J Mol Model (2009) 15:959–969
DOI 10.1007/s00894-008-0446-1

ORIGINAL PAPER

Low-frequency vibrational modes and infrared absorbance of red, blue and green opsin

Saravana Prakash Thirumuruganandham · Herbert M. Urbassek

Received: 24 November 2008 / Accepted: 23 December 2008 / Published online: 3 February 2009

Abstract Vibrational excitations of low-frequency collective modes are essential for functionally important conformational transitions in proteins. We carried out an analysis of the low-frequency modes in the G protein coupled receptors (GPCR) family of cone opsins based on both normal-mode analysis and molecular dynamics (MD) simulations. Power spectra obtained by MD can be compared directly with normal modes. In agreement with existing experimental evidence related to transmembrane proteins, cone opsins have functionally important transitions that correspond to approximately 950 modes and are found below 80 cm^{-1}. This is in contrast to bacteriorhodopsin and rhodopsin, where the important low-frequency transition modes are below 50 cm^{-1}. We find that the density of states (DOS) profile of blue opsin in a solvent (e.g. water) has increased populations in the very lowest frequency modes (<15 cm^{-1}); this is indicative of the increased thermostability of blue opsin. From our work we found that, although light absorption behaves differently in blue, green and red opsins, their low-frequency vibrational motions are similar. The similarities and differences in the domain motions of blue, red and green opsins are discussed for several representative modes. In addition, the influence of the presence of a solvent is reported and compared with vacuum spectra. We thus demonstrate that terahertz spectroscopy of low-frequency modes might be relevant for identifying those vibrational degrees of freedom that correlate to known conformational changes in opsins.

S. P. Thirumuruganandham · H. M. Urbassek (✉)
Fachbereich Physik und Forschungszentrum OPTIMAS,
Universität Kaiserslautern,
Erwin-Schrödinger-Straße,
67663 Kaiserslautern, Germany
e-mail: urbassek@rhrk.uni-kl.de
URL: http://www.physik.uni-kl.de/urbassek

Keywords Normal modes · Molecular dynamics · Vibrational modes · Proteins · Infrared absorbance

PACS numbers 36.20.Ng · 87.15.M- · 87.14.ep · 87.15.hp

Introduction

The nature of the internal motions of proteins is a subject of considerable interest, particularly because some of these motions are known to play an important role in protein functions such as protein dynamics and structural transitions [1]. The frequencies of the $3N-6$ vibrational degrees of freedom existing in a protein are broadly distributed from approximately 3,700 cm^{-1} (O–H stretching) to a few per centimetre in the mid- to far-infrared regions [2, 3]. Stretching vibrations of the O–H, N–H and C–H bonds of a protein contribute high frequencies due to their strong force constants and small masses, whereas low-frequency modes occur due to weak force-constant groups (such as methyl groups) as well as global collective motions of a large number of atoms.

These latter low-frequency modes in particular manifest as bending, deformation, and twisting motions involving changes in bond angles [1, 4]. The importance of these collective mode dynamics comes from two fundamental aspects of protein action: (1) low-frequency modes of proteins are related to functional properties, and (2) it is believed that low-frequency collective modes are responsible for the direct flow of conformational energy in many biological processes. Conformational changes can be achieved by activating these modes to cross low-energy barriers by ligand binding or environmental changes [5, 6].

On the experimental side, a number of techniques have been employed to study the large-scale motions of

biomolecules; here we mention small-angle X-ray scattering, [7] far-infrared (FIR) absorbance spectra [7–17], and FIR transmission spectra [5, 18]. Also, ultrafast OHD-RIKES spectroscopy has been applied to investigate low-frequency motions of biological systems [19–22].

Since the early 1980s, both normal-mode analysis (NMA) and molecular dynamics (MD) simulations have proved to be useful for the theoretical study of the collective motions of biological molecules. Recent investigations on low-frequency spectra and mode shapes obtained by quasi-harmonic and normal-mode analysis include an investigation of the conformational heterogeneity in the enzyme dihydrofolate reductase (DHFR), [23] the low-frequency vibrational spectra of amino acids, [24] and a study of the four structurally similar but functionally different proteins—CysB, lysine/arginine/ornithine-binding protein (LAO), the enzyme porphobilinogen deaminase (PBGD), and ribose-binding protein (RBP)—representative of four distinct classes [25]. Applications of MD simulations and other computational techniques for analysing the dynamics of biomolecules with a focus on areas most closely connected to IR spectroscopy [26, 27] included the calculation of the spectral density [28] and the energy flow in biomolecules, [22] the analysis of biomolecules–water interactions, and water–water interactions near a biomolecule, as well as investigations on biomolecular geometry [29].

The aim of the present work was to investigate the lowest-frequency vibrational motions of three geometrically similar proteins, viz., red, blue and green opsin. These opsins serve as photoreceptors in the human eye and are specialised nerve cells for light reception and transduction. They are coupled to G proteins, which signal through the cAMP-PKA-CREB pathway and hence are also called G protein coupled receptors (GPCR). In total, the amino acid sequence identity between green and red opsin amounts to 96%, while the changes between blue and red or green opsin are larger, allowing only for 43% identity [30].

The availability of high-resolution crystal structures of cone opsin, and advances in computational methodology provide us with the opportunity to examine the details of their low frequency modes. We are motivated by recent work on bacteriorhodopsin (bR) and rhodopsin (Rh) [10, 31–33], which investigated the low-frequency spectra of these two proteins and analysed their similarities and differences. Because of their biological and physiological importance, we investigated the low-frequency vibrational motions of blue, red and green opsin using normal-mode analysis as well as FIR absorbance; we concentrate on the spectral region between 0 and 80 cm^{-1}.

Methods

Molecules

The structures of the three opsins were taken from the X-ray crystal structure data obtained by Stenkamp et al. [RCSB PDB entries 1kpn (blue), 1kpx (green) and 1kpw (red)] [34, 35]; these served as the starting structures for setting up our simulations (Fig. 1). The program CHARMM33 with the CHARMM22 parameter set [36, 37] was used for all our simulations. Opsin-retinal force-field parameters were taken from the stream file toppar_all22_prot_retinol.str. [38]

