J|A|C|S
ARTICLES

Published on Web 08/19/2010

Side-Chain Conformational Changes of Transcription Factor PhoB upon DNA Binding: A Population-Shift Mechanism

Tsutomu Yamane,[†] Hideyasu Okamura,[‡] Yoshifumi Nishimura,[†] Akinori Kidera,[†,§] and Mitsunori Ikeguchi*[,†]

Department of Supramolecular Biology, Graduate School of Nanobioscience, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045 Japan, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan, and Research Program for Computational Science, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

Received April 16, 2010; E-mail: ike@tsurumi.yokohama-cu.ac.jp

Abstract: Using molecular dynamics (MD) simulations and analyses of NMR relaxation order parameters, we investigated conformational changes of side chains in hydrophobic cores upon DNA binding for the DNA binding/transactivation domain of the transcription factor PhoB, in which backbone conformational changes upon DNA binding are small. The simulation results correlated well with experimental order parameters for the backbone and side-chain methyl groups, showing that the order parameters generally represent positional fluctuations of the backbone and side-chain methyl groups. However, topological effects of the side chains on the order parameters were also found and could be eliminated using normalized order parameters for each amino acid type. Consistent with the NMR experiments, the normalized order parameters from the MD simulations showed that the side chains in one of the two hydrophobic cores (the soft core) were highly flexible in comparison with those in the other hydrophobic core (the hard core) before DNA binding and that the flexibility of the hydrophobic cores, particularly of the soft core, was reduced upon DNA binding. Principal component analysis of methyl group configurations revealed strikingly different side-chain dynamics for the soft and hard cores. In the hard core, side-chain configurations were simply distributed around one or two average configurations. In contrast, the side chains in the soft core dynamically varied their configurations in an equilibrium ensemble that included binding configurations as minor components before DNA binding. DNA binding led to a restriction of the side-chain dynamics and a shift in the equilibrium toward binding configurations, in clear correspondence with a population-shift model.

Introduction

Conformational changes in proteins upon ligand binding play important functional roles, such as in molecular motors, enzymatic reactions, material transport, and signal transduction. The mechanism underlying conformational changes has been explained by the induced-fit model, in which ligands induce conformational changes in proteins to make their binding sites fit the ligands.[1] An alternative model is the population-shift or pre-existing equilibrium model,[2,3] in which protein conformations significantly fluctuate in the ligand-free state and the equilibrium ensemble before ligand binding includes specific conformations complementary to the ligands as minor components. The ligands then bind selectively to the binding conformations, shifting the equilibrium toward these conformations.

The population-shift model has been supported by both experimental and theoretical studies,[4−6] which have mainly focused on backbone changes only. However, a protein structural database analysis[7] showed that 91% of proteins in the data set used in the analysis have less than 2 Å of C_α root-mean-square deviations (RMSDs) between the ligand-bound and -free conformations, and therefore, most proteins do not experience considerable conformational changes in their backbones upon ligand binding.[7] The analysis also showed that even in the case of small changes in backbones, the side chains had significantly altered conformations upon ligand binding.[7]

In this paper, we address the mechanism underlying side-chain conformational changes, focusing in particular on side-chain dynamics before ligand binding and how they are affected by ligand binding. NMR relaxation experiments allowed for quantitative characterization of side-chain internal dynamics, the amplitude of which was evaluated as a model-free order parameter. To date, side-chain dynamics has been investigated using NMR order parameters for aliphatic[8−13] and aromatic[14,15] groups as well as hydrophilic groups, including amide,[16−18]

[†] Yokohama City University.
[‡] National Institute of Agrobiological Sciences.
[§] RIKEN.

