Investigating protein dynamics in collective coordinate space

Akio Kitao* and Nobuhiro Go†

Currently, collective coordinates are commonly employed in order to examine protein dynamics. In recent studies, they have been successfully applied to finding functionally relevant motions, to investigating the physical nature of protein dynamics, to sampling of the conformational space and to the analysis of experimental data. Collective coordinates also have other possible applications.

Addresses
Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
*e-mail: kitao@qchem.kuchem.kyoto-u.ac.jp
†e-mail: go@qchem.kuchem.kyoto-u.ac.jp

Current Opinion in Structural Biology 1999, **9**:164–169

http://biomednet.com/elecref/0959440X00900164

© Elsevier Science Ltd ISSN 0959-440X

Abbreviations

BPTI	bovine pancreatic trypsin inhibitor
ED	essential dynamics
JAM	jumping-among-minima
MC	Monte Carlo
MD	molecular dynamics
MODC	molecule optimal dynamic coordinate
MSF	mean square fluctuation
NM	normal mode
NMA	NM analysis
PC	principal component
PCA	PC analysis

Introduction

Dynamics is essential for a protein to function. Recent advancements in computers and computational techniques enable us to simulate protein dynamics more accurately in a more realistic environment; however, it is not a trivial problem to extract a functionally relevant motion from simulation results. A key solution to this problem is the use of collective coordinates. Recent studies have shown that functionally relevant motions occur along the direction of a few collective coordinates, which dominantly contribute to atomic fluctuation. One example is the molecular dynamics (MD) simulation of CRP:(cAMP)$_2$, the cAMP receptor protein bound to cAMP, carried out by García and Harman [1]. Two conformations, open and closed, exist in the crystal environment, depending on the occurrence of binding to DNA. The transition from the open to the closed conformation found during a MD simulation involves only two collective coordinates. Another example is the case of the T4 lysozyme [2•], whereby a large variation of the hinge-bending angle has been seen in the crystal environment. The first two largest amplitude collective modes determined from the MD trajectory, which describes the domain motion, are identified as being the closure and twisting modes.

More importantly, collective coordinates can be used to predict a low-dimensional subspace in which significant protein motion is expected to take place. If this concept is true, the number of degrees of freedom can be extremely reduced. In a Cartesian coordinate space protein, the number of internal degrees of freedom is (3N–6), where N is the number of atoms considered. In the case of lysozyme, where N ≅ 2000, the number of degrees of freedom can be reduced to 5% if 300 appropriately chosen collective coordinates span the subspace. It is well known that atomic fluctuations within such a subspace dominantly contribute to total fluctuation. A subspace that is spanned by a small number of low-frequency normal modes is termed an 'important subspace' [3]. This concept has been successfully employed to refine crystal structures [3–5]. Later, Amadei *et al.* [6] termed a subspace that is spanned by a subset of anharmonic collective coordinates determined by MD an 'essential subspace'. A small number of large amplitude collective modes determined by simulations depends on the simulation length in some cases [7], because large amplitude transitions occur stochastically. Teeter and Case [8] showed that the very low frequency modes are sensitive to the force-field; however, the frequency distributions and fluctuations were remarkably similar. In addition, the subspace spanned by a small number of collective coordinates is invariant in many cases and is almost independent of the treatment of the degrees of freedom [9,10], the solvent effect [11,12], the simulation length [13••] and the initial and other conditions [14]. In this paper, we review how the concept of these subspaces is a powerful tool for investigating protein dynamics.