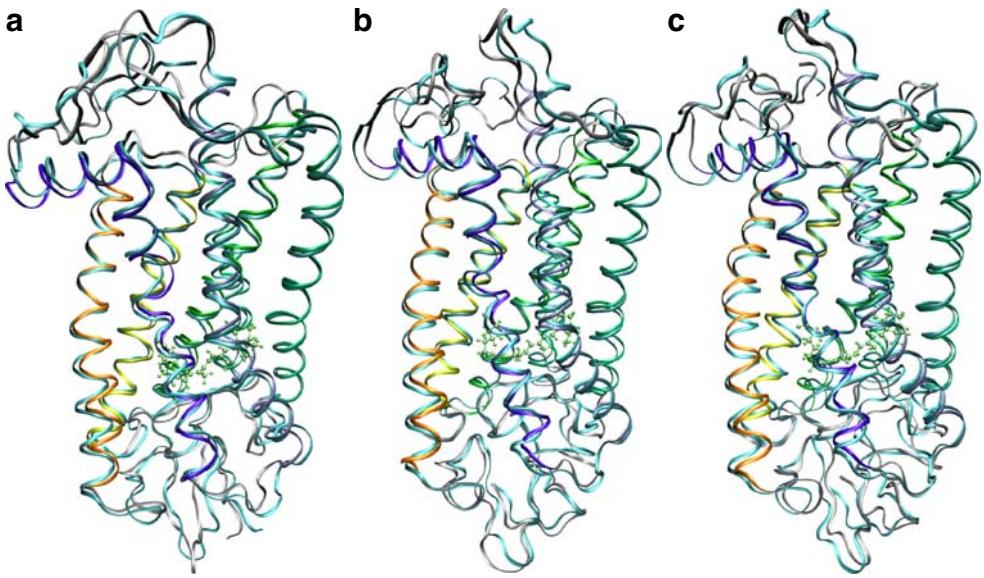

Fig. 1 Overlay of molecular structures of **a** blue (lysine 293 bound to retinal), **b** green (lysine 312 bound to retinal), and **c** red (lysine 312 bound to retinal) opsin before (*cyan*) and after minimisation. Helices: *Orange* I, *yellow* II, *light green* III, *green* IV–V, *ice blue* VI, *violet* VII–VIII. *Silver* Cytoplasmic and extracellular loops, *lime* retinal

Simulation

Normal-mode analysis

Normal-mode calculations were performed to analyse the normal-mode density, atomic fluctuations and FTIR absorbance spectra of blue (1 kpn), green (1 kpw) and red (1 kpx) opsins. The crystal structures of these proteins were generally inappropriate for direct harmonic analysis due to poor nonbonding contacts and unoptimised internal geometries. To apply a normal-mode calculation, it is necessary to first relax the molecule to an energy minimum. For all structures, the mass-weighted Hessian was diagonalised to obtain the harmonic normal-mode frequencies. The resulting eigenvalues from this diagonalisation were all greater than or nearly equal to zero, indicating that the structures were sufficiently energy minimised to have reached a stable minimum. A series of energy minimisations was carried out to satisfy this requirement. This was done using cycles of steepest-descent and the adopted basis Newton-Raphson algorithms [36, 39]. Harmonic constraints were applied on the protein backbones (25 kcal mol^{-1} $Å^{-2}$), and were progressively reduced at each cycle (15, 10, 5 kcal mol^{-1} $Å^{-2}$); finally, the constraints were removed and free minimisation was carried out for an additional 20,000 iteration steps using the adopted basis Newton- Raphson algorithm, until the average gradient of the potential energy reached 10^{-6} kcal mol^{-1} $Å^{-1}$. The nonbonded interactions were truncated at 13 Å and the van der Waals interactions were switched off between 10 Å and 12 Å. Normal modes were obtained using iterative diagonalisation in a mixed basis set (DIMB), previously developed for large molecular systems [23, 40–42]. This method has previously been applied to haemoglobin and phosphoglycerate kinase [40]. It uses an iterative procedure that converges from a set of initial trial modes to the refined lowest-frequency modes within a given error tolerance.

FIR absorbance

We calculated theoretical terahertz FIR absorbance for all three proteins based on the absorbance of the fundamental vibrational transition $\upsilon'=\upsilon''+1$ of each normal mode, Q_i, as proportional to the square of the transition dipole moment derivative, $\delta\mu_i/\delta Q_i$. As is common, we do not include anharmonicities in this approach. The absorbance is numerically evaluated according to [43–46]:

$$I_i = \left|\left\langle \upsilon' \left| \frac{\delta\mu_i}{\delta Q_i} \right| \upsilon'' \right\rangle\right|^2 \cong \left| \sum_{j=1}^{3N} \frac{e_j^0 \left(q_{i,j} - q_{i,j}^0\right)}{\Delta Q_i} \right|^2, \quad (1)$$

where the sum is over all $3N$ cartesian coordinate displacements, $q_j - q_j^0$, derived from the eigenvectors of the Hessian by dividing by the square root of the atomic masses, and ΔQ_i represents the displacement of atoms in mode Q_i, arbitrarily set to 1 in the approximation of Eq. 1. Since CHARMM assigns a static partial charge, e_j^0, according to atom type regardless of bond length, the electrical anharmonicity is by definition zero, making this approximation necessary. Equation 1 yields the integrated intensity and, therefore, for comparison with experimental data, we used a Lorentzian function to describe each line shape, as given by the following formula:

$$I(\nu) = I_i \frac{\Gamma^2}{\left[\pi(\nu - \nu_i)^2 + \Gamma^2\right]}. \quad (2)$$

Here, I_i is taken from Eq. 1, and Γ is defined as the full width at half maximum (FWHM). In the present case, the lines are broadened uniformly to a width of $\Gamma=0.2$ cm^{-1}; no anharmonicities are included [7, 44, 45, 47].

Root-mean-square fluctuations

From the calculated eigenfrequencies and eigenvectors, statistical measures of motion can be derived. These measures are useful in identifying which residues move the most and which remain stationary. At temperature T, the fluctuations of atom j (in one Cartesian direction) are given in the harmonic approximation by

$$\left\langle \Delta x_j^2 \right\rangle = \frac{k_B T}{M_j} \sum_{i=1}^{3N-6} \frac{y_{ji}}{\omega_i^2}, \quad (3)$$

where k_B is Boltzmann's constant, i spans the modes of interest and y_{ji} is the jth component in the eigenvector corresponding to mode i [48–50].

MD simulation: protein in vacuum

These simulations were performed by starting from the minimised structure of all three opsins. Subsequently, the system was heated to the simulation temperature of 300 K in steps of 5 K over 3,000 integration time steps of 1 fs. In the ensuing equilibration phase, the particle velocities were rescaled to bring the temperature back to 300 K, whenever it deviated by more than 10 K from this value. This was continued until such rescaling steps were no longer necessary, which was after 100 ps. Finally, the production run, in which the system configuration was stored every 10 fs over a total simulated time of 500 ps, was started. The cutoff distance for the nonbonded list was set to 13 Å, while the cutoff distances for the van der Waals smoothing function were 10 Å at the start and 12.5 Å at the end. These

values were found to be an appropriate compromise to increase the accuracy of the simulation without greatly affecting the computational effort per simulation step. To allow a simulation time step of 1 fs, the bonds between the hydrogen molecules and the heavy atoms were kept rigid using the SHAKE algorithm. The rate of sampling corresponds to a maximum frequency of 3,335 cm^{-1}. Only a few bond-stretching vibrations occur above this frequency. Since we are interested in the low-frequency heavy-atom motions, this rate of sampling is adequate. The last 20,000 history points (approximately 200 ps) of the simulation were used for the analysis [32]. We chose the length of our MD simulation runs to be sufficiently long to identify the lowest-frequency eigenmodes. For longer runs, the velocity autocorrelation, and hence the vibrational spectra, might possibly be subject to further changes as a consequence of slow protein conformation changes induced by the finite temperature of 300 K.

