES (Xu et al., 1997). (Note the salt bridge shifting mechanism involved in GroEL is common among many other allosteric proteins [Lipscomb, 1983], including hemoglobin [Perutz, 1969a, 1969b].) The energy for breaking those salt bridges is provided by ligand binding. Therefore, from the viewpoint of the transition of the entire complex, the existence of the intersubunit latches imposes a huge energy barrier that separates the closed state from the open state. This is exactly the reason that harmonic NMA of the entire complex in the closed state would not reveal the ligand-induced opening motions of the ring. However, the normal modes calculated for an isolated subunit properly captured the intrinsic flexibility of the subunit itself, and the low-frequency ones are precisely the ones that the GroEL subunit follows during conformational transition. In fact, the study (Ma et al., 2000) showed that the GroEL complex has indeed evolved in such a way that the motion of the entire

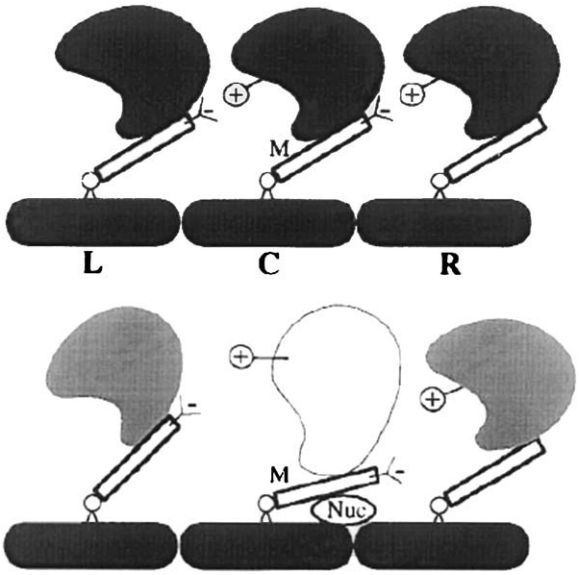

Figure 4. Schematic Illustration of the Intersubunit Interactions in GroEL

For the purpose of illustration, the nucleotide is assumed to be bound only to the central subunit (C). The salt bridge between Arg197 and Glu386 and the effect of the steric interactions are indicated. The intermediate domain is represented by the M helix. This figure is adopted from Figure 4 in Ma et al. (2000).

complex can seamlessly accommodate the motions of individual subunits in order to achieve a concerted opening of the ring for positive cooperativity (Horovitz et al., 2001; Horovitz and Yifrach, 2000), the so-called "vegetable steamer" model (Thirumalai and Lorimer, 2001).

In general, the case of GroEL shows that it is easy for one to overlook the hidden harmonic nature of conformational changes if the analysis is performed on a different level of structural components. If the analysis was performed on the system including the interested structural components plus their interacting neighbors, the modes of this system may not reveal the desired motions for the interested components, as in the cases of those interlocked subunits in the GroEL complex (Ma and Karplus, 1998; Ma et al., 2000) and those permanently deformed α helices (Emberly et al., 2003). However, if one analyzes the modes of the structural components in isolation, the harmonic modes will match the expected conformational changes very well.

Of course, the above arguments by no means imply that the modes of the entire complex are not meaningful (Keskin et al., 2002). They play a role on a larger scale, which is certainly relevant to biomolecular functions. In most cases, the intrinsic modes of smaller building blocks are closely coupled to the modes of the larger complex in fulfilling its functions. As mentioned above, it is one of the primary tasks in the evolution of supramolecular complexes to optimize such a coupling. This viewpoint is a "dynamic" analog to the "principle of minimal frustration" originally proposed to explain the energy landscape for protein folding (Onuchic et al., 1997). Nature designs complexes in such a way that the motions of the individual subunits are consistent with the motions of overall complexes with a minimum of frustration.

Therefore, when one uses NMA to study the behavior of a complex, a hierarchical application of NMA at a different level of complexity, in accord with the hierarchical design of molecular structures, is usually the best way. It is particularly important to remember at what level the injection of ligand binding energy drives the conformational changes.