Determining a collective coordinate set

A collective coordinate set can be determined as a set of eigenvectors that diagonalizes a second moment matrix. This is based on what is called principal component analysis (PCA). In Cartesian coordinate space, an element of the second moment matrix **A**, a_{ij}, is defined as:

$$a_{ij} \equiv \left\langle \left(x_i - \langle x_i \rangle\right)\left(x_j - \langle x_j \rangle\right) \right\rangle \tag{1}$$

where $\langle ... \rangle$ is the average over the instantaneous structures sampled during the period of simulation and x_i is a mass-weighted atomic coordinate. A set of eigenvalues and eigenvectors is obtained by solving the standard eigenvalue problem:

$$\mathbf{AW} = \mathbf{W}\zeta \tag{2}$$

where $\zeta = diag(\zeta_m)$ is a diagonal matrix whose mth diagonal element is the eigenvalue ζ_m, **W** is a matrix whose mth column vector w_m is the eigenvector of ζ_m. The vector w_m, given by a linear combination of Cartesian coordinates, represents the axis of the mth collective coordinate in the conformational space. The eigenvalue ζ_m represents the

mean square fluctuation (MSF) along this axis. The second moment matrix is calculated directly from a MD or Monte Carlo (MC) trajectory. The motions along these coordinates are termed 'modes', although they are not necessarily vibrational. In normal mode analysis (NMA), the second moment matrix is given by:

$$\mathbf{A} = k_B T \mathbf{F}^{-1} \quad (3)$$

where $\mathbf{F}$ is a matrix of second derivatives of the potential energy at a minimum point, k_B is the Boltzmann constant and T is the absolute temperature. In dihedral angle space, the number of degrees of freedom is reduced to about an eighth. For established methods in dihedral angle space, see the review by Hayward and Go [15].

Further explanation may be required concerning the term 'PCA'. When PCA was first applied to a MD trajectory, it was termed 'quasi-harmonic', because the focus was on the anharmonic aspects of protein dynamics [16]. The term 'principal component analysis' was first used in [11] because the projection of the MD trajectory onto a subspace was more highlighted. The term 'essential dynamics' (ED) is also used to represent a few anharmonic collective modes [6]. García and Harman [1] called principal coordinates 'molecule optimal dynamic coordinates' (MODCs). These analyses share essentially the same concept, apart from some details, for example, the atoms considered, whether it is mass weighted or not and so on. We prefer the use of mass-weighted coordinates because the relationship between PCA and NMA is clearer.

Practically, the most time consuming aspect of PCA and NMA is the diagonalization of the second moment matrix. In PCA, the reduction of the matrix can be done easily by selecting solely heavy atoms or Cα atoms. As such a reduction cannot be done, diagonalization is more involved in NMA. For recent advancements in both NMA and diagonalization techniques, see the review by Hayward [17]. Simplified models have been proposed and have been successful in predicting atomic fluctuations [18,19••]. These methods are designed to reproduce a few large-amplitude collective modes using simplified force-fields.

As an alternative to the second moment matrix, the metric matrix has been employed to determine collective coordinates in a method termed 'multidimensional scaling' [20].

Projection of protein motion onto collective coordinate space

One of the first attempts to project MD and MC trajectories onto the normal mode (NM) space was carried out on the human lysozyme [21]. This was successful in explaining the anharmonicity in a single conformational substate. Projection onto the NM space is not, however, an efficient way of observing intersubstate transitions [22]. Two important features, transitions between substates and solvent effects, were clearly revealed by the first attempt at projecting a MD trajectory onto the principal coordinate space [9]. By projecting a MD trajectory of melittin in water onto the two-dimensional space spanned by the largest and second largest amplitude principal modes, a transition from one substate to another is very clearly seen. The motion along these coordinates involves a bending of the molecule. Soon after, conformational transitions in the subspace spanned by a few principal components (PCs) were reported in the cases of Met-enkephaline [23], crambin [24], bovine pancreatic trypsin inhibitor (BPTI) [12] and hen-egg lysozyme [6]. PCA has now become a common tool for investigating important motions of biological molecules [25–29].