(1) Koshland, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.
(2) Ma, B.; Kumar, S.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **1999**, *12*, 713–720.
(3) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789–796.
(4) Kern, D.; Zuiderweg, E. R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 748–757.
(5) Lange, O. F.; Lakomek, N. A.; Fares, C.; Schroder, G. F.; Walter, K. F.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; de Groot, B. L. *Science* **2008**, *320*, 1471–1475.
(6) Okazaki, K.; Takada, S. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11182–11187.
(7) Gutteridge, A.; Thornton, J. *J. Mol. Biol.* **2005**, *346*, 21–28.
(8) Palmer, A. G.; Hochstrasser, R. A.; Millar, D. P.; Rance, M.; Wright, P. E. *J. Am. Chem. Soc.* **1993**, *115*, 6333–6345.

guanidino,[19,20] carboxyl, and carbonyl[21] groups. Among these, order parameters for methyl groups (S^2_{axis}) have often been used to investigate the side-chain dynamics in hydrophobic cores.[8−12,22] Despite a clear definition of S^2_{axis} as the fluctuation of the bond vector between a methyl carbon and an adjacent non-hydrogen atom, S^2_{axis} has been shown to be only weakly correlated with simple geometrical properties such as depth of burial, packing density, and crystallographic temperature factors of methyl groups.[23] Therefore, the side-chain dynamics implied by S^2_{axis} is not trivial but should be investigated in combination with more sophisticated tools, such as all-atom molecular dynamics (MD) simulations providing atomic details of protein dynamics.[24,25]

Our target was the DNA binding/transactivation domain of the transcription factor PhoB ($PhoB^{DBTD}$). PhoB is a transcriptional regulator protein in *Escherichia coli* that is involved in the two-component signal transaction system for phosphate uptake. For both the DNA-free and -bound forms, atomic structures of $PhoB^{DBTD}$ have been determined using NMR spectroscopy and X-ray crystallography.[26−28] The NMR solution structures have been refined by us using all-atom MD simulations coupled with NMR-derived restraints and including explicit waters and full electrostatic interactions.[28] In the resulting solution structures, the average backbone RMSD between the DNA-free and -bound forms is 1.0 Å for all residues excluding the terminal residues and 0.65 Å for secondary structure elements, indicating that the backbone conformational changes of $PhoB^{DBTD}$ upon DNA binding are small. In contrast to the very small conformational changes in the backbone structures, the side-chain dynamics exhibits significant changes upon DNA binding, as revealed by our previous NMR relaxation experiments.[22] The NMR order parameters for side-chain methyl groups suggested that before DNA binding, the side chains in one of the two hydrophobic cores (called the "soft" core) in $PhoB^{DBTD}$ are more fluidlike than those in the other hydrophobic core (called the "hard" core) (see Figure 1). Upon DNA binding, the side-chain packing appears to be enhanced in the hydrophobic cores, particularly in the soft core. However, the detailed mechanism of the DNA-induced changes in the side-chain dynamics remains unclear.

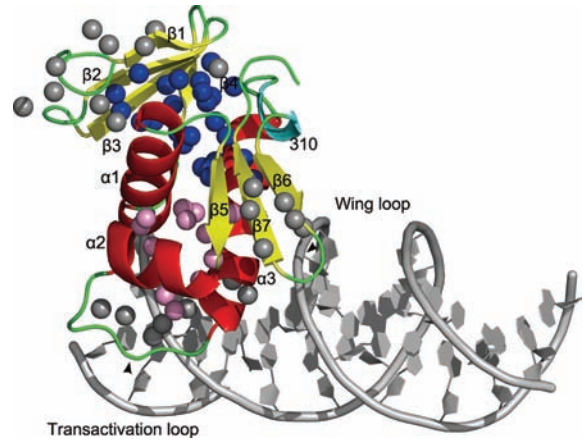

Figure 1. Cartoon representation of the average structure of $PhoB^{DBTD}$ complexed with DNA (PDB entry 2Z33). Methyl carbons (depicted as spheres) in the hard and soft hydrophobic cores are colored blue and pink, respectively. This figure was drawn using PyMOL.[40]

In this study, all-atom MD simulations for $PhoB^{DBTD}$ in both DNA-free and -bound forms were carried out to elucidate atomically detailed mechanisms of the DNA-induced changes in side-chain dynamics observed in the NMR experiments. Simulation results were assessed by comparison with the backbone and side-chain order parameters from the NMR experiments. In particular, we examined whether the simulations reproduced the difference in the side-chain dynamics of the soft and hard cores observed in the NMR experiments. The structural origins of the flexibility in the soft core were investigated using principal component analysis and clustering. Finally, by analyzing side-chain rotamers in the soft core, we arrived at an atomically detailed mechanism underlying the DNA-induced changes in the side-chain dynamics.