MD simulation: protein in water

In order to immerse the opsin proteins in water, we made use of the CHARMM crystal facility, employing 25,000 TIP3 water molecules in a cubic box of side length 94.28 Å with periodic boundary conditions. Water molecules with their oxygen atom closer than 2.8 Å to a protein were discarded. The water molecules that remain in the system are now in a perturbed environment, i.e. surrounding a protein molecule. Therefore, before starting the simulation of the protein–water system, a short energy minimisation using harmonically reduced force constraints on the protein as well as steepest descent and adopted basis Newton-Raphson steps was performed; this helps to remove any strong van der Waals contacts between the water molecules and the protein.

All simulation systems were equilibrated with the following protocol. In the first 50 ps, the system was heated to the working temperature of 300 K using velocity scaling. For the next 50–250 ps, a proper equilibration was carried out, in which the pressure was kept constant using the extended-pressure algorithm. Finally, a production phase of 500 ps followed. Here, an (NpT) ensemble (temperature T=300 K, pressure p=1 atm) was realised using the extended-pressure algorithm implemented in a leap-frog Verlet integrator with default Nose-Hoover thermostat. In addition, long-range electrostatic interactions were modelled with the particle-mesh Ewald (PME) method. The width of the Gaussian distribution K used in this method was set to 0.34. The grid size of the PME method was $108\times108\times108$ $Å^3$ and a cut off of 12.0 Å was employed. During the simulation, the SHAKE algorithm was used to keep all bond lengths fixed at their equilibrium values. The integration time step was 1 fs; only every 50th step was stored in a trajectory file for further evaluation.

Power spectra

To analyse the influence of a solvent environment on the internal dynamics of opsins, we calculated the vibrational density of states (DOS; power spectrum) as a function of frequency for both the opsin in vacuum and in water. This power spectrum $S(\nu)$ can be calculated from the Fourier transform $\widetilde{C}_{\upsilon\upsilon}(\nu)$ of the velocity auto-correlation function $C_{\upsilon\upsilon}(t)$ via

$$S(\nu) = \frac{2}{k_B T} \sum_{j=1}^{3N} M_j \widetilde{C}_{\upsilon\upsilon}(\nu), \tag{4}$$

where M_j is the mass of atom j. The velocity auto-correlation function is calculated from a MD simulation as

$$C_{\upsilon\upsilon}(t) = \frac{\left\langle \vec{\upsilon}_i(t+\tau) \cdot \vec{\upsilon}_i(\tau) \right\rangle}{\left\langle \vec{\upsilon}_i(\tau) \cdot \vec{\upsilon}_i(\tau) \right\rangle}, \tag{5}$$

where $\vec{\upsilon}_i(t)$ is the velocity vector of atom i in a protein molecule at time t and the angular brackets denote averaging over all atoms of this type in the system and over initial times. The Fourier transform of the velocity auto-correlation function provides a direct measure of the vibrational DOS as a function of the frequency [51]. Sampling of the velocities was done as frequently as every 1 fs to obtain the high-frequency modes, and less often to obtain the low-frequency modes.

Results

Normal-mode analysis

Figure 1 illustrates an overlay of the energy-minimised structure of each protein with the corresponding crystallographic structure. The root-mean-square (RMS) deviation of each structure, compared with the corresponding X-ray crystallographic structure, is 2.25 Å, 2.49 Å and 2.43 Å for blue, green and red opsin, respectively. The relaxed structure is indeed quite close to the homology model; we share this finding with Ref. [52]. Since the amplitude of a mode is inversely proportional to the frequency, the most significant motion is produced by the low-frequency modes. The conformational motion, which is of importance for biological events, is expected to involve small energetic changes. Thus, we have concentrated on the low-frequency motion of the proteins as determined from normal-mode dynamics, and extracted several characteristic modes.

In the recent past, the structure, dynamics and energetics of bR have been studied extensively by MD; [53–56] the conformational modes of bR have also been studied experimentally using inelastic neutron scattering [57, 58]. The lowest few normal modes of bR wild type (WT) and

the D96N mutant have been obtained using CHARMM-based FIR spectral measurements starting at a frequency of 10 cm^{-1} [10]. Recently, the sliding-block iterative-diagonalisation method [32] was applied to study the same system; here it was found that the lowest frequency is at 1.442 cm^{-1}; the corresponding motion is almost completely localised on the loop that connects helices B and C, with particular involvement of five residues (68–72) GLY-GLY-GLU-GLN-ASN. Theoretical terahertz spectral studies on the same samples found even smaller frequencies of −0.02 cm^{-1} and −0.11 cm^{-1} for WT and the D96N mutant, respectively [31]. Although the photoactivation path is different in Rh and bR, their vibrational spectra share significant similarities [48]. In addition to the studies on Rh, investigations on homology modelling are able to explain the sequence similarities between bovine rhodopsin and the coloured opsins; these similarities have been found to amount to around 40% [35].

Figure 2 presents our data on the vibrational spectra of the cone opsins. In particular, the frequencies of the normal modes up to 80 cm^{-1} (850 modes) are plotted in Fig. 2a for blue opsin and compared to the other opsins in Fig. 2b. There were no negative frequencies in the calculation, and the six modes corresponding to translational and rotational motion were all present. In Fig. 2a, we also plotted the bR spectrum for comparison; we note that our result is in close agreement with literature data [33]. Apart from the difference in mode density—blue opsin consists of 348 amino acids, while bR only of 222—the opsin spectrum looks qualitatively similar to the bR spectrum. This is understandable from the equivalent geometrical form of these two proteins. We note that the frequency distribution of opsins as calculated by MD will be shown in Fig. 6a for blue opsin and in Fig. 6b for comparison of all three opsins, as discussed below.

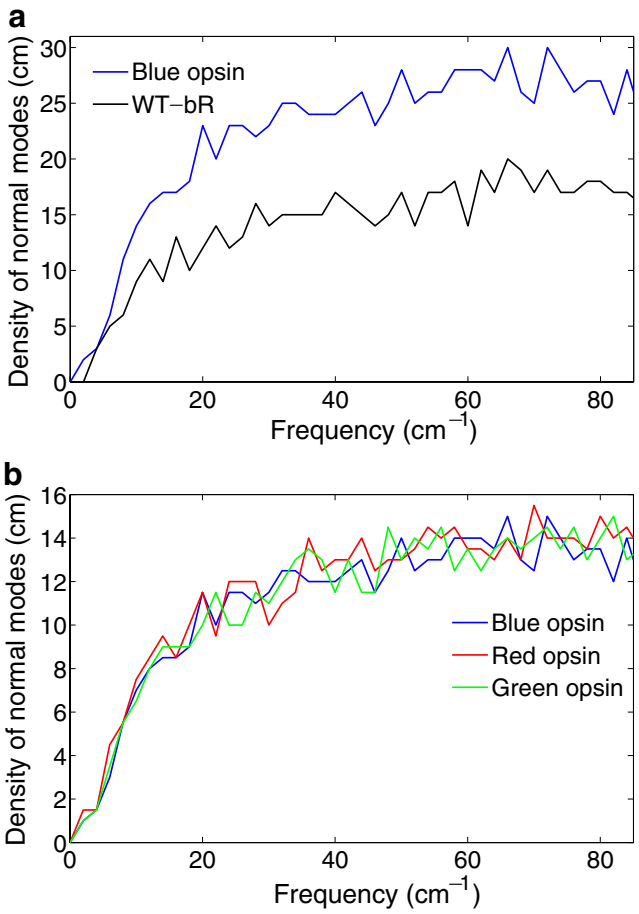

Fig. 2 Vibrational density of states (DOS) as calculated by normal-mode analysis (NMA) for blue opsin and bR (**a**) and for the three opsins studied here (**b**)

