Theor Chem Acc (2011) 130:1115–1129
DOI 10.1007/s00214-011-0992-y

REGULAR ARTICLE

Incipient structural and vibrational relaxation process of photolyzed carbonmonoxy myoglobin: statistical analysis by perturbation ensemble molecular dynamics method

Masayoshi Takayanagi · Masataka Nagaoka

Received: 28 March 2011 / Accepted: 5 July 2011 / Published online: 10 September 2011

Abstract The incipient structural and vibrational energy relaxation process of photolyzed carbonmonoxy myoglobin was analyzed by the perturbation ensemble molecular dynamics (PEMD) method, in which many pairs of perturbed and unperturbed MD simulations are executed for ensemble-averaging to obtain statistically significant results by canceling out thermal fluctuations. First, we have shown that the experimentally reported anisotropic expansion can be detected within a picosecond after photolysis. The good agreement between the experimental and computational results indicates that the PEMD method can predict legitimately those changes driven by perturbations even if the changes might be subtle and smaller than thermal fluctuations. Second, the structural relaxation including the "clamshell rotation" in E and F helices was successfully analyzed. The high time resolution analysis has clarified the incipient structural dynamics on a subpicosecond timescale: the clamshell rotation starts at His64, Val68, and His93 following both the heme doming and the dissociated CO ligand collision. Third, the vibrational energy relaxation from the heme to the globin matrix is elucidated not only temporally but also spatially. This is the first "thorough" report of the spacetime-resolved excess kinetic energy redistribution of photolyzed MbCO in the globin matrix with a statistically significant precision, ±1 K. The incipient anisotropic vibrational relaxation occurs clearly within a picosecond in the direction perpendicular to the heme plane by the "through-bond" and "through-projectile" pathways, and the isotropic relaxation then follows by the "through-space" pathway. Finally, it is concluded that the PEMD method is a powerful tool to understand the incipient relaxation process driven by the perturbation.

Keywords Hemeprotein · Ensemble molecular dynamics method · Perturbation ensemble · Vibrational relaxation

Dedicated to Professor Akira Imamura on the occasion of his 77th birthday and published as part of the Imamura Festschrift Issue.

M. Takayanagi · M. Nagaoka (✉)
Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
e-mail: mnagaoka@is.nagoya-u.ac.jp
URL: http://www.ncube.human.nagoya-u.ac.jp/

M. Takayanagi · M. Nagaoka
Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Honmachi, Kawaguchi 332-0012, Japan

1 Introduction

To understand protein functions from the atomistic point of view, it is necessary to investigate relaxation processes of proteins, induced by external perturbations, such as light absorption, ligand binding, electron transport, and so forth [1]. For the purpose, nonequilibrium molecular dynamics (MD) simulations can be used to computationally analyze the relaxation processes. However, it is often difficult to analyze the changes driven by the perturbation because of large thermal fluctuations at room temperature. To resolve this thermal noise problem, we have proposed a new nonequilibrium MD procedure, perturbation ensemble MD (PEMD) method, in which many pairs of perturbed and unperturbed MD simulations are executed for ensemble-averaging to obtain statistically significant results by canceling out thermal fluctuations [1–3].

As a first application of the PEMD method, we have analyzed a relaxation process of photolyzed carbonmonoxy myoglobin (MbCO) (Fig. 1). Myoglobin (Mb) is a small

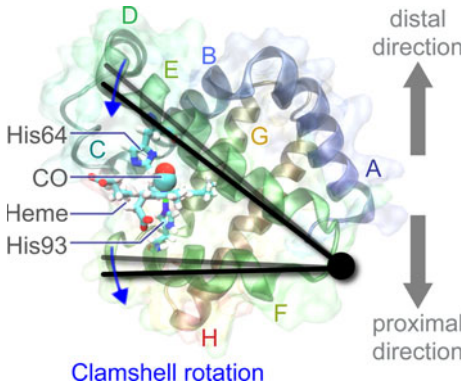

Fig. 1 Schematic illustration of the clamshell rotation driven by the ligand photolysis of MbCO. We define the direction from the heme iron atom to the heme-bound carbon atom of the ligand CO as "distal direction" and that from the iron atom to the heme-bound nitrogen atom of His93 as "proximal direction"

globular protein consisting of a globin matrix (153 amino acids) with a prosthetic heme. The active site of Mb, the heme, binds various kinds of small gaseous molecules such as molecular oxygen (O_2), carbon monoxide (CO), nitric oxide (NO), and cyanide ion (CN^-), i.e., MbO_2, MbCO, MbNO, and $MbCN^-$. The photolysis of MbCO has long been studied since after absorption of visible light the CO ligand dissociates with a very high quantum yield. This photolysis triggers structural changes in the heme from a six-coordinated planar form to a five-coordinated domed one. In addition, a part of the photon energy absorbed is injected into the heme as excess vibrational energy. These events bring about not only the structural relaxation process of the globin matrix, but also excess vibrational energy relaxation from the hot heme to the surrounding globin and solvent matrices. A lot of investigations have been done to date on this phenomenon since it is a good model system for studying relaxation processes in biomacromolecules [1–21].

With respect to the structural relaxation process, incipient anisotropic shape changes in MbCO following ligand photolysis were studied by optically heterodyne-detected (OHD) transient grating (TG) spectroscopy [4] and molecular dynamics (MD) simulations [2]. X-ray crystallography [5] and its picosecond time-resolved applications [6–8] have been clarifying the precise differences in atomic positions between the deoxy and the CO-ligated state and detecting the ongoing structural relaxation after the CO photolysis. A recent study of the Mb mutants by time-resolved UV resonance Raman spectroscopy [9] demonstrated the hydrophobicity changes in environments around Trp and Tyr residues. The most striking structural differences between MbCO and deoxyMb occur in the E and F helices, where it is assumed that the heme doming motion pushes the F helix toward the proximal direction (we define the direction from the heme iron atom to the heme-bound carbon atom of the ligand CO as "distal direction" and that from the iron atom to the heme-bound nitrogen atom of His93 as "proximal direction", see Fig. 1) and, at the same time, the collapse of E helix toward the heme occurs by the removal of heme-bound CO ligand. These changes can be interpreted as a rotation of the two helices whose pivot is at the EF corner, called "clamshell rotation". This rotation was observed by the high-resolution X-ray crystallography [5] and modeled and studied with a theoretical structure prediction by the QM/MM method [10].

From the viewpoint of the excess energy relaxation, since the pioneering molecular dynamics (MD) simulation by Henry et al. [11], not only the cooling process of the vibrationally hot heme but also the concomitant heating of the surroundings has been studied. The most illuminating experiments for the heme relaxation, in which the heme ν_4 band was monitored by time-resolved resonance Raman spectroscopy, showed the double exponential decay of the heme with time constants of 1.9 ± 0.6 ps and 16 ± 9 ps [12, 13]. Several theoretical studies using MD simulation demonstrated a similar double exponential decay of the heme [14, 15]. Furthermore, the important role of the propionate side chains of heme in the relaxation process from the heme to the solvent [15] was also indicated by MD simulations [14–17] and spectroscopic studies [18, 19]. Frequency-resolved local energy diffusivities of Mb were theoretically analyzed, and it was found that rapid energy flow between the heme and other parts of the protein occurs mainly at frequencies below $\sim 350\ cm^{-1}$ [20]. With regard to the solvent heating, the femtosecond IR spectroscopy of D_2O solvent was utilized to reveal the double relaxation pathway from the heme to the solvent [21]. However, there is no direct experimental measurement of the heating in the globin matrix. Theoretical studies using MD simulations indicated that there is an anisotropic redistribution of the excess kinetic energy from the heme to the surrounding protein [14, 16].

Since the thermal energy fluctuations at room temperature are often larger than the amount of transferred energy per degree of freedom, it is difficult to investigate energy transfer pathways in proteins. To resolve the difficulty, several approximate methods have been proposed. For example, in the anisotropic thermal diffusion method, all atoms in the system are initially minimized and cooled to 10 K to drastically reduce the thermal noise [22]. Another example is the pump–probe MD method in which selected pump atoms are oscillated by the artificial force [23]. Although such methods could detect energy transport channels in proteins, the approximations (cooling or artificial oscillation) may influence the intrinsic dynamics of proteins at room temperature and, as a result, careful evaluation of the results is necessary. On the contrary, as a straightforward method, hundreds of nonequilibrium MD

simulations were carried out to analyze the transient flow of vibrational energy along a photoswitchable Aib peptide helix [24, 25]. Their ensemble-averaged results could detect energy transfer from the photoexcited chromophore to the adjacent residues. However, it was hard to detect energy transfer to residues far away from the chromophore because of the poor signal-to-noise ratio.

In the present paper, we analyze the incipient relaxation process of photolyzed MbCO by the PEMD method on the picosecond timescale. In Sect. 2, we describe the theoretical framework of the PEMD method, in which a large number of pairs of perturbed MD (PMD) and unperturbed MD (UMD) simulations are executed for ensemble-averaging to cancel out thermal fluctuations [1–3]. In Sect. 3, the computational modeling of MbCO in aqueous solution and CO ligand photolysis and the computational analysis of the PEMD method are explained. Then, in Sect. 4, the structural and vibrational energy relaxation process of photolyzed MbCO are discussed with a high statistically significant precision and high time resolution.

2 Theoretical methods

2.1 Ensemble molecular dynamics method to cancel out thermal fluctuations

To obtain statistically significant information from MD simulations, it is often necessary to deal with large thermal fluctuations at room temperature. Our approach is a straightforward ensemble MD (EMD) method in which many MD trajectories are ensemble-averaged to cancel out the thermal fluctuations.

