Protein Tertiary Structural Changes Visualized by Time-Resolved X-ray Solution Scattering

Sena Ahn, Kyung Hwan Kim, Youngmin Kim, Jeongho Kim and Hyotcherl Ihee*

Center for Time-Resolved Diffraction, Department of Chemistry, Graduate School of Nanoscience & Technology (WCU), KAIST, Daejeon 305-701, Korea

Supporting Information

Data collection

Details of data collection are described in a previous publication.[1] Briefly, TR-WAXS data were collected on beamline ID09B at the ESRF using a pump-probe scheme. Nanosecond laser pulses (4 ns fwhm, 532 nm, 1 mJ/mm^2) were used to initiate CO photolysis in sperm whale MbCO and delayed X-ray pulses (5×10^8 photons per pulse, 100 ps (fwhm), 3% bandwidth around the wavelength of 0.83 Å) from the synchrotron, were used to investigate the evolving structures in the sample. Scattering data were collected for the time delay of 10 ns and were interleaved by a measurement at -50 μs, which served as a reference for the unperturbed sample.

X-ray scattering curve calculation

CRYSOL is a program to evaluate solution scattering from macromolecules with known atomic structures.[2] We used this program to calculate the scattering intensity in the MD-restrained rigid-body modeling. This program uses multiple expansion (spherical harmonic expansion) of the scattering amplitudes to calculate the spherically averaged scattering pattern. The scattering intensity by CLYSOL has three components, namely scattering of atoms in vacuo $A_a(\mathbf{q})$, scattering from the excluded volume $\rho_s A_s(\mathbf{q})$, and scattering from hydration shell $\delta\rho_b A_b(\mathbf{q})$.

The scattering intensity evaluate

$$I(q) = < \left| A_a(\mathbf{q}) - \rho_s A_s(\mathbf{q}) + \delta\rho_b A_b(\mathbf{q}) \right|^2 >_\Omega \qquad (1)$$

Here, the electron density of the bulk solvent is ρ_s, that of the hydration shell is ρ_b, the contrast of the shell and bulk solvent is $\delta\rho_b = (\rho_b - \rho_s)$, and $< >_\Omega$ stands for the average over all orientations for the momentum transfer vector $\mathbf{q}$ and Ω is the solid angle in reciprocal space.

Actually, there are two common approaches to solid angle averaging over the exponential terms of the scattering amplitude, one is by the application of spherical harmonics expansion and the other is by the Debye formula.[3] The scattering amplitudes of hydration shell are difficult to describe by the Debye formula, and instead the expression of the molecular surface shape by the expansion of spherical harmonics is often used. In spite of this weakness the Debye formula is often used without including the

hydration shell contribution because of the mathematical simplicity and the relatively very small contribution of border layer to the total scattering intensity.[4-7]

The scattering from solute atoms $A_a(\mathbf{q})$ and solvent by the excluded volume $A_s(\mathbf{q})$ for N particles is written as

$$A_a(\mathbf{q}) = \sum_{j=1}^{N} f_j(q) e^{i\mathbf{q}\cdot\mathbf{r_j}},$$
$$A_s(\mathbf{q}) = \rho_s \sum_{j=1}^{N} g_j(q) e^{i\mathbf{q}\cdot\mathbf{r_j}} \quad (2)$$

where $f_j(q)$ is the form factor of the *j-th* atom and $g_j(q)$ is the form factor of the solvent displaced by the *j-th* atom. If we use Debye formula with solvent-corrected form factor

$$f_j^s(q) = f_j(q) - \rho_s g_j(q), \quad (3)$$

then the scattering intensity can be expressed as follows (ignoring the hydration shell):

$$I(q) = <\left|A_a(\mathbf{q}) - A_s(\mathbf{q})\right|^2>_\Omega = \sum_{i=1}^{N}\sum_{j=1}^{N} f_i^s(q) f_j^s(q) \frac{\sin(qr_{ij})}{qr_{ij}} \quad (4)$$

Here, we introduced the Debye formula to calculate the differential of the χ^2 potential (see below) which contains a theoretical scattering function in the MD-restrained rigid-body modeling, because we used the Debye formula instead of CRYSOL for the simplicity assuming the hydration layer does not fluctuate in the solution phase.