Methods

For both DNA-free and -bound forms, 10 initial structures for MD simulations were taken from the previous NMR refinement study for $PhoB^{DBTD}$.[28] The initial structures were immersed in a water box, and counterions were added to neutralize the system. All of the simulations were carried out with the MD program MARBLE[29] using CHARMM22/CMAP for proteins,[30] CHARMM27 for DNA,[31] and TIP3P for water[32] as the force-field parameters. Electrostatic interactions were calculated using the particle-mesh Ewald method.[33] The Lennard-Jones potential was smoothly switched to zero over the range 8−10 Å. The symplectic integrator for rigid bodies was used for constraining the bond lengths and angles involving hydrogen atoms.[29] The time step was 2.0 fs. In the equilibrations, the systems were gradually heated to 310 K for 40 ps with the NMR-derived restraints, after which the 10 ns product

(9) Muhandiram, D. R.; Yamazaki, T.; Sykes, B. D.; Kay, L. E. *J. Am. Chem. Soc.* **1995**, *117*, 11536–11544.
(10) Millet, O.; Muhandiram, D. R.; Skrynnikov, N. R.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 6439–6448.
(11) Brath, U.; Akke, M.; Yang, D.; Kay, L. E.; Mulder, F. A. *J. Am. Chem. Soc.* **2006**, *128*, 5718–5727.
(12) Skrynnikov, N. R.; Mulder, F. A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 4556–4566.
(13) LeMaster, D. M.; Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255–9264.
(14) Teilum, K.; Brath, U.; Lundstrom, P.; Akke, M. *J. Am. Chem. Soc.* **2006**, *128*, 2506–2507.
(15) Boyer, J. A.; Lee, A. L. *Biochemistry* **2008**, *47*, 4876–4886.
(16) Boyd, J. *J. Magn. Reson., Ser. B* **1995**, *107*, 279–285.
(17) Buck, M.; Boyd, J.; Redfield, C.; MacKenzie, D. A.; Jeenes, D. J.; Archer, D. B.; Dobson, C. M. *Biochemistry* **1995**, *34*, 4041–4055.
(18) Mulder, F. A. A.; Skrynnikov, N. R.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 967–975.
(19) Berglund, H.; Baumann, H.; Knapp, S.; Ladenstein, R.; Haerd, T. *J. Am. Chem. Soc.* **2002**, *117*, 12883–12884.
(20) Wilkinson, T. A.; Botuyan, M. V.; Kaplan, B. E.; Rossi, J. J.; Chen, Y. *J. Mol. Biol.* **2000**, *303*, 515–529.
(21) Paquin, R.; Ferrage, F.; Mulder, F. A.; Akke, M.; Bodenhausen, G. *J. Am. Chem. Soc.* **2008**, *130*, 15805–15807.
(22) Okamura, H.; Makino, K.; Nishimura, Y. *J. Mol. Biol.* **2007**, *367*, 1093–1117.
(23) Igumenova, T. I.; Frederick, K. K.; Wand, A. J. *Chem. Rev.* **2006**, *106*, 1672–1699.
(24) Best, R. B.; Clarke, J.; Karplus, M. *J. Am. Chem. Soc.* **2004**, *126*, 7734–7735.
(25) Best, R. B.; Clarke, J.; Karplus, M. *J. Mol. Biol.* **2005**, *349*, 185–203.
(26) Okamura, H.; Hanaoka, S.; Nagadoi, A.; Makino, K.; Nishimura, Y. *J. Mol. Biol.* **2000**, *295*, 1225–1236.
(27) Blanco, A. G.; Sola, M.; Gomis-Ruth, F. X.; Coll, M. *Structure* **2002**, *10*, 701–713.
(28) Yamane, T.; Okamura, H.; Ikeguchi, M.; Nishimura, Y.; Kidera, A. *Proteins* **2008**, *71*, 1970–1983.
(29) Ikeguchi, M. *J. Comput. Chem.* **2004**, *25*, 529–541.
(30) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
(31) MacKerell, A. D., Jr.; Banavali, N.; Foloppe, N. *Biopolymers* **2001**, *56*, 257–265.
(32) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(33) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