In PCA, solvent effects are evident in two ways. One is detected as small-amplitude conformational transitions that are not seen in a vacuum [11,12]. In the case of melittin, this type of transition involves a change of hydration structure [30]. Fine structures in the free energy surface, which are not seen in a vacuum, exist in water [31]. In the case of myoglobin, global transitions were observed more frequently in a fully solvated model, compared with those in a partially solvated model [32]. This may be due to the fact that the presence of solvent lowers the barriers of large conformational transitions. The other solvent effect appears as a damping of collective motion. In the case of melittin [11] and BPTI [12], time correlation functions of the large-amplitude PCs show overdamping behavior. In the case of a harmonic energy surface, this damping effect can be elegantly incorporated into NMA. 'Langevin mode' analysis was first proposed by Lamm and Szabo [33] and was applied to crambin and DNA by Kottalam and Case [34]. The overdamping and underdamping behaviors of collective modes were elucidated by this method; however, the problem as to how to estimate the matrix of friction coefficients remains. From a back calculation of this matrix from an MD trajectory, it has been shown that simple hydrodynamic models for the friction matrix fail to accurately predict off-diagonal elements of the friction matrix [11]. If friction coefficients are correctly given, Langevin mode analysis can explain the spectral shape determined by an inelastic neutron scattering experiment in a low-frequency region [12].

It should be also noted that there is an important feature that is not affected by solvent. Using the cross-correlations and co-variances of atomic fluctuations, Ichiye and Karplus [35] showed that the correlations of internal motions are similar in different environments. This is consistent with the finding that a subspace spanned by a small number of collective coordinates is almost invariant, as described above.

By projecting an MD trajectory onto principal coordinates, Caves *et al.* [36••] reported a problem of sampling using MD. Using 10 independent MD trajectories of crambin with different initial velocities, they found four major conformational regions. In this case, the use of multiple trajectories is more efficient in sampling compared to a single, long trajectory. An alternative way of solving the

Table 1

Characteristics of collective modes determined by PCA and JAM*.

	Multiply hierarchical mode	Singly hierarchical mode	Harmonic mode
Free energy surface			
Probability distribution	Non-Gaussian multiple peak	Gaussian-like single peak	Gaussian
Anharmonicity factor ρ_m†	> 2	2–1	= 1
Packing topology	Significant change	No change	No change
Fluctuation of dihedral angles	Large change in both mainchains and sidechains	Large change in sidechains	Small change
Possible relationship with other classifications	Essential subspace [6] Relaxation dynamics [1]	Fluctuation dynamics [1]	Unimodal dynamics [1]
	Equivalent to NM important subspace [3] (spans multiply and singly hierarchical modes)		
Specific data for the human lysozyme			
Principal mode number	1–30	31–300	301–6117
% in mode number	0.5	4.5	95.0
% in total MSF	82.1	15.0	2.9
Transition timescale (s)	> 10^{-13} (various transition timescales co-exist in one mode)	10^{-13}–10^{-11}	–
RMSF (Å)	> 0.08	0.08–0.015	0.015 >

*A newly generated table based on [13••]. †The factor ρ_m is defined as $\rho_m = \left(\zeta_m / \zeta_m^{har}\right)^{1/2}$, where ζ_m^{har} is the expected MSF from NM along the direction [6]. RMSF, root mean square fluctuation.

sampling problem is the use of efficient sampling methods, for example, multicanonical MC [37] and MD [38]; however, these still need enormous amounts of computational time, even in vacuum [39]. Recently, Becker [40] examined the possibility of using PCA for the mapping of the potential energy surface of small peptides.

Energy landscapes in collective coordinate space

After pioneering work on the energy landscapes of myoglobin [41] and BPTI [42–46] in their native states, most of the recent attempts to investigate energy landscapes have been carried out using collective coordinates. From non-Gaussian probability distribution functions of projections of an MD trajectory onto principal coordinates, large-amplitude nonlinear motions were identified [24]. Similarly, Amadei *et al.* [6] termed the space spanned by these anharmonic coordinates 'essential subspace'. In order to further understand harmonic and anharmonic aspects of protein dynamics, NMs have been employed [22,47]. Using what is called the 'anharmonicity factor', anharmonic modes and harmonic modes are clearly separated by principal mode numbers [22]. In the case of BPTI, only 11.7% of the modes are anharmonic (equivalent to 7.4% in an all-atom model) and, in the case of human lysozyme [13••], only 5.0% are anharmonic. By comparing the NMs of 201 different minima, Janežič *et al.* [48] found that low-frequency NMs extensively overlap in angular frequency when minima are separated by a relatively short time interval. Using different time windows, Troyer and Cohen [20] found different types of conformational transitions. Transitions among seven major clusters in the binding loop of BPTI were suggested to be functionally important as the inhibitor. García and Harman [1] pointed out three types of dynamics — unimodal, fluctuation and relaxation. In the case of crambin in a crystal environment, a motion whose characteristic time is slower than 2 ns was found [49••].

