J|A|C|S
ARTICLES
Published on Web 09/21/2004

Cooperative Motions of Protein and Hydration Water Molecules: Molecular Dynamics Study of Scytalone Dehydratase

Noriaki Okimoto,*,† Takashi Nakamura,‡ Atsushi Suenaga,§ Noriyuki Futatsugi,§ Yoshinori Hirano,† Isamu Yamaguchi,‡ and Toshikazu Ebisuzaki†

Contribution from the Computational Astrophysics Laboratory, Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan, Laboratory for Remediation Research, Environmental Plant Research Group, Plant Science Center, RIKEN, 1-7-22 Suehiro-cho, Tsurumi, Yokohama, Kanagawa 230-0045, Japan, and High Performance Biocomputing Research Team, Bioinformatics Group, Genomic Sciences Center, RIKEN, 61-1 Ono-cho, Tsurumi, Yokohama, Kanagawa 230-0046, Japan

Received April 5, 2004; E-mail: okimoto@gsc.riken.go.jp

Abstract: Two molecular dynamics (MD) simulations totaling 25 ns of simulation time of monomeric scytalone dehydratase (SD) were performed. The enzyme has a ligand-binding pocket containing a cone-shaped $\alpha+\beta$ barrel, and the C-terminal region covers the binding pocket. Our simulations clarified the difference in protein dynamics and conformation between the liganded protein and the unliganded protein. The liganded protein held the ligand molecule tightly and the initial structure was maintained during the simulation. The unliganded protein, on the other hand, fluctuated dynamically and its structure changed largely from the initial structure. In the equilibrium state, the binding pocket was fully solvated by opening of the C-terminal region, and the protein dynamics was connected with hydration water molecules entry into and release from the binding pocket. In addition, the cooperative motions of the unliganded protein and the hydration water molecules produced the path through the protein interior for ligand binding.

Introduction

Proteins fluctuate dynamically in an aqueous solution, and protein function is related to the fluctuation of the protein. Protein dynamics and conformation are influenced by various factors such as the binding of ligand molecules and dynamics of the solvent. Many X-ray studies on proteins in liganded and unliganded states have provided evidence that protein conformational change and its function are closely linked. These studies suggest that functional large-scale conformational change, involving protein domains and secondary structures, occurs upon ligand binding. Although there have been many studies on ligand binding, it has not been fully elucidated. For soluble proteins such as enzymes and transport proteins, the effect on protein dynamics and conformation of solvent water molecules that have unusual physical properties originating from the hydrogen bond network is very important and cannot be ignored. The effects of the water solution on protein dynamics and conformation have therefore been investigated using both experimental and theoretical methods.[1−7] A recent study has shown that migration of hydration water molecules occurred in coupling with domain movement, and this finding is a typical example that motions of protein domains and hydration water molecules are closely related.[8] Study of protein dynamics in association with the behavior of hydration water molecules is therefore necessary for a better understanding of protein function.

In this work, MD simulations of a soluble protein, focusing on the dynamics of the protein and solvent water molecules, were performed. We selected SD as a research target. SD is one of the enzymes involved in fungal melanin biosynthesis in the phytopathogenic fungus, *Pyricularia oryzae*, which causes rice blast disease. Melanin is essential for infection of host rice leaves, and SD is therefore an attractive target for fungicide design.[9] The enzyme, which is composed of 172 amino acid residues, has a unique folded conformation including a cone-shaped $\alpha+\beta$ barrel motif (Figure 1A) and forms a trimer in X-ray crystal structure. SD has a hydrophobic ligand-binding pocket. Several active-site residues in the pocket play important roles in the catalytic reaction of SD.[10] Two short helices (H4

† Computational Astrophysics Laboratory.
‡ Laboratory for Remediation Research.
§ High Performance Biocomputing Research Team.

(1) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(2) Ferrand, M.; Dianoux, A. J.; Petry, W.; Zaccai, G. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9668−9672.
(3) Gerstein, M.; Chothia, C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10167−10172.
(4) Hayward, S.; Kitao, A.; Hirata, F.; Go, N. *J. Mol. Biol.* **1993**, *234*, 1207−1217.
(5) Higo, J.; Sasai, M.; Shirai, H.; Nakamura, H.; Kugimiya, T. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5961−5964.
(6) Kitao, A.; Hirata, F.; Go, N. *J. Phys. Chem.* **1993**, *97*, 10223−10230.
(7) Okimoto, N.; Tsukui, T.; Kitayama, K.; Hata, M.; Hoshino, T.; Tsuda, M. *J. Am. Chem. Soc.* **2000**, *122*, 5613−5622.
(8) Nakasako, M.; Fujisawa, T.; Adachi, S.; Kudo, T.; Higuchi, S. *Biochemistry* **2001**, *40*, 3069−3079.
(9) Chen, J. M.; Xu, S. L.; Wawrzak, Z.; Basarab, G. S.; Jordan, D. B. *Biochemistry* **1998**, *37*, 17735−17744.