If a number of trajectories $N_{\rm traj}$ is large enough to ensure that $\sqrt{N_{\rm traj}-1}/\sqrt{N_{\rm traj}} \cong 1$, the mean of observable A is

$$\langle A\rangle_{N_{\rm traj}} = \frac{1}{N_{\rm traj}}\sum_{I=1}^{N_{\rm traj}} A^I, \tag{1}$$

and the standard error of this mean is

$$\sigma_{\rm M} = \frac{s}{\sqrt{N_{\rm traj}}}, \tag{2}$$

where s is the standard deviation (SD) defined by

$$s^2 = \frac{1}{N_{\rm traj}}\sum_{I=1}^{N_{\rm traj}} \left(A^I - \langle A\rangle_{N_{\rm traj}}\right)^2. \tag{3}$$

Assuming the confidence level is c (%), the real mean $\bar{A}$ should be within the confidence interval (CI) as

$$\langle A\rangle_{N_{\rm traj}} - t(c)\cdot\sigma_{\rm M} \le \bar{A} \le \langle A\rangle_{N_{\rm traj}} + t(c)\cdot\sigma_{\rm M}, \tag{4}$$

where $t(c)$ is a number depending on both the degrees of freedom $N_{\rm traj} - 1$ and the confidence level c. We normally assume $c = 95\%$, and the $t(95\%)$ would become ~ 1.96 with a sufficiently large $N_{\rm traj}$, resulting in the inequality

$$\langle A\rangle_{N_{\rm traj}} - 1.96\sigma_{\rm M} \le \bar{A} \le \langle A\rangle_{N_{\rm traj}} + 1.96\sigma_{M}. \tag{5}$$

Let us assume that the statistical precision required for analyzing the quantity A is known numerically as h, where h is a real number assigned, for example, by comparing with the precision in the experimental measurement. Then, the required number of trajectories $N_{\rm traj}$ can be estimated by the relation

$$t(c)\cdot\sigma_{\rm M} \le h. \tag{6}$$

We call this the c-CI EMD method, or simply the EMD method (usually, $c = 95\%$). For such a method to estimate the ensemble average of an observable A one must use a sufficiently large number of trajectories so that the inequality in Eq. 6 might be satisfied.

2.2 Perturbation ensemble MD method: ensemble MD simulations with and without perturbation

Focusing on the responses of proteins to perturbations, we have used the PEMD method. This is a combination of the EMD method and a perturbative treatment. The perturbative treatment is indispensable to remove protein dynamics that undergoes even without the perturbation [2].

In the PEMD method, $N_{\rm traj}$ pairs of MD simulations are executed with and without the perturbation, i.e., perturbed MD (PMD) and unperturbed MD (UMD), respectively. A pair of PMD and UMD trajectories is schematically presented in Fig. 2. Hereafter, we use the term "snapshot" to refer to the atomic coordinates and velocities of the entire MD system (**r, v**). For each pair of PMD and UMD calculations, the initial snapshot is identical except for those in the perturbed region (in Fig. 2, depicted as "S"). Next, observable A is calculated from the PMD and UMD simulations as

$$(A^{{\rm PMD},I}(t), A^{{\rm UMD},I}(t)) \quad I = 1, 2, \ldots, N_{\rm traj}, \tag{7}$$

where $A^{{\rm PMD},I}$ and $A^{{\rm UMD},I}$ are those values of the physical quantity A calculated through the Ith PMD and UMD trajectory, respectively. Then, instantaneous deviation of A between the Ith PMD and UMD simulations is calculated:

$$\delta A^I(t) = A^{{\rm PMD},I}(t) - A^{{\rm UMD},I}(t). \tag{8}$$

At short times after perturbation, the difference in Eq. 8 provides high-precision measurements by effectively canceling out the thermal fluctuations because the time evolution of the thermal fluctuations in the MD system is common in each pair of PMD and UMD trajectories

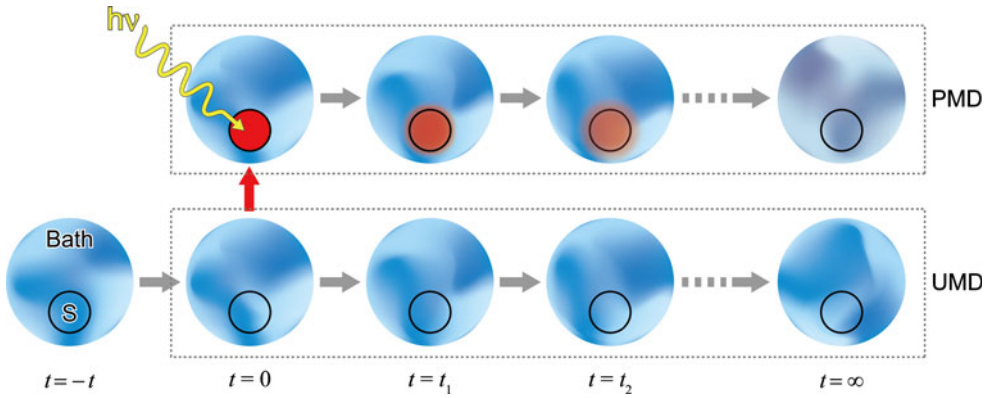

Fig. 2 Schematic view of a pair of PMD and UMD trajectories in the PEMD method. The MD system is composed of "S", where we give perturbation $h\nu$ at time $t = 0$, and other region "Bath". The thermal fluctuations in the system are represented as *blue* while color mapping. The PMD and UMD simulations are executed from the same initial equilibrium snapshot except for "S". In the UMD trajectory, the system maintains equilibrium for all $t \cdot 0$. On the other hand, in the PMD trajectory, the perturbation relaxes from "S" to the surrounding "Bath" and finally, the system reaches equilibrium. Because the time evolution of the thermal fluctuations is common in the PMD and UMD trajectories at short times after perturbation, the comparison between PMD and UMD gives high precision by effectively canceling out the thermal fluctuations

(Fig. 2). The deviation is ensemble-averaged over the N_{traj} trajectory pairs:

$$\langle \delta A(t) \rangle = \frac{1}{N_{\mathrm{traj}}} \sum_{I=1}^{N_{\mathrm{traj}}} \delta A^I(t) = \frac{1}{N_{\mathrm{traj}}} \sum_{I=1}^{N_{\mathrm{traj}}} \left(A^{\mathrm{PMD},I}(t) - A^{\mathrm{UMD},I}(t) \right). \quad (9)$$

In addition, the time change in A at time t with respect to that at the initial time ($t = 0$) is defined as

$$\Delta A^I(t) = A^I(t) - A^I(0), \quad (10)$$

and the ensemble-averaged time change and its perturbative variation are thus calculated as

$$\langle \Delta A(t) \rangle = \frac{1}{N_{\mathrm{traj}}} \sum_{I=1}^{N_{\mathrm{traj}}} \Delta A^I(t) = \frac{1}{N_{\mathrm{traj}}} \sum_{I=1}^{N_{\mathrm{traj}}} \left(A^I(t) - A^I(0) \right) \quad (11)$$

and

$$\langle \delta \Delta A(t) \rangle = \frac{1}{N_{\mathrm{traj}}} \sum_{I=1}^{N_{\mathrm{traj}}} \delta \Delta A^I(t) = \frac{1}{N_{\mathrm{traj}}} \sum_{I=1}^{N_{\mathrm{traj}}} \{ \left(A^{\mathrm{PMD},I}(t) - A^{\mathrm{PMD},I}(0) \right) - \left(A^{\mathrm{UMD},I}(t) - A^{\mathrm{UMD},I}(0) \right) \}. \quad (12)$$

Because of the effective cancelation of the thermal fluctuations by the differentiation between each pair of PMD and UMD trajectories ($\delta A^I(t)$ in Eq. 8), precisions of the averages $\langle \delta A(t) \rangle$ and $\langle \delta \Delta A(t) \rangle$ are very high at short times after perturbation (this will be discussed later in Sect. 3.2.4, see Table 1). This is an improvement over the previous analysis with hundreds of nonequilibrium MD simulations [24, 25].

Table 1 Min, max, and average of the two-sided 95% confidence intervals (95% CI) of the ensemble-averaged residue kinetic temperature changes calculated from 20,000 pairs of PMD and UMD trajectories in each time window

Time window	Min (K)		Max (K)		Average (K)
$L = 0.0$ ps	Leu9	±0.0001	His64	±0.1764	±0.0074
$L = 0.3$ ps	Leu9	±0.0042	His64	±0.3019	±0.0604
$L = 0.6$ ps	Val10	±0.0449	His64	±0.4419	±0.1968
$L = 1.0$ ps	Trp14	±0.2651	Gly153	±0.8165	±0.4481
$L = 2.9$ ps	Trp14	±0.5516	Gly124	±1.0720	±0.7381

It should be noted that, as was discussed in the previous section, the number of MD trajectory pairs N_{traj} must satisfy the relation Eq. 6 for the target change $\Delta A(t)$ or $\delta \Delta A(t)$. With such a large N_{traj}, we call the calculation procedure the c-CI PEMD method, or simply the PEMD method (usually, $c = 95\%$).

3 Application to the analysis of relaxation processes of photolyzed MbCO

As an application of the PMED method, we focus on the MbCO · Mb* + CO process, that is, the photolysis of MbCO and the subsequent energy flow (cooling) of the vibrationally excited heme, provided the electronic state of heme stays in the (d, d) state after the photolysis.