Timescales of Harmonic Motions

The presence of solvent damping dramatically slows down the large-amplitude motions for biomolecules. Therefore, the timescales of real molecular motions are much longer than what can be estimated from eigenvalues of normal modes that are almost always computed in vacuum. A commonly asked question is why short-timescale normal modes can successfully describe real molecular motions that occur on much larger timescales? The answer is that many large conformational changes effectively follow harmonic trajectories (only in an overall sense) in space as described by normal modes, but they do not follow the timescales of free harmonic oscillators. A rather naïve, but informative, analog can be found in daily life. With a fixed amount of energy, one can swing a man's arm for the same amount of distance in directions intrinsically allowed by a human body both in air and in a swimming pool. But it is much slower in a swimming pool because of water resistance. Therefore, the presence of solvent has a much larger impact on the effective frequencies of harmonic oscillations than on the intrinsically allowed directionality that is determined by the potential surface only. In a theoretical study of neutron scattering spectra (Hinsen et al., 2000), the large-amplitude motions of a protein were described by Brownian motion in an effective harmonic potential. The major anharmonic contributions were found to arise from rigid-body diffusion of the side chains. As seen in many activated processes, for the "overdamped" harmonic oscillators (where the motions are slaverized by friction), the rate of motion only depends on the friction, not on normal mode frequencies. The latter accounts for the "entropy" effect (potential well volume) rather than a timescale for inertial motion.

Generally speaking, the information provided by eigenvectors for the directionality of conformational transitions has wider applications than the information provided by eigenvalues. In the cases of elastic NMA (Atilgan et al., 2001; Hinsen, 1998; Ming et al., 2002a, 2003b; Tama et al., 2002; Tirion, 1996), only the eigenvectors are immediately meaningful. The eigenvalues, other than for sorting the order of modes, however, can be used to compute mechanical properties if they are scaled by macroscopic elastic constants such as Young's modulus. A successful example is found in a study of the differential flexibility of two key components of bacterial flagellum, filament and hook (Flynn and Ma, 2004).

In many large systems, as shown by the recent success of elastic NMA, the directionality is closely related to the overall shape of the molecules (Ma, 2004).

Is Coarse Graining Still Useful for Systems with Atomic Structures?

Using coarse-grained NMA to study large molecular complexes is a major trend in recent years, as exemplified by the elastic NMA (Atilgan et al., 2001; Hinsen, 1998; Ming et al., 2002a, 2003b; Tama et al., 2002; Tirion, 1996). For large systems with only low-resolution structural data available, the usefulness of coarse-grained NMA has been demonstrated beyond any doubt in numerous cases, especially in experimental structural refinement (Brink et al., 2004; Carazo, 2004; Tama et al., 2004; Wu and Ma, 2004a, 2004b). However, a lingering question is why one would still adopt a coarse-grained model for systems with atomic or near-atomic structures available, e.g., the ribosome complexes (Ban et al., 2000; Tama et al., 2003; Wang et al., 2004).

First of all, due to the omission of high-frequency components and anharmonic motions, using low-frequency normal modes as dynamic parameters to describe biomolecular motions is a coarse-grained approach even when the modes are calculated from realistic molecular mechanics potential. The advantage of doing NMA in general is to gain sampling efficiency because, at the current stage, molecular dynamics simulation based on all-atom models cannot completely sample conformational motions on biologically relevant timescales and lengthscales. Coarse-grained NMA has the advantage of allowing one to faithfully extract essential dynamic information with a much extended capacity so as to make it possible to study systems with sizes completely beyond conventional NMA. Interestingly, the coarse-grained NMA offers a unique way to simplify the calculation without losing too much dynamic information at the level one is interested in. Therefore, the coarse-grained NMA will remain an important approach for studying supramolecular complexes regardless of the resolution of the structural data. Ultimately, it is desirable to combine the NMA sampling with molecular dynamics sampling so that smooth coupling between low-frequency global motions and high-frequency side chain dynamics can be achieved. Successful examples can already be found in literature (Tatsumi et al., 2004; Zhang et al., 2003).

