ELSEVIER

Journal of Molecular Graphics and Modelling 21 (2003) 309–319

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Correlation between normal modes in the 20–200 cm^{-1} frequency range and localized torsion motions related to certain collective motions in proteins

Z.W. Cao, X. Chen, Y.Z. Chen*

Department of Computational Science, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore

Received 3 July 2002; accepted 10 October 2002

Abstract

In certain biologically relevant collective motions, such as protein domain motions and sub-domain motions, large amplitude movements are localized in one or a few flexible regions consisting of a small number of residues. This paper explores the possible use of normal mode analysis in probing localized vibrational torsion motions in these flexible regions that may be related to certain collective motions. The normal modes of 10 structures of five proteins in different conformation (TRP repressor, calmodulin, calbindin D_{9k}, HIV-1 protease and troponin C), known to have shear or hinge domain or sub-domain motion, respectively, are analyzed. Our study identifies, for each structure, unique normal modes in the 20–200 cm^{-1} frequency range, whose corresponding motions are primarily concentrated in the region where large amplitude torsion movements of a known domain or sub-domain motion occur. This suggests possible correlation between normal modes at 20–200 cm^{-1} frequency range and initial fluctuational motions leading to localized collective motions in proteins, and thus the potential application of normal mode analysis in facilitating the study of biologically important localized motions in biomolecules.

Keywords: Fluctuational motion; Hinge motion; Molecular modeling; Shear motion; Torsion motion; Vibrational motion; Flexible region; Linker

1. Introduction

Collective motions of biomolecules, which involve a number of continuous residues moving in a concerted fashion, play important roles in biological processes. Examples are hinge and shear domain motions in proteins [1–4] and single base pair opening motions in nucleic acids [5,6]. These types of collective motions normally occur in proteins or nucleic acids with half-rigid domains, constrained sub-domains, functional groups linked by short flexible regions and having few packing constraints and are, therefore, free to undergo conformational change. A few large torsion angle changes in these flexible regions were found to be sufficient to produce almost the entire collective motion in proteins [7]. Localized torsional motions have been shown to be involved in the initial stages of biologically important low energy collective motions [8,9]. A database of protein motions has been published in which possible flexible regions and large amplitude torsion angle changes responsible for these motions are described [7].

Normal mode analysis has been used as one of the methods in the study of collective motions in biomolecules [10–16]. Given its advantage of requiring only one structure and of providing detailed description of long-time motions deviating from equilibrium positions [17], normal mode analysis has been considered a useful approach [18] to complement other methods, such as molecular dynamics [19–22] and Monte Carlo simulation [23], in the study of protein motions. So far, attention has mainly been focused on very low-frequency modes (<20 cm^{-1}) in proteins and RNAs [24–30]. These modes have significant anharmonic and collective characters, and they are dominant in biomolecules [4]. However, the underlying motions of these modes are delocalized [4,31]. For instance, two modes below 20 cm^{-1} in *ras p21* were found to involve relatively large amplitude motions in a number of residues scattered across different parts of the protein [11]. The lowest frequency modes in ribonuclease and lysozyme contain large amplitude motions covering almost the entire protein [32]. A mode at 0.60 cm^{-1} in a yeast tRNA was shown to involve relatively large amplitude

* Corresponding author.
E-mail address: yzchen@cz3.nus.edu.sg (Y.Z. Chen).
URL: http://www.cz3.nus.edu.sg/~yzchen

1093-3263/02/$ – see front matter © 2002 Elsevier Science Inc. All rights reserved.
PII: S1093-3263(02)00185-7

motions in more than 50% of the residues in the RNA [28].

The modes relevant to the localized large amplitude torsion displacement of domain or sub-domain collective motions appear to be located in the frequency range between 20 and 200 cm^{-1}. A study showed that the torsional vibrations in α-helices occur between 50 and 110 cm^{-1}, in β-sheets between 5 and 30 cm^{-1}, and in turns between 5 and 75 cm^{-1} [32]. Vibrational modes in this frequency range have been observed experimentally. For example, two Raman peaks were identified at 30 and 70 cm^{-1} for the protein chymotrypsin [33,34]. For lysozyme, two Raman peaks were found at 25 and 75 cm^{-1} [35]. Moreover, several peaks were observed from 30 to 140 cm^{-1} in polynucleotides [36–38], which are considered to be related to the torsion vibrational modes [32].

As for the modes above 200 cm^{-1}, it has been reported that the underlying motions primarily involve stretching and angle-bending motions of individual bonds in both metal-free biomolecules [39–41] and those complexed with metals [42,43]. Several studies have shown good correlation between computed normal modes and observed Raman and IR spectral lines in biomolecules at frequencies above 600 cm^{-1} [44–48].

These studies seem to suggest that normal modes in the frequency range of 20–200 cm^{-1} might be useful for probing localized fluctuational motions leading to large amplitude displacements in flexible regions that are related to collective protein motions of biological significance. The potential use of normal mode analysis for investigating localized torsional motions in flexible regions can be tested on proteins with known domain or sub-domain motions. In this work, we study five proteins known to have shear or hinge domains or sub-domain motion, respectively. These motions primarily involve torsional movements confined to the flexible regions within proteins such as hinge or linker residues. Localized vibrational motions are examined and identified for these proteins from the computed normal modes, since they may initiate large amplitude torsional movements in specific flexible regions responsible for a domain or sub-domain motion. The computed modes of each protein are searched for one or more such modes whose relatively large amplitude motions are primarily concentrated in the known flexible regions.

The proteins studied in this work are the TRP repressor, calmodulin, calbindin D_{9k}, HIV-1 protease and troponin C. The TRP repressor is known to have domain shear motion [49]; calmodulin and calbindin D_{9k} are known to have domain hinge motion [50,51]; HIV-1 protease and troponin C exhibit sub-domain hinge motion [52,53]. All these motions involve large amplitude torsional angle displacements restricted to a small number of residues, and they are representatives of typical domain or sub-domain motions in proteins. The crystal or NMR structure for one or more conformations of each protein is also available. Therefore, these proteins are suitable candidates for testing the usefulness of normal mode analysis in probing localized motions.

2. Methods

2.1. Structures

The 3D structure of different forms of the five proteins were obtained from the Protein Databank (PDB). Specifically, the holo form of the TRP repressor (1WRP.pdb; Lawson et al., 1988 [49]) was used for this work. Calmodulin of both the open (4CLN.pdb) [54] and the closed conformations (2BBM.pdb) [55] were studied. Three structures of the apo form (1CLB.pdb) [56], half-cadmium saturated state (1CDN.pdb), and fully-calcium saturated state (4ICB.pdb) [58] were examined for calbindin D_{9k}. For HIV-1 protease, both the ligand-free structure (1HHP.pdb) [59] and ligand-bound state (1HIV.pdb) [60] were investigated. Finally, two structures of the metal-free state (1TNW.pdb) [61] and metal-bound half-saturated state (1NCY.pdb) [62] were calculated for troponin C. Each of these structures is optimized by 10,000 iterations of energy minimization via the Powell method in the SYBYL program. The Amber 4.1 force field and Kollman partial charges were selected in the structural optimization. A distance-dependent dielectric constant was applied [63], and the final termination gradient was set at 0.05 kcal/(mol A) for the minimization procedure.