10.1021/ja048053u CCC: $27.50 **© 2004 American Chemical Society**

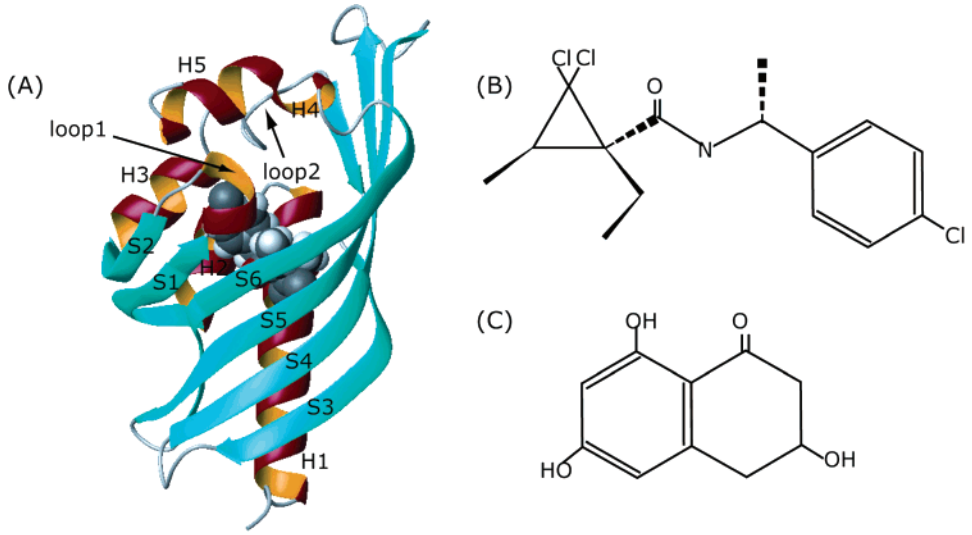

Figure 1. Schematic drawing of the monomeric structure of scytalone dehydratase and structural formulas of carpropamid and scytalone molecules. (A) The enzyme consists of four α-helices (H1−H4) and six β-strands (S1−S6). The ligand molecule (carpropamid), shown in the CPK model, is positioned in a hydrophobic ligand-binding pocket. The C-terminal region containing two short helices (H4 and H5) covers the binding pocket. (B) Structural formula of inhibitor of SD, carpropamid, and (C) structural formula of scytalone, which is the substrate of SD.

and H5) in the C-terminal region cover the binding pocket and are thought to function as a lid for the pocket. A study on the C-terminal region has shown that truncated SD lacking the C-terminal region after F158 has little enzymatic activity.[11] Thus, investigation of the dynamics of the C-terminal region seems to be important for elucidation of the enzyme reaction. The starting structures for our simulations were liganded and unliganded structures. X-ray crystallographic studies have revealed the three-dimensional structures of SD complexed with inhibitors.[9,12−14] The structure of wild-type unliganded SD has not been determined by X-ray crystallographic study, but the structure of an SD mutant, F162A, in which the C-terminal region was truncated has been determined.[15] It was shown in that structural study that the flexibility of the C-terminal region in wild-type SD hindered the crystallization and that crystallization of the mutated SD was successfully achieved by reducing the flexibility of the C-terminal region. Therefore, the starting structure of the unliganded protein was made by removing the ligand molecule from the crystal structure of the liganded protein. Since the size of monomeric SD is not so large, we considered that MD study could reveal the dynamics of the C-terminal region and that of the binding pocket, which are thought to relate to ligand binding. In the present study, 10- and 15-nsec MD simulations of monomeric liganded and monomeric unliganded SDs were carried out to determine (i) the hydration water structure in the binding pocket of the unliganded protein, the three-dimensional structure of which is unknown, and (ii) the difference in protein dynamics and conformation between the liganded protein and the unliganded protein. The results revealed not only the dynamic behavior of liganded SD and unliganded SD but also the path for binding of the ligand molecule to the binding pocket.

Materials and Methods

MD Simulations. The starting structures for the simulations were liganded and unliganded structures. Both structures were constructed on the basis of the X-ray crystallographic structure (PDB entry 7STD[14]) of the trimeric SD-inhibitor complex. The inhibitor was a carpropamid that was bound tightly to the SD (Figure 1B). The N-terminal eight residues were absent in the crystal structure. However, it was thought that the absence of these N-terminal residues would not affect our results since these residues are not essential for its function. A monomer of SD-inhibitor complex, which was extracted from the crystal structure of the trimeric complex, was used as the liganded structure. Two water molecules bound to the inhibitor were included, and their locations were set at the oxygen positions of crystal water molecules because these water molecules were thought to contribute to the stabilization of the inhibitor conformation in the binding pocket. The unliganded structure was made by removing the inhibitor and two bound water molecules from the liganded structure. Therefore, the C-terminal regions covered the binding pockets for both of the initial conformations of the liganded and the unliganded proteins. This conformation is called a closed form. A conformation in which the C-terminal region opens up the top of the binding pocket is called an open form. The atomic charges of the inhibitor carpropamid, which are not included in the standard parm96 parameter set,[16] were calculated using RHF/6-31G* with Gaussian98[17] and the restrained electrostatic potential (RESP) method.[18]

The amino acid residues in proteins were set at pH 7, and four Na+ ions were added to keep the whole systems neutral. The liganded and unliganded structures were surrounded by TIP3P water molecules[19] spherically. The size of each sphere was chosen so that the distance of the protein atoms from the wall was greater than 15.0 Å. Each water drop was restrained by a half-harmonic potential (1.5 kcal/mol·Å^2). The solvated systems of the liganded and the unliganded proteins contained 50 158 and 49 439 atoms, respectively. The unliganded protein contained only one water molecule in the binding pocket. The energies of systems were minimized using 4000 steps of steepest descent followed by 6000 steps of the conjugate gradient method.

(10) Basarab, G. S.; Steffens, J. J.; Wawrzak, Z.; Schwartz, R. S.; Lundqvist, T.; Jordan, D. B. *Biochemistry* **1999**, *38*, 6012−6024.
(11) Motoyama, T.; Imanishi, K.; Yamaguchi, I. *Biosci. Biotechnol. Biochem.* **1998**, *62*, 564−566.
(12) Lundqvist, T.; Rice, J.; Hodge, C. N.; Basarab, G. S.; Pierce, J.; Lindqvist, Y. *Structure* **1994**, *2*, 937−944.
(13) Nakasako, M.; Motoyama, T., Kurahashi, Y.; Yamaguchi, I. *Biochemistry* **1998**, *37*, 9931−9939.
(14) Wawrzak, Z.; Sandalova, T.; Steffens, J. J.; Basarab, G. S.; Lundqvist, T.; Lindqvist, Y.; Jordan, D. B. *Proteins* **1999**, *35*, 425−439.
(15) Motoyama, T.; Nakasako, M.; Yamaguchi, I. *Acta Crystallogr., Sect. D Biol. Crystallogr.* **2002**, *58*, 148−150.
(16) Kollman, P. A.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. In *Computer Simulation of Biomolecular System*; Wilkinson, A., Weiner, P., Van Gunsteren, W., Eds.; Elsevier: The Netherlands, 1997; pp 83−96.