The jumping-among-minima (JAM) model elucidates the above findings more clearly [13••]. In this model the second moment matrix is divided into two terms:

$$a_{ij} = \sum_{k=1}^{M} f_k \left(\langle x_i \rangle_k - \langle x_i \rangle\right)\left(\langle x_j \rangle_k - \langle x_j \rangle\right) + \sum_{k=1}^{M} f_k \left\langle \left(x_i - \langle x_i \rangle_k\right)\left(x_j - \langle x_j \rangle_k\right) \right\rangle_k \quad (4)$$

where M is the number of conformational substates, k is the index of the substate, $\langle ... \rangle_k$ is an average over the period during which the state point stays in the kth substate and f_k is a fraction of this period to the total simulation time. The first and second terms of Equation 4 represent contributions from atomic fluctuations arising from JAM and arising from fluctuations within each substate, respectively. This model was applied to the analysis of a 1 ns MD trajectory of the human lysozyme in water. From the analysis of intrasubstate motion,

each conformational subspace is shown to be nearly harmonic and mutually similar. Using the anharmonicity factor and analyzing the shapes of the probability distribution functions, three types of modes are identified: multiply hierarchical modes, singly hierarchical modes and harmonic modes. Along the axis of multiply hierarchical modes, different levels of energy barriers co-exist; see Table 1 for details.

Beyond the native state, we think that collective coordinate descriptions of protein dynamics would still be useful. A test calculation shows that, even in the wider conformational space beyond the native state, a small number of PCs still dominantly contribute to the total fluctuation (A Kitao, H Hirose, S Nakamura, M Ikeguchi, K Shimizu, N Go, unpublished data). This suggests the possibility of using PCA to examine the hypotheses of the folding pathway and the folding funnel by projecting the trajectories of a folding/unfolding simulation onto a low-dimensional subspace.

Functional importance of collective motion

As already described above, PCA is a powerful tool for determining functionally relevant motions. Recently, in order to further understand the characteristics of collective motions, the concept of a 'dynamic domain' has been introduced, whereby collective motion is reduced to rigid-body domain motion [2•,50,51,52••]. Using well-defined criteria, dynamic domains and rotation axes have been determined. The domain motion is quantified by a rotation angle around the axis. The concept of a dynamic domain is expected to give some insights into understanding protein structure.

There are some cases for which the functional importance of collective motion is not evident from the motion itself. Using NM variables, an analytical model for formulating a protein electron transfer reorganization energy spectrum was proposed [53•] and applied to cytochrome c [54•]. This is an extension of the work done by Schulten and Tesch [55] to include a collective coordinate description so that the structural aspects of the phenomenon can be discussed explicitly. This model enables the identification of NMs that contribute significantly to the reorganization energy. As described in the original studies [53•,54•], it can be easily extended to a more realistic model using PCA.

Using collective coordinate as a basis set for efficient sampling

As protein dynamics is highly anisotropic, MC simulations with isotropic sampling step sizes are inefficient. In order to overcome this, NMs are used as variables and the sampling step size is scaled by NM amplitudes [56]. In MD, collective coordinates that span essential space are used as variables [57]. These methods remarkably improve the sampling efficiency, although the problem still remains as to how the solvent effect is incorporated into the former method. One possible way is the use of the XRISM theory. Although extensive calculation is needed, XRISM has been applied to a 436 site (atom) model of melittin [31]. A recent, efficient algorithm [58] could be incorporated into these efficient sampling methods.