For the three opsins, there are about 880 modes with frequencies below 80 cm^{-1}, which constitute about 5.4%, 5.1%, and 5.3% of the total number of vibrational modes of blue (16,623 modes), green (17,259), and red (17,265) opsins, respectively; these low-frequency modes are known to describe the displacements of the backbone atoms, which are most important for describing large-scale conformational changes. The calculated spectra are very similar and have the same overall frequency dependence for the three opsins: for all the three opsins we can see an initial sharp increase up to around 25 cm^{-1}, followed by a somewhat slower increase. In the low-frequency region (0–80 cm^{-1}) of blue opsin, a maximum mode density of 14/cm^{-1} appears near the frequency 65 cm^{-1}. Beyond this frequency, the density of the modes stays roughly constant with little fluctuation. A similar behaviour was found in the green and red opsins. In normal mode, with frequencies higher than 65 cm^{-1} (~180 modes), movements of atoms in the protein molecule are localised to one or a few side chains as well as in several sets of residues, cf. also Fig. 4b below; the displacement vectors of atoms close to each other are similar, and their spatial changes are continuous [6, 59]. At lower frequencies, the protein molecule behaves rather like a nonlocal and continuous elastic body. For all our systems, there are about 474 modes with frequencies from 2 to 55 cm^{-1}; in these, the atomic motions are delocalised-i.e., the majority of atoms in the molecules move in a concerted way.

We note that previous studies [31, 60] have indicated that particularly the low-frequency vibrational structure of proteins will change if metastable structures (i.e. not fully minimised structures, or local minima different from the global minimum) are analysed. The origin of these changes was attributed to the hydrogen bonding network of the protein, which is more compact for the fully minimised structure. In our study, we did not investigate the effect of not-fully minimised structures, and consider only our fully energy-minimised opsin molecules.

From analysis of the RMS displacements calculated by Eq. 3, we found that their main contribution stems from low-frequency modes. Figure 3 displays the RMS fluctuations averaged over all atoms in each mode. It shows which modes contribute most significantly to the atomic fluctuations, and thus presents a measure of the overall mobility of the three opsins. It can be seen that RMS fluctuations are maximum for the first few modes (below around 5 cm^{-1}) indicating that low-frequency modes induce most of the large-scale motion. Conversely, high-frequency modes result in very small atomic fluctuations [45]. From the contributions to the mean-square displacement of all atoms from low-frequency vibrational modes, we see clearly that the very-low-frequency wave numbers smaller than 35 cm^{-1} exhibit dominant contributions. As shown from the NMA, these modes generally involve nonlocalised (collective) motions of the molecule. Even though it is likely that these modes are coupled to each other non-linearly, the concept of collective vibrational variables corresponding to the very-low-frequency vibrational modes is expected to be useful. The dominant contribution mentioned above indicates that the motion in these collective variables determines the basic dynamic structure of the opsins. It is of interest to examine the form of the normal modes themselves. This is of particular importance for evaluation of the correlation between the motions of different atoms and different groups of atoms. A closer analysis of the vibrational dynamics indicates that the large-scale motions have a collective character that may involve several neighbouring atoms, a residue, or even groups of many atoms in a given region of the opsin [23, 61]. The magnitude of the displacement of atoms is determined mainly by modes with frequencies lower than 35 cm^{-1}, and there are about 320 modes in this range.

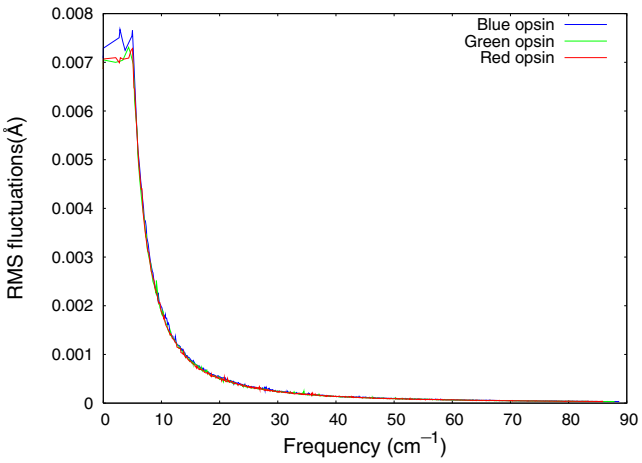

Fig. 3 Average root-mean-square (RMS) fluctuations of all atoms of the three opsins studied as a function of vibrational frequency of the normal modes at 300 K (Eq. 3)

Figure 4 displays a vector representation of the normal-mode motion of the three opsins. We selected two modes, the second eigenmode at roughly 3 cm^{-1}, and mode number 589 at around 60 cm^{-1}. Mode number 2 is typical of very-low-frequency motion: it shows a fully delocalised motion encompassing the entire molecule. While mode number 1 (not shown here, since its motion is mainly perpendicular to the drawing plane) shows a twisting motion in which the upper half of the opsin turns clockwise and the lower half anti-clockwise. Mode 2 (Fig. 4a) shows a torsional motion, in which the upper and lower half of the molecule rotate in cyclonal (and anti-cyclonal) motion around an axis perpendicular to the drawing plane. Figure 4b presents the vector presentation of a (rather arbitrarily selected) higher mode; we chose mode number 589 with a frequency of around 60 cm^{-1}. Here, the motion has become more confined to the cytoplasmic region; both the cytoplasmic side of the helices and the cytoplasmic loops are engaged in stronger motion while the extracellular loops show only little engagement. Thus, in summary, Fig. 4 exemplifies the amount of agreement in the mode patterns and in the mode frequencies that exists between the three different opsins. Clearly, the agreement is not complete; however, we found it impossible to assign any dissimilarities found in the vibration patterns to deviations in the opsin sequences.

From a further analysis of the mode structure, we found that the general normal-mode motion is similar for the blue, green and red opsins. For example, in the lowest terahertz region analysed, 2–14 cm^{-1}, vibrational motion is confined mainly to the cytoplasmic loop and in up- and down-ward motion of the helix, while the extracellular regions exhibit only little motion. In the mid-terahertz region, 22–40 cm^{-1}, the motion is dominant in all-helix vibrations. Here, different combinations, like inward–outward to the membrane and protein interior, and up–downward to the extracellular and cytoplasmic region, can be discerned. At higher frequencies, 40–65 cm^{-1}, the movement slowly passes towards the extracellular region. Finally, at the highest frequencies studied here, 65–85 cm^{-1}, we observe only small-amplitude motion over all the protein; this demonstrates that these modes all contribute to localized motion.