runs were performed at 1 atm and 310 K for each of 10 initial structures in the DNA-free and -bound forms.

The order parameters S^2_{NH} and S^2_{axis} were calculated from MD trajectories aligned by mass-weighted least-squares fitting using the equation $S^2_{NH/axis} = {}^3/_2\ \mathrm{tr}(\langle\mathbf{\Phi}\rangle^2) - {}^1/_2(\mathrm{tr}\langle\mathbf{\Phi}\rangle)^2$, where $\Phi_{ij} = r_i r_j/\mathbf{r}^2$ and the r_i ($i = 1, 2, 3$) are the x, y, and z components of the instantaneous vector $\mathbf{r}$ of the N−H bond and the C−C or S−C bond for S^2_{NH} and S^2_{axis}, respectively.[25] The order parameters were averaged over 10 trajectories in the DNA-free and -bound forms.

Results and Discussion

Backbone Dynamics. We carried out 10 independent MD simulations each for DNA-free and -bound PhoB^{DBTD}, starting from different NMR structures. All of the simulations were stable, with average backbone RMSDs from the initial NMR structures of 1.33 ± 0.25 Å for the DNA-free form and 0.96 ± 0.18 Å for the DNA-bound form.

Backbone order parameters calculated from simulation results ($S^2_{NH,calc}$) were in agreement with the experimental results ($S^2_{NH,exp}$) (Figure 2A). The correlation coefficients between $S^2_{NH,calc}$ and $S^2_{NH,exp}$ were 0.91 and 0.88 for the DNA-free and -bound forms, respectively. However, a single run of 10 ns was insufficient for sampling (e.g., the smallest correlation coefficient for a single run was 0.64). Averaging $S^2_{NH,calc}$ over 10 independent runs improved correlations with experimental data. Variations of $S^2_{NH,calc}$ for each run are indicated by the error bars in Figure 2.

Except for the wing loop connecting β6 and β7 (Figure 1), most of the S^2_{NH} remained unchanged upon DNA binding, consistent with slight backbone conformational changes upon DNA binding. Significant flexibility of the wing loop in the DNA-free form was suppressed upon DNA binding as a result of direct interactions of Arg219 and Thr223 with DNA.[28]

Side-Chain Order Parameters. The order parameters for side-chain methyl groups calculated from the simulations ($S^2_{axis,calc}$) were also well-correlated with the experimental data ($S^2_{axis,exp}$). The correlation coefficients between $S^2_{axis,calc}$ and $S^2_{axis,exp}$ were 0.82 and 0.79 for the DNA-free and -bound states, respectively. As for S^2_{NH}, averaging $S^2_{axis,calc}$ over 10 independent runs improved the correlations with the experimental data.