Spectroscopic studies have revealed that the MbCO ligand dissociation by the electronically excited heme occurs in a very short time (within the instrument response time $\sim$2 ps) and recombination of the CO to the heme takes place in the time regime of microseconds to milliseconds [12, 13]. Therefore, we assume that the electronically excited heme appears instantly after the photon absorption and the excited state is retained during the MD simulations.

3.1 Computational modeling

3.1.1 MbCO in aqueous solution

To investigate the structural relaxation of the photolyzed MbCO (Mb* + CO) system in aqueous solution, we used PEMD method. MD simulations were all carried out with the AMBER 7 or 9 PMEMD programs using the parm99 force field [26]. For the heme and the heme-bound CO ligand parameters, we used bond, angle, and dihedral parameters as developed by Henry et al. [27], while for the electrostatic and van der Waals parameters, we used those determined by Giammona [28]. The atomic charges of the dissociated CO molecule were calibrated to mimic its dipole moment 0.11 D [29], i.e., $-0.021e$ and $+0.021e$ for the carbon and oxygen atom, respectively. All the MD simulations were performed using periodic boundary conditions. We used the particle mesh Ewald (PME) procedure to handle long-range electrostatic interactions. The MbCO structure determined by neutron diffraction (PDB ID#: 2MB5 [30]) was adopted as an initial structure. The protonation states of histidine residues were assigned as described in the PDB file (His24, His64, and His93 are singly protonated and others are doubly protonated). Because all the doubly protonated histidines are exposed to solvent, their protonation states affect little to the relaxation dynamics of photolyzed MbCO. The MbCO was solvated with $\sim$3,000 water molecules and 9 chloride counterions yielding a periodic box size of $\sim 51 \times 47 \times 47$ Å^3. All the NPT ensemble (constant pressure and temperature) simulations were executed at ambient condition (300 K, 1 atm).

We should note that the initial crystal structure of MbCO, especially in the vicinity of the heme, is stable under the simulation conditions in this work. We executed two long 5-ns MD simulations of MbCO and found that the RMSD of the whole MbCO and that of residues in the vicinity of the heme group (Thr39, Lys42, Phe43, Arg45, His64, Thr67, Val68, Ala71, Leu72, Leu89, Ser92, His93, His97, Ile99, Tyr103, Leu104, Ile107, Phe138) increased to 2.0 and 1.0 Å within 100 ps and then fluctuated around 2.0–2.5 and 1.0–1.5 Å, respectively. This stability indicates that the lengths of our equilibrium MD simulations (i.e., for several hundreds of ps as will be described in Sect. 3.2) are decent enough.

3.1.2 CO ligand photolysis and the perturbation

In the PEMD method, we first prepared a large number of initial equilibrium snapshots of MbCO in aqueous solution. Hereafter, we describe the snapshot of the MD system as $(\mathbf{r}, \mathbf{v}) = (\mathbf{r}^{\mathrm{Mb}}, \mathbf{r}^{\mathrm{CO}}, \mathbf{r}^{\mathrm{solvent}}, \mathbf{v}^{\mathrm{Mb}}, \mathbf{v}^{\mathrm{CO}}, \mathbf{v}^{\mathrm{solvent}})$, where $\mathbf{r}^{\mathrm{Mb}}$ and $\mathbf{v}^{\mathrm{Mb}}$ are atomic coordinates and velocities of Mb including the heme, $\mathbf{r}^{\mathrm{CO}}$ and $\mathbf{v}^{\mathrm{CO}}$ are those of CO ligand, and $\mathbf{r}^{\mathrm{solvent}}$ and $\mathbf{v}^{\mathrm{solvent}}$ are those of solvent including water and counterions, respectively. Then, from each of the snapshots, we executed a pair of PMD and UMD simulations. In the UMD simulation, no perturbation was introduced and an equilibrium MD simulation of MbCO was executed from the equilibrium snapshot. On the other hand, in the PMD simulation, a perturbation was introduced at $t = 0$ to simulate the relaxation process after CO ligand photolysis by a 353-nm photon (81 kcal/mol) absorption as follows. The force-field parameters of the heme were changed from those of the ground state ("Liganded" in ref [27]) to those of the excited state ("Unliganded hard" in ref [27]). With the "Unliganded hard" parameter, the heme deformation energy, or the difference in potential energy for the iron in its in-plane and out-of-plane positions, is 10 kcal/mol. Moreover, a repulsive potential function between the iron atom and the carbon atom of the CO ligand [15] was introduced as a function of the distance between the iron and ligand carbon atoms $r^{\mathrm{Fe\text{-}C}}$ as $V(r^{\mathrm{Fe\text{-}C}}) = A/r^{12} + B/r^6$ with the parameters $A = 180.507$ Å^{12} kcal/mol and $B = 244.054$ Å^6 kcal/mol by replacing the 6–12 Lennard-Jones potential between the iron and carbon atoms. These parameters were determined by fitting the energy curve obtained in the molecular orbital (MO) study [15] for the $^1B_2(3d_\pi \rightarrow 3d_{z^2})$ excited state of the C_{2v} symmetrical FeP(py)CO (FeP: porphyrinato-iron; py: pyridine). In this Super-CI CASSCF calculation, 14 electrons were distributed in all possible ways among 12 active MOs. The ligand dissociation energy from its equilibrium distance in the "Liganded" potential is $V(1.75\ \text{Å}) - V(\infty) = 8.7$ kcal/mol. In addition, we deposited the excess kinetic energy, 54 kcal/mol, into 24 atoms of the porphyrin ring by adding 24 randomly generated excess velocity vectors to the original velocity ones as explained in our previous work [15]. It should be noted here that the 24 random velocity vectors were generated to conserve the same total linear and angular momenta of the 24 atoms as before the photon absorption. The remaining energy of the photon (27 kcal/mol) was assumed to be used to change the electronic state of the heme.

3.2 Computational analysis

3.2.1 Radius of gyration for structural relaxation analysis

In this section, the integration time-step was 2 fs and the SHAKE method was used to constrain the hydrogen–heavy atom bond distances. TIP3P water model [31] was adopted. After an energy minimization and a 400-ps equilibration MD in the NPT ensemble, a 600-ps sampling MD of MbCO in the NPT ensemble was executed saving snapshots every 1 ps to generate 600 equilibrium snapshots. Then, 600 pairs of PMD and UMD simulations for 100 ps were executed for analysis.

To detect the anisotropic structural change after photolysis, the total radius of gyration R_g and its three Cartesian components, R_{gx}, R_{gy}, and R_{gz}, are calculated from $\mathbf{r}^{Mb}$, i.e., the atomic coordinates of Mb (including the heme) as

$$R_g(t) = \sqrt{\frac{\sum_{i\in Mb} m_i |\mathbf{r}_i(t) - \mathbf{r}_c(t)|^2}{\sum_{i\in Mb} m_i}} \tag{13}$$

and

$$R_{ga}(t) = \sqrt{\frac{\sum_{i\in Mb} m_i [a_i(t) - a_c(t)]^2}{\sum_{i\in Mb} m_i}} \quad a = x, y, z \tag{14}$$

where m_i is the mass of ith atom, and $\mathbf{r}_i = (x_i, y_i, z_i)$ and $\mathbf{r}_c = (x_c, y_c, z_c)$ are the Cartesian coordinates of the ith atom and the center of mass (COM) of the Mb, respectively. For the sake of convenience, the superscript Mb is omitted at $\mathbf{r}^{Mb}$ in Eqs. 13 and 14. Here, the same coordinate system was used in Ref. [14]: the xy plane is taken to coincide with the heme plane defined by the Fe atom and the two nitrogen atoms of the pyrrole heme rings A and B, and the z axis is perpendicular to it. For each MD trajectory analysis, the coordinate system was fixed at that defined at each initial structure.

3.2.2 Inter-residue distance change for clamshell rotation analysis

In this section, the SHAKE method was not used with the integration time-step 0.2 fs and SPC/F flexible water model [32] was adopted to include the vibrations of bonds including hydrogen atoms. From the initial MbCO structure, we executed a 200-ps equilibration MD in the NPT ensemble. Next, a 500-ps sampling MD in the NPT ensemble was performed saving snapshots every 5-ps interval. By repeating such equilibration and sampling MDs for 20 times starting from different initial velocities, we obtained 2,000 snapshots as a whole. However, during the 20 sets of equilibration and sampling MD trajectories, we detected three distinct conformational substates of His64, in/N_ε, in/N_δ, and out conformations. This conformational instability is attributed to the N_δ protonated His64, whose protonation state was determined by neutron diffraction. We adopted the in/N_δ conformation snapshots (1,232 out of the 2,000 snapshots) to analyze the relaxation process of photolyzed MbCO at neutral pH. Detailed discussion of this conformational instability is given in our previous work [3]. From the 1,232 snapshots of the in/N_δ conformation, we executed 1,232 pairs of PMD and UMD simulations for 20 ps.