2.2. Normal mode computation

Under the harmonic approximation, internal motions of a biomolecule can be modeled by the following Hamiltonian:

$$\begin{aligned} H_0 = & \sum_{\text{atoms}} \frac{P^2}{2m} + \sum_{\text{bond-stretch}} \frac{1}{2} K_{\text{r}} (r - r_{\text{eq}})^2 \\ & + \sum_{\text{bond-angle}} \frac{1}{2} K_\theta (\theta - \theta_{\text{eq}})^2 + \sum_{\text{bond-rotation}} \frac{1}{2} K_\phi (\phi - \phi_{\text{eq}})^2 \\ & + \sum_{\text{S-bonds}} \frac{1}{2} K_{\text{S}} (r - \langle r \rangle)^2 + \sum_{\text{H-bonds}} \frac{1}{2} K_{\text{H}} (r - \langle r \rangle)^2 \\ & + \sum_{\text{non-bonded}} \frac{1}{2} K_{ij} (r - \langle r_{ij} \rangle)^2 + V_{\text{st}} \end{aligned} \quad (1)$$

The first term is the kinetic energy term. V_{st} is the static part of the Hamiltonian, i.e. the potentials at equilibrium positions. K_{r} and r_{eq} are covalent bond (other than disulfide bond) force constant and equilibrium bond length, respectively; K_θ and θ_{eq} are bond angle bending force constant and equilibrium bond angle, respectively. Since changes in torsional and non-bonded forces are small, K_ϕ and K_{ij} can be expressed as the second derivative of the relevant potential. These parameters are from the AMBER force field [21]. K_{S}, K_{H} and K_{ij} are force constants of disulfide, H-bond and non-bonded interactions between ith and jth atoms and $\langle r \rangle$ is the equilibrium distance between atoms. These force constants can also be derived from the second derivative of the relevant potential [21,64,65]. AMBER [21] is used for the non-bonded Van der Waals and electrostatic interaction

potentials. A distance-dependent dielectric constant is used in the electrostatic potential in consideration of partial shielding effect of water [63]. Morse potential [66], which is a function of the hydrogen bond donor–acceptor distance, is used to represent hydrogen bonding interactions [64]. This potential has been shown to give reasonable descriptions of hydrogen bonding energy and the dynamics of hydrogen bonding disruption in biomolecules [64,67].

The coordinates of the atoms, in the study, proteins were obtained from the respective PDB file. The equation of motion derived from Eq. (1) is:

$$\left(\sum_k b_k^+ \phi_k b_k - \omega_\sigma I\right) q^\sigma = 0 \quad (2)$$

where σ is the index of the set of parameters ω_σ that satisfy this equation. The collective harmonic motions corresponding to each parameter are termed normal modes and ω_σ is the frequency of these modes. Another parameter k specifies the type of harmonic motion (bond stretching, angle bending, bond rotation, disulfide bond motion, hydrogen bonding motion and the motion induced by non-bonded interactions). The ω_σ and q^σ are termed the eigenfrequency and the eigenvector of Eq. (2) in mass-weighted Cartesian (MWC) coordinates, respectively. The variable q^σ describes the amplitude and direction of instantaneous atomic displacements. Variable b_k is a transition matrix (B matrix) of the particular type of harmonic motion from MWC coordinates to internal coordinates. The formulation of the B matrix for both bonded and non-bonded atom–atom interactions can be found in the literature [44,65]. The variable ϕ_k is the matrix of the respective force fields divided by the square roots of the masses. The actual computation of the eigenfrequencies and eigenvectors in Eq. (2) was carried out by using the LAPACK general real symmetric matrix diagonalization routine dsyev [68] on a DEC alpha workstation.

3. Results and discussion

Collective domain motions and sub-domain motions, such as hinge, shear or combinations of both, are known to be involved in low energy conformational changes between closed and open forms of a number of proteins [2]. Structural analysis of the open, closed and half-open conformational states of several proteins showed that the related large conformational change is often associated with motions confined to a small number of torsions of specific flexible regions [2,49,50,54,55]. The rigid secondary structures and compact domains in these proteins are often connected by flexible regions, which allow them to move as quasi-rigid bodies. The initial fluctuational motions concentrated in these flexible regions are expected to be detectable from normal mode analysis that might eventually lead to large amplitude domain or intra-domain motions. As discussed later, specific unique normal modes with motions primarily confined to the flexible region of the known localized motion in the structure of each of the five proteins studied in this work were identified.

3.1. TRP repressor

TRP repressor, a 25 kDa DNA binding protein, tandemly binds to DNA as a dimmer [69]. It regulates three operons involved in tryptophan synthesis. Each of the subunits contains six helices, forming a central core and a reading-head domain [70]. Between the core and reading-head domain, there is an L-tryptophan binding site. The binding of L-tryptophan at this site is required in the recognition of the targeted DNA sequence by this protein. The shear motions in the helix-turn-helix (HTH) reading head domain were found to be responsible for the conformational change necessary to enable the protein to bind to the targeted DNA sequence [49].

Structural comparison between the holo and apo forms of the TRP repressor indicated that the shear motions seem to primarily involve torsional changes in the HTH reading head domain [49]. The α-carbon atoms in some of these residues have been found to shift by up to 4 Å, resulting from separate shear motions of the two helices (helix D and E) in the reading-head domain (from 0.75 to 1.5 Å and 5 to 20°, respectively). Further study of the torsional angle changes between the holo and apo forms of this protein suggests that the two shear motions involve torsional displacements in two separate sections. The shear motion in helix D is associated with movements in the flexible region between GLY64 and MET66, and that in helix E involves displacements in the section between ASN72 and ARG82 [7].

Analysis of the computed normal modes of the holo form of the TRP repressor found two unique modes at frequencies of 88.8 and 96.6 cm^{-1} for shear motion of helix D. They contain relatively large amplitude eigenvectors primarily concentrated in the section between GLY64 and SER67, which approximately overlap the identified flexible region within helix D for large amplitude shear motion [7]. Here, the relatively large amplitude eigenvectors refer to those substantially larger in amplitude than the rest in a particular mode. The mode at 88.8 cm^{-1} is shown in Fig. 1, which indicates that the amplitude of eigenvectors between GLY64 and SER67 is substantially larger than the rest outside the section. The other mode at 96.6 cm^{-1} has similar characteristics. In all figures presented in this paper, the computed eigenvectors are displayed as thick solid lines as in the previous papers [11,32]. For clarification purposes, the magnitudes are amplified to the extent that the largest eigenvectors can be easily recognized. Normal mode eigenvectors are representative of the amplitude and direction of vibrational motions. Thus, the concentration of large amplitude eigenvectors in this section is indicative of concentration of flucuational motions and energy that might lead to the shear motion in helix D.

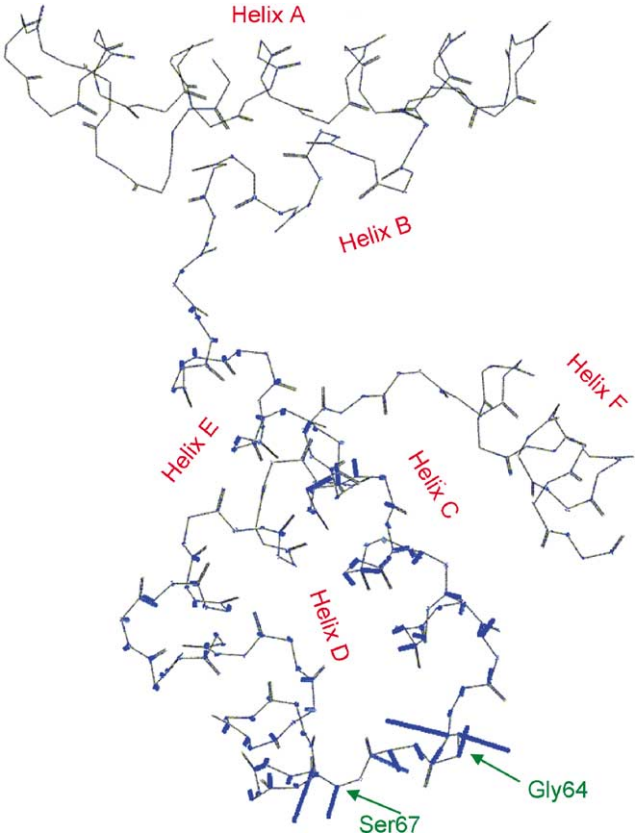

Fig. 1. The normal mode of the holo form of TRP repressor (PDB ID: 1WRP) at the frequency of 88.8 cm^{-1}. The eigenvectors of these mode are shown as the blue thick lines. In order to have a clear view of the eigenvectors, only the main chain of the protein is displayed (gray lines). The region where large amplitude eigenvectors are primarily concentrated is marked by the residue names at both the ends. The same legend is also applied to Figs. 2–11.