Analysis of experimental data by means of collective variables

In B-factor refinement of X-ray crystal structures, atomic motion is assumed to be isotropic. Although anisotropic B-factor refinement has been tried, it requires that more parameters be determined. Using the NM important subspace, the number of parameters is highly reduced in NM refinement [3–5,59]. In addition, external and internal motions can be clearly separated. The internal motion of the human lysozyme in its crystal environment is highly anisotropic and atomic fluctuations are in good agreement with those determined by NM refinement [3]. The refinement of the NMR spectroscopy structure was also carried out using the same concept [60]. Using PCA, Abseher *et al.* [61] found a significant overlap between those conformations sampled by unconstrained MD and those determined in order to satisfy NMR constraints. This is as expected from the concept discussed in the Introduction. Collective coordinates have also been used to estimate NMR order parameters [62,63].

The JAM model can also be used to analyze experimental data obtained by X-ray and neutron scattering, laser spectroscopy and NMR spectroscopy. If intrasubstate motion is given by NM and is invariant in all conformation substates, the atomic coordinate $x_i(t)$ is given by the sum of the contributions from external motion $x_i^{EXT}(t)$, JAM motion $x_i^{JAM}(t)$ and NM vibration $x_i^{NM}(t)$ as:

$$x_i(t) = x_i^{EXT}(t) + x_i^{JAM}(t) + x_i^{NM}(t). \quad (5)$$

For example, in the case of inelastic neutron scattering, the incoherent dynamic structure factor $S_{inc}(Q,\omega)$ is given by the convolution of three structure factors — $S_{inc}^{EXT}(Q,\omega)$, $S_{inc}^{JAM}(Q,\omega)$ and $S_{inc}^{NM}(Q,\omega)$ as:

$$S_{inc}(Q,\omega) = S_{inc}^{EXT}(Q,\omega) \otimes S_{inc}^{JAM}(Q,\omega) \otimes S_{inc}^{NM}(Q,\omega) \quad (6)$$

if these three types of motion are uncorrelated. Here, the symbol ⊗ denotes the convolution product over angular frequency. In the case of staphylococcal nuclease, the structure factor measured by TFXA was in remarkably good agreement with the calculated $S_{inc}^{NM}(Q,\omega)$ factor, except in a few high-frequency peaks and very low-frequency regions [64••]. It may also be possible to characterize the JAM motion from a high-resolution structure factor.

Conclusions

Collective coordinates are employed to investigate protein dynamics. They are useful not only in finding functionally relevant motions, but also in understanding the physical nature of the protein energy landscape, efficient sampling and the analysis of experimental data.

Acknowledgements

We are very grateful to Steven Hayward for reading this manuscript carefully. From April 1998 to March 1999, AK was a Monbusho sponsored Japanese Overseas Research Fellow in Gerhard Wagner's group at the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Garcia AE, Harman JG: **Simulations of CRP:(cAMP)$_2$ in noncrystalline environments show a subunit transition from the open to the closed conformation.** *Protein Sci* 1996, **5**:62-71.

2. de Groot BL, Hayward S, van Aalten DMF, Amadei A,
• Berendsen HJC: **Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data.** *Proteins* 1998, **31**:116-127.
Using the concept of a dynamic domain, the large hinge-bending motion observed from MD was compared with and was found to be similar to variations of conformations seen in various crystal environments and mutants.

3. Kidera A, Inaka K, Matsushima M, Go N: **Normal mode refinement: crystallographic refinement of protein dynamic structure. II. Application to human lysozyme.** *J Mol Biol* 1992, **225**:477-486.

4. Kidera A, Go N: **Refinement of protein dynamic structure: normal mode refinement.** *Proc Natl Acad Sci USA* 1990, **87**:3718-3722.

5. Kidera A, Go N: **Normal mode refinement: crystallographic refinement of protein dynamic structure. I. Theory and test by simulated diffraction data.** *J Mol Biol* 1992, **225**:457-475.

6. Amadei A, Linssen ABM, Berendsen HJC: **Essential dynamics of proteins.** *Proteins* 1993, **17**:412-425.

7. Balsera MA, Wriggers W, Oono Y, Schulten K: **Principal component analysis and long time protein dynamics.** *J Phys Chem* 1996, **100**:2567-2572.