The following R factor can be also used as a criterion of agreement between the experimental and theoretical signals.

$$R = \frac{\sum_q \frac{1}{\sigma_q}\left|I_{exp}(q) - I_{cal}(q)\right|}{\sum_q \frac{1}{\sigma_q}\left|I_{exp}(q)\right|} \quad (5)$$

MD-restrained rigid-body modeling

In MD-restrained rigid-body modeling, a crystal structure (pdb file) is used as a starting point for MD simulations. The crystal structure is divided into a number of rigid bodies which consist of one or two alpha helices and hemes. The rigid bodies are allowed to move under the influence of the chemical and χ^2 force fields.[4,8] Since the atomic structure within a rigid body is constrained to be the same as that of the crystal structure, the force field within the rigid body is not needed, but only the Van der Waals interactions between rigid bodies are included in the chemical force field. The χ^2 force field is introduced to drive the molecular structure generated by MD simulations toward a structure that gives a difference scattering curve that matches the experimental difference scattering curve. In other words, the total potential (U) on the rigid bodies has the Van der Waals term (U_{LJ}) and the χ^2 term as follows:

$$U = c_1 U_{LJ} + c_2 \chi^2 \tag{6}$$

where c_1 and c_2 are weight parameters that scale the magnitude of the two terms appropriately. U_{LJ} and χ^2 which defines the quality of agreement between experimental data and theoretical values are as follows with the scaling factor c_s and reduced units:.

$$U_{LJ} = \sum_{i=1}^{N} \sum_{j \neq i}^{N} 4\left(\frac{\sigma^{12}}{r_{ij}^{12}} - \frac{\sigma^{6}}{r_{ij}^{6}} \right) \tag{7}$$

$$\chi^2 = \sum_{q=1}^{N_q} \left(\frac{I_{\exp}(q) - c_s I_{cal}(q)}{\sigma_q} \right)^2 \tag{8}$$

The force field is the gradient of the total potential, and therefore the total force acting on the *i*th particle among total number of particles, *N* is as follows:

$$\begin{aligned} \mathbf{f}_i &= -\nabla_i U \\ &= 24 c_1 \sum_{j \neq i}^{N} \left(2\frac{\sigma^{12}}{r_{ij}^{13}} - \frac{\sigma^{6}}{r_{ij}^{7}} \right) \\ &+ 2 c_2 c_s \sum_{q=1}^{N_q} \left(\frac{I_{\exp}(q) - c_s I_{cal}(q)}{\sigma_q^2} \right) \sum_{j \neq i}^{N} f_i^s(q) f_j^s(q) \left(\cos(q r_{ij}) - \frac{\sin(q r_{ij})}{q r_{ij}} \right) \frac{\mathbf{r}_{ij}}{r_{ij}^2} \end{aligned} \tag{9}$$

In the L-J potential, σ is defined as the r value where the corresponding potential is zero and the $\mathbf{r}_{ij}$ (= $\mathbf{r}_i$ - $\mathbf{r}_j$) is the distance between particle *i* and particle *j*. If the distance between two particles is smaller than σ, the repulsion increases steeply. We defined the following three types of σ values for the L-J potential between rigid bodies as follows: For the N-C bond between two rigid bodies consisting of helices, $\sigma_{N\text{-}C}$ = 1.28Å, for the bond between the heme domain and the histidine, $\sigma_{Hem\text{-}N}$ = 2Å , and for the atoms between two rigid bodies, $\sigma_{a\text{-}a}$ = 1.2Å.