All of the MD simulations in this work were carried out using a special-purpose computer for MD, the MD machine (MDM), that can accelerate the computation of all nonbonded interactions.[20−25] The MD software was a modified Amber 6.0[26] for the MDM. The parm96 force field[16] was adopted, and the time step was set at 1 fsec. The bond lengths involving hydrogen atoms were constrained to equilibrium lengths using the SHAKE method.[27] The temperature of each system was kept constant according to Berendsen's algorithm with a coupling time of 0.2 ps.[28] The temperatures of respective systems were gradually increased to 310 K over a period of 100 ps, and additional 10- and 15-nsec MD simulations of the liganded and the unliganded proteins at 310 K were performed for data collection.

Principal Component Analysis. Principal component analysis (PCA) projects multidimensional data onto low-dimensional subspaces.[29,30] If the distribution of the multidimensional data is nonisotropic, the PCA will identify a low-dimensional subspace that describes it very well. In this work, we used PCA to divide the trajectory into several groups, such as preequilibrium and equilibrium states. The PCA was carried out by diagonalizing the mass-weighted positional covariance matrix constructed from the trajectory after a least-squares fit to an averaged structure. The larger the eigenvalue, obtained by diagonalization of the matrix, the more efficient the projection onto that axis becomes. For our simulations, all backbone heavy atoms were applied to the PCA.

Molecular Mechanics Poisson−Boltzmann/Surface Area Method. The coordinates from each trajectory were saved every 5 ps, and each of them was used to evaluate free energy by the molecular mechanics Poisson−Boltzmann/surface area (MM-PBSA) method. The MM-PBSA free energy of each snapshot (G_{TOT}) is approximated as the sum of two terms: the internal energy of the protein (E_{MM}) and the solvation free energy (ΔG_{SOLV}) (eq 1). E_{MM} is the sum of internal strain energy (E_{int}), van der Waals energy, and electrostatic energy. E_{int} is the energy associated with vibration of covalent bonds and rotation of valence bond angles and torsional angles.

$$G_{TOT} = E_{MM} + \Delta G_{SOLV} \quad (1)$$

The vibrational component of protein molecule entropy (S_{solute}) can be calculated by normal-mode analysis on an energy-minimized structure.

$$\Delta G_{SOLV} = \Delta G\text{sol:np} + \Delta G\text{sol:pol} = (\gamma \cdot \text{SASA} + b) + \Delta G\text{sol:pol} \quad (2)$$

The solvation energy, ΔG_{SOLV}, is divided into two parts, electrostatic contributions, ΔGsol:pol, and all other contributions, ΔGsol:np (eq 2). The nonpolar solvation energy, ΔGsol:np, is calculated as a function of solvent-accessible surface area. In eq 2, by which the nonpolar solvation contribution is calculated, SASA is the solvent-accessible surface area calculated using the MSMS program,[31] and γ and b are 0.00542 kcal/mol Å^2 and 0.92 kcal/mol, respectively. The probe radius was 1.4 Å. ΔGsol:pol was calculated by solving the Poisson−Boltzmann equation using the delphi program.[32,33] PARSE atomic radii[34] and parm96 charges[16] were used in this calculation. The grid spacing used was 1.0 Å. The dielectric constants inside and outside the molecule were 4.0 and 80.0, respectively.

For the MM-PBSA calculations of the unliganded protein, the water molecules in the binding pocket were considered explicitly. The ΔGsol:np was not calculated in the binding pocket because the effect of the water molecules in the pocket to the ΔGsol:np was taken into consideration in the calculation of the E_{MM}. The dielectric constant of the explicit water molecules in the binding pocket for ΔGsol:np was set at 4.0.

Table 1. Backbone RMSDs Averaged over the Last 2.5-ns Periods

	averaged backbone RMSD (Å)[a]	standard deviation (Å)[a]
liganded protein	1.07	0.07
unliganded protein	1.75	0.11

[a] In each system, 500 sets of coordinates (every 5 ps) were considered.

Results

Checking of the Simulation. To check whether the simulations of the liganded and unliganded proteins had been performed properly, the backbone RMSDs and *B*-factors were analyzed. Table 1 shows the averaged backbone RMSDs and their standard deviations over the last 2.5 ns of simulation. The RMSD of the unliganded protein became larger than that of the liganded protein, suggesting that the unliganded structure was somewhat different from the liganded structure. The analysis of RMSDs indicated that each simulation did not contain obvious artifacts and had reached an equilibrium state. Figure 2 shows *B*-factor values of the main chain atoms against each residue. The *B*-factor values of the liganded protein show a tendency similar to those of the X-ray crystal structure, suggesting that there is little difference between the monomer and trimer in protein dynamics (Figure 2). On the other hand, some differences in protein fluctuation were found between the