Although S^2_{axis} was previously shown to be only weakly correlated with static structural properties such as depth of burial and packing density, we found that the positional root-mean-square fluctuations (RMSFs) of the methyl carbons estimated from the MD simulations generally were well-correlated with S^2_{axis} (see the Supporting Information text and Figure S1). Nevertheless, the topology of the side chains also affected S^2_{axis}, as shown in Figure 3. Some residues exhibited significantly different S^2_{axis} values despite their low side-chain RMSFs. For example, the S^2_{axis} values for Met167 and Thr168 were very different (0.43 and 0.93, respectively), but the RMSFs of both side chains were <1 Å (Figure 3). As described in the Supporting Information, in branched side chains (e.g., threonine), the bond-vector directions are restricted when the positional deviations of the side-chain atoms are small (Figure 3A−D). In contrast, unbranched side chains (e.g., methionine) retain small positional deviations despite a wide variety of bond-vector directions in the side chain (Figure 3E−H). When the side-chain RMSFs of a methionine are large, the methionine exhibits a very low S^2_{axis} (e.g., 0.09 for Met147; Figure 3I−K). Therefore, for correct interpretations of S^2_{axis}, the topological effects of each amino acid type should be taken into account. The dependence of S^2_{axis} on side-chain types is shown in Figure 3M. To eliminate the amino acid-type dependence of S^2_{axis}, Mittermaier et al.[34] empirically proposed the normalized order parameter S^2_{norm} [defined as $S^2_{norm} = (S^2_{axis} - \mu_m)/\sigma_m$, where μ_m and σ_m are the average and standard deviation, respectively, of S^2_{axis} for each amino acid] on the basis of the experimental data for five proteins. Despite the fact that only one protein in the DNA-free and -bound states was simulated in this study, the averages of $S^2_{axis,calc}$ for each amino acid accurately reproduced the averages of the experimental values for five proteins (Figure 3L). Employing S^2_{norm} instead of S^2_{axis} reduced the dependence on the amino acid type (Figure 3N). Thus, these results indicate that S^2_{norm} is more suitable as an indicator of side-chain flexibility.

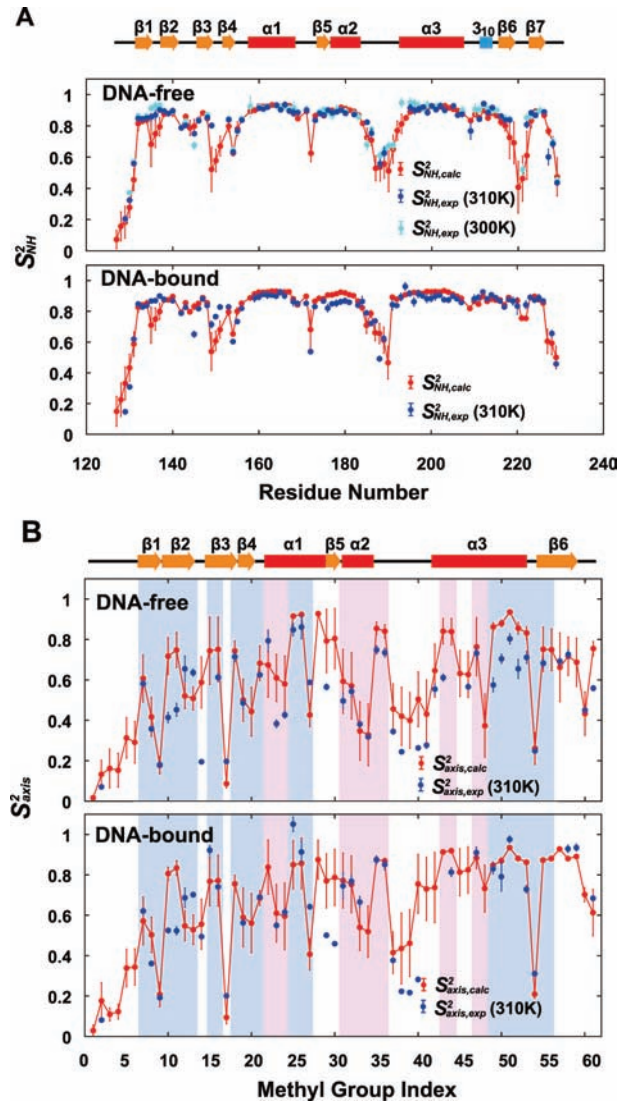

Figure 2. Comparison of the model-free order parameters obtained from the NMR experiments[22] and the MD simulations. The values from the MD simulations are averages over the results of 10 trajectories, and the error bars show the standard errors. In (A), the model-free order parameters for backbone amide groups calculated from the MD simulation results ($S^2_{NH,calc}$) and NMR experimental results ($S^2_{NH,exp}$) are plotted. In (B), the model-free order parameters for side-chain methyl groups calculated from the MD simulation results ($S^2_{axis,calc}$) and NMR experimental results ($S^2_{axis,exp}$) are presented. The methyl groups in the hard and soft hydrophobic cores are colored cyan and pink, respectively. The methyl group indices are listed in Table S1 in the Supporting Information.