How to Identify Biologically Relevant Modes

In all cases of NMA, there is always a question about which mode(s) is (are) functionally relevant. The calculated normal modes, as an orthonormal mathematical basis set, only provide information on all the possible ways that a structure can move. They do not immediately suggest how the structure actually moves for a particular protein. This means that one usually cannot tell which mode is (or modes are) functionally relevant from a set of normal modes calculated for a structure, regardless of the resolution. Additional experimental data are needed in order to identify the biologically relevant modes. Figure 5 schematically illustrates a typical example. For a protein with two globular domains linked by a flexible hinge, the shape of the molecule gives three universal low-frequency modes at whatever resolution the NMA is performed. They are a bending mode, a twisting mode (looking from the top), and a wobbling mode (looking from the back). The active site for such a protein architecture is universally located in the cleft between the two domains. However, for each individual case, the ligand-induced conformational change can follow one single bending mode or a combination of some, or all, of the three modes depending on the detailed atomic interactions between the protein and ligand.

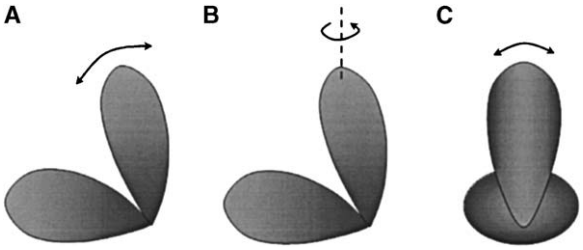

Figure 5. Schematic Illustration of Three Universal Rigid-Body Modes for Bilobate Structures

(A–C) These structures are (A) a hinge-bending mode, (B) a twisting mode (best visualized by a top view), and (C) a wobbling mode (best viewed from the back of the molecule).

In general, although it is hard to identify biologically relevant modes without additional experimental information, it is almost always true that the functionally important modes, once identified, are one or several of the low-frequency modes contained in the normal mode basis set. That is simply because, mechanically, it takes the least amount of energy to achieve conformational transitions along low-frequency normal modes. The pathways that move against normal modes are probably washed out during evolution. This is again a manifestation of the dynamic extension of the principle of minimal frustration.

Finally, the author would like to close this review by reminding ourselves and readers that the types of architecture of biological molecules differ widely, and this variation, in turn, modulates the nature of the molecules' conformational dynamics in a diversified way. Consequently, the suitability of NMA to model conformational dynamics also varies widely. For some systems, NMA appears to be ideal, but, for some others, it is not a very effective method. It is therefore extremely important for one to bear in mind the usefulness and limitations of NMA for various applications. It is also important to remember that the potential of NMA is maximized when it is combined with other experimental data so as to go beyond a "cartoon-like" description of molecular motions. This is precisely the spirit of normal mode-based structural refinement, which has shown great promise in many large complex structures at a wide range of resolutions.

Acknowledgments

The author thanks insightful suggestions from John Straub, Ivet Bahar, and two anonymous reviewers. The author acknowledges support from the National Institutes of Health (R01-GM067801) and the National Science Foundation Career Award (MCB-0237796). The author is also a recipient of the Award for Distinguished Young

Scholars Abroad from the National Natural Science Foundation of China.

References

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

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*, 173–181.

Balog, E., Becker, T., Oettl, M., Lechner, R., Daniel, R., Finney, J., and Smith, J.C. (2004). Direct determination of vibrational density of states change on ligand binding to a protein. Phys. Rev. Lett. *93*, 028103.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science *289*, 905–920.

Beuron, F., Flynn, T.C., Ma, J., Kondo, H., Zhang, X., and Freemont, P.S. (2003). Motions and negative cooperativity between p97 domains revealed by cryoelectron microscopy and quantized elastic deformational model. J. Mol. Biol. *327*, 619–629.

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

Brink, J., Ludtke, S.J., Kong, Y., Wakil, S.J., Ma, J., and Chiu, W. (2004). Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. Structure (Camb.) *12*, 185–191.

Brooks, B., and Karplus, M. (1985). Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. Proc. Natl. Acad. Sci. USA *82*, 4995–4999.

Brooks, C.L., III, Karplus, M., and Pettitt, B.M. (1988). Proteins: a theoretical perspective of dynamics, structure, and thermodynamics. Adv. Chem. Phys. *71*, 1–249.

Bruccoleri, R.E., Karplus, M., and McCammon, J.A. (1986). The hinge-bending mode of a lysozyme-inhibitor complex. Biopolymers *25*, 1767–1802.

Carazo, J.M. (2004). Accessing information on the conformational flexibility of molecular machines. Structure *12*, 170–171.