(17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Mont-gomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millan, J. M.; Daniel, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morlkuma, K.; Malick, D. K.; Robuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Oritiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M.; Peng, C. Y.; Nanayakkara, A.; Gonzakez, C.; Challacombe, M.; Gill, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98*, Revision A.1x; Gaussian, Inc.: Pittsburgh, PA, 2001.
(18) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *7*(40), 10269−10280.
(19) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926−935.
(20) Narumi, T.; Susukita, R.; Ebisuzaki, T.; McNiven, G..; Elmegreen, B. *Mol. Simul.* **1999**, *21*, 401−415.
(21) Susukita, R.; Ebisuzaki, T.; Elmegreen, B. G..; Furusawa, H.; Kato, K.; Kawai, A.; Kobayashi, Y.; Koishi, T.; McNiven, G. D.; Narumi, T.; Yasuoka, K. *Comput. Phys. Comm.* **2003**, *155*, 115−131.
(22) Fujii, Y.; Okimoto, N.; Hata, M.; Narumi, T.; Yasuoka, K.; Susukita, R.; Suenaga, A.; Futatsugi, N.; Koishi, T.; Furusawa, H.; Kawai, A.; Ebisuzaki, T.; Neya, S.; Hoshino, T. *J. Phys. Chem. B* **2003**, *107*, 10274−10283.
(23) Okimoto, N.; Yamanaka, K.; Suenaga, A.; Hirano, Y.; Futatsugi, N.; Narumi, T.; Yasuoka, K.; Susukita, R.; Koishi, T.; Furusawa, H.; Kawai, A.; Hata, M.; Hoshino, T.; Ebisuzaki, T. *Chem-Bio. Informatisc. J.* **2003**, *3*, 1−11.
(24) Suenaga, A.; Hatakeyama, M.; Ichikawa, M.; Yu, X.; Futatsugi, N.; Narumi, T.; Fukui, K.; Terada, T.; Taiji, M.; Shirouzu M.; Yokoyama, S.; Konagaya, A. *Biochemistry* **2003**, *42*, 5195−5200.
(25) Suenaga, A.; Kiyatkin, A. B.; Hatakeyama, M.; Futatsugi, N.; Okimoto, N.; Hirano, Y.; Narumi, T.; Kawai, A.; Susukita, R.; Koishi, T.; Furusawa, H.; Yasuoka, K.; Takada, N.; Ohno, Y.; Taiji, M.; Ebisuzaki, T.; Hoek, J. B.; Konagaya, A.; Kholodenko, B. N. *J. Biol. Chem.* **2004**, *279*, 4657−4662.
(26) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 6;* University of California, San Francisco: San Francisco, CA, 1999.
(27) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(28) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(29) Kitao, A.; Hirata, F.; Go, N. *Chem. Phys.* **1991**, *158*, 447−472.
(30) Kitao. A.; Go, N. *Curr. Opin. Struct. Biol.* **1999**, *9*, 164−169.
(31) Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*, 305−320.
(32) Rocchia, W.; Alexov E.; Honig, B. *J. Phys. Chem. B* **2001**, *105*, 6507−6514.
(33) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128−137.
(34) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.

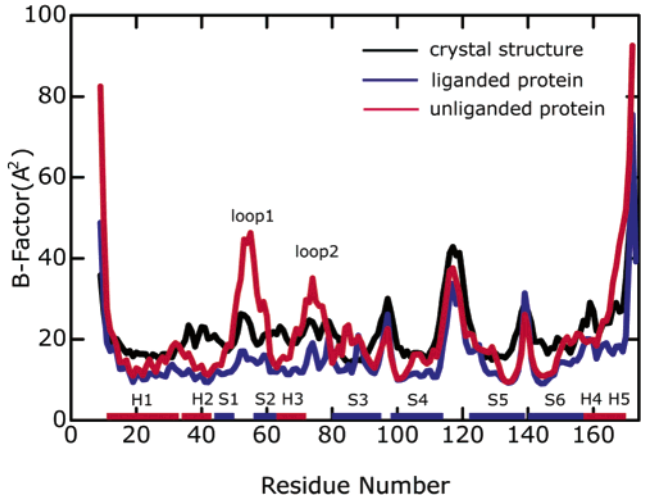

Figure 2. *B*-factor values of the main chain atoms against residue number. Black, blue, and red lines indicate values from the experimental data on crystal structure (PDB entry 7STD) and simulations of the liganded and unliganded proteins, respectively. The colored bars on the abscissa show the secondary structural elements in Figure 1.

liganded and unliganded proteins. The *B*-factor curve of the unliganded protein deviated in four regions from that of the liganded protein. Prominent differences were observed in the loop between S1 and S2 (loop1), the loop between H3 and S3 (loop2), and the C-terminal region. Since the first two regions are parts around the C-terminal region, these observations indicate that the C-terminal region fluctuated greatly. A difference was also found in the β-strand S3. Figure 3 shows the MD trajectories projected onto the plane defined by the first two principal components. At a glance, it is clear that the trajectory of the unliganded protein was more diffusive and more complicated than the trajectory of the liganded protein. These data also confirm that the conformational transition of the unliganded protein was much larger than that of the liganded protein. From the results of the PCA, we regarded 2.5−10- and 9−15-ns trajectories of the liganded and unliganded proteins (red dots shown in Figure 3) as equilibrium states. Hereafter, we will describe the results for unliganded and liganded proteins in that order.

Simulation of Unliganded Protein. (a) Process before Reaching an Equilibrium State. In the period of the first 5 ns (black and green dots shown in Figure 3), the conformation of the unliganded protein changed from a closed form to an open form (Figure 4A and B). Namely, the C-terminal region, which covers the binding pocket in the starting structure, opened up the entrance of the binding pocket that we named a "front gate". As a consequence, the hydrogen bonds between the C-terminal region and loop1, a linker of S1 and S2 (Lys56Nϵ−Phe169O, Phe53O−Arg166Nη1, and Phe53O−Arg166η2), which were formed at the starting structure, disappeared from 4 ns (Figure 5) and the helical loop1 (Figure 1) was unlaced. Once these hydrogen bonds were lost, they did not return to their former state; thus, the C-terminal region fluctuated dynamically just before reaching the equilibrium state (Figure 5). At this time, a few water molecules entered the binding pocket from the front gate (Figure 4B). In the next 2.5 ns (blue dots shown in Figure 3), intake of solvent water molecules inside the protein occurred suddenly, resulting in the localization of 7−8 water molecules in the binding pocket at 7.5 ns (Figure 4C). At this stage, the intake of solvent water molecules occurs not only from the front gate but also from the space between α-helix H1 and the β-barrel, which we named a "rear gate". Finally, for the 7.5−9-ns trajectory, the water cluster inside the binding pocket was arranged to form a large water cluster with nine water molecules (Figure 4D). Hence, the inside of the binding pocket for the unliganded proteins was solvated fully.