Once the total forces between rigid bodies are determined, MD simulation runs based on Newtonian equations. The coordinate of the center of mass (COM) of the rigid body is:

$$\mathbf{R} = \frac{\sum m_i \mathbf{r}_i}{M} \ , \tag{10}$$

where m_i and $\mathbf{r}_i$ are the mass and position of the *i*th atom of the rigid-body and M is the mass of a rigid body ($M = \sum_i m_i$).

The COM moves translationally and the relative rotational motion with respect to the COM determines the relative position ($\mathbf{r}_i' = \mathbf{r}_i - \mathbf{R}$) of the *i*-th atom referenced to **R**, $\mathbf{r}_i'$. At each step of the MD simulation, **R,** $\mathbf{r}_i$ and $\mathbf{r}_i'$ are updated. The equation for the translational motion of the COM is as follows:

$$\ddot{\mathbf{R}} = \frac{\mathbf{F}}{M} \tag{11}$$

where $\mathbf{F} = \sum_i \mathbf{f}_i$ is total force on the rigid-body as the sum of $\mathbf{f}_i$ acting on the *i*-th atom within a rigid-body. The total torque acting on each rigid body with respect to the COM is as follows:

$$\mathbf{N} = \sum_i \mathbf{r}_i' \times \mathbf{f}_i \,. \tag{12}$$

The velocity of $\mathbf{r}_i'$ relative to **R** is as follows:

$$\dot{\mathbf{r}}_i' = \boldsymbol{\omega} \times \mathbf{r}_i' \tag{13}$$

where $\boldsymbol{\omega}$ is the angular velocity with respect to the COM of the rigid body.

The moment of inertia with respect to the COM is as follows:

$$\mathbf{I} = \begin{bmatrix} \sum m_i(y_i'^2 + z_i'^2) & -\sum m_i x_i' y_i' & -\sum m_i x_i' z_i' \\ -\sum m_i x_i' y_i' & \sum m_i(x_i'^2 + z_i'^2) & -\sum m_i y_i' z_i' \\ -\sum m_i x_i' z_i' & -\sum m_i y_i' z_i' & \sum m_i(x_i'^2 + y_i'^2) \end{bmatrix} \tag{14}$$

The rotational equation of motion of a rigid body around **R** can be expressed as[9]:

$$\dot{\boldsymbol{\omega}} = \mathbf{I}^{-1} \cdot (\mathbf{N} - \boldsymbol{\omega} \times (\mathbf{I} \cdot \boldsymbol{\omega})) \tag{15}$$

Eq. (10) updates **R** and combining eqns. (12)~(14) updates $\mathbf{r}_i'$.

References

(1) Cammarata, M.; Levantino, M.; Schotte, F.; Anfinrud, P. A.; Ewald, F.; Choi, J.; Cupane, A.; Wulff, M.; Ihee, H. *Nat. Methods* **2008**, *5*, 881.
(2) Svergun, D.; Barberato, C.; Koch, M. H. J. *J. Appl. Crystal.* **1995**, *28*, 768.
(3) Guinier, A. *X-Ray Diffraction: In Crystals, Imperfect Crystals, and Amorphous Bodies*; Dover Publications: New York, 1994.
(4) Grishaev, A.; Wu, J.; Trewhella, J.; Bax, A. *J. Am. Chem. Soc.* **2005**, *127*, 16621.
(5) Tiede, D. M.; Zhang, R.; Seifert, S. *Biochemistry* **2002**, *41*, 6605.
(6) Zhang, R.; Thiyagarajan, P.; Tiede, D. M. *J. Appl. Cryst.* **2000**, *33*, 565.
(7) Chacon, P.; Moran, F.; Diaz, J. F.; Pantos, E.; Andreu, J. M. *Biophys. J.* **1998**, *74*, 2760.
(8) Kojima, M.; Timchenko, A. A.; Higo, J.; Ito, K.; Kihara, H.; Takahashi, K. *J. Appl. Cryst.* **2004**, *37*, 103.
(9) Goldstein, H. *Classical Mechanics*; Addison-Wesley Publishing Company, Inc., 1980.