The ensemble-averaged atomic positions of Mb were, in general, calculated from the ensembles of PMD and UMD trajectories as

$$\langle \mathbf{r}_i(t)\rangle^{P}_{N_{traj}} = \frac{1}{N_{traj}} \sum_{n=1}^{N_{traj}} \mathbf{r}_i^{P,n}(t) \tag{15}$$

and

$$\langle \mathbf{r}_i(t)\rangle^{U}_{N_{traj}} = \frac{1}{N_{traj}} \sum_{n=1}^{N_{traj}} \mathbf{r}_i^{U,n}(t). \tag{16}$$

It should be noted that, in Eqs. 15 and 16, $\mathbf{r}_i^{P,n}(t)$ and $\mathbf{r}_i^{U,n}(t)$ are RMS fitted to the initial crystal structure to remove the contribution of translational and rotational motion of Mb. The ensemble-averaged COM position of the residue r was calculated from the PMD and UMD structures as

$$\langle \mathbf{R}_r(t)\rangle^{P}_{N_{traj}} = \frac{\sum_{i\in r} m_i \langle \mathbf{r}_i(t)\rangle^{P}_{N_{traj}}}{\sum_{i\in r} m_i} \tag{17}$$

and

$$\langle \mathbf{R}_r(t)\rangle^{U}_{N_{traj}} = \frac{\sum_{i\in r} m_i \langle \mathbf{r}_i(t)\rangle^{U}_{N_{traj}}}{\sum_{i\in r} m_i}. \tag{18}$$

In addition, the ensemble-averaged residue displacement between the PMD and UMD structures was calculated as

$$\left|\langle \delta\mathbf{R}_r(t)\rangle_{N_{traj}}\right| = \left|\langle \mathbf{R}_r(t)\rangle^{P}_{N_{traj}} - \langle \mathbf{R}_r(t)\rangle^{U}_{N_{traj}}\right|. \tag{19}$$

We also calculated the ensemble-averaged inter-residue distance change between the COM position of residue r and that of residue s as

$$\delta\left|\langle \mathbf{R}_{rs}(t)\rangle_{N_{traj}}\right| = \left|\langle \mathbf{R}_{rs}(t)\rangle^{P}_{N_{traj}}\right| - \left|\langle \mathbf{R}_{rs}(t)\rangle^{U}_{N_{traj}}\right| \tag{20}$$

$$= \left|\langle \mathbf{R}_r(t)\rangle^{P}_{N_{traj}} - \langle \mathbf{R}_s(t)\rangle^{P}_{N_{traj}}\right| - \left|\langle \mathbf{R}_r(t)\rangle^{U}_{N_{traj}} - \langle \mathbf{R}_s(t)\rangle^{U}_{N_{traj}}\right|. \tag{21}$$

3.2.3 *Residue kinetic temperature for vibrational energy relaxation pathway*

The integration time-step was 0.2 fs, and SPC/F flexible water model [32] was adopted to include the vibrations of those chemical bonds including hydrogen atoms. The protonation state for the distal histidine (His64) was changed from the N_δ atom to the N_ε atom to avoid the conformational instability as discussed in our previous work [3].

The calculation procedure of the PEMD described in this section is shown in Fig. 3. Following a 500-ps equilibration MD in the NPT ensemble, we executed a 1,000-ps sampling MD in the NPT ensemble saving snapshots every 5 ps. We generated 20,000 equilibrium snapshots as a whole by repeating the equilibration and sampling MDs 100 times starting from the different initial velocities. Then, 20,000 pairs of 3-ps PMD and UMD simulations for 3 ps were executed for analysis.

To analyze the vibrational energy relaxation process of photolyzed MbCO, time-dependent kinetic temperatures [33] were calculated from atomic velocities of Mb and the CO ligand. The instantaneous atomic kinetic temperature of the ath atom is defined as

$$T_a^{\mathrm{atom}}(t) = \frac{2}{3k_\mathrm{B}}\frac{1}{2}m_a\mathbf{v}_a^2(t), \tag{22}$$

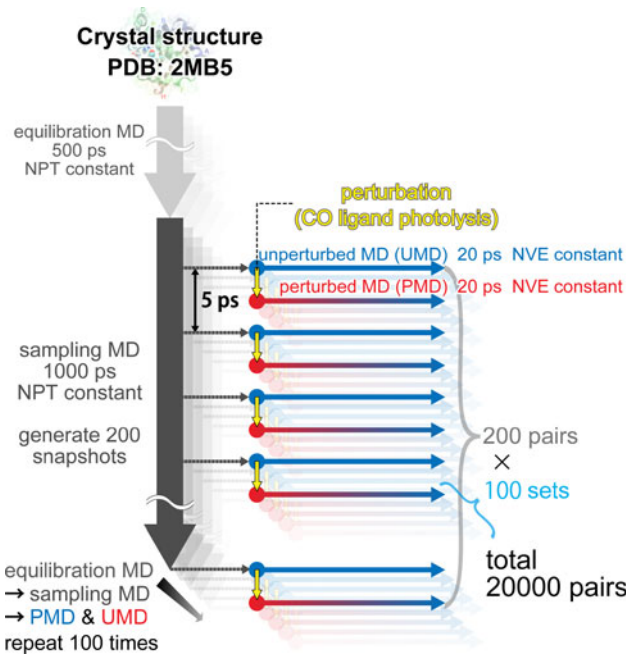

Fig. 3 Schematic illustration of the PEMD procedure for calculating 20,000 pairs of PMD and UMD trajectories

where k_B is the Boltzmann's constant, and m_a and $\mathbf{v}_a$ are the mass and velocity vector of the ath atom, respectively. The instantaneous residue kinetic temperature of the rth residue is defined as

$$T_r(t) = \frac{1}{N_a}\sum_{i\in\text{residue }r}^{N_r} T_r^{\mathrm{atom}}(t) = \frac{2}{3N_r k_B}\sum_{a\in\text{residue }r}^{N_r}\frac{1}{2}m_a\mathbf{v}_a^2(t), \tag{23}$$

where N_r is the number of atoms of the rth residue. In addition to the 153 globin residues, we regarded the heme and CO as 154th and 155th residues, respectively, to calculate the residue kinetic temperatures.

In the sense of the PEMD method, we calculated the change in the residue kinetic temperature from each pair of PMD and UMD simulations as

$$\delta T_r^n(t) = T_r^{\mathrm{P},n}(t) - T_r^{\mathrm{U},n}(t), \tag{24}$$

where $T_r^{\mathrm{P},n}(t)$ and $T_r^{\mathrm{U},n}(t)$ are the instantaneous residue kinetic temperatures of the rth residue at time t in the nth PMD and UMD trajectories, respectively. We averaged the changes over 0.1-ps time window (50 values with a 2-fs interval) as

$$\delta T_r^{n,L} = \frac{1}{50}\sum_{t=L}^{L+0.1\,\mathrm{ps}}\delta T_r^n(t), \tag{25}$$

$$L = 0.0, 0.1, \ldots, 2.9\,\mathrm{ps}. \tag{26}$$

Finally, the changes in residue kinetic temperature are ensemble-averaged over N_traj trajectory pairs as

$$\langle\delta T_r\rangle_{N_\mathrm{traj}}^{L} = \frac{1}{N_\mathrm{traj}}\sum_{n=1}^{N_\mathrm{traj}}\delta T_r^{n,L}. \tag{27}$$

Here, we give a brief summary of the computational resources required to calculate the 20,000 pairs of MD trajectories described in this section. We executed 500-ps equilibration MDs for 100 times, 1-ns sampling MDs for 100 times, and 3-ps PMD and UMD pair MDs for 20,000 times. The total simulation time is 500 ps × 100 + 1,000 ps × 100 + 3 ps × 20,000 × 2 = 270 ns. It takes about 400 s to simulate 1-ps trajectory using a single-thread AMBER 9 PMEMD program on a dual Xeon E5430 quad core CPU machine. By executing MD simulations concurrently using the single-thread PMEMD program on a PC cluster, which consists of 13 dual Xeon E5430 CPU machines (in total 104 CPU cores), the whole simulation took about 12 days (270 ns × 400 s/ps/104 (cores) = 12 days). It should be noted that all the MD simulations can be executed independently, and thus, calculation efficiency of the concurrent execution of 104 simulations is almost 100%. The total output data size of 20,000 pairs of PMD and UMD trajectories is 3.73 TB in the NetCDF format.

3.2.4 Number of MD trajectory pairs determined by residue kinetic temperature fluctuation

Before we start MD simulations using the PEMD method, we should first estimate how many trajectory pairs, N_{traj}, are required to obtain statistically significant results. Here, we provide an example of how the required number N_{traj} can be estimated to analyze the residue kinetic temperature in Eq. 27.

To simulate the perturbation of 353-nm photon absorption, we deposited initially 54 kcal/moles of excess vibrational energy into 24 atoms of the heme porphyrin ring as the perturbation in PMD simulations. At the initial condition, the energy is localized only in the 24 porphyrin atoms and there is a $\sim$750 K increase in kinetic temperature. Then, the vibrational energy relaxes into the surrounding globin and solvent atoms. It was estimated by the classical diffusion picture that the temperature increase at 5 Å from the heme never rises more than 20 K after 580-nm photon absorption [21]. Assuming that the precision in kinetic temperature should be smaller than the amount of temperature increase by one order of magnitude, we adopted a target precision of 1 K to analyze the vibrational energy relaxation pathway. Accordingly, we set $h = 1$ K in Eq. 6.

To estimate the magnitude of thermal fluctuations, SDs of $\delta T_r^{n,L}$ in Eq. 25 were estimated from a set of preliminary MD simulations with 50 pairs of 3-ps PMD and UMD trajectories. The largest SD was evaluated to be 70 K at $L = 2.9$ ps in glycine residues. Therefore, the required number N_{traj} of MD trajectory pairs to achieve a 95% confidence interval (95% CI) of $\pm$ 1 K must satisfy Eq. 6 as

$$t(95\%) \cdot \sigma_{\text{M}} = t(95\%) \cdot \frac{s}{\sqrt{N_{\text{traj}}}} = 1.96 \frac{70\,\text{K}}{\sqrt{N_{\text{traj}}}} \leq 1\,\text{K}, \tag{28}$$

and then is approximately estimated,

$$N_{\text{traj}} \geq (1.96s)^2 \cong 18,800, \tag{29}$$

where σ_{M} and s (= 70 K) are the standard error of the mean and the sample SD, respectively. Accordingly, we adopted a considerable number of trajectory pairs, $N_{\text{traj}} = 20{,}000$. As far as we know, this is the largest sampling of trajectories for analyzing the nonequilibrium protein dynamics.