(b) Equilibrium State. In the equilibrium state, protein dynamics and conformation were closely related to the dynamic behavior of hydration water molecules in the binding pocket. Figure 6 shows the time course of the number of water molecules in the binding pocket. In the equilibrium state (9 ns∼), the water molecules went in and out of the binding pocket. The amino acid residues in the binding pocket are so hydrophobic that the water cluster formed few hydrogen bonds with the residues in the binding pocket. The water cluster was in flux

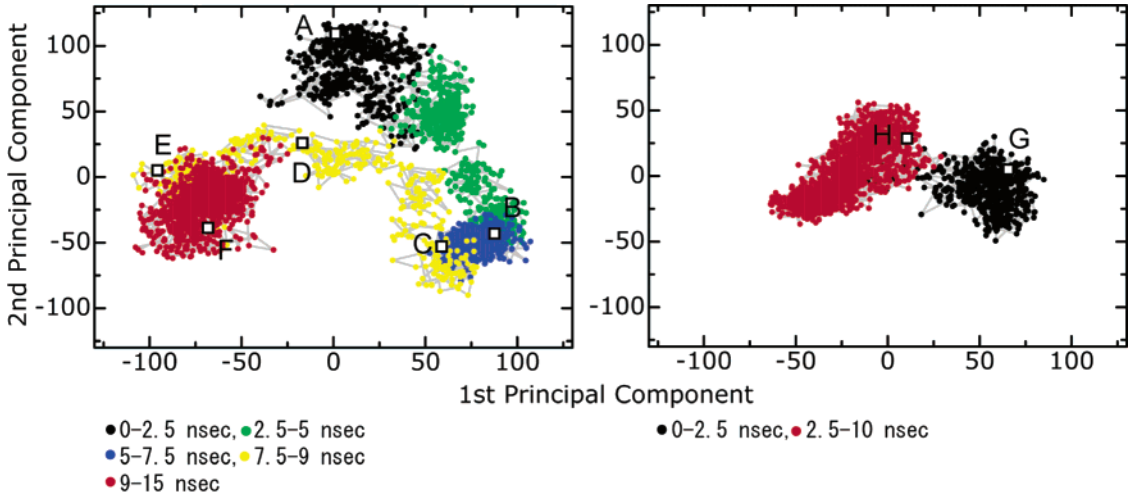

Figure 3. Whole trajectories projected onto the plane formed by the first two largest principal components calculated from PCA. The first two principal component accounts of the unliganded and liganded proteins contributed to 61.1% and 46.0%, respectively, of the total mass-weighted mean-square fluctuations of proteins. The figures on the left and right show the trajectories of the unliganded and liganded proteins, respectively. Black, green, blue, yellow, and red dots correspond to 0−2.5-ns, 2.5−5-ns, 5−7.5-ns, 7.5−9-ns, and 9−15-ns trajectories in the unliganded protein. The 9−15-ns trajectory was regarded as the equilibrium state. Open boxes from A to F correspond to the loci of 1.0 ns, 4.5 ns, 7.5 ns, 8.5 ns, 12.2 ns, and 14.8 ns. For the liganded protein, black and red dots correspond to 0−2.5-ns and 2.5−10-ns trajectories, respectively, and the 2.5−10-ns trajectory was regarded as the equilibrium state. Open boxes of G and H correspond to the loci of 1.0 ns and 10.0 ns.

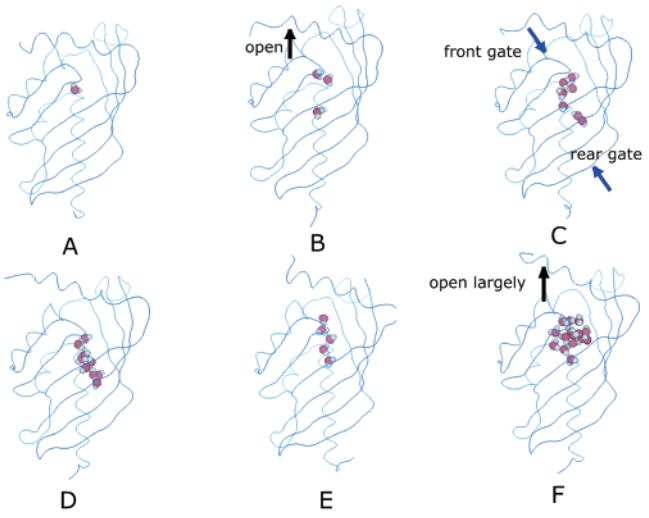

Figure 4. Hydration water molecules in the binding pocket of the unliganded protein. The structures from A to F correspond to the open boxes from A to F in Figure 3. A, snapshot at 1 ns; B, 4.5 ns; C, 7.5 ns; D, 8.5 ns; E, 12.2 ns; F, 14.8 ns. Water molecules are shown in the CPK model.

and it was fragmented and reconstructed by gathering water molecules. This is illustrated in Figure 4D−F, in which water molecules going in and out of the binding pocket can be seen. A comparison of the solvation process of the binding pocket (Figure 6) and whole structural change (Figure 7A) for the unliganded protein revealed that these results were well correlated, especially for the period from P1 to P3 in Figure 6. This correlation between dynamics of the protein and hydration water molecules in the binding pocket indicates that the protein structure changed just at the time when the fragments derived from the water cluster went out of the binding pocket through the front or rear gates (Figures 6 and 7A). Thus, the fluctuations of the regions involving the front gate (loop1, loop2, and C-terminal region) and the region involving the rear gate, S3, deviated from those of the liganded protein (Figure 2). For the period from P′3 to P4 in Figure 6, an adequate correlation was not obtained from the RMSD values of the whole structure (Figure 7A). The C-terminal region, however, began to move dynamically again after P′3 as observed in the preequilibrium state (Figures 5 and 7B). It is thought that the fluctuation of the C-terminal region promoted the intake of solvent water molecules into the pocket since the largest water cluster with 15 water molecules was formed at nearly 14.8 ns (P4 of Figure 6).

In addition, we performed energetic analyses to obtain a better understanding of the conformational change of the unliganded protein in the equilibrium state. Figure 8A shows the free-energy contour map of the trajectory of the equilibrium state of the unliganded protein projected onto the plane of the first two principal components. The free energies were estimated by using the MM-PBSA method. The trajectory of the unliganded protein moves from the local minimum on the right of the map to the ground minimum on the left via a local high-energy region. It is important for the relationship between dynamics of the protein and hydration water molecules that the unliganded protein formed a large water cluster with 14−15 water molecules in the binding pocket in the vicinity of the global minimum. Figure 8C and D shows the protein structures at the points of the highest and the lowest energy levels on the contour map. Differences between these two structures were found in the C-terminal region and the binding pocket. For the structure at the point of the lowest energy level, the C-terminal region opened up largely and the binding pocket contained the largest water clusters, as mentioned above. These structural differences were clearly reflected in the values of E_{MM} and ΔG_{SOLV} of free energy.