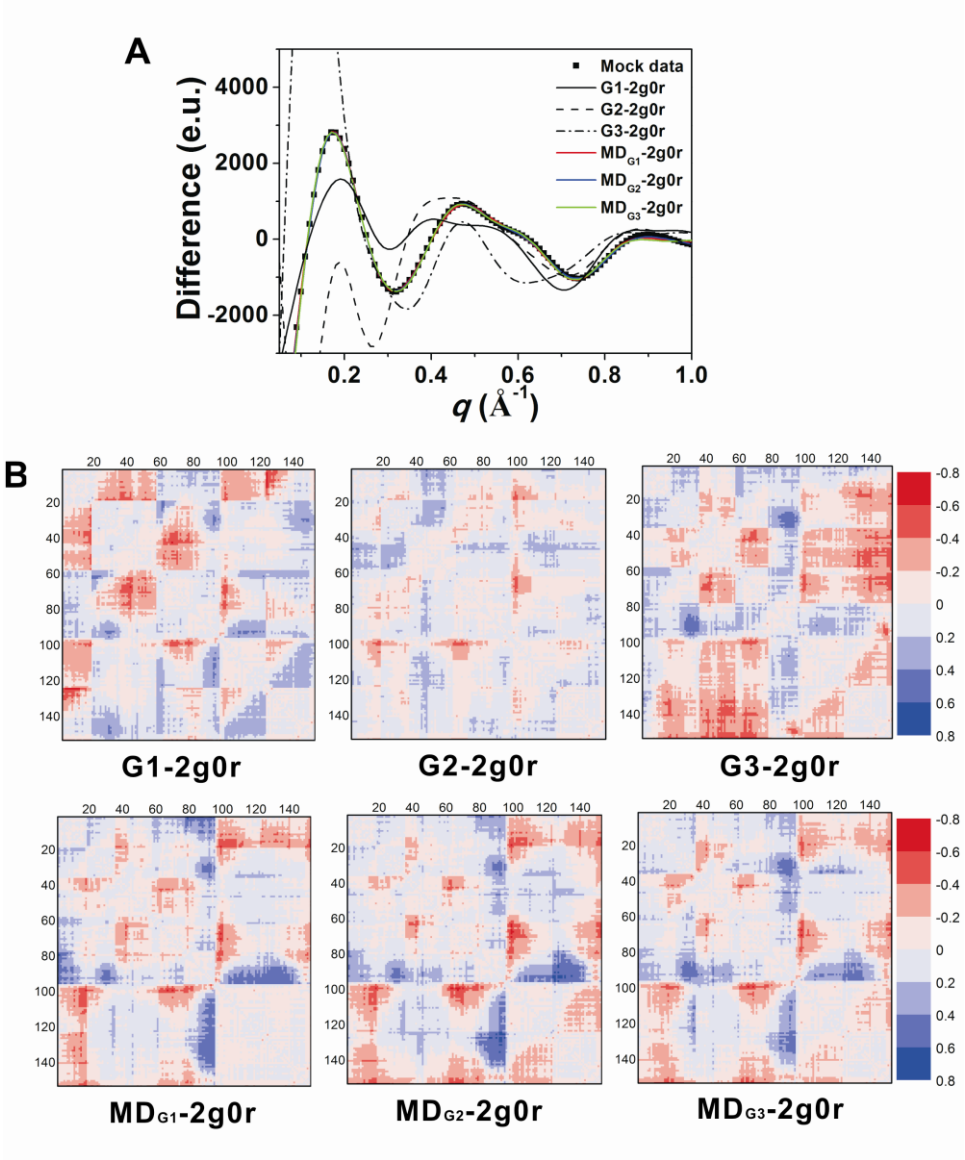

Figure S1. (A) Difference scattering curves before and after rigid-body-restrained MD simulations. The mock data was generated by subtracting the scattering curve of a crystallographic model for MbCO (2g0r) from the scattering curve of a Mb structure modified from 2g0v. We also generated structural variants (G1, G2 and G3) by modifying 2g0v as starting structures. After experiment-restrained MD simulations the final structures (MD_{G1}, MD_{G2} and MD_{G3}) converged to the correct structure with RMSD values less than 0.1 Å. (B) (Top) Comparison of the difference distance maps of G1, G2, and G3 structures. The 2g0r structure is the reference. (Bottom) Comparison of the difference distance maps of MD_{G1}, MD_{G2} and MD_{G3} structures. The