Using the 20,000 trajectory pairs described in the previous section, 95% CIs of $\langle \delta \Delta T_r \rangle_{20,000}^{L}$ were calculated and summarized in Table 1. Immediately after photolysis at $L = 0.0$ ps, the average CI is very small, $\pm$0.0074 K. The high precision was accomplished by the subtraction procedure of UMD trajectory from PMD one, whose initial atomic velocities, or residue kinetic temperature distribution, were the same as those in the UMD ones except for the perturbed heme atoms as explained in Fig. 2 (in this case, "S" is the heme, and "Bath" is other globin and solvent matrices). The CIs monotonically increased with time, and the average reached $\pm$0.74 K at $L = 2.9$ ps. In this time window, the CI of each residue depends on the number of atoms composing the residue. The smallest residue, glycine, has the largest CI ($\sim\pm$1.05 K), while the largest residue, tryptophan, does the smallest one ($\sim\pm$0.55 K). Because the largest CI was found $\pm$1.07 K in Gly124, it was ascertained that the target precision $\pm$1 K was achieved.

4 Results and discussion

4.1 Anisotropic structural relaxation

Since an immediate anisotropic expansion of photolyzed MbCO in the direction perpendicular to the heme plane was reported by transient grating (TG) spectroscopy [4], we have investigated the structural relaxation of MbCO subsequent to ligand photolysis by calculating the radius of gyration R_{g} (Eq. 13) and its three Cartesian components, $R_{\text{g}x}$, $R_{\text{g}y}$, and $R_{\text{g}z}$ (Eq. 14). Then, as a first application of the PEMD method, we show the anisotropic structural relaxation of photolyzed MbCO by using 600 pairs of PMD and UMD trajectories (explained in Sect. 3.2.1) [2].

In Fig. 4, we plotted the change in R_{g} of a single PMD trajectory (black dots) and those changes ensemble-averaged over 5, 25, 125, and 600 PMD trajectories (black thin line, blue line, red line, and black thick line) as

$$\langle \Delta R_{\text{g}}(t) \rangle = \frac{1}{N_{\text{traj}}} \sum_{I=1}^{N_{\text{traj}}} \left(R_{\text{g}}^{I}(t) - R_{\text{g}}^{I}(0) \right), \quad N_{\text{traj}} = 1, 5, 25, 125, 600. \tag{30}$$

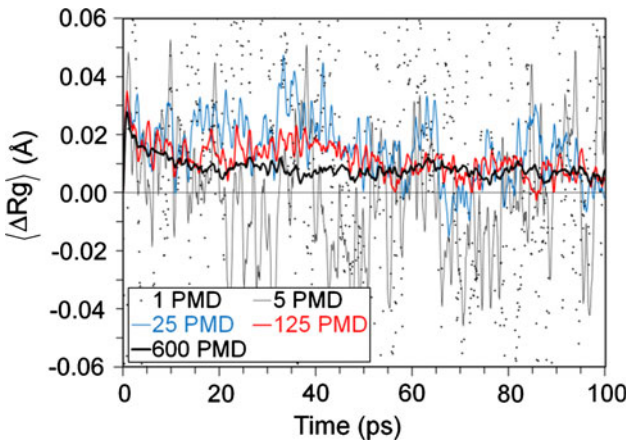

Fig. 4 R_{g} changes in a single PMD trajectory (1 PMD), ensemble-averaged over 5, 25, 125, and 600 PMD trajectories (5 PMD, 25 PMD, 125 PMD, and 600 PMD)

It can be understood that the fluctuation was significantly reduced by increasing the number of trajectories for the average. By averaging R_g over the ensemble of 600 values, the R_g change was naturally obtained with small fluctuation, and we can detect subtle changes to a precision of 0.03 Å. Such changes were almost completely obscured in a single PMD value due to fluctuations larger than 0.1 Å.

Next, to investigate the anisotropic structural relaxation after photolysis, three Cartesian components of the radius of gyration, R_{gx}, R_{gy}, and R_{gz}, were calculated by ensemble-averaging over 600 PMD trajectories (Fig. 5). Although they seem to exhibit just small fluctuations and different behaviors were clearly indicated for each component, those calculated by ensemble-averaging over the 600 UMD trajectories also exhibited varying behaviors even without photolysis: Mb expands in the *x*-direction and contracts in the *y*- and *z*-directions.

We need, therefore, a perturbative treatment additionally to cancel out the changes in the UMD trajectories as in Eq. 8. By applying the ensemble-averaging with the perturbative treatment, we were able to obtain statistically significant anisotropic structural dynamics of photolyzed MbCO (Fig. 6).

Comparing the three component values in Fig. 6, it was found that the behaviors of in-plane (R_{gx} and R_{gy}) and out-of-plane (R_{gz}) components are specifically different. While those in the former showed contractions after the short-time small expansions within 1 ps (peak at about 0.7 ps), the latter component clearly increased during the simulation time with a rapid increase up to 0.04 Å immediately after photolysis. This anisotropic structural change, which mainly occurs in the *z*-direction, was induced by two mechanisms: one is the displacement of the heme iron atom induced by the heme conformational change from a planar form to a domed one and the other is the photolyzed CO ligand collision with the distal

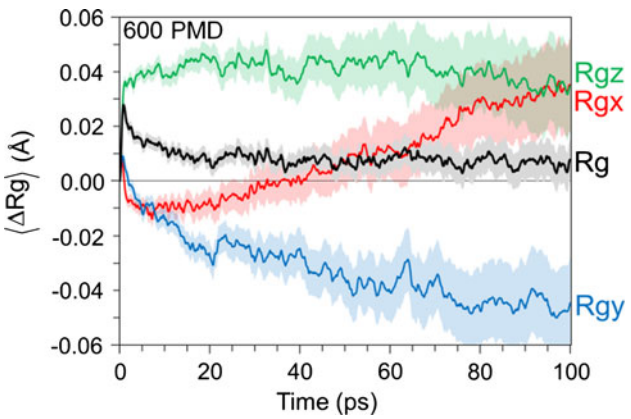

Fig. 5 R_g changes ensemble-averaged over 600 PMD trajectories. The 95% CIs of R_g changes are shown as painted regions

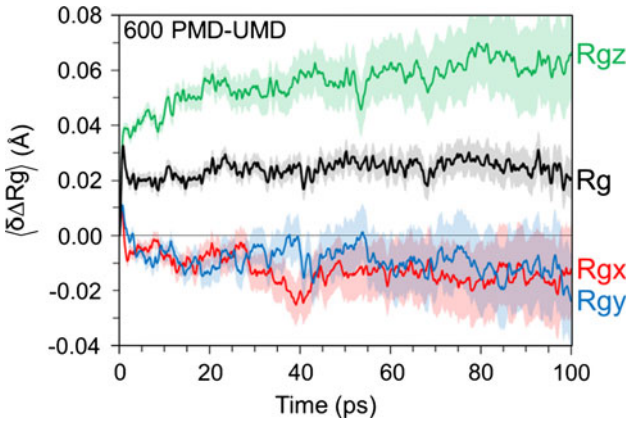

Fig. 6 R_g changes calculated by the EPMD method from 600 pairs of PMD and UMD trajectories. The 95% CIs of Rg changes are shown as painted regions

pocket residues. These push the proximal histidine or distal pocket residues outward along the *z*-direction and result in the *z*-direction expansion. Details of these displacements of heme pocket residues are provided in the next section.

It was found that most of the obtained anisotropic changes occur within 1 ps, followed by relatively small changes. This agrees with the results of TG spectroscopy by Goodno et al. [4]. They also demonstrated anisotropic conformational relaxation in a protein within 500 fs, whose out-of-plane (*z*-direction) expansion is larger than the in-plane (*x*- and *y*-directions) changes. However, their analysis could not definitively determine whether the in-plane changes were expansion or contraction because the observed signals were the sum of the contributions from the globin strain and the solvent heating. Since our calculations, meanwhile, include only the contribution from the structural change of Mb and reveal contraction in the *x*- and *y*-directions, the present observation may be reasonably considered affirmative evidence that Mb partially shows the physical properties of an elastic medium. That is to say, if Mb expands along one direction, it contracts at the same time along perpendicular directions to keep its atomic number density constant.

4.1.1 Clamshell rotation

By high-resolution X-ray crystallography [5], the most striking structural differences between MbCO and deoxy-Mb were detected in the E and F helices. That is, the F helix is displaced toward the proximal direction, and, at the same time, the E helix is displaced toward the heme by the removal of heme-bound CO ligand. These changes can be interpreted as a rotation of the two helices whose pivot is at the EF corner, called "clamshell rotation" (Fig. 1). As a

second application of the PEMD method [3], we investigated the clamshell rotation of photolyzed MbCO by analyzing the 1,232 in/N_δ conformation trajectory pairs described in Sect. 3.2.1.