Chacon, P., Tama, F., and Wriggers, W. (2003). Mega-dalton biomolecular motion captured from electron microscopy reconstructions. J. Mol. Biol. *326*, 485–492.

Cui, Q., Li, G., Ma, J., and Karplus, M. (2004). A normal mode analysis of structural plasticity in the biomolecular motor F(1)-ATPase. J. Mol. Biol. *340*, 345–372.

Cusack, S., Smith, J., Finney, J., Tidor, B., and Karplus, M. (1988). Inelastic neutron scattering analysis of picosecond internal protein dynamics. Comparison of harmonic theory with experiment. J. Mol. Biol. *202*, 903–908.

Deak, J., Chiu, H.L., Lewis, C.M., and Miller, R.J.D. (1998). Ultrafast phase grating studies of heme proteins: observation of the low-frequency modes directing functionally important protein motions. J. Phys. Chem. *102*, 6621–6634.

Diamond, R. (1990). On the use of normal modes in thermal parameters refinement: theory and application to the bovine pancreatic trypsin inhibitor. Acta Crystallogr. A *46*, 425–435.

Doruker, P., Jernigan, R.L., and Bahar, I. (2002). Dynamic of large proteins through hierarchical levels of coarse-grained structures. J. Comput. Chem. *23*, 119–127.

Emberly, E.G., Mukhopadhyay, R., Wingreen, N.S., and Tang, C. (2003). Flexibility of alpha-helices: results of a statistical analysis of database protein structures. J. Mol. Biol. *327*, 229–237.

Flynn, T.C., and Ma, J. (2004). Theoretical analysis of twist/bend ratio and mechanical moduli of bacterial flagellar hook and filament. Biophys. J. *86*, 3204–3210.

Gao, Y.Q., Yang, W., Marcus, R.A., and Karplus, M. (2003). A model for the cooperative free energy transduction and kinetics of ATP hydrolysis by F1-ATPase. Proc. Natl. Acad. Sci. USA *100*, 11339–11344.

Goh, C.S., Milburn, D., and Gerstein, M. (2004). Conformational changes associated with protein-protein interactions. Curr. Opin. Struct. Biol. *14*, 104–109.

Haliloglu, T., Bahar, I., and Erman, B. (1997). Gaussian dynamics of folded proteins. Phys. Rev. Lett. *79*, 3090–3093.

Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. Proteins *33*, 417–429.

Hinsen, K., Petrescu, A.J., Dellerue, S., Bellissent-Funel, M.C., and Kneller, G.R. (2000). Harmonicity in slow protein dynamics. Chem. Phys. *261*, 25–37.

Horovitz, A., and Yifrach, O. (2000). On the relationship between the hill coefficients for steady-state and transient kinetic data: A criterion for concerted transitions in allosteric proteins. Bull. Math. Biol. *62*, 241–246.

Horovitz, A., Fridmann, Y., Kafri, G., and Yifrach, O. (2001). Review: allostery in chaperonins. J. Struct. Biol. *135*, 104–114.

Karplus, M., and Gao, Y.Q. (2004). Biomolecular motors: the F1-ATPase paradigm. Curr. Opin. Struct. Biol. *14*, 250–259.

Kern, D., and Zuiderweg, E.R. (2003). The role of dynamics in allosteric regulation. Curr. Opin. Struct. Biol. *13*, 748–757.

Keskin, O., Bahar, I., Flatow, D., Covell, D.G., and Jernigan, R.L. (2002). Molecular mechanisms of chaperonin GroEL-GroES function. Biochemistry *41*, 491–501.

Kidera, A., Inaka, K., Matsushima, M., and Go, N. (1992). Normal mode refinement: crystallographic refinement of protein dynamic structure applied to human lysozyme. Biopolymers *32*, 315–319.

Kong, Y., Ming, D., Wu, Y., Stoops, J.K., Zhou, Z.H., and Ma, J. (2003). Conformational flexibility of pyruvate dehydrogenase complexes: a computational analysis by quantized elastic deformational model. J. Mol. Biol. *330*, 129–135.

Kottalam, J., and Case, D.A. (1990). Langevin modes of macromolecules: applications to crambin and DNA hexamers. Biopolymers *29*, 1409–1421.