8. Teeter MM, Case DA: **Harmonic and quasiharmonic descriptions of crambin.** *J Phys Chem* 1990, **94**:8091-8097.

9. Kitao A, Go N: **Conformational dynamics of polypeptides and proteins in the dihedral angle space and in the Cartesian coordinate space: normal mode analysis of deca-alanine.** *J Comput Chem* 1991, **12**:359-368.

10. Kitao A, Hayward S, Go N: **Comparison of normal mode analyses on a small globular protein in dihedral angle space and Cartesian coordinate space.** *Biophys Chem* 1994, **52**:107-114.

11. Kitao A, Hirata F, Go N: **The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulation of melittin in water and in vacuum.** *Chem Phys* 1991, **158**:447-472.

12. Hayward S, Kitao A, Hirata F, Go N: **Effect of solvent on collective motions in globular proteins.** *J Mol Biol* 1993, **234**:1207-1217.

13. Kitao A, Hayward S, Nobuhiro G: **Energy landscape of a native**
•• **protein: jumping-among-minima model.** *Proteins* 1998, **33**:496-517.
The physical nature of an energy landscape was investigated using a jumping-among-minima model. The mutual similarity of the conformational substates, the distribution of conformational substates, the multidimensional space and the hierarchy of conformational substates are discussed.

14. de Groot BL, van Aalten DMF, Amadei A, Berendsen HJC: **The consistency of large concerted motions in proteins in molecular dynamics simulations.** *Biophys J* 1996, **71**:1707-1713.

15. Hayward S, Go N: **Collective variable description of native protein dynamics.** *Annu Rev Phys Chem* 1995, **46**:223-250.

16. Levy RM, Srinivasan AR, Olson WK, McCammon JA: **Quasi-harmonic method for studying very low frequency modes in proteins.** *Biopolymers* 1984, **23**:1099-1112.

17. Hayward S: **Normal mode analysis of biological molecules.** In *Computational Biochemistry and Biophysics.* Edited by Becker OM, MacKerell AD Jr, Roux B, Watanabe M. New York: Marcel Dekker; 1999:in press.

18. Tirion MM: **Large amplitude elastic motions in proteins from a single-parameter, atomic analysis.** *Phys Rev Lett* 1996, **77**:1905-1908.

19. Bahar I, Atilgan AR, Erman B: **Direct evaluation of thermal**
•• **fluctuations in proteins using a single-parameter harmonic potential.** *Fold Des* 1997, **2**:173-181.
An elastic network model was proposed and was applied to predict X-ray B-factors. The atomic fluctuations of 12 different molecules calculated using this simplified model are in good agreement with the B-factors.

20. Troyer JM, Cohen FE: **Protein conformational landscapes: energy minimization and clustering of a long molecular dynamics trajectory.** *Proteins* 1995, **23**:97-110.

21. Horiuchi T, Go N: **Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: human lysozyme.** *Proteins* 1991, **10**:106-116.

22. Hayward S, Kitao A, Go N: **Harmonicity and anharmonicity in protein dynamics: a normal mode analysis and principal component analysis.** *Proteins* 1995, **23**:177-186.

23. Abagyan R, Argos P: **Optimal protocol and trajectory visualization for conformational searches of peptides and proteins.** *J Mol Biol* 1992, **225**:519-532.

24. Garcia AE: **Large-amplitude nonlinear motions in proteins.** *Phys Rev Lett* 1992, **68**:2696-2699.

25. Basu G, Kitao A, Hirata F, Go N: **A collective motion description of the 3_{10}-/α-helix transition: implications for a natural reaction coordinate.** *J Am Chem Soc* 1994, **116**:6307-6315.

26. Garcia AE, Soumpasis DM, Jovin TM: **Dynamics and relative stabilities of parallel- and antiparallel-stranded DNA duplexes.** *Biophys J* 1994, **66**:1742-1755.

27. van Aalten DMF, Amadei A, Linssen ABM, Eijsink VGH, Vriend G, Berendsen HJC: **The essential dynamics of thermolysin: confirmation of the hinge-bending motion and comparison of simulations in vacuum and water.** *Proteins* 1995, **22**:45-54.