While thermal fluctuations of each atomic position are about 1 Å at room temperature, atomic displacements between deoxyMb and MbCO are around 0.1 Å and smaller than the fluctuations. For instance, the heme iron atom displaced 0.3 Å, and the E and F helices translated 0.12 Å after a CO ligand binds to the heme [5]. It is, therefore, indispensable to cancel out the thermal fluctuations by the statistical averaging procedure over an ensemble of MD simulations. The number of trajectory pairs, 1,232, is sufficient to detect 0.1 Å structural changes brought about by the photolysis because most of the 95% CIs of the atomic displacements ensemble-averaged over the 1,232 trajectories range within ±0.04 Å and even the largest does within ±0.1 Å.

In Fig. 7, we superposed a pair of ensemble-averaged PMD and UMD structures calculated from Eqs. 16 and 17, respectively, at 20 ps after ligand photolysis (i.e., averaging 1,232 MD simulations). The observed structural changes are consistent with those obtained by taking the differences between the time-resolved X-ray diffraction structures determined after a 100-ps time delays reported in [6,7] and the high resolution structure reported in [5]. His93 was pushed away from the heme in the proximal direction, Leu29, Ile107, and Leu104 were displaced in the distal direction, and His64 moved into the void formed by the ligand dissociation. The fact that most of the residues around the heme displace in the direction perpendicular to the heme plane is also consistent with the anisotropic shape changes in photolyzed MbCO as discussed in the previous section.

Our PEMD method has a significant advantage over the time-resolved X-ray crystallography. The time resolution of the PEMD is the integration time-step for solving the equations of motion, 0.2 fs in this simulation. This is substantially higher than the resolution of the time-resolved X-ray crystallography, 100 ps. With this advantage, our simulation provides new insight into how Mb structure is perturbed immediately after ligand photolysis.

To investigate the clamshell rotation after ligand photolysis with high time resolution, the ensemble-averaged residue displacement as defined by Eq. 19 at 0.2, 0.6, and 20 ps after photolysis is plotted in Fig. 8. At the early stage of the relaxation at 0.2 ps, only the heme pocket residues (Leu29, Phe43, His64, Val68, and His93) exhibited large displacements, meaning that the structural perturbation is confined in the vicinity of the heme. After 0.6 and 20 ps, the perturbation propagated mainly around the initially displaced residues and, at last, displacements were found over the entire region of the protein.

This plot reflects the occurrence of the clamshell rotation in the E and F helices. The plot after 20 ps shows the V-shaped pattern from Val68 in the E helix to His93 in the F helix with the bottom at the EF corner. The displacement directions of both helices are consistent with the clamshell rotation (Fig. 1). The rotation pivot at the EF corner is also confirmed by the appearance of the bottom at the corner.

In the E helix, only His64 and Val68 exhibited large displacements at 0.2 ps (Fig. 8). The initial behaviors of the two residue side chains were similar. They were pushed away from the heme by the steric collision with the dissociated CO ligand, but the structure was transient with a short life time less than 0.5 ps. Although there is no direct experimental observation in this transient structure, the asymmetric conformational relaxation observed by TG spectroscopy in photolyzed MbCO within 500 fs [4] may reflect this transient escape of His64 and Val68.

Next, after 0.6 ps and later, the side chains of the two residues approached the heme to fill in the void formed by

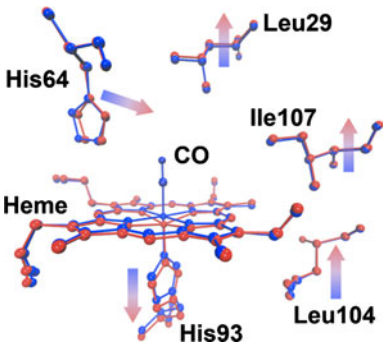

Fig. 7 Structures ensemble-averaged over 1,232 PMD (*red*) and UMD (*blue*) trajectories at 20 ps after ligand photolysis

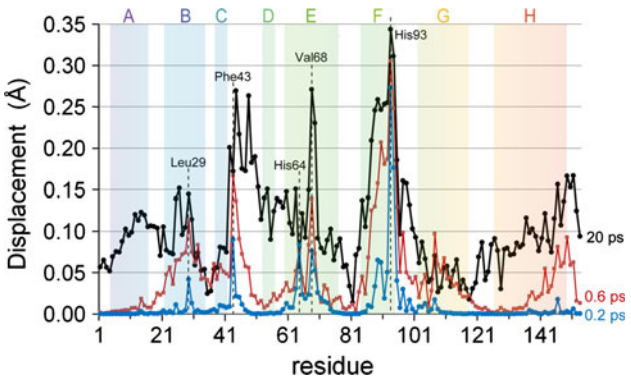

Fig. 8 Ensemble-averaged residue displacement of each residue calculated from 1,232 PMD and UMD trajectories at 0.2 (*blue*), 0.6 (*red*), and 20 ps (*black*) after ligand photolysis. Helix units (*A*–*H*) are drawn by *colored bands*

the ligand dissociation and kept their position during the 20-ps simulation. Subsequent to the two residue displacements toward the heme, the C-terminal side of the E helix (from Gly65 to Lys78) also moved toward the proximal direction. This collapse motion of the E helix driven by His64 and Val68 displacements is consistent with the clamshell rotation.

Moreover, our result is also consistent with the immediate decrease (within the instrument response time) of hydrophobicity around Trp14 observed by UV resonance Raman spectroscopy [9]. The inter-residue distance change (defined by Eq. 20) between Trp14 and Leu69 exhibited initial 0.036 ± 0.009 Å contraction at 0.2 ps and subsequent expansion up to 0.06 ± 0.02 Å within 0.8 ps.

4.2 Vibrational energy relaxation pathway of photolyzed MbCO

4.2.1 Immediate anisotropic relaxation pathway through projectile and through bond

In Fig. 9, using the 20,000 pairs of PMD and UMD trajectories described in Sect. 3.2.3, we plotted the time-resolved changes in the residue kinetic temperatures $\langle \delta T_r \rangle^L_{20,000}$ calculated by Eq. 27 accompanied by their 95% CIs. It should be noted that this is the first report of the time-resolved excess kinetic energy redistribution of photolyzed MbCO in the globin matrix with a statistically significant precision (±1 K) and high time resolution (0.1-ps time window). Because of the high precisions, we can

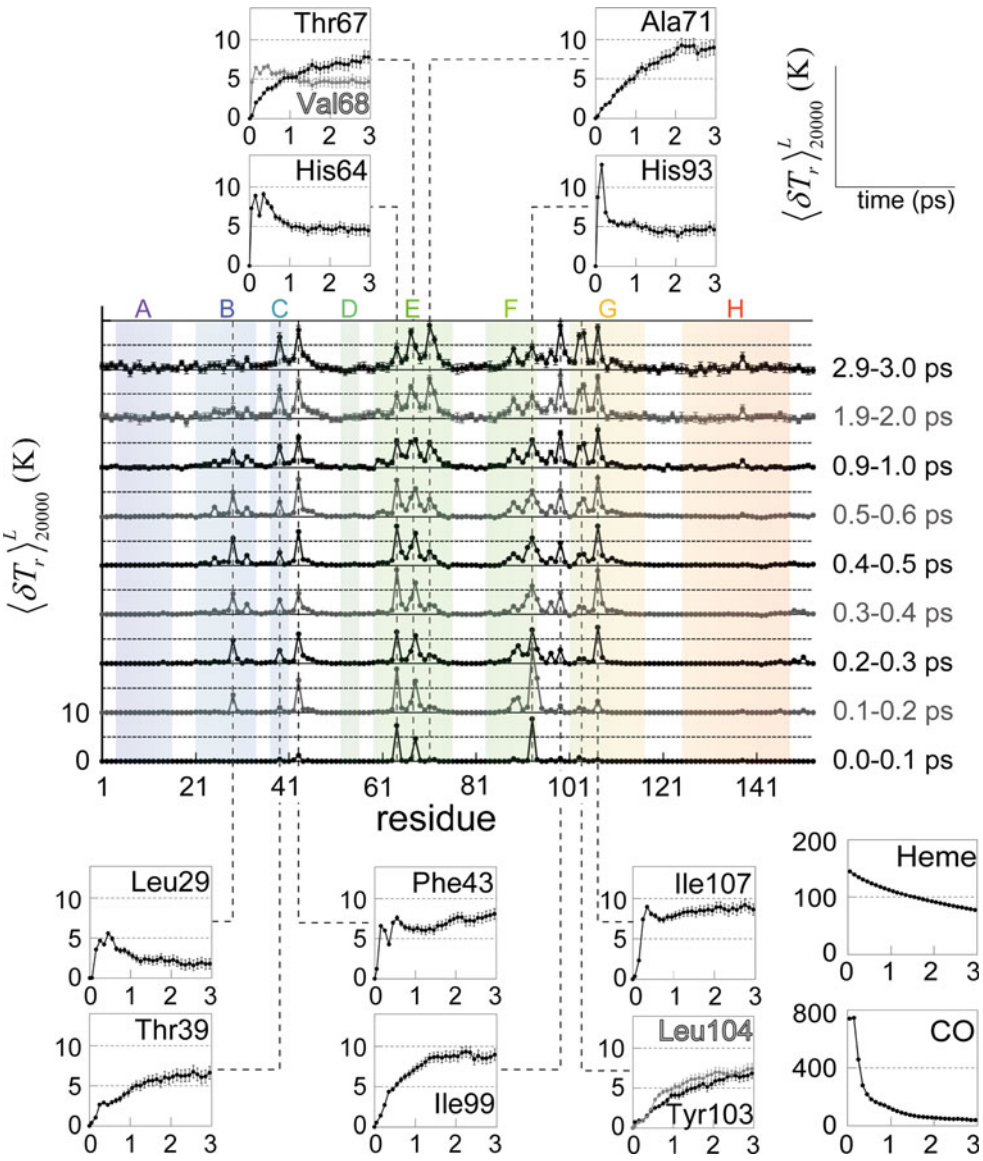

Fig. 9 (*Middle*) Distributions of the changes in the residue kinetic temperature calculated by Eq. 27 in each time window. The temperatures are time averaged over 0.1-ps time window and ensemble-averaged over 20,000 pairs of PMD and UMD trajectories; 95% CIs are shown as *error bars*. Helix units (*A*–*H*) are drawn by *colored bands* (*top* and *bottom*). Time-dependent residue kinetic temperature changes in several important residues including the heme and CO. The *horizontal axis* is time in picosecond, and the vertical one is the residue kinetic temperature changes in K