In order to obtain a quantitative measure of the engagement of the various domains of the proteins in a vibrational mode, for each C_α atom we calculated the normalised squared atomic displacements [62, 63],

$$D_i = \frac{\vec{d}_i^{\,2}}{\Sigma_j \vec{d}_j^{\,2}}. \qquad (6)$$

Here, $\vec{d}_i$ is the component of the eigenvector corresponding to the *i*th C_α atom. The normalised atomic displacements for mode numbers 2 and 589 are represented

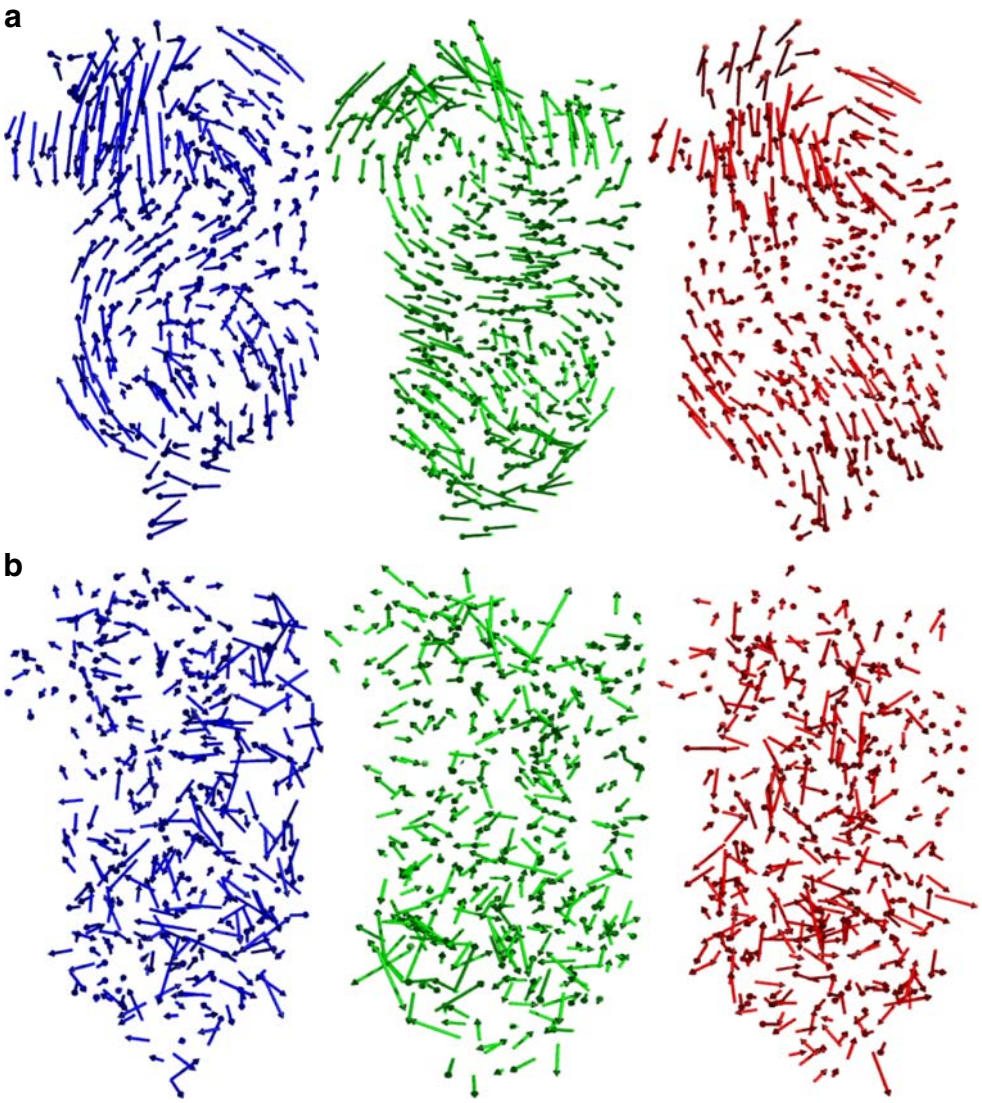

Fig. 4 a–c Vector representation of normal mode motions for the three opsins. **a** Mode 2: Blue opsin 3.30 cm^{-1}, green opsin 2.92 cm^{-1}, red opsin 2.77 cm^{-1}. **b** Mode 589: Blue opsin 59.12 cm^{-1}, green opsin 58.89 cm^{-1}, red opsin 57.13 cm^{-1}

in Fig. 5 as a function of the amino acid sequence. The highest peaks in the plots of Fig. 5 correspond to the most displaced residues in the corresponding mode.

In mode number 2, both the eigenvector representation (Fig. 4) and the atomic displacements show a strong concentration of the vibrational activity in the cytoplasmic part of the opsins, as evidenced by the peaks visible in the cytoplasmic loops between helices V and VI, and beyond helix VII. This feature is found for all three proteins. The remaining vibrational activity is quite uniformly distributed among the other helices. In mode number 589, in contrast, stronger differences between the three opsins show up. The green opsin is characterised by strong motion localised in the cytoplasmic loops between helices V and VI; here, no or only little activity is found for the two other opsins. Red opsin shows strongest activity between helices IV and V. Several other distinctive features may be read off Fig. 5b. Comparing Fig. 5a and b, we see that the vibrational activity is shifted from the cytoplasmic to the extracellular part of the proteins. We conclude that this plot quantifies the idea that, with increasing frequency, vibrational modes become more and more specific to the amino acid sequence of the various opsins.

Infrared absorbance

The study of opsins has become an area of considerable interest in biochemical research seeking information about the dynamics and function of these proteins [3]. There are three main underlying reasons: (1) these proteins are unusually stable. (2) They exhibit strong spectral shifts in the 400–600 nm region, which are connected to reaction intermediates of the different opsins. (3) It is possible to measure vibrational spectra, characteristic geometries as well as protonated states [32]. References [35, 64–66] present experimentally measured FIR spectra; however, these data were taken at higher frequencies. To our knowledge, no experimental data are available on the low-