difference distance maps for G1, G2 and G3 show clearly different patterns whereas those for the final structures (MD_{G1}, MD_{G2} and MD_{G3}) are almost identical, indicating that after experiment-restrained MD simulations the final structures (MD_{G1}, MD_{G2} and MD_{G3}) converge to the correct structure.

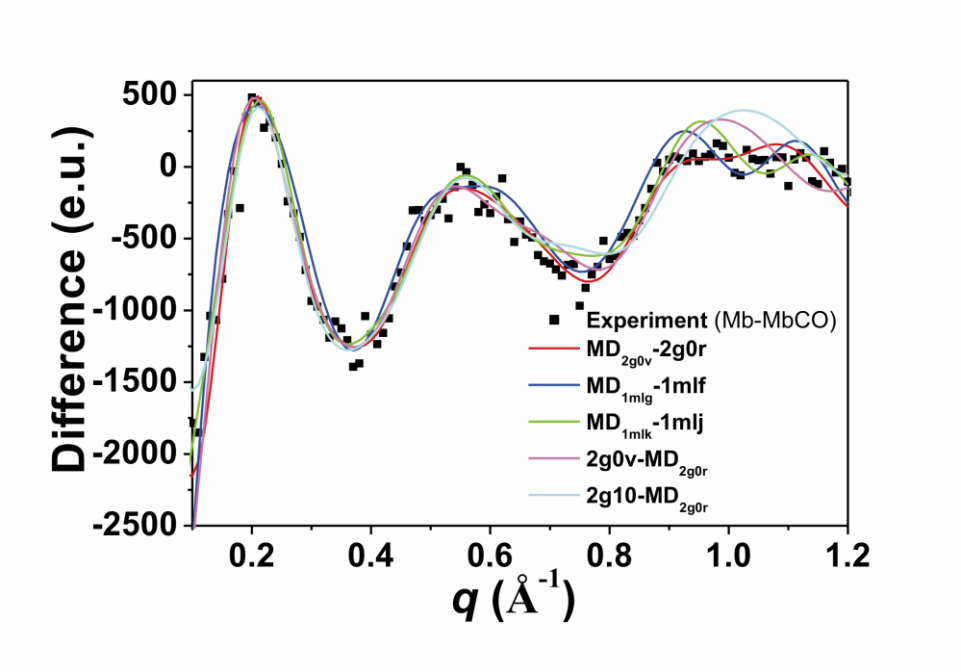

Figure S2. The result curves after MD-restrained rigid-body modeling to the experimental data using various crystal structures as starting structures. Each legend designates the fixed structure and varied structure in each modeling process. The best agreement was obtained for the case where MbCO was fixed at 2g0r and Mb was fit starting from 2g0v. *R* factors are 0.20, 0.30, 0.24, 0.27, 0.33 from to red and cyan.