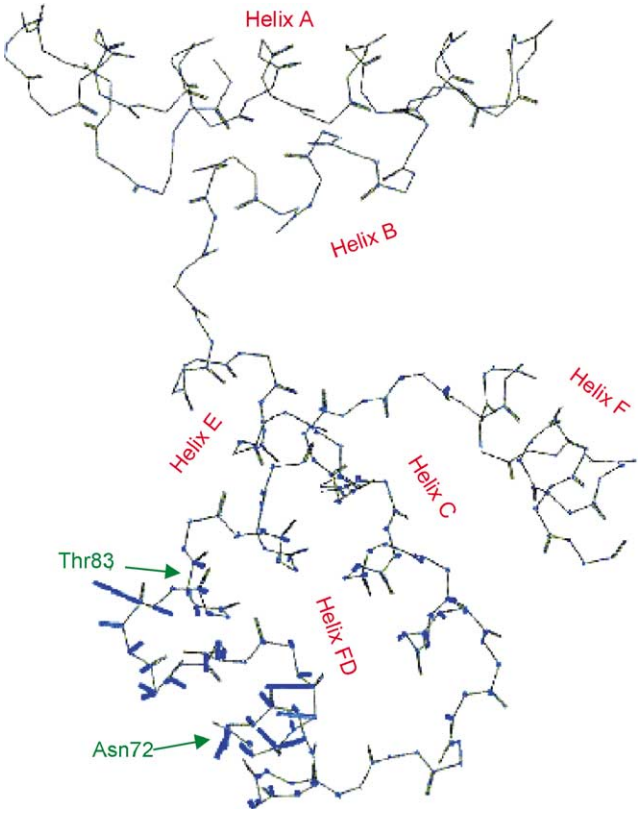

Fig. 2. The normal mode of the holo form of TRP repressor (PDB ID: 1WRP) at the frequency of 85.7 cm^{-1}.

Two unique modes, located at 85.7 and 86.2 cm^{-1}, that might be linked to the shear motion in helix E were identified. One of these modes (at 85.7 cm^{-1}) is shown in Fig. 2 and the others possess similar characteristics. From Fig. 2, it can be determined that the large amplitude eigenvectors are primarily concentrated in the section between ASN72 and THR83, which essentially overlaps the flexible region known to be associated with large amplitude movements for the shear motion in domain E.

3.2. Calmodulin

Calmodulin (CaM) is a highly flexible protein that binds to various proteins. It is also a target of a number of peptide ligands and toxins [50,54,55]. This protein is composed of two domains connected by a central flexible chain, which adopts a helical conformation in the ligand-free open state and a coiled conformation in the ligand-bound closed state. An examination of NMR and X-ray structures revealed that the uncoiling of the central helix in the ligand-free structure enables the two domains to recognize and engulf a target peptide or toxin [50,55]. Analysis of the electron density map of calmodulin showed very low electron density in the section between ALA73 and GLU83, which suggests the origin of the high flexibility of this region. Structural study concluded that the region between MET72 and GLU82 acts as an "expansion joint" to allow the two domains to position themselves differently as different targets bind [50]. Therefore, the flexible region is likely located in the section between MET72 and GLU83.

Three unique modes that might be related to the domain hinge motion are identified for each closed and open form of calmodulin. These identified modes are located at 139.1, 132.7, and 131.6 cm^{-1} for the closed form, and 111.8, 124.1, and 97.8 cm^{-1} for the open conformation, respectively. One such mode at 139.1 cm^{-1} for the closed form is shown in Fig. 3 and, one for the open form at 111.8 cm^{-1} is shown in Fig. 4. The other identified modes share similar characteristics with one of these two modes. From Figs. 3 and 4, it can be noted that the large amplitude eigenvectors are primarily concentrated in the section between Met72 and Ser81 in the closed form, and in the section between MET72 and GLU83 in the open form of this protein. These are within the flexible region, where the large amplitude movements for the domain hinge motion occur.

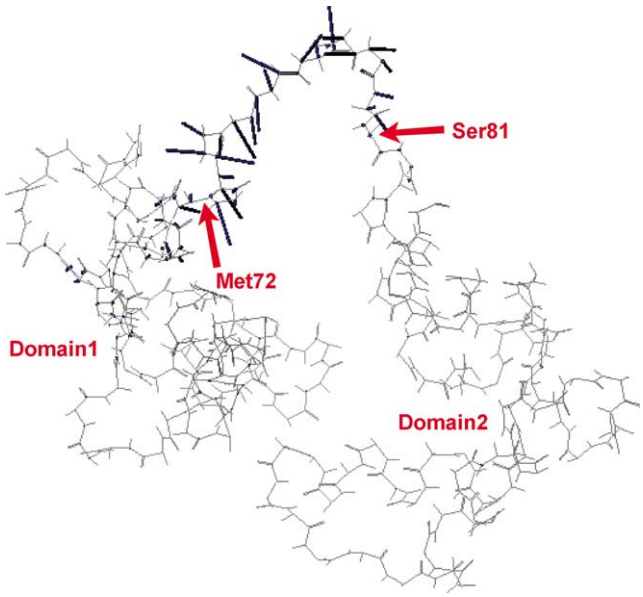

Fig. 3. The normal mode of the closed form of calmodulin (PDB ID: 2BBM) at the frequency of 139.1 cm^{-1}.

3.3. *Calbindin D_{9k}*

Calbindin D_{9k} is a Vitamin D-dependent calcium-binding protein, which is involved in numerous calcium-mediated intra- and extra-cellular processes [71]. Its structure consists of two EF-hands, joined by a flexible linker and a short anti-parallel beta-type interaction between the calcium-binding loops [72]. Examination of the available 3D structure of the apo, half-saturated, and fully-saturated forms of this protein revealed the stepwise conformational changes along the Apo → (Ca^{2+})II1 → (Ca^{2+})I,II2 pathway [57]. Upon binding of calcium ions, the two binding loops in the protein are moved closer to each other by about 1 Å, while the two helices are shifted towards each other by about 0.45 Å [73,74]. The largest change appears to be primarily confined to the middle of helix IV and in the packing of helix III onto the remainder of the protein [72]. Structural analysis of different forms of this protein identified a flexible region between Phe36 and Gly57 which is likely responsible for the observed conformation change [7].

An analysis of the respective normal modes for the apo, cadmium half-saturated, and fully-saturated forms identified several modes with localized motion primarily concentrated in the flexible region and these may be related to the motions leading to the hinge domain motion for apo form. For the apo form, three modes were identified at 92.9, 103.3, and 112.8 cm^{-1}. For the cadmium half-saturated form, three modes were identified at 94.9, 103.3, and 112.7 cm^{-1}. For the fully-saturated form, two modes were identified at 103.8 and 93.5 cm^{-1}. Fig. 5 shows an identified mode for the apo form at 92.9 cm^{-1} whose motions are primarily located in the region between Phe36 and Glu48. A mode for the cadmium half-saturated form at 94.9 cm^{-1} is shown in Fig. 6 in