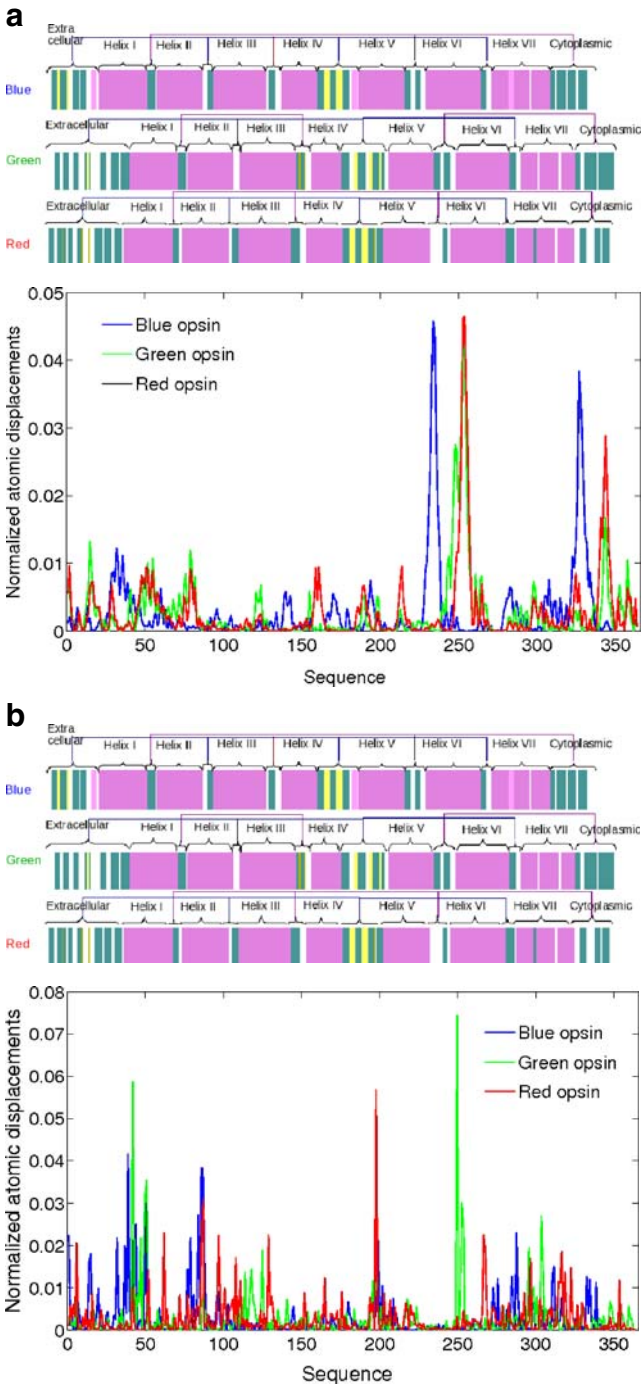

Fig. 5 Normalised squared atomic displacement of mode numbers 2 (**a**) and 589 (**b**) for the three opsins

frequency FIR spectra of the cone opsins—the focus of interest in the present work.

The IR absorbance spectra displayed in Fig. 6 show a rather steady increase with frequency over the entire range of frequencies studied; superimposed we find several sharp lines, e.g. at around 40 cm^{-1}, and then (somewhat broadened) at 70 cm^{-1} and above. The general increase is reminiscent of the low-frequency behaviour of both the normal mode DOS and the power spectrum; however, the rather flat structureless density found in Fig. 2 above around 40 cm^{-1} is not found so prominently in the FIR spectrum. These similarities and deviations are due to the dipole derivatives entering Eq. 1, and hence also to the underlying charge assignment made.

Previous work on low-frequency vibrations (below 100 cm^{-1}) of bR and Rh revealed the existence of normal modes that were localised almost completely on the loop regions (in particular, the loop connecting helices B and C) [10, 31]. Such modes were not found in the present study for the coloured opsins; rather, we observed normal modes with a high activity in the cytoplasmic and extracellular domains, but not restricted to a single loop.

Power spectra in vacuo

Figure 7 displays the power spectrum as determined via the velocity auto-correlation function by MD simulation; Fig. 7 may be compared to Fig. 2, which shows the vibrational DOS as calculated in the NMA. Both calculation schemes give results in close agreement with each other. In particular, the positions of the peaks in both distributions nearly coincide. This indicates that harmonic analysis is a good approximation, even at a finite temperature (300 K),

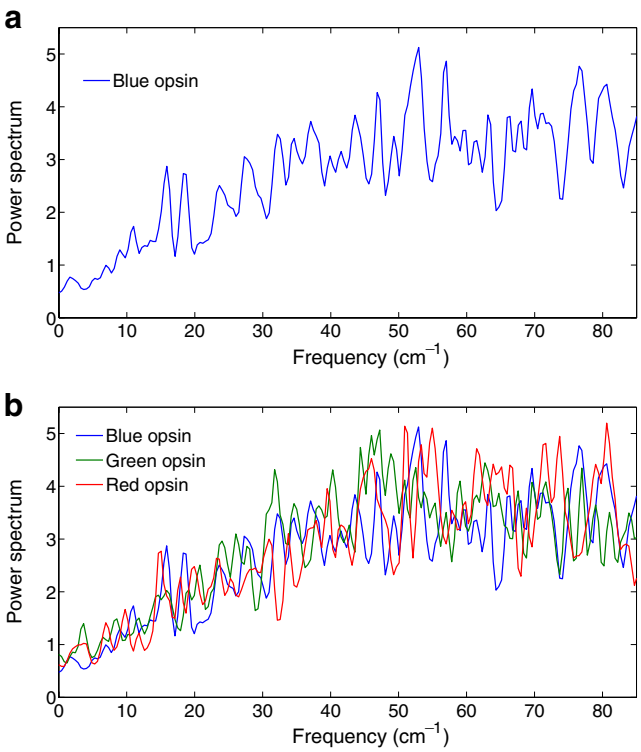

Fig. 6 Power spectra of blue opsin (**a**) and for the three opsins studied (**b**), as obtained by molecular dynamics (MD) simulation

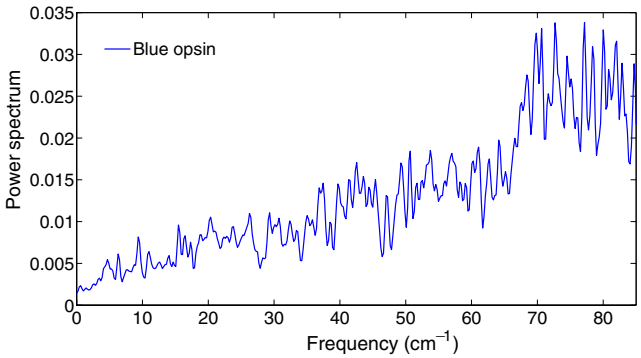

Fig. 7 Power spectrum of blue opsin in water as obtained by MD simulation

where the anharmonicity in the potential surface may not be ignored.

Unlike the discrete frequencies in the normal-mode dynamics, the frequency distribution obtained from the MD simulation is continuous (cf. Fig. 7). This is a consequence of the fact that conformational variations during the simulation broaden the frequencies of the individual modes, causing them to overlap. In order to better compare the discrete NM spectrum with the continuous MD distributions, the NM distribution has been expressed as a mode density by convolution with a Lorentzian function generating a FWHM resolution of 2.0 cm^{-1}. The density of modes in the MD simulation increases steadily up to around 45 cm^{-1}. A corresponding increase is observed for normal-mode dynamics.

We have extracted sample modes from the MD simulation at intervals of 0.4 cm^{-1}—the highest resolution possible for the available simulation time span—up to 50 cm^{-1}, yielding a total of 125 modes for all three opsins. Thus, all motions with periods between 200 ps and 2 ps are considered. At the low frequency end of the spectrum, the lowest normal mode is 2.37 cm^{-1}, while the MD trajectory resolves motions down to 0.4 cm^{-1}. The trajectory is, therefore, longer than needed to cover all relevant motions.