Simulation of the Liganded Protein. Figure 8B shows the free-energy contour map of the equilibrium state of the liganded protein. The trajectory moves from the highest region on the free-energy surface, via a local minimum on the left of the map to the local minimum on the right on the contour map. Figure 8E and F shows the structures at the points of the highest and lowest energy levels. A comparison of these structures shows that the structures of loop1 and the C-terminal region are different, but both structures are very similar on the whole. This structural difference mainly contributed to the difference in the value of E_{MM} of free energy. As the period of its stay in the high-energy region is very short, the liganded structure did not

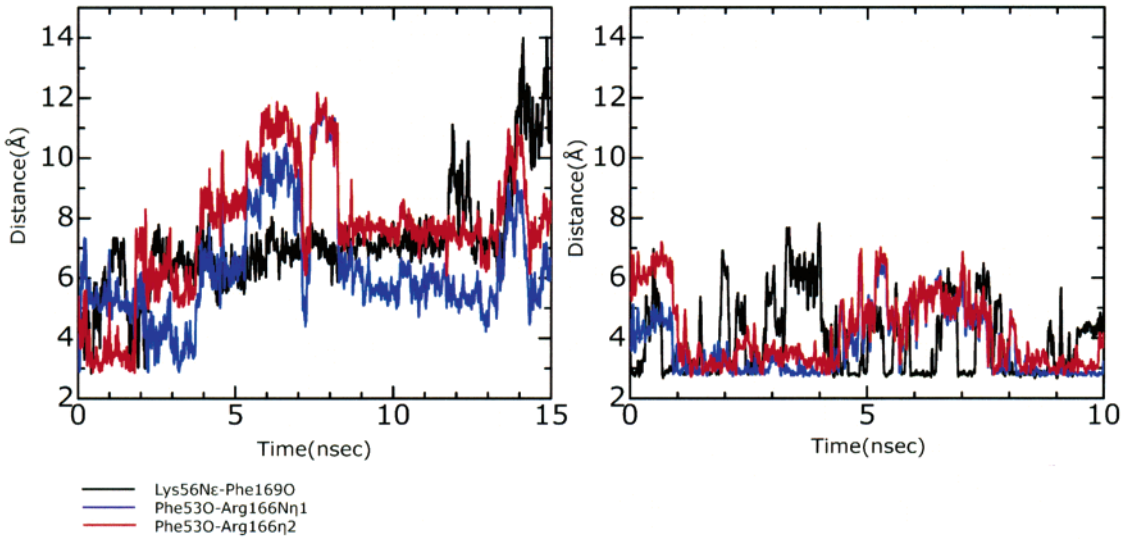

Figure 5. Time courses of hydrogen bonds between the C-terminal region and loop1. The figures on the left and right show changes in the hydrogen bond distances of the unliganded and the liganded proteins, respectively. Black, blue, and red lines indicate the distances of Lys56Nε−Phe169O, Phe53O−Arg166Nη1, and Phe53O−Arg166η2, respectively.

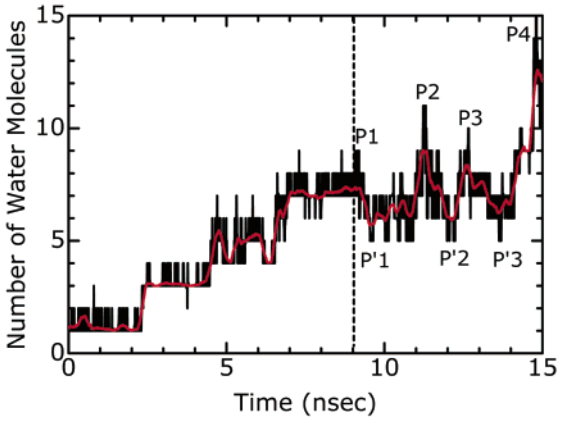

Figure 6. Time course of the number of water molecules in the binding pocket. The black line indicates the number of the water molecules and the red line indicates the averaged values over intervals of 0.25 ns. A broken line indicates the beginning of the equilibrium state. P1−P4 are maximal points and P′1−P′3 are minimal points.

change greatly and could stably retain a closed form during the simulation of the equilibrium state. In this regard, the interactions between the ligand and the C-terminal region, in addition to the interactions between loop1 and the C-terminal region, would prevent the C-terminal region from opening (Figure 5). Therefore, the transition of the liganded protein is thought to be a process for slowly eliminating strain of the protein structure.

Discussion

MD simulations of liganded and unliganded SDs were performed to determine the effect of the ligand binding on protein dynamics and the cooperative motions of the protein and hydration water molecules. The results of this study provided the information not only on the dynamics of the protein and hydration water molecules but also on the binding process of the ligand molecule. Judging from several experimental data on protein structure and enzyme activity, we consider that the results of this work are helpful for an understanding of the experimental data.

In the simulation of the liganded protein, the protein held the inhibitor, carpropamid, tightly in the hydrophobic binding pocket. The interactions on the inhibitor are mainly hydrophobic, but several hydrogen bonds with Tyr50, His85, His110, Ser 129, and Asn131 in the binding pocket, as were observed in the crystal structures,[13,14] were formed. Moreover, the C-terminal region forms three hydrogen bonds with loop1 (Lys56Nϵ−Phe169O, Phe53O−Arg166Nη1, and Phe53O−Arg166η2) and covers the binding pocket tightly (Figure 5). The existence of these hydrogen bonds on the C-terminal region was supported by the results of an experiment on the enzyme activity. Motoyama et al. reported that deletion of the C-terminal 10 residues caused loss of 98.5% of activity.[11] These computational and experimental results suggest that the C-terminal portion of SD is very important in catalysis or ligand binding. The conformation of the protein-inhibitor complex was stabilized through these interactions, and a large conformational change, as was observed in the simulation of the unliganded protein, did not occur. Therefore, the liganded protein retained a closed form over a simulation period of 10 ns. This computational result also supports the experimental observation that carpropamid worked as a tight binding inhibitor.[13]