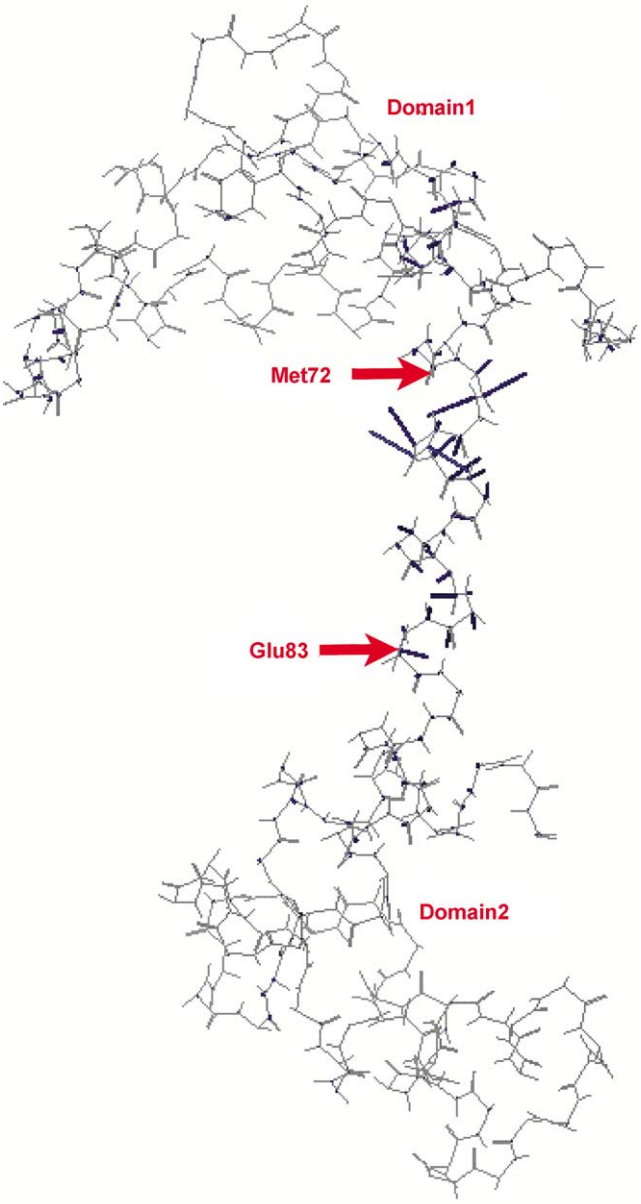

Fig. 4. The normal mode of the open form of calmodulin (PDB ID: 4CLN) at the frequency of 111.8 cm^{-1}.

which large amplitude eigenvectors primarily concentrated in the region between Phe36 and Glu51. Fig. 7 shows an identified mode for the fully-saturated form at 103.8 cm^{-1}, in which the large amplitude motions are concentrated in two sections within the flexible region, one at Gly57 and another between Lys41 and Ser44. The large amplitude motions in these modes are all located within the region where hinge motions occur. Each of the other identified modes, not shown in Figs. 5–7, shares similar characteristics with the corresponding mode shown.

From Figs. 5–7, a trend is apparent in that the number of residues with large amplitude motions decrease from the apo form to the half-saturated form and to the fully-saturated form, which is indicative of reduced flexibility towards the

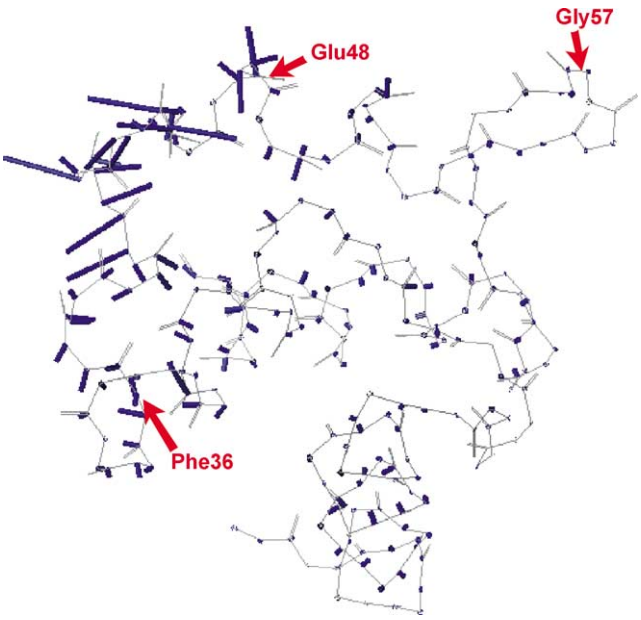

Fig. 5. The normal mode of the apo form for calbindin D_{9k} (PDB ID: 1CLB) at the frequency of 92.9 cm^{-1}.

ligand-bound state and is consistent with the observation that the ligand-bound state is more stable than the ligand-free states [75,76].

3.4. HIV-1 protease

HIV-1 protease (HIVP) is a homodimer consisting of 99 residues and has long been a primary target for anti-AIDs drug development. The two flaps of this protein, which together comprise one quarter of the total structure, are found to be highly flexible, an essential characteristic for its function [77,78]. The function of the two flaps is to control access

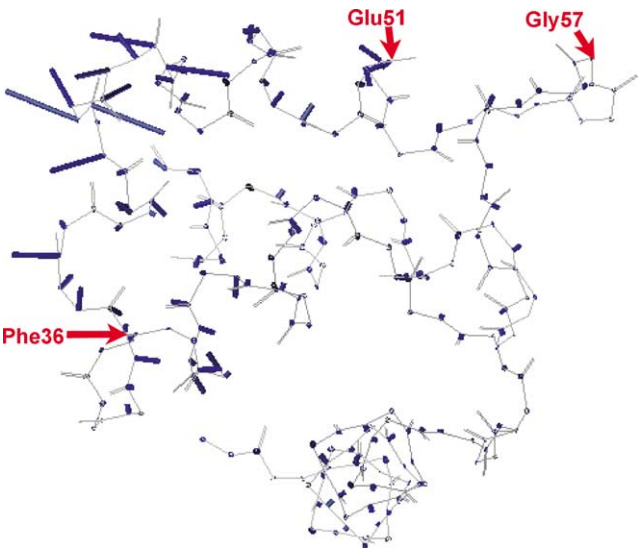

Fig. 6. The normal mode of the half-calcium saturated form of calbindin D_{9k} (PDB ID: 1CDN) at the frequency of 94.9 cm^{-1}.

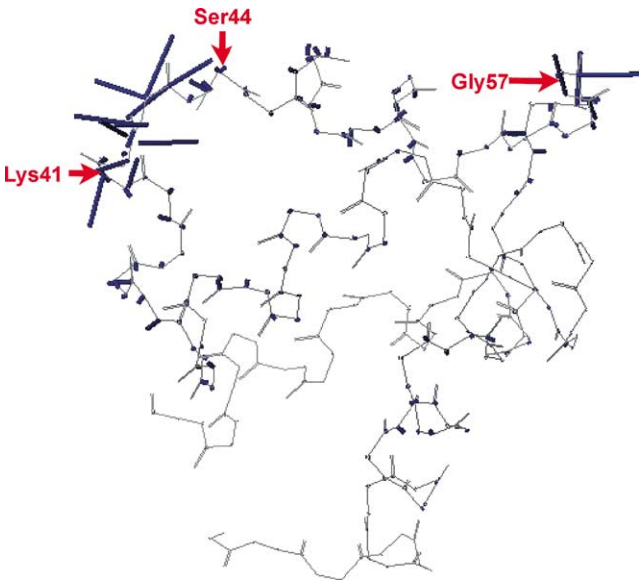

Fig. 7. The normal mode of the fully-calcium saturated state of calbindin D_{9k} (PDB ID: 4ICB) at the frequency of 103.8 cm^{-1}.

to the active site [79] and to exclude water from the binding site [80]. This protein adopts an open form in the native or ligand-free state, and a closed form in the ligand-bound state [81]. The motion in the two flaps induces the conformational change between the two forms. Structural analysis indicated that each flap motion is a fragment hinge motion within the respective chain. The flexible region is located in the section between Lys45 and Gly68 in both A- and B-chains [7].