Power spectrum: solvent effects

Opsins are transmembrane proteins; their natural environment is the cell membrane. However, computational studies of transmembrane proteins in their (natural) environment are rare [27]. We rather follow the example of Refs. [18, 54], in which changes induced in the vibrational spectra were calculated by solvating the molecules in water; these changes can be taken to give an indication of the influence of the molecular environment on the vibrational spectra. Actually, the effects can be expected to be much more pronounced in polar solvents such as methanol, acetonitrile, and water than in the natural lipid-matrix protein environment [27].

Thus, previous MD simulations of the dynamics of Rh and free opsins have been carried out where the proteins were immersed in TIP3 water molecules (in the absence of lipids) to study the 11-cis-retinal initiation on conformational rearrangements of the protein part of the molecule in the chromophore region, and peripheral parts of the Rh molecule [67].

Figures 6a and 7 show the power spectra of blue opsin in vacuum and in a solvent (water), respectively. Figures 6b and 8 show the analogous data for all 3 opsins. The overall shapes of the spectra are similar to the power spectra of other proteins, including those calculated from MD simulations [26, 41] and from normal-mode dynamics analysis. They are also similar to neutron scattering spectra obtained for globular proteins [45]. Significant differences can be discerned in the DOS profiles at the lower wavenumbers. Fig. 8b shows an expanded view of the DOS profiles from 2.0 cm^{-1} to 25 cm^{-1}. Relative to the red and green opsins, blue opsin has a substantially increased population of the lowest frequency modes (below ~15 cm^{-1}).

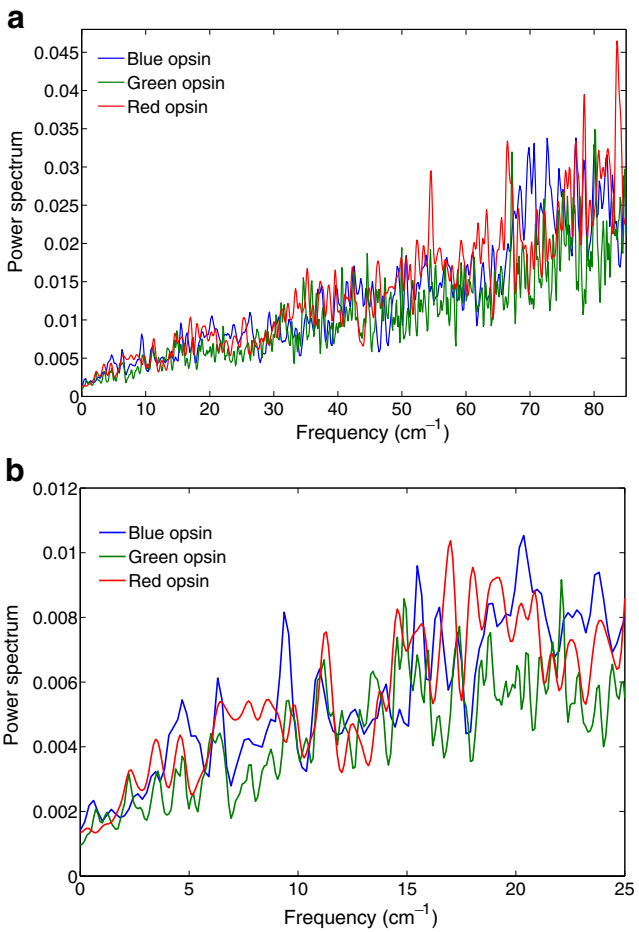

Fig. 8 a,b Power spectra of the three opsins in solvent. **a** Full low-frequency range. **b** Expanded view from 2–25 cm^{-1}

Theoretical studies have established that the modes in this region correspond to collective motions, in which many atoms in a given protein region move in a concerted manner. Blue opsin thus appears to undergo concerted motions more often then green and red opsin. The degree to which these low-frequency modes have been increased varies among the red and green opsins, with red opsin showing a larger increase than green opsin. The order in which the low-frequency modes are increased is seen to match the order of thermostability for the blue and red opsins [23, 43, 68]. We might predict that the thermostability factor and the RMS fluctuations (Fig. 3) of blue opsin are higher than that of the other opsins, as concerted motions occur more frequently in blue opsin than in red or green opsin.

Conclusions

In this paper, we present the calculated terahertz vibration spectra for the three geometrically similar proteins, viz. blue, green and red opsin.

1. The low-frequency vibrational DOS of the three cone opsins were calculated using both NMA and MD simulation at 300 K. Good agreement between the results of both methods was observed. This proves that any vibrational anharmonicities, which might affect molecular motion at this temperature, are moderate. Anharmonic effects, which MD has shown to be important in protein motions, are neglected in the normal mode calculations. In particular, a MD simulation samples many minima in the multidimensional configuration space, whereas the harmonic model is restricted to a single minimum. The present results suggest that, for certain dynamical properties, especially those involving averages over many modes, [36] the RMS atomic fluctuations calculated from the normal modes behave similarly to those from MD simulations. Furthermore, the frequency distribution of the modes that make the dominant contribution to the fluctuations (0.2–65 cm^{-1}) is in accord with the estimates from MD. This agreement provides evidence for the utility of a normal-mode treatment for the analysis of the internal motions of opsins [36].
2. The vibrational spectra calculated for the three opsins show good agreement. This may appear trivial, since the geometrical structure of the three opsins is identical; each protein consists of seven transmembrane helices connected by extracellular and cytoplasmic loops. However, their amino acid sequence differs by up to 40 %. We found it impossible to trace back the origin of any slight differences in the vibrational spectra or in the mode motions found to differences in the amino acid sequence of the three opsins.
3. The IR absorbance was determined using the charge assignments of the CHARMM code. Below a frequency of 50 cm^{-1}, the absorbance is qualitatively similar to the vibrational DOS, demonstrating that FIR spectroscopy may be used to gain information on the low-frequency modes of these molecules.
4. The mode structure of several representative low-frequency vibrational modes was determined and analysed. The change from large scale motion relevant for conformational changes to relative motion of molecular subunits against each other was documented.
5. We investigated the effect of the protein environment on the vibrational spectra of the opsins. Due to computational convenience, instead of embedding the opsins in a lipid membrane, we solvated them in water; due to the polar nature of water, this will exaggerate any environment effect. A strong influence of the environment on the damping of these vibrations was observed, while the frequencies of the vibrations were affected only slightly.

Acknowledgements The authors acknowledge financial support by the Deutsche Forschungsgemeinschaft via the Graduiertenkolleg 792, and thank R. Diller for valuable discussions on the subject.