For the unliganded protein, it took a long simulation time of 9 ns to reach the equilibrium state.This long period of the preequilibrium state would be valid since the simulation time before reaching the equilibrium state was about the same length in the simulation of the unliganded protein under a different condition. A drastic conformational change occurred in the preequilibrium state; consequently, the unliganded protein changed from a closed form to an open form. This would be caused by the disappearance of interactions between residues in the protein pocket and the C-terminal region, which is due to the absence of the ligand molecule. After the opening of the C-terminal region, solvent water molecules were taken in the binding pocket, and the number of water molecules in the pocket rose to 7−8 just before the equilibrium state was reached (Figure 6). At this time, intake of water molecules occurred not only from the front gate but also from the rear gate. Probably, a substrate or inhibitor that is larger than a water molecule would be taken in from the front gate since the rear gate is too narrow for ligand binding.

In the equilibrium state of the unliganded protein, the inside of the binding pocket was fully solvated (Figure 6) and there was a correlation between the conformational change of the protein and the dynamics behavior of hydration water molecules in the binding pocket. The water molecules went in and out of the binding pocket through the front or rear gate in accordance with the protein dynamics. Thus, the fluctuation of the unliganded protein become larger than that of the liganded protein, especially in the parts on the front gate, that is, loop1, loop2, and the C-terminal region, and the part on the rear gate, S3 (Figure 2). These motions of the unliganded protein would be closely related to the unsuccessful results in crystallization trials for the unliganded wild-type protein.[15] Since the β-barrel and α-helix H1 are trimer interfaces in the crystal structure of the trimer complex of ligand-bound protein, they are thought to play

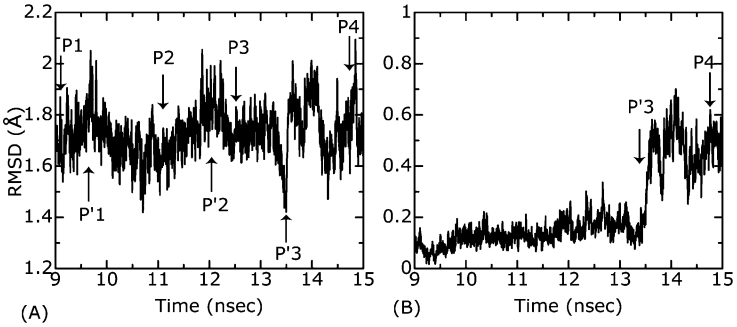

Figure 7. Time courses of RMSD of (A) the whole structure of the unliganded protein and (B) the C-terminal region (residues 158−172) of the unliganded protein. P1−P4 and P′1−P′3 correspond to those in Figure 6.