Our study identifies two modes at 111.5 and 137.3 cm^{-1} for the ligand-free structure, and two modes at 124.9 and 93.6 cm^{-1} for the ligand-bound structures. It was noted that all the available ligand-free structures of HIV-1 protease in PDB are single-chain monomers [82,83]. To examine whether the characteristics of the identified normal modes are significantly affected by the lack of a second chain, we conduct a separate computation of the normal modes of the A-chain in ligand-bound structure with the B-chain completely removed. Our computation shows that the two identified modes concentrated on the same flexible regions are still present. The motions in these modes are very similar to those in the original dimmer structures, the only difference being that the frequency of each mode is shifted to a slightly higher frequency when the B-chain is removed. In this case, the frequency of the two modes became 107.9 and 142.5 cm^{-1}, respectively. Therefore, our study seems to suggest that the lack of the second chain appears to have very limited effect on the characteristics of normal modes in the frequency range of 90 cm^{-1} and above.

Fig. 8 shows the mode at 115.5 cm^{-1} for the ligand-free structure. The large amplitude motions are seen to be concentrated in the section between Gly49 and Lys55 within the flexible region. An identified mode at 124.9 cm^{-1} for the ligand-bound structure is shown in Fig. 9 from which one can see that the large amplitude motions are localized

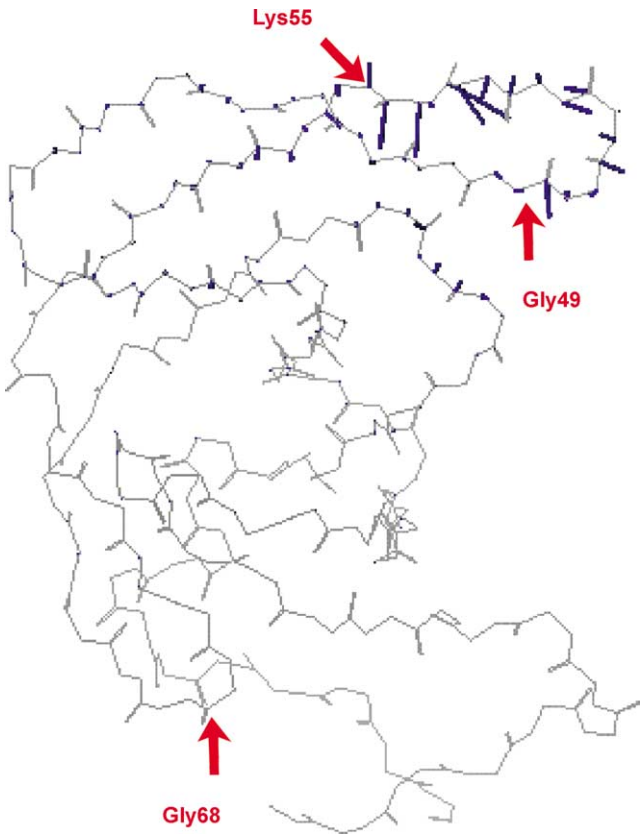

Fig. 8. The normal mode of the apo form of HIV-1 protease (PDB ID: 1HHP) at the frequency of 115.5 cm^{-1}.

in the section between ARG57 and Gly68 within the flexible region in the A-chain. The not shown corresponding mode for the B-chain is at 111.5 cm^{-1} and its large amplitude motions are concentrated in the section between Asp60 and Gly68. These modes show characteristics of "cantilever" motion described in the literature, such that the binding ligand appears to rigidify the flap while make other section of protein more flexible [83,84]. Computations on two other HIV-1 protease structures (3PHV for ligand-free and 1DIF for ligand-bound) indicated that their normal modes show similar behavior to that presented here.

3.5. Troponin C

Troponin C is vital to contraction of skeletal muscles [85]. Similar to calmodulin, the structure is composed of two globular sections (the N- and C- domains) joined by a long linker [86]. The N-domain is comprised of a NAD unit (N-, A- and D-helices) and a BC unit (B- and C-helices). The C-domain consists of four helices of E-, F-, G-, and H-helices. Each domain appears to be structurally independent, which allows the protein to adopt different orientations to interact with a variety of substrates [87]. Following nerve stimulation, the influx of Ca^{2+} irons from the sarcoplasmic reticulum induces muscle contraction, which is triggered by the calcium-induced "opening" of the regulatory N-domain of troponin C [86,87]. It was found that this "opening" leads to the movement of the BC unit away from the NAD unit [73], which presents itself to its target protein troponin I [87–89]. A flexible region between Leu14 and Gly35 is responsible for the "opening" hinge motion [65].

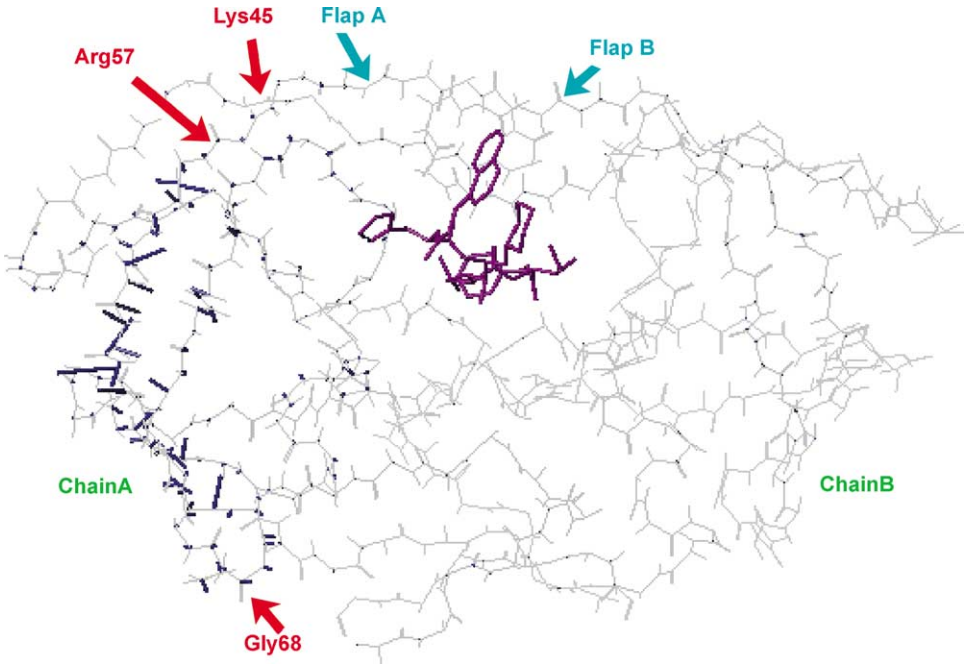

Fig. 9. The normal mode of the ligand-bound state of HIV-1 protease with U75875 (PDB ID: 1HIV) at the frequency of 124.9 cm^{-1}. The ligand U75875 is shown as purple sticks.