References

1. Dauber-Osguthorpe P, Osguthorpe DJ, Stern PS, Moult J (1999) J Comput Phys 151:169–189
2. Farantos SC (2007) CPS-IEEE Computer Society, ICCSA2007, pp 444–450
3. Xie A, van der Meer AF, Austin RH (2002) Phys Rev Lett 88:018102–018104
4. Moritsugu K, Miyashita O, Kidera A (2000) Phys Rev Lett 85:3970–3973
5. Xu J, Plaxco KW, Allen SJ (2006) J Phys Chem B 110:24255–24259
6. Ming D, Wall ME (2005) Proteins 59:697–707
7. Leitner M, David, Havenith M, Gruebele M (2006) Int Rev Phys Chem 25:553–582
8. Fanconi B (1973) Biopolymers 12:2759–2776
9. Shotts WJ, Sievers AJ (1973) Chem Phys Lett 21:586–588
10. Whitmire SE, Wolpert D, Markelz AG, Hillebrecht JR, Galan J, Birge RR (2003) Biophys J 85:1269–1277
11. Yamamoto K, Tominaga K, Sasakawa H, Tamura A, Murakami H, Ohtake H, Sarukura N (2002) Bull Chem Soc Jpn 75:1083–1092
12. Smith J, Kuczera K, Karplus M (1990) Proc Natl Acad Sci USA 87:1601–1605
13. Martel P, Calmettes P, Hennion B (1991) Biophys J 59:363–374
14. Diehl M, Doster W, Petry W, Schober H (1997) Biophys J 73:2726–2732
15. Markelz AG, Roitberg A, Heilweil EJ (2000) Chem Phys Lett 320:42–48
16. Xu J, Plaxco KW, Allen JS (2006) Protein Sci 15:1175–1181
17. Zhang C, Durbin SM (2006) J Phys Chem B 110:23607–23613

18. Hinsen K, Kneller GR (2008) Proteins Struct Funct Bioinform 70:1235–1242
19. Hinsen K, Kneller GR (1999) J Chem Phys 111:10766–10769
20. Krishnan M, Balasubramanian S (2003) Phys Rev B 68:064304
21. Nina M, Roux B, Smith JC (1995) Biophys J 68:25–39
22. Wang Q, Wong CF, Rabitz H (1998) Biophys J 75:60–69
23. Balog E, Smith J, Perahia D (2006) Phys Chem Chem Phys 8:5543–5548
24. Korter TM, Balu R, Campbell MB, Beard MC, Gregurick SK, Heilweil EJ (2006) Chem Phys Lett 418:65–70
25. Keskin O, Jernigan RL, Bahar I (2000) Biophys J 78:2093–2106
26. Mathias G, Marx D (2007) Proc Natl Acad Sci USA 104:6980–6985
27. Rajamani R, Gao J (2002) J Comp Chem 23:96–105
28. Pleiss J, Jähnig F (1991) Biophys J 59:795–804
29. Šiber A (2004) Phys Rev B 70:075407
30. Nathans J, Thomas D, Hogness DS (1986) Science 232:193–202
31. Balu R, Zhang H, Zukowski E, Chen JY, Markelz AG, Gregurick SK (2008) Biophys J 94:3217–3226
32. Kaledin AL, Kaledin M, Bowman JM (2006) J Chem Theory Comput 2:166–174
33. Markelz A, Whitmire S, Hillebrecht J, Birge R (2002) Phys Med Biol 47:3797–3805
34. Stenkamp RE, Filipek S, Driessen CAGG, Teller DC, Palczewski K (2002) Biochim Biophys Acta 1565:168–182
35. Trabanino RJ, Vaidehi N, Goddard WA (2006) J Phys Chem B 110:17230–17239
36. Brooks B, Karplus M (1983) Proc Natl Acad Sci USA 80:6571–6575
37. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comp Chem 4:187–217
38. Loccisano AE (2007) PhD thesis, Bayer School of Natural and Environmental Sciences, Duquesne University
39. Baudry J, Hayward RL, Middendorf HD, Smith JC (1997) In: Cusack S, Buttner H, Ferrand M, Langan P, Timmins P (eds) Biological macromolecular dynamics, vol 7. Adenine, pp 49–54
40. Mouawad L, Perahia D (2004) Biopolymers 33:599–611
41. Kitao A, Hirata F, Go N (1991) Chem Phys 158:447–472
42. Tama F, Sanejouand YH (2001) Protein Eng 14:1–6
43. Maiti PK, Pascal TA, Vaidehi N, William I, Goddard A (2004) Nucleic Acids Res 32:6047–6056
44. Dacey DM, Lee BB (1994) Nature 367:731–735
45. Nadler W, Brunger AT, Schulten K, Karplus M (1987) Proc Natl Acad Sci USA 84:7933
46. Hayward S, Kitao A, Go N (1995) Proteins 23:177–186
47. Walther M, Plochocka P, Fischer B, Helm H, Jepsen UP (2002) Biopolymers 67:310–313
48. Brooks B, Janezic D, Karplus M (1995) J Comp Chem 16:1522–1542
49. Janezic D, Venable R, Brooks B (1995) J Comp Chem 16:1554–1566
50. Janezic D, Brooks B (1995) J Comp Chem 16:1543–1553
51. Pal S, Balasubramanian S, Bagchi B (2003) Phys Rev E 67:61502
52. Wintrode PL, Zhang D, Vaidehi N, Arnold FH, Goddard WA (2003) J Mol Biol 327:745–757
53. Schlegel B, Sippl W, Höltje H-DD (2005) J Mol Model 12:49–64
54. Kholmurodov K, Fel'dman T, Ostrovskii M (2007) Neurosci Behav Physiol 37:161–174
55. Terakita A (2005) Genome Biol 6:213
56. Terstegen F, Kolster K, Falzewski S, Buß V (2000) In: Entel P, Wolf DE (eds) Structure and dynamics of heterogeneous systems. World Scientific, Singapore, pp 26–35
57. Fitter J, Heberle J (2000) Biophys J 79:1629–1636
58. Fitter J, Herrmann R, Hau T, Lechner R, Dencher N (2001) Physica B 30:1
59. van Vlijmen HWT, Karplus M (1999) J Phys Chem B 103:3009–3021
60. Siegrist K, Bucher CR, Mandelbaum I, Walker ARH, Balu R, Gregurick SK, Plusquellic DF (2006) J Am Chem Soc 128:5764–5775
61. Go N, Noguti T, Nishikawa T (1983) Proc Natl Acad Sci USA 80:3696–3700
62. Reuter N, Hinsen K, Lacapere J-J (2003) Biophys J 85:2186–2197
63. Gaillard T, Martin E, San Sebastian E, Cossio FP, Lopez X, Dejaegere A, Stote RH (2007) J Mol Biol 374:231–249
64. Jacobs GH (1996) Proc Natl Acad Sci USA 93:577–581
65. Kochendoerfer GG, Lin SW, Sakmar TP, Mathies RA (1999) Trends Biochem Sci 24:300–305
66. Yuan C, Kuwata O, Liang J, Misra S, Balashov SP, Ebrey TG (1999) Biochemistry 38:4649–4654
67. Zhang D, McCammon JA (2005) PLoS Comput Biol 1:e62
68. Hope AJ, Partridge JC, Dulai KS, Hunt DM (1997) Proc Biol Sci R Soc 264:155–163