28. Nadig G, Vishveshwara S: **Effects of substrate binding on the dynamics of RNase A: molecular dynamics simulations of UpA bound and native RNase A.** *Biopolymers* 1997, **42**:505-520.

29. Laatikanien R, Saarela J, Tuppurainen K, Hassinen T: **Internal motions of native lysozyme are more organized than those of mutants: a principal component analysis of molecular dynamics data.** *Biophys Chem* 1998, **73**:1-5.

30. Kitao A, Hirata F, Go N: **Effects of solvent on the conformation and the collective motions of a protein. 2. Structure of hydration in melittin.** *J Phys Chem* 1993, **97**:10223-10230.

31. Kitao A, Hirata F, Go N: **Effects of solvent on the conformation and the collective motions of a protein. 3. Free energy analysis by the extended RISM theory.** *J Phys Chem* 1993, **97**:10231-10235.

32. Andrews BK, Romo T, Clarage JB, Pettitt BM, Phillips GN Jr: **Characterizing global substates of myoglobin.** *Structure* 1998, **6**:587-594.

33. Lamm G, Szabo A: **Langevin modes of macromolecules.** *J Chem Phys* 1986, **85**:7334-7348.

34 Kottalam J, Case DA: **Langevin modes of macromolecules: applications to crambin and DNA hexamers.** *Biopolymers* 1990, **29**:1409-1421.

35. Ichiye T, Karplus M: **Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations.** *Proteins* 1991, **11**:205-217.

36. Caves LSD, Evanseck JD, Karplus M: **Locally accessible**
•• **conformations of proteins: multiple molecular dynamics simulations of crambin.** *Protein Sci* 1998, **7**:649-666.
The problems of sampling were discussed using multiple molecular dynamics trajectories. As simulation length is limited, it is important to find an effective way of sampling conformational space, as discussed in this paper.

37. Hansmann UHE, Okamoto Y: **Prediction of peptide conformation by multicanonical algorithm: new approach to the multiple-minima problem.** *J Comput Chem* 1993, **14**:1333-1338.

38. Nakajima N, Nakamura H, Kidera A: **Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides.** *J Phys Chem B* 1997, **101**:807-824.

39. Shirai H, Nakajima N, Higo J, Kidera A, Nakamura H: **Conformational sampling of CDR-H3 in antibodies by multicanonical molecular dynamics simulation.** *J Mol Biol* 1998, **275**:481-496.

40. Becker OM: **Principal coordinate maps of molecular potential energy surfaces.** *J Comput Chem* 1998, **19**:1255-1267.

41. Elber R, Karplus M: **Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin.** *Science* 1987, **235**:318-321.

42. Noguti T, Go N: **Structural basis of hierarchical multiple substates of a protein. I: introduction.** *Proteins* 1989, **5**:97-103.

43. Noguti T, Go N: **Structural basis of hierarchical multiple substates of a protein. II: Monte Carlo simulation of native thermal fluctuations and energy minimization.** *Proteins* 1989, **5**:104-112.

44. Noguti T, Go N: **Structural basis of hierarchical multiple substates of a protein. III: side chain and main chain local conformations.** *Proteins* 1989, **5**:113-124.

45. Noguti T, Go N: **Structural basis of hierarchical multiple substates of a protein. IV: rearrangements in atom packing and local deformations.** *Proteins* 1989, **5**:125-131.

46. Noguti T, Go N: **Structural basis of hierarchical multiple substates of a protein. V: nonlocal deformations.** *Proteins* 1989, **5**:132-138.

47. Hayward S, Kitao A, Go N: **Harmonic and anharmonic effects in the dynamics of BPTI: a normal mode analysis and principal component analysis.** *Protein Sci* 1994, **3**:936-943.

48. Janežič D, Veneble RM, Brooks BR: **Harmonic analysis of large systems. III. Comparison with molecular dynamics.** *J Comput Chem* 1995, **16**:1554-1566.