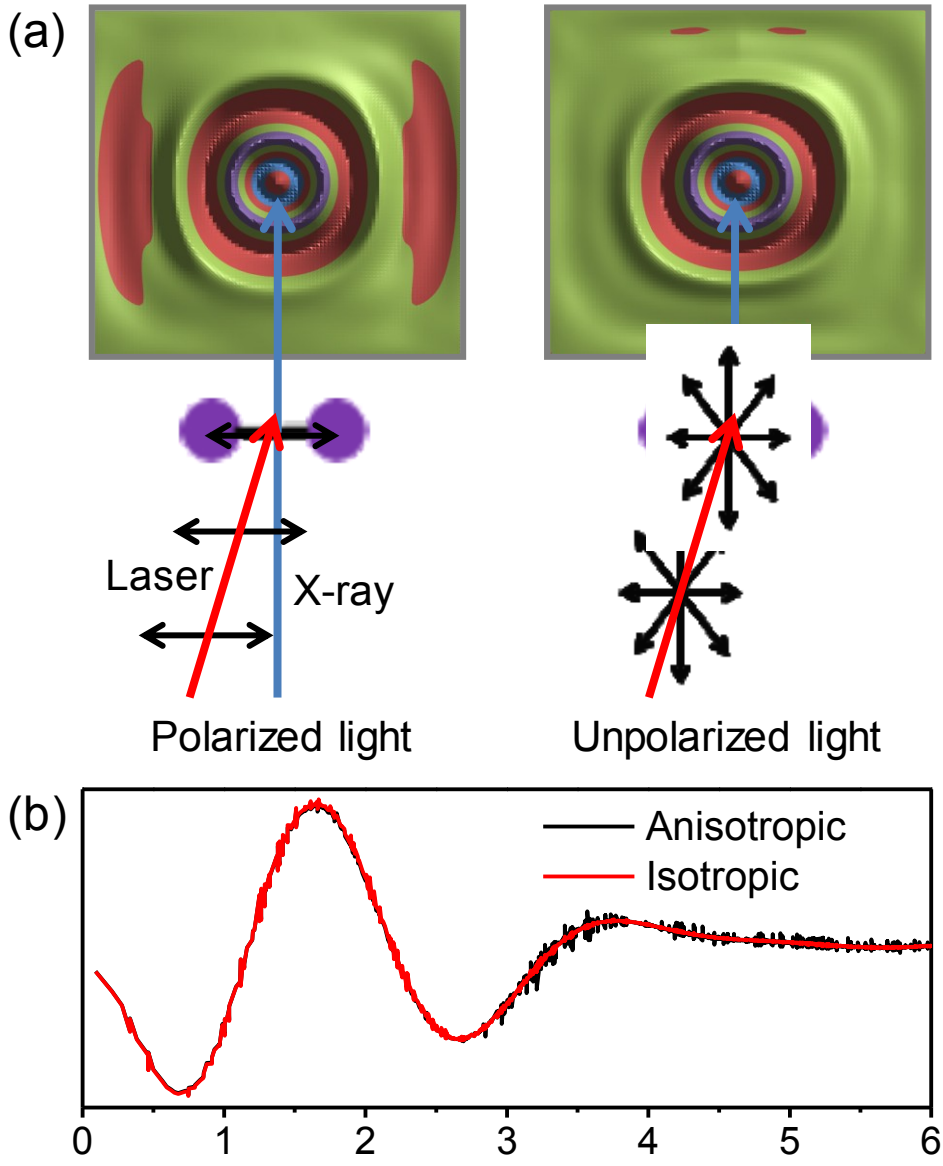

Figure S3. (a) The 2D X-ray diffraction images are obtained after photoexcitation of iodine molecule by either linearly polarized (left) or unpolarized (right) laser light. In the case of polarized excitation, the anisotropy is distinct along the direction of polarization, while no anisotropy is observed with unpolarized excitation. (b) Comparison of 1D diffraction curves obtained by ring-integrating the two 2D images shown in (a) along the perimeter. The 1D diffraction curves obtained with linearly polarized (black) and unpolarized (red) photoexcitation exhibit no difference, indicating that the anisotropic information is wiped out by integration. However, the reduction of dimensionality to 1D curve will simplify and facilitate the analysis of measured dynamics.