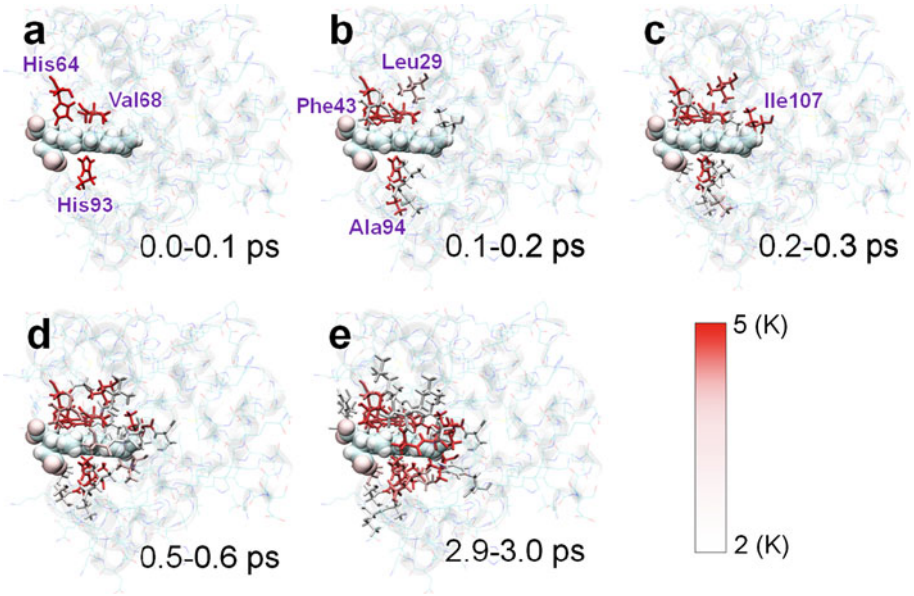

Fig. 10 Distributions of the changes in the residue kinetic temperature in the globin matrix. Those residues whose temperature increases were more than 2.0 degrees were highlighted in the *white-red* color scale

clearly detect "hot" residues to which the perturbation provides excess kinetic energy. The "hot" residues were visualized in Fig. 10 at $L = 0.0$, 0.1, 0.2, 0.5, and 2.9 ps (Eq. 27). To accentuate highly heated residues, those residues whose temperature increases were more than 2.0 K are highlighted in the white-red color scale.

To compare with previous studies of photolyzed MbCO, we first analyze the relaxation rate of the heme. The obtained high-precision kinetic temperature curve of the heme can be adequately modeled by a single exponential with a time constant of 2.16 ps. The constant is not different so significantly from those previously reported by MD simulation (i.e., single exponential with a time constant of $\sim$5.9 ps [14] and double exponential with time constants of 0.84 and 18.7 ps [15]) and TR^3 spectroscopy (double exponential with time constants of 3.0 ± 1.0 ps and 25 ± 14 ps [12]), indicating that our PE method should reproduce the relaxation process of photolyzed heme so reasonably.

Now, we report the time-dependent kinetic energy relaxation process. In the first time window $L = 0.0$ ps, there were only three highly heated residues, His64, Val68, and His93. In the next window $L = 0.1$ ps, Leu29, Phe43, Leu89, Ala90, and Ala94 additionally followed. In the third window $L = 0.2$ ps, further, Ile107 showed a significant increase. In contrast, those residues far from the heme do not show naturally so marked increase up to $L = 0.2$ ps window. For example, since A and H helices are far from the heme, there are very small increases (the largest increase is 0.43 ± 0.05 K at Tyr146 at $L = 0.2$ ps).

In the heme distal side, the heating process strongly correlates with the dissociated CO ligand behavior. After the photolysis, CO is pushed away from the heme iron atom and receives large kinetic energy (the increase in the residue kinetic temperature is 741 ± 5 K at $L = 0.0$ ps). The initially heated two residues, His64 and Val68, are both very close to the heme-bound CO (before photolysis, the nearest distances are 2.89 ± 0.02 Å and 3.30 ± 0.03 Å, respectively), and thus, the accelerated CO immediately collides with them. Other three residues, Leu29, Phe43, and Ile107, are farther from the CO (before photolysis, the nearest distances are 4.68 ± 0.03 Å, 3.67 ± 0.03 Å, and 5.34 ± 0.03 Å, respectively), and the collisions occur later than those of His64 and Val68. These collisions brought about the rapid drop in the residue kinetic temperature of CO to 182 ± 3 K at $L = 0.5$ ps (Fig. 9).

On the other hand, in the heme proximal side, the excess energy transfers through the covalent bond between the heme iron atom and the nitrogen atom of His93 side chain. By switching the heme force field from the six-coordinated planar structure to the five-coordinated domed one, the heme doming is induced within 0.2 ps after the photolysis, and, simultaneously, the covalently bonded His93 is pushed away, absorbing a part of the excess energy. Actually, the peak of the His93 residue kinetic temperature appears at $L = 0.1$ ps to show 12.84 ± 0.20 K (Fig. 9). This is the largest increase among all the residue kinetic temperature of amino acid residues. Then, several residues surrounding His93 (e.g., Leu89, Ala90, Ser92, Ala94, His97, and Ile99) show gradual heating.

As Sagnella and Straub suggested, it is convenient to categorize the relaxation pathway into the three, by which the direct energy transfer occurs differently from the heme to the protein [14]: (1) "through-projectile" (i.e., through collisions of the dissociated CO ligand molecule), (2) "through-bond" (i.e., through the covalent bond between the heme iron atom and the imidazole of His93), and (3)

"through-space" (i.e., through nonbonded collisions of the heme with other residues). According to the pathway categorization, the heated residues up to $L = 0.2$ ps can obviously be categorized into two groups. One group is located in the proximal side of the heme (His93, Leu89, Ala90, and Ala94) and the other is in the distal heme pocket (Leu29, Phe43, His64, Val68, and Ile107). The former residues showed a characteristic to absorb excess energy by "through-bond" pathway, while the latter did by "through-projectile" pathway.

As visualized in Fig. 10a–c, the anisotropic excess kinetic energy relaxation within 0.3 ps mainly occurs in the direction perpendicular to the heme plane by both "through-bond" and "through- projectile" pathways. This is consistent with the previous finding that MbCO shows the anisotropic expansion after the CO photolysis: Mb largely expands in the direction perpendicular to the heme plane, while slightly contracts in the horizontal plane [2, 5].

4.2.2 *Subsequent isotropic relaxation pathway through space*

After $L = 0.3$ ps, several residues (Thr39, Thr67, Ala71, Tyr103, and Leu104) revealed gradual temperature increases (Fig. 9). Because these residues are adjacent to the heme but do not directly collide with the dissociated CO, we can conclude that such temperature increases were driven by the "through-space" relaxation pathway. In addition, with respect to the residues significantly heated by the "through-projectile" pathway (Leu29, Phe43, His64, Val68, and Ile107), four out of the five residues (Phe43, His64, Val68, and Ile107) are also adjacent to the heme (i.e., the nearest distance between the residue and the heme are shorter than 3.8 Å) and should receive excess energy directly from the heme. In fact, they showed more than 4 K increase during the 3-ps MD simulation. The only exception is Leu29, which is 6.8 Å away from the heme, and thus, there is no direct energy relaxation by the "through-space" pathway, resulting in cooling from the peak 5.58 ± 0.23 K at $L = 0.4$ ps to 1.79 ± 0.66 K at $L = 2.9$ ps.

The energy relaxation by the "through-bond" and "through-projectile" pathways occurs very fast within 0.3 ps as discussed in the previous section. On the contrary, the "through-space" relaxation is relatively slow. The most heated residues at $L = 2.9$ ps are Ala71 (9.00 ± 0.93 K) and Ile99 (8.96 ± 0.71 K) and receive excess energy via the "through-space" pathway.

In contrast to the anisotropic relaxation by the "through-projectile" and "through-bond" pathways, the "through-space" relaxation is isotropic around the heme. From a comparison between Fig. 10c and e, we can recognize that the residues adjacent to the heme and coplanar to the heme plane (Thr39, Thr67, Ala71, Ile99, Tyr103, and Leu104) are significantly heated from 0.3 to 2.9 ps. It is worth mentioning that this relaxation to the adjacent residues was predicted by calculating the frequency-resolved communication maps for Mb [20]. While the maps at low frequencies (below ~ 350 cm^{-1}) indicated significant energy flow between the heme and residues of C, E, F, and G helices, those at high frequencies (above ~ 350 cm^{-1}) did no efficient energy transport from the heme to other residues. According to this prediction, it is likely that low-frequency vibrations play a dominant role in the isotropic "through-space" relaxation.