49. •• Garcia AE, Blumenfeld R, Hummer G, Krumhasl JA: **Multi-basin dynamics of a protein in a crystal environment.** *Physica D* 1997, **107**:225-239.

A high quality and lengthy (5.1 ns) molecular dynamics simulation of crambin was carried out using a particle mesh ewald summation. Two different types of transitions between minima have been observed: relatively short dynamics occurred in 5 to 800 ps and longer timescale dynamics occurred in 2 ns and longer.

50. Hayward S, Kitao A, Berendsen HJC: **Model-free methods of analyzing domain motions in proteins from simulations: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme.** *Proteins* 1997, **27**:425-437.

51. Wriggers W, Schulten K: **Protein domain movements: detection of rigid domains and visualization of hinges in comparisons of atomic coordinates.** *Proteins* 1997, **29**:1-14.

52. •• Hayward S, Berendsen HJC: **Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme.** *Proteins* 1998, **30**:144-154.

A new and consistent methodology was introduced in order to analyze conformational changes in proteins for dynamic domains, hinge axes and hinge-bending residues. Its application has led to new insights into the motions of T4 lysozyme and citrate synthase.

•53. Basu G, Kitao A, Kuki A, Go N: **Protein electron transfer reorganization energy spectrum from normal mode analysis. 1. Theory.** *J Phys Chem B* 1998, **11**:2076-2084.

A new analytical model for describing the coupling of protein fluctuations to electron transfer is proposed using normal modes. It allows a direct connection between actual protein motion and electron transfer to be made.

•54. Basu G, Kitao A, Kuki A, Go N: **Protein electron transfer reorganization energy spectrum from normal mode analysis. 2. Application to Ru-modified cytochrome c.** *J Phys Chem B* 1998, **11**:2085-2094.

The application of the method described in [53•] to Ru-modified cytochrome c. There are no special modes that make a dominant contribution to the electron transfer reorganization energy. This may be an important feature of the electron transfer process, in contrast to other cases [1,2•], in which a few collective modes are functionally relevant.

55. Schulten K, Tesch M: **Coupling of protein motion to electron transfer – molecular-dynamics and stochastic quantum-mechanics study of photosynthetic reaction centers.** *Chem Phys* 1991, **158**:421-446.

56. Noguti T, Go N: **Efficient Monte Carlo method for simulation of fluctuating conformations of native proteins.** *Biopolymers* 1985, **24**:527-546.

57. Amadei A, Linssen ABM, de Groot BL, van Aalten DMF, Berendsen HJC: **An efficient method for sampling the essential subspace of proteins.** *J Biomol Struct Dyn* 1996, **13**:615-625.

58. Kinoshita M, Okamoto Y, Hirata F: **Calculation of hydration free energy for a solute with many atomic sites using the RISM theory: a robust and efficient algorithm.** *J Comput Chem* 1997, **18**:1320-1326.

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

60. Bruschweiler R, Case DA: **Collective NMR relaxation model applied to protein dynamics.** *Phys Rev Lett* 1994, **72**:940-943.

61. Abseher R, Horstink L, Hilbers CW, Nilges M: **Essential spaces defined by NMR structure ensembles and molecular dynamics simulation show significant overlap.** *Proteins* 1998, **31**:370-382.

62. Sunada S, Go N, Koehl P: **Calculation of nuclear magnetic resonance order parameters in proteins by normal mode analysis.** *J Chem Phys* 1996, **104**:4768-4775.

63. Sunada S, Go N: **Calculation of nuclear magnetic resonance order parameters in proteins by normal mode analysis. II. Contribution from localized high frequency motions.** *J Chem Phys* 1996, **105**:6560-6564.

64. •• Lamy AV, Smith JC, Yunoki J, Parker SF, Kataoka M: **High-resolution vibrational inelastic neutron scattering: a new spectroscopic tool for globular proteins.** *J Am Chem Soc* 1997, **119**:9268-9273.

A high-resolution spectrum measured by inelastic neutron scattering was shown to be in good agreement with the spectrum calculated using normal mode analysis. This indicates the possibility of using neutron scattering as a tool for investigating protein dynamics in atomic detail.