Krebs, W.G., Alexandrov, V., Wilson, C.A., Echols, N., Yu, H., and Gerstein, M. (2002). Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. Proteins *48*, 682–695.

Kumar, S., Ma, B., Tsai, C.J., Sinha, N., and Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci. *9*, 10–19.

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

Lamm, G., and Szabo, A. (1986). Langevin modes of macromolecules. J. Chem. Phys. *85*, 7334–7348.

Leo-Macias, A., Lopez-Romero, P., Lupyan, D., Zerbino, D., and Ortiz, A.R. (2005). An analysis of core deformations in protein superfamilies. Biophys. J. *88*, 1291–1299.

Levitt, M., Sander, C., and Stern, P.S. (1985). Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. J. Mol. Biol. *181*, 423–447.

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

Lipscomb, W.N. (1983). Structure and catalysis of enzymes. Annu. Rev. Biochem. *52*, 17–34.

Ma, J. (2004). New advances in normal mode analysis of supermolecular complexes and applications to structural refinement. Curr. Protein Pept. Sci. *5*, 119–123.

Ma, J., and Karplus, M. (1997). Ligand-induced conformational changes in *ras* p21: a normal mode and energy minimization analysis. J. Mol. Biol. *274*, 114–131.

Ma, J., and Karplus, M. (1998). The allosteric mechanism of the chaperonin GroEL: a dynamic analysis. Proc. Natl. Acad. Sci. USA *95*, 8502–8507.

Ma, J., Sigler, P.B., Xu, Z., and Karplus, M. (2000). A dynamic model for the allosteric mechanism of GroEL. J. Mol. Biol. *302*, 303–313.

Ma, J., Flynn, T.C., Cui, Q., Leslie, A., Walker, J.E., and Karplus, M. (2002). A dynamic analysis of the rotation mechanism for conformational change in F1-ATPase. Structure *10*, 921–931.

MacKerell, A.D., Bashford, D., Jr., Bellott, M., Dunbrack, R.L., Jr., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. *B102*, 3586–3616.

McQuarrie, D.A. (1976). Statistical Mechanics (New York: Harper & Row).

Ming, D., Kong, Y., Lambert, M., Huang, Z., and Ma, J. (2002a). How to describe protein motion without amino-acid sequence and atomic coordinates. Proc. Natl. Acad. Sci. USA *99*, 8620–8625.

Ming, D., Kong, Y., Wakil, S.J., Brink, J., and Ma, J. (2002b). Domain movements in human fatty acid synthase by quantized elastic deformational model. Proc. Natl. Acad. Sci. USA *99*, 7895–7899.

Ming, D., Kong, Y., Wu, Y., and Ma, J. (2003a). Simulation of F-actin filaments of several microns. Biophys. J. *85*, 27–35.

Ming, D., Kong, Y., Wu, Y., and Ma, J. (2003b). Substructure synthesis method for simulating large molecular complexes. Proc. Natl. Acad. Sci. USA *100*, 104–109.

Miyashita, O., Onuchic, J.N., and Wolynes, P.G. (2003). Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. Proc. Natl. Acad. Sci. USA *100*, 12570–12575.

Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. *247*, 536–540.

Onuchic, J.N., Luthey-Schulten, Z., and Wolynes, P.G. (1997). Theory of protein folding: the energy landscape perspective. Annu. Rev. Phys. Chem. *48*, 545–600.

Perutz, M.F. (1969a). The Croonian Lecture, 1968. The haemoglobin molecule. Proc. R. Soc. Lond. B. Biol. Sci. *173*, 113–140.

Perutz, M.F. (1969b). Structure and function of hemoglobin. Harvey Lect. *63*, 213–261.

Reat, V., Dunn, R., Ferrand, M., Finney, J.L., Daniel, R.M., and Smith, J.C. (2000). Solvent dependence of dynamic transitions in protein solutions. Proc. Natl. Acad. Sci. USA *97*, 9961–9966.

Seno, Y., and Go, N. (1990a). Deoxymyoglobin studied by the conformational normal mode analysis II. The conformational change upon oxgenation. J. Mol. Biol. *216*, 111–126.

Seno, Y., and Go, N. (1990b). Deoxymyoglobin studied by the conformational normal mode analysis. I. Dynamics of globin and the heme-globin interaction. J. Mol. Biol. *216*, 95–109.