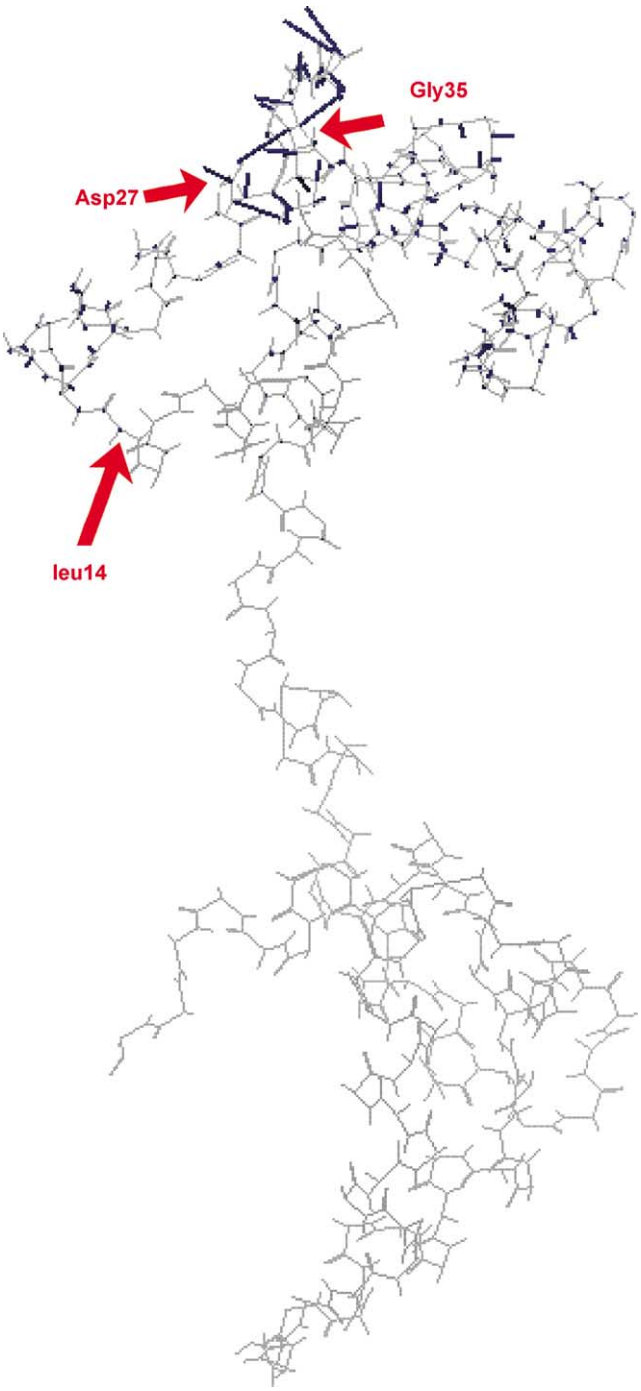

Fig. 10. The normal mode of the metal-free troponin C (PDB ID: 1TNW) at the frequency of 137.8 cm^{-1}.

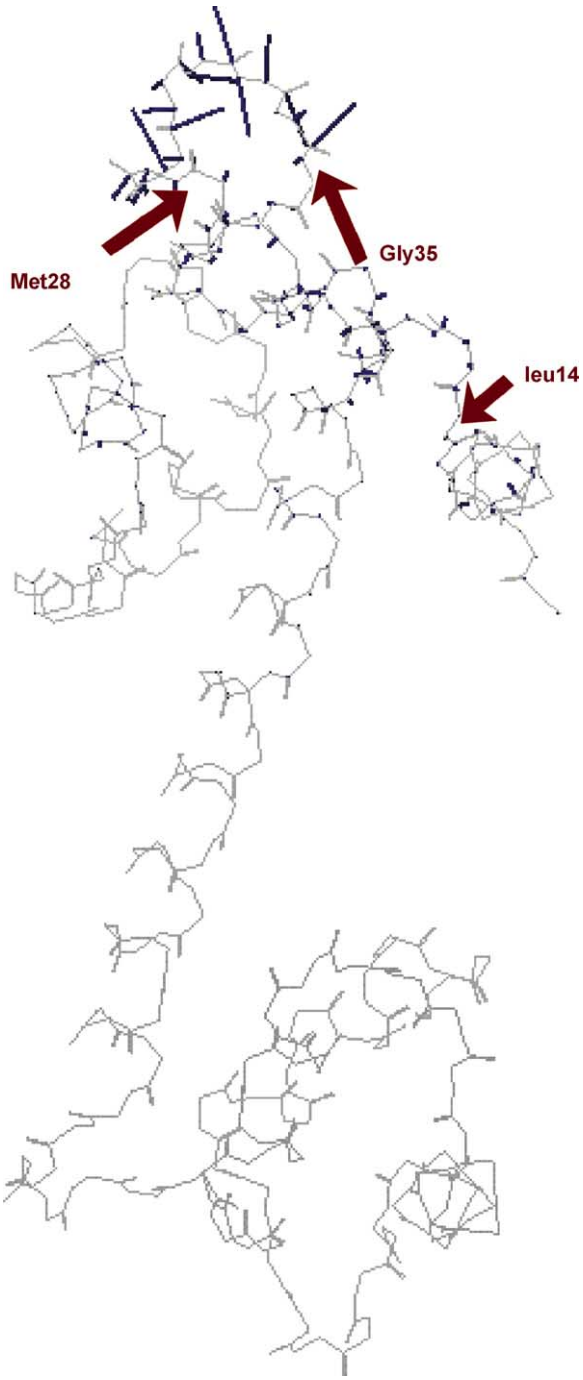

Fig. 11. The normal mode of the metal-bound half-saturated state of Troponin C (PDB ID: 1NCY) at the frequency of 103.7 cm^{-1}.

For the metal-free structure, only one mode, at 137.8 cm^{-1}, with large amplitude motions concentrated in the flexible region was found. On the other hand, for the metal-bound half-saturated state of Troponin C, two such modes at 118.0 and 103.7 cm^{-1} were identified. Fig. 10 shows the mode at 137.8 cm^{-1} for the metal-free structure, in which the large amplitude motions are found to concentrate in the section between Asp27 and Gly35. The mode at 103.7 cm^{-1} for the metal-bound structure is shown in Fig. 11, in which the large amplitude motions are concentrated in the section between Met28 and Gly35. The large amplitude motions in these modes are all located within the flexible region.

4. Concluding remarks

Normal modes contain a rich set of useful information about localized vibrational motions that might lead to

biologically relevant torsional motions. Our study suggests possible correlation between localized collective motions in protein flexible regions and individual vibrational normal mode in the frequency range between 20 and 200 cm^{-1}. In this work, knowledge about the location of the flexible regions related to a specific collective motion is required in the identification of the unique normal modes that might be associated with that motion. In order to fully explore the potential of normal mode analysis in facilitating the study of localized collective motions in biomolecules, a method for analysis of motions from a normal mode without prior knowledge of the location of a flexible region needs to be developed.

Acknowledgements

This work was supported in part by Singapore Academic Research Fund.