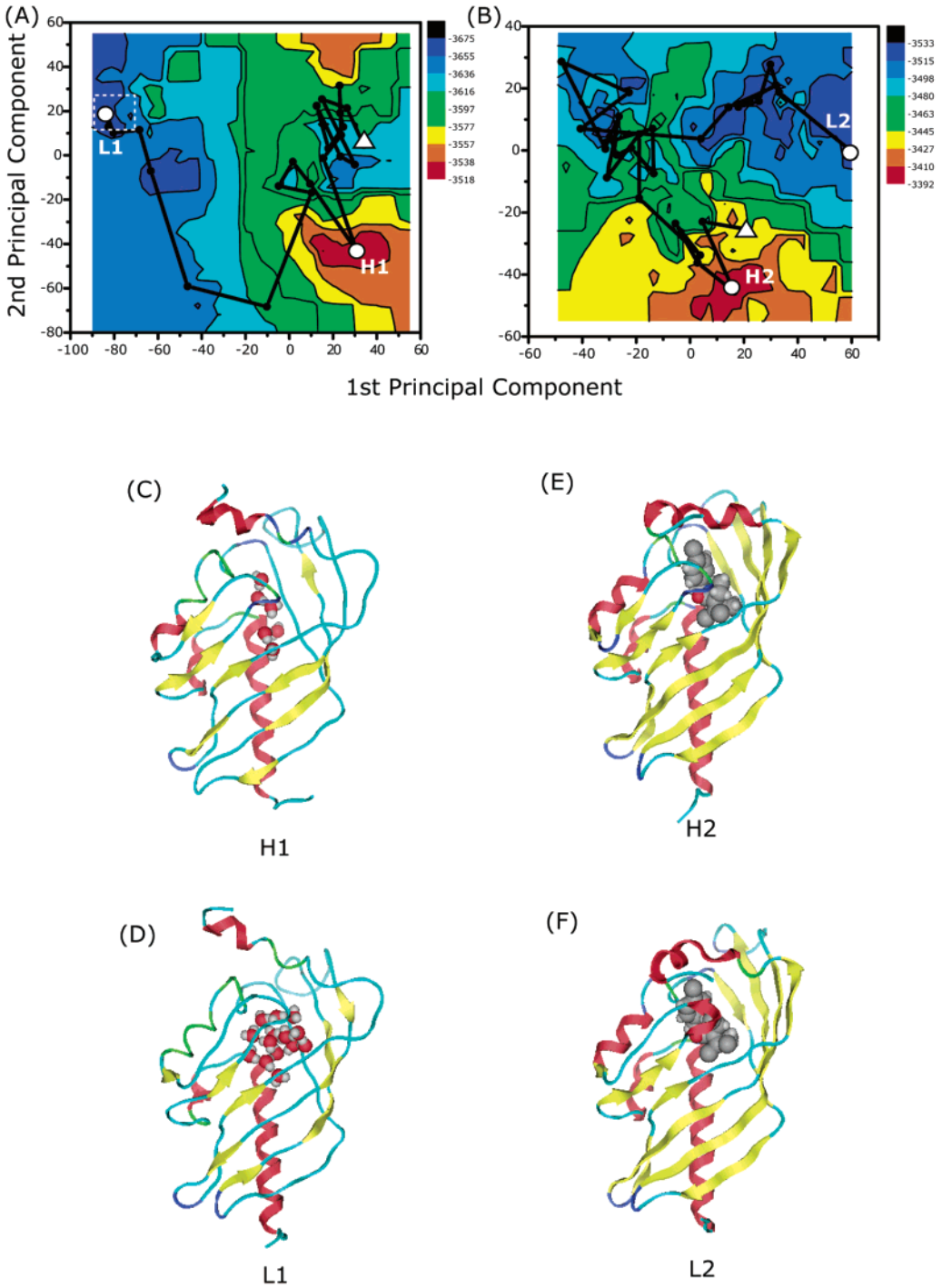

Figure 8. Free-energy contour map of each trajectory of the equilibrium state and protein structures. (A and B) These two maps (A) and (B) indicate the free-energy contour maps of the trajectories of the unliganded and liganded proteins. The free energies of each trajectory were projected onto the plane formed by the first two largest principal components calculated from PCA of the trajectories of the equilibrium states. Respective black lines are loci of the points on trajectories at intervals of 0.25 ns, and triangles indicate starting points. H1 and L1 are the loci of the highest and the lowest energy levels for the unliganded protein, respectively, and H2 and L2 are the loci of the highest and the lowest energy levels for the liganded protein, respectively. The rectangular box indicates the region in which more than 14 water molecules are localized in the binding pocket. Energies are in kcal/mol. (C and D) The two structures correspond to L1 and H1 in Figure 8A. (E and F) The two structures correspond to L1 and H1 in Figure 8B.

an important role in the trimer aggregation. Therefore, it is thought that the fluctuation of S3 in the β-barrel, induced by the movement of the hydration water molecules, inhibits trimer aggregation and hinders crystallization. With regard to the C-terminal region, it would control the water intake into and release from the binding pocket, judging from the relationship between its flexibility and the solvation process of the binding pocket (see Figures 6 and 7). Although we did not perform simulation for the mutant F162A in the current work, this mutation seems to reduce the flexibility of the C-terminal region, and, as a result, it is thought that the intake and release of the hydration water molecules and the fluctuation of S3 were restrained.

From observation of a series of dynamics and conformations of the unliganded protein, we found the path leading the water molecules from the front gate into the binding pocket. This path

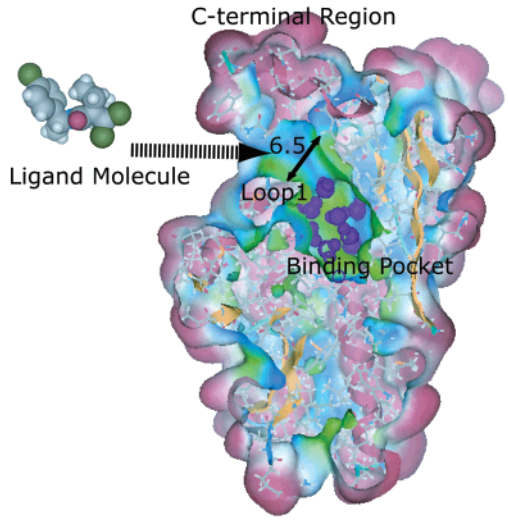

Figure 9. Path leading ligand molecule from the front gate into the binding pocket. This figure shows the 14.8-ns structure of the unliganded protein in sections. The purple balls in the binding pocket indicate water molecules. The width of the path is nearly 6.5 Å.

is illustrated in Figure 9. In the conformation that contained the largest water cluster in the binding pocket and existed in the ground-energy minimum in Figure 8A, this path is nearly 6.5 Å in width, which roughly corresponds to the width of the inhibitor, carpropamid (Figure 9). Therefore, this path is expected to be used for the ligand binding. For the binding of ligand molecule, removal of the water molecules from the binding pocket would be necessary. Hydration water molecules in the binding pocket would go out through the rear gate. Movements of the hydration water molecules in the active-site cleft have been examined in some studies, and it has been suggested that hydration water molecules move dynamically in the substrate-free enzyme.[8,22]

Furthermore, we expect that the protein dynamics of SD may apply to those of other proteins containing an $\alpha+\beta$ barrel motif. Adipocyte fatty acid binding protein (AFABP), for example, consists of a 10-stranded antiparallel β-barrel capped by an α-helical domain consisting of two short α-helices, and its function is fatty acid transfer from AFABP to phospholipid membranes.[35] The α-helical domain covers one end of the binding cavity and participates in a lid for ligand entry into and exit from the binding cavity. The results of this study on SD should contribute to a better understanding of the protein function and dynamics of AFABP.

Conclusion

We clarified the dynamics and conformations of liganded and unliganded SDs. The liganded protein held the ligand molecule tightly and is in a favorable state for enzymatic reaction. The unliganded protein, the three-dimensional structure of which is unknown, fluctuated dynamically with intake of water molecules into the pocket and release of those from it. Namely, the hydration water molecules and the unliganded protein moved cooperatively. These cooperative motions produced ligand-binding pathway through protein interior in the vicinity of the ground minimum on the free-energy surface of the unligand protein. Therefore, the unliganded protein explores the conformational free-energy surface extensively and is in a waiting state for ligand binding. The failure in crystallization of unliganded wild-type protein in the structural study seems to result from the flexibility of loop1, loop2, S3, and the C-terminal region, which are involved with the water intake into and release from the pocket.

Acknowledgment. We are grateful to Dr. T. Narumi, Dr. R. Susukita, Dr. A. Kawai, Dr. T. Koishi, Mr. H. Furusawa, and Dr. K. Yasuoka for their support in the use of MDGRAPE-2. This work was supported by a Special Postdoctoral Research Program from RIKEN (to N.O.) and by the contracted research "Protein 3000 Project" by the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to A.S. and N.F.). We thank Dr. T. Takahashi, Dr. T. Hoshino, and Dr. S. Hirono for their fruitful discussions and comments on this work.

JA048053U

(35) Liou, H. L.; Storch, J. *Biochemistry* **2001**, *40*, 6475−6485.