Smith, J.C. (1991). Protein dynamics: comparison of simulations with inelastic neutron scattering experiments. Q. Rev. Biophys. *24*, 227–291.

Smith, J., Kuczera, K., and Karplus, M. (1990). Dynamics of myoglobin: comparison of simulation results with neutron scattering spectra. Proc. Natl. Acad. Sci. USA *87*, 1601–1605.

Suhre, K., and Sanejouand, Y.H. (2004). On the potential of normal-mode analysis for solving difficult molecular-replacement problems. Acta Crystallogr. D Biol. Crystallogr. *60*, 796–799.

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*, 733–747.

Tama, F., Wriggers, W., and 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.

Tama, F., Valle, M., Frank, J., and Brooks, C.L., III. (2003). Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. Proc. Natl. Acad. Sci. USA *100*, 9319–9323.

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*, 315–326.

Tatsumi, R., Fukunishi, Y., and Nakamura, H. (2004). A hybrid method of molecular dynamics and harmonic dynamics for docking of flexible ligand to flexible receptor. J. Comput. Chem. *25*, 1995–2005.

Thirumalai, D., and Lorimer, G.H. (2001). Chaperonin-mediated protein folding. Annu. Rev. Biophys. Biomol. Struct. *30*, 245–269.

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

Thomas, A., Field, M.J., and Perahia, D. (1996b). Analysis of the low-frequency normal modes of the R state of aspartate transcarbamylase and a comparison with the T state modes. J. Mol. Biol. *261*, 490–506.

Thomas, A., Hinsen, K., Field, M.J., and Perahia, D. (1999). Tertiary and quaternary conformational changes in aspartate transcarbamylase: a normal mode study. Proteins *34*, 96–112.

Tirion, M.M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys. Rev. Lett. *77*, 1905–1908.

Tirion, M.M., ben-Avraham, D., Lorenz, M., and Holmes, K.C. (1995). Normal modes as refinement parameters for the F-actin model. Biophys. J. *68*, 5–12.

Trakhanov, S., Vyas, N.K., Kristensen, D.M., Ma, J., and Quiocho, F.A. (2005). The ligand free and bound structures of the binding protein (LivJ) of the *E. coli* ABC Leu/Ile/Val transport system: trajectory and dynamics of interdomain rotation and polar versus nonpolar interactions in ligand binding. Biochemistry, in press.

Wang, Y., Rader, A.J., Bahar, I., and Jernigan, R.L. (2004). Global ribosome motions revealed with elastic network model. J. Struct. Biol. *147*, 302–314.

Wu, Y., and Ma, J. (2004a). Normal-mode-based refinement of an F-actin model against fibre diffraction data. Fibre Diff. Rev. *12*, 25–28.

Wu, Y., and Ma, J. (2004b). Refinement of F-actin model against fibre diffraction data by long-range normal modes. Biophys. J. *86*, 116–124.

Xu, Z., and Sigler, P.B. (1998). GroEL/GroES: structure and function of a two-stroke folding machine. J. Struct. Biol. *124*, 129–141.

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

Xu, C., Tobi, D., and Bahar, I. (2003). Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T↔R2 transition. J. Mol. Biol. *333*, 153–168.

Yang, W., Gao, Y.Q., Cui, Q., Ma, J., and Karplus, M. (2003). The missing link between thermodynamics and structure in F1-ATPase. Proc. Natl. Acad. Sci. USA *100*, 874–879.

Zhang, Z., Shi, Y., and Liu, H. (2003). Molecular dynamics simulations of peptides and proteins with amplified collective motions. Biophys. J. *84*, 3583–3593.

Zheng, W., and Doniach, S. (2003). A comparative study of motor-protein motions by using a simple elastic-network model. Proc. Natl. Acad. Sci. USA *100*, 13253–13258.

Zhou, Z.H., Liao, W., Cheng, R.H., Lawson, J.E., McCarthy, D.B., Reed, L.J., and Stoops, J.K. (2001). Direct evidence for the size and conformational variability of the pyruvate dehydrogenase complex revealed by three-dimensional electron microscopy. The "breathing" core and its functional relationship to protein dynamics. J. Biol. Chem. *276*, 21704–21713.