On the other hand, both experimental [18] and theoretical [16] studies of H93G mutants concluded that the covalent bond between the heme and His93 does not play a primary role in the process of heme cooling for Mb because the mutation of His93 little affected the heme relaxation time. It is, from the present statistically significant residue temperature data, that we can directly investigate the efficiency of the "through-bond pathway" in transferring the excess vibrational energy by comparing the temperature increase in His93 with those of other residues. Despite the direct covalent bond, it is then found that the temperature increase in His93 at $L = 2.9$ ps was only 4.60 ± 0.70 K, which is about half those of Ala71 and Ile99 (Fig. 9). This directly indicates that the energy transfer by the "through-bond" pathway is inefficient compared with that by "through-space". The "through-bond" relaxation is only important immediately after photolysis.

5 Concluding remarks

In this study, we outlined the PEMD method, in which many pairs of PMD and UMD simulations are executed for ensemble-averaging to cancel out thermal fluctuations. Then, as an application of the PEMD method, the incipient relaxation process of photolyzed MbCO was analyzed using such statistically significant results.

First, the PEMD method was used to elucidate the anisotropic structural change in photolyzed MbCO. With 600 pairs of PMD and UMD trajectory pairs, the changes in R_g and its three Cartesian components R_{gx}, R_{gy}, and R_{gz} induced by the ligand photolysis were obtained with a satisfactory precision 0.03 Å. The results drawn from the method agree very well with the previous observation of the transient grating experiment [5], indicating the anisotropic expansion within 500 fs; that is, Mb largely expands in the direction perpendicular to the heme plane and slightly contracts in the other directions.

Second, we investigated the structural relaxation process focusing on the "clamshell rotation" in the E and F helices with 1,232 trajectory pairs. As for the residues in the

vicinity of the heme, the ensemble-averaged structures of PMD and UMD after 20 ps of the ligand photolysis are consistent with those by time-resolved X-ray diffraction at 100-ps delay time [6, 7]. The structural change at 20 ps is consistent with the clamshell rotation: E and F helices displaced in the proximal direction and the rotation pivot at the EF corner is confirmed. Moreover, with the advantage of our PEMD method, we could detect the structural change with a time resolution higher than that of the time-resolved X-ray crystallography, 100 ps. At the early stage of the relaxation after 0.2 ps of photolysis, only the heme pocket residues (Leu29, Phe43, His64, Val68, and His93) exhibit large displacements. In collision with the dissociated CO ligand, His64 and Val68 experience the transient structures whose life time is less than 0.5 ps.

Third, we analyzed the excess energy relaxation pathway in photolyzed MbCO. This is the first report of the time-resolved excess kinetic energy redistribution of photolyzed MbCO in the globin matrix with a statistically significant precision (± 1 K) and high time resolution (0.1-ps time window). In fact, the largest difficulty in analyzing the relaxation is the thermal fluctuations (~ 70 K), which are much larger than the kinetic temperature increase by the photolysis (<10 K). In this case, we needed a considerable number of MD trajectory pairs, 20,000, to adequately cancel out the thermal fluctuations. It was confirmed that the number, 20,000, is adequate enough for ensemble-averaging, resulting in 95% CIs less than 1.07 K for residue kinetic temperatures. The statistically significant results demonstrated that incipient excess energy dissipation occurred mainly in the direction perpendicular to the heme plane. The incipient relaxation is dominated by the "through-bond" and "through-projectile" pathways. Then, the isotropic relaxation follows by the "through-space" pathway. From the fact that the residue temperature increase in His93 is about half of those of other residues that efficiently absorb vibrational energy by the "through-space" pathway, it was directly confirmed that the covalent bond between the heme and His93 does not play a primary role in the heme cooling process.

In this work, the relaxation process of photolyzed MbCO was analyzed by MD simulations, and thus, the obtained results should be influenced by the force-field parameters. We adopted "Unliganded hard" potential [27] for the electronically excited heme whose deformation energy is 10 kcal/mol. With respect to the interaction between the electronically excited heme and CO ligand, the repulsive potential with 8.7 kcal/mol dissociation energy was used. These deformation and dissociation energies bring about the fast kinetic energy relaxation to the heme pocket residues by "through-bond" and "through-projectile" pathways, respectively. Therefore, the amount of energy distributed in the heme pocket residues immediately after photolysis (<0.3 ps in Fig. 9) must be sensitive to the force-field parameters. A careful calibration of the parameters, especially for those of the electronically excited heme, by using high-level MO calculation must improve the quantitative accuracy of the amount of the distributed energy. Although the parameters used in this work might not be optimal in the context of quantitative accuracy, it can be concluded that our results decently unveiled the timescale and pathways of the energy relaxation because such qualitative aspects would not be strongly affected by the amount of energy.

Finally, we summarize the advantages of the PEMD method in analyzing the relaxation processes of proteins. One significant advantage is the high time resolution that is limited only by the integration time-step for solving the equations of motion, usually less than a few femtoseconds. As was shown in the clamshell rotation analysis, this resolution is much higher than those of experiments, making shed light theoretically on the subdynamics in the shorter timescale. Another important advantage is that we can control the precisions of ensemble-averaged quantities by adopting an appropriate number of trajectory pairs depending on the desirable precisions. Moreover, because the calculation of each MD simulation is independent of all others, it is possible to calculate the simulations concurrently on many computers and, therefore, we have utilized a large number of MD trajectory pairs (20,000 pairs) to analyze the vibrational energy redistribution with a high precision of ± 1 K. It is anticipated in the future that, from large samplings of MD trajectories, the PEMD method should extract some hidden protein dynamics unrecognized in the present experimental observation.

Acknowledgments This work was partially supported by a grant-in-aid for the 21st Century COE program "Frontiers of Computational Science" at Nagoya University and also by a grant-in-aid for Science Research from the Ministry of Education, Culture, Sport, Science and Technology in Japan and the Core Research for Evolutional Science and Technology (CREST) "High Performance Computing for Multiscale and Multi-physics Phenomena" from the Japan Science and Technology Agency.

References

1. Leitner DM, Straub JE (2009) Proteins: energy, heat and signal flow (computation in Chemistry). CRC Press, Boca Raton
2. Takayanagi M, Okumura H, Nagaoka M (2007) J Phys Chem B 111:864
3. Takayanagi M, Iwahashi C, Nagaoka M (2010) J Phys Chem B 114:12340
4. Goodno GD, Astinov V, Miller RJD (1999) J Phys Chem A 103:10630
5. Kachalova GS, Popov AN, Bartunik HD (1999) Science 284:473
6. Schotte F, Lim M, Jackson TA, Smirnov AV, Soman J, Olson JS, Phillips GN Jr, Wulff M, Anfinrud PA (2003) Science 300:1944

7. Schotte F, Soman J, Olson JS, Wulff M, Anfinrud PA (2004) J Struct Biol 147:235
8. Hummer G, Schotte F, Anfinrud PA (2004) Proc Natl Acad Sci USA 101:15330
9. Sato A, Gao Y, Kitagawa T, Mizutani Y (2007) Proc Natl Acad Sci USA 104:9627
10. Guallar V, Jarzecki AA, Friesner RA, Spiro TG (2006) J Am Chem Soc 128:5427
11. Henry ER, Eaton WA, Hochstrasser RM (1986) Proc Natl Acad Sci USA 83:8982
12. Mizutani Y, Kitagawa T (1997) Science 278:443
13. Mizutani Y, Kitagawa T (2002) Bull Chem Soc Jpn 75:623
14. Sagnella DE, Straub JE (2001) J Phys Chem B 105:7057
15. Okazaki I, Hara Y, Nagaoka M (2001) Chem Phys Lett 337:151
16. Bu L, Straub JE (2003) J Phys Chem B 107:10634
17. Zhang Y, Fujisaki H, Straub JE (2007) J Phys Chem B 111:3243
18. Ye X, Demidov A, Rosca F, Wang W, Kumar A, Ionascu D, Zhu L, Barrick D, Wharton D, Champion PM (2003) J Phys Chem A 107:8156
19. Koyama M, Neya S, Mizutani Y (2006) Chem Phys Lett 430:404
20. Leitner DM (2009) J Chem Phys 130:195101
21. Lian T, Locke B, Kholodenko Y, Hochstrasser RM (1994) J Phys Chem 98:11648
22. Ota N, Agard DA (2005) J Mol Biol 351:345
23. Sharp K, Skinner JJ (2006) Proteins Struct Funct Bioinform 65:347
24. Botan V, Backus EHG, Pfister R, Moretto A, Crisma M, Toniolo C, Nguyen PH, Stock G, Hamm P (2007) Proc Natl Acad Sci USA 104:12749
25. Nguyen PH, Park S-M, Stock G (2010) J Chem Phys 132:025102
26. Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA (2006) AMBER 9. University of California, San Francisco
27. Henry ER, Levitt M, Eaton WA (1985) Proc Natl Acad Sci USA 82:2034
28. Giammona DA (1984) Ph.D. thesis, University of California, Davis
29. The Chemical Society of Japan (ed) (1993) Kagaku-binran (Handbook of Chemistry) Basic volume, 4th ed. II. Maruzen, Tokyo (in Japanese)
30. Cheng BX, Schoenborn BP (1990) Acta Cryst B46:195
31. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926
32. Tironi IG, Brunne RM, van Gunsteren WF (1996) Chem Phys Lett 250:19
33. Evans DJ, Morriss G (2008) Statistical mechanics of nonequilibrium liquids, 2nd edn. Cambridge University Press, Cambridge