References

[1] I. Bahar, B. Erman, T. Haliloglu, R.L. Jernigan, Efficient characterization of collective motions and inter-residue correlations in proteins by low-resolution simulations, Biochemistry 36 (1997) 13512–13523.
[2] M. Gerstein, A.M. Lesk, C. Chothia, Structural mechanisms for domain movements in proteins, Biochemistry 33 (1994) 6739–6749.
[3] M. Zhang, T. Yuan, Molecular mechanisms of calmodulin's functional versatility, Biochem. Cell Biol. 76 (1998) 313–323.
[4] B. Brooks, M. Karplus, Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 6571–6575.
[5] D.J. Patel, S. Ikuta, S. Kozlowski, K. Itakura, Sequence dependence of hydrogen exchange kinetics in DNA duplexes at the individual base pair level in solution, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 2184–2188.
[6] M. Gueron, M. Kochoyan, J.L. Leroy, A single mode of DNA base-pair opening drives imino proton exchange, Nature 328 (1987) 89–92.
[7] M. Gerstein, W.G. Krebs, A database of macromolecular motions, Nucleic Acids Res. 26 (1998) 4280–4290.
[8] Y.Z. Chen, V. Mohan, R.H. Griffey, The opening of a single base without perturbations of neighboring nucleotides: a study on crystal B-DNA duplex d(CGCGAATTCGCG)2, J. Biomol. Struct. Dyn. 15 (1998) 765–777.
[9] D.M. Van Aalten, D.A. Erlanson, G.L. Verdine, L. Joshua-Tor, Nucleotide, structure: a structural snapshot of base-pair opening in DNA, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 11809–11814.
[10] T. Nishikawa, N. Go, Normal modes of vibration in bovine pancreatic trypsin inhibitor and its mechanical property, Proteins 2 (1987) 308–329.
[11] J. Ma, M. Karplus, Ligand-induced conformational changes in *ras p21*: a normal mode and energy minimization analysis, J. Mol. Biol. 274 (1997) 114–131.
[12] S. Hayward, A. Kitao, H.J. Berendsen, Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics of lysozyme, Proteins 27 (1997) 425–437.
[13] S. Jaaskelainen, C.S. Verma, R.E. Hubbard, P. Linko, L.S. Caves, Conformational change in the activation of lipase: an analysis in terms of low-frequency normal modes, Protein Sci. 76 (1998) 1359–1367.
[14] T. Simonson, D. Perahia, Normal modes of symmetric protein assemblies: application to the tobacco mosaic virus protein disk, Biophys. J. 61 (1992) 410–427.
[15] R.P. Sheridan, R.M. Levy, S.W. Englander, Normal mode paths for hydrogen exchange in the peptide ferrichrome, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 5569–5572.
[16] F. Tama, Y.H. Sanejouand, Conformational change of proteins arising from normal-mode calculations, Protein Eng. 14 (2001) 1–6.
[17] K. Hinsen, A. Thomas, M.J. Field, Analysis of domain motions in large proteins, Proteins 34 (1999) 369–382.
[18] D.A. Case, Molecular dynamics and normal mode analysis of biomolecular rigidity. in: M.F. Thorpe, P.M. Duxbury (Eds.), Rigidity Theory and Application. Kluwer Academic/Plenum, New York, 1999.
[19] B. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217.
[20] W.F. Van Gunsteren, Molecular dynamics and stochastic dynamics simulation: a primer. in: Computer Simulation of Biomolecular Systems Theoretical and Experimental Applications, ESCOM, Leiden, 1993.
[21] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[22] A. Kitao, F. Hirata, N. Go, The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulations of melittin in water and vacuum, Chem. Phys. 158 (1991) 447–472.
[23] D. Frenkel, Monte Carlo simulations: a primer. in: Computer Simulation of Biomolecular Systems Theoretical and Experimental Applications, ESCOM, Leiden, 1993.
[24] A. Matsumoto, M. Tomimoto, N. Go, Dynamical structure of transfer RNA studied by normal mode analysis, Eur. Biophys. J. 28 (1999) 369–379.
[25] H. Ishida, Y. Jochi, A. Kidera, Dynamic structure of Subtilisin–Eglin c complex by normal mode analysis, Proteins 32 (1998) 324–333.
[26] M.M. Tirion, Large amplitude elastic motions in proteins from a single-parameter, atomic analysis, Phys. Rev. Lett. 77 (1996) 1905–1908.
[27] M.M. Tirion, D. ben-Avraham, Normal mode analysis of G-actin, J. Mol. Biol. 230 (1993) 186–195.
[28] S. Nakamura, J. Doi, Dynamics of transfer RNAs analyzed by normal-mode calculation, Nucleic Acids Res. 22 (1994) 514–521.
[29] W.M. David, A.A. David, Enzyme specificity under dynamic control: a normal mode analysis of α-lytic protease, J. Mol. Biol. 286 (1999) 267–278.
[30] A. Thomas, M.J. Field, L. Mouawad, D. Perahia, Analysis of low-frequency normal modes of the T-state of aspartate *trans*-carbamylase, J. Mol. Biol. 257 (1996) 1070–1087.
[31] B. Brooks, M. Karplus, Normal mode for specific motions of macromolecules: application to the hinge-bending mode of lysozyme, Proc. Natl. Acad. Sci. U.S.A. 82 (1985) 4995–4999.
[32] M. Levitt, C. Sander, P.S. Stern, Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysosyme, J. Mol. Biol. 181 (1985) 147–423.
[33] K.G. Brown, S.C. Erfurth, E.W. Small, W.L. Peticolas, Conformationally dependent low-frequency motions of proteins by laser Raman spectroscopy, Proc. Nat. Acad. Sci. U.S.A. 69 (1972) 1467–1469.
[34] W.L. Peticolas, Low-frequency vibrations and the dynamics of proteins and polypeptides, Methods Enzymol. 61 (1979) 425–458.
[35] H.D. Bartunik, P. Jolles, J. Berthou, A.J. Dianoux, Intra-molecular low-frequency vibrations in lysozyme by neutron time-of-flight spectroscopy, Biopolymers 21 (1982) 43–50.

[36] J.W. Powell, G.S. Edwards, L. Genzel, F. Kremer, A. Wittlin, W. Kubasek, W. Peticolas, Investigation of far-infrared vibrational modes in polynucleotides, Phys. Rev. A 35 (1987) 3929–3939.
[37] T. Weidlich, S.M. Lindsay, R. Qi, A. Rupprecht, W.L. Peticolas, G.A. Thomas, A Raman study of low-frequency intra-helical modes in A-, B-, and C-DNA, J. Biomol. Struct. Dyn. 8 (1990) 139–171.
[38] G. Edwards, C. Liu, Sequence dependence of low-frequency Raman-active modes in nucleic acids, Phys. Rev. A 44 (1991) 2709–2717.
[39] W.C. Reisdorf, S. Krimm, Infrared amide I' band of coiled coil, Biochemistry 35 (1996) 1383–1386.
[40] Y. Guan, C.J. Wurrey, G.J. Thomas Jr., Vibrational analysis of nucleic acids. I. The phosphodiester group in dimethyl phosphate model compounds: $(CH_3O)_2PO_2-$, $(CD_3O)_2PO_2-$, and $(13CH_3O)_2PO_2-$, Biophys. J. 66 (1994) 225–235.
[41] B. Melchers, E.W. Knapp, F. Parak, L. Cordone, A. Cupane, M. Leone, Structural fluctuations of myoglobin from normal-modes, Mossbauer, Raman, and absorption spectroscopy, Biophys. J. 70 (1996) 2092–2099.
[42] J. Wang, S. Takahashi, D.L. Rousseau, Identification of the overtone of the Fe–CO stretching mode in heme proteins: a probe of the heme active site, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 9402–9406.
[43] J.G. Duguid, V.A. Bloomfield, J.M. Benevides, G.J. Thomas Jr., Raman spectroscopy of DNA–metal complexes. Part II. The thermal denaturation of DNA in the presence of Sr^{2+}, Ba^{2+}, Mg^{2+}, Ca^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, and Cd^{2+}, Biophys. J. 69 (1995) 2623–2641.
[44] Y.Z. Chen, E.W. Prohofsky, Normal-mode calculation of a netropsin–DNA complex: effect of structural deformation on vibrational spectrum, Biopolymers 35 (1995) 657–666.
[45] Y.Z. Chen, E.W. Prohofsky, Sequence and temperature dependence of the inter-base hydrogen bond breathing modes in B-DNA polymers: comparison with low-frequency Raman peaks and their role in helix melting, Biopolymers 35 (1995) 573–582.
[46] S.A. Lee, A. Rupprecht, Y.Z. Chen, Drug binding to DNA: observation of the drug–DNA hydrogen bond-stretching modes of netropsin bound to DNA via Raman spectroscopy, Phys. Rev. Lett. 80 (1998) 2241–2244.
[47] B. Tidor, M. Karplus, The contribution of vibrational entropy to molecular association: the dimerization of insulin, J. Mol. Biol. 238 (1994) 405–414.
[48] T. Miura, G.J. Thomas Jr., Structure and dynamics of inter-strand guanine association in quadruplex telomeric DNA, Biochemistry 34 (1995) 9645–9654.
[49] C.L. Lawson, R. Zhang, R.W. Schevitz, Z. Otwinowski, A. Joachimiak, P.B. Sigler, Flexibility of the DNA-binding domains of the TRP repressor, Proteins 3 (1988) 18–31.
[50] W.E. Meador, A.R. Means, F.A. Quiocho, Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures, Science 262 (1993) 1718–1721.
[51] Z. Jia, D. Barford, A.J. Flint, N.K. Tonks, Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B, Science 268 (1995) 1754–1758.
[52] P.M.D. Fitzgerald, B.M. McKeever, J.F. VanMiddlesworth, J.P. Springer, J.C. Heimbach, C.-T. Leu, W.K. Herber, R.A.F Dixon, P.L. Darke, Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl–pepstatin at 2.0-Å resolution, J. Biol. Chem. 265 (1990) 14209–14219.
[53] R.F. Service, Flexing muscle with just one amino acid, Science 271 (1996) 31–33.
[54] D.A. Taylor, J.S. Sack, J.F. Maune, K. Beckingham, F.A. Quiocho, Structure of a recombinant calmodulin from *Drosophila melanogaster* refined at 2.2-Å resolution, J. Biol. Chem. 266 (1991) 21375–21380.
[55] M. Ikura, G.M. Clore, A.M. Gronenborn, G. Zhu, C.B. Klee, A. Bax, Solution structure of a calmodulin-target peptide complex by multidimensional NMR, Science 256 (1992) 632–638.
[56] N.J. Skelton, J. Kordel, W.J. Chazin, Determination of the solution of Apo calbindin D_{9k} by NMR spectroscopy, J. Mol. Biol. 249 (1995) 441–462.
[57] M. Akke, S. Forsen, W.J. Chazin, Solution structure of (Cd^{2+})-1-calbindin D_{9k} reveals details of the stepwise structural changes along the Apo → (Ca^{2+})II1 → (Ca^{2+})I,II2 binding pathway, J. Mol. Biol. 252 (1995) 102–121.
[58] L.A. Svensson, E. Thulin, S. Forsen, Proline *cis–trans* isomers in calbindin D_{9k} observed by X-ray crystallography, J. Mol. Biol. 223 (1992) 601–606.
[59] S. Spinelli, Q.Z. Liu, P.M. Alzari, P.H. Hirel, R.J. Poljak, The 3D structure of the aspartyl protease from the HIV-1 isolate BRU, Biochimie 73 (1991) 1391–1396.
[60] N. Thanki, J.K. Rao, S.I. Foundling, W.J. Howe, J.B. Moon, J.O. Hui, A.G. Tomasselli, R.L. Heinrikson, S. Thaisrivongs, A. Wlodawer, Crystal structure of a complex of HIV-1 protease with a dihydroxyethylene-containing inhibitor: comparisons with molecular modeling, Protein Sci. 1 (1992) 1061–1072.
[61] C.M. Slupsky, B.D. Sykes, NMR solution structure of calcium-saturated skeletal muscle troponin C, Biochemistry 34 (1995) 15953–15964.
[62] M. Sundaralingam, R. Bergstrom, G. Strasburg, S.T. Rao, P. Roychowdhury, M. Greaser, B.C. Wang, Protein, structure molecular structure of troponin C from chicken skeletal muscle at 3-Å resolution, Science 227 (1985) 945–948.
[63] J.A. McCammon, M. Karplus, Internal motions of antibody molecules, Nature 268 (1977) 765–766.
[64] Y.Z. Chen, E.W. Prohofsky, The role of a minor groove spine of hydration in stabilizing poly(dA)·poly(dT) against fluctuational inter-base H-bond disruption in the premelting temperature regime, Nucleic Acids Res. 20 (1992) 415–419.
[65] Z.W. Cao, Y.Z. Chen, Hydrogen bond disruption probability in proteins by a modified self-consistent harmonic approach, Biopolymers 58 (2001) 319–328.
[66] N.C. Baird, Simulation of hydrogen bonding in biological systems: ab initio calculations for NH_3–NH_3 and NH_3–NH^{4+}, Int. J. Quantum. Chem. Symp. 1 (1974) 49–53.
[67] Y.Z. Chen, E.W. Prohofsky, Premelting base pair opening probability and drug binding constant of a daunomycin—Poly d(GCAT)–Poly d(ATGC) complex, Biophys. J. 66 (1994) 820–826.
[68] E. Anderson, Z. Bai, C. Bischof, C. Demmel, LAPACK User's Guide, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
[69] C.L. Lawson, J. Carey, Tandem binding in crystals of a TRP repressor/operator half-site complex, Nature 366 (1993) 178–182.
[70] R.W. Schevitz, Z. Otwinowski, A. Joachimiak, C.L. Lawson, P.B. Sigler, The 3D structure of TRP repressor, Nature 317 (1985) 782–786.
[71] L.C. McCary, M. Staun, H.F. DeLuca, A characterization of Vitamin D-independent intestinal calcium absorption in the osteopetrotic (op/op) mouse, Arch. Biochem. Biophys. 368 (1999) 249–256.
[72] S. Marchand, B. Roux, Molecular dynamics study of calbindin D_{9k} in the apo, singly and doubly calcium-loaded states, Proteins 33 (1998) 265–284.
[73] B.B. Kragelund, M. Jonsson, G. Bifulco, W.J. Chazin, H. Nilsson, B.E. Finn, S. Linse, Hydrophobic core substitutions in calbindin D_{9k}: effects on Ca^{2+} binding and dissociation, Biochemistry 37 (1998) 8926–8937.
[74] A. Ababou, J.R. Desjarlais, Solvation energetics and conformational change in EF-hand proteins, Protein Sci. 10 (2001) 301–312.
[75] L. Marler, J. Blankenship, M. Rance, W.J. Chazin, Site–site communication in the EF-hand Ca^{2+} binding protein calbindin D_{9k}, Nat. Struct. Boil. 7 (2000) 245–250.
[76] S. Linse, O. Teleman, T. Drakenberg, Ca^{2+} binding to calbindin D_{9k} strongly affects backbone dynamics: measurements of exchange rates of individual amide protons using 1H NMR, Biochemistry 29 (1990) 5925–5934.
[77] D.M. York, T.A. Darden, L.G. Pedersen, M.W. Anderson, Molecular dynamics simulation of HIV-1 protease in a crystalline environment and in solution, Biochemistry 32 (1993) 1443–1453.

[78] J.R. Collins, S.K. Burt, J.W. Erickson, Flap opening in HIV-1 protease simulated by 'activated' molecular dynamics, Nat. Struct. Biol. 2 (1995) 334–338.
[79] A.M. Skalka, Retroviral proteases: first glimpses at the anatomy of a processing machine, Cell 56 (1989) 911–913.
[80] T. Blundell, L. Pearl, Retroviral proteinases: a second front against AIDS, Nature 337 (1989) 596–597.
[81] A.F. Wilderspin, R.J. Sugrue, Alternative native flap conformation revealed by 2.3-Å resolution structure of SIV proteinase, J. Mol. Biol. 239 (1994) 97–103.
[82] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[83] D.J. Jacobs, A.J. Rader, L.A. Kuhn, M.F. Thorpe, Protein flexibility prediction using graph theory, Proteins 44 (2001) 150–165.
[84] W.E. Harte Jr., S. Swaminathan, M.M. Mansuri, J.C. Martin, I.E. Rosenberg, D.L. Beveridge, Domain communication in the dynamical structure of human immunodeficiency virus 1 protease, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 8864–8868.
[85] K.A. Satyshur, S.T. Rao, D. Pyzalska, W. Drendel, M. Greaser, M. Sundaralingam, Refined structure of chicken skeletal muscle troponin C in the two-calcium states at 2-Å resolution, J. Biol. Chem. 263 (1998) 1628–1647.
[86] J. Digel, O. Abugo, T. Kobayashi, Z. Gryczynski, J.R. Lakowicz, J.H. Collins, Calcium- and magnesium-dependent interactions between the C-terminus of troponin I and the N-terminal, regulatory domain of troponin C, Arch. Biochem. Biophys. 387 (2001) 243–249.
[87] S.M. Ngai, R.S. Hodges, Structural and functional studies on troponin I and troponin C interactions, J. Cell Biochem. 83 (2001) 33–46.
[88] H.C. Li, K. Hideg, P.G. Fajer, The mobility of troponin C and troponin I in muscle, J. Mol. Recognit. 10 (1997) 194–201.
[89] M.C. Moncrieffe, S. Eaton, Z. Bajzer, C. Haydock, J.D. Potter, T.M. Laue, F.G. Prendergast, Rotational and translational motion of troponin C, J. Biol. Chem. 274 (1999) 17464–17470.