(34) Mittermaier, A.; Davidson, A. R.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 9004–9005.

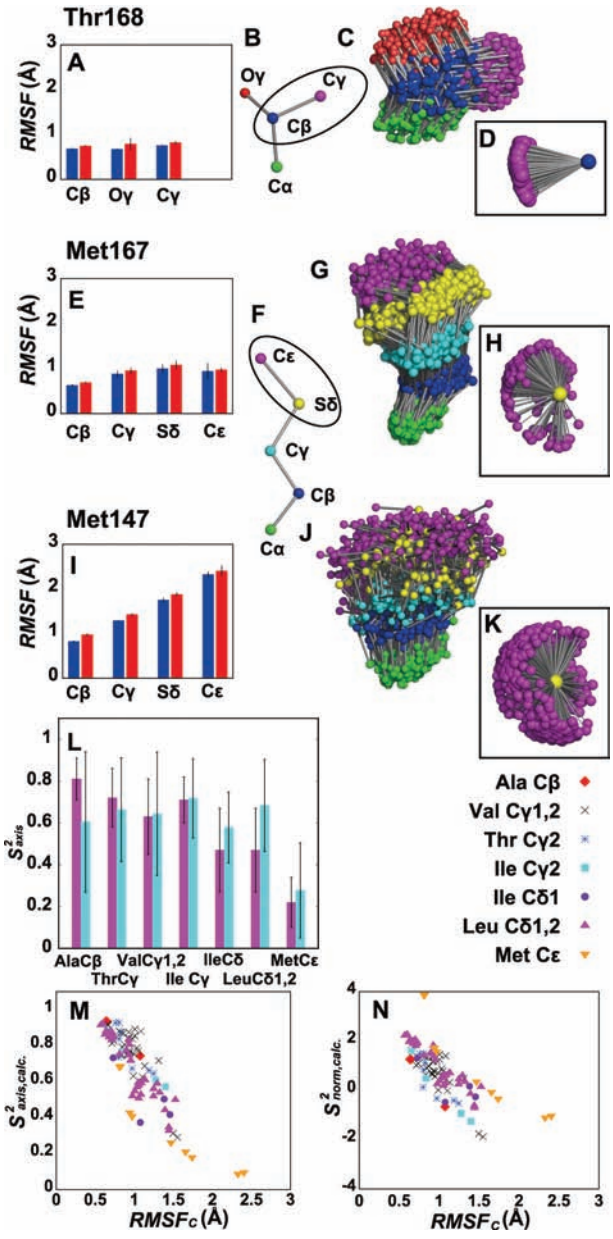

Figure 3. Topological effects of side chains on S^2_{axis}. As an example, residues showing largely different S^2_{axis} values were selected (0.93 for Thr168; 0.43 for Met167; 0.09 for Met147). The RMSFs of side-chain atoms of (A) Thr168, (E) Met167, and (I) Met147 in the DNA-free (blue bars) and -bound (red bars) forms are compared. Snapshots taken every 10 ps of a trajectory in the DNA-bound form are aligned for (C) Thr168, (G) Met167, and (J) Met147 to show fluctuations of the side chains. Colors are assigned to the side-chain atoms for (B) threonine and (F) methionine. Variations of bond vectors between the methyl carbon and the adjacent non-hydrogen atom are depicted for (D) Thr168, (H) Met167, and (K) Met147. In (D), (H), and (K), snapshots are translated in order to match the non-hydrogen atom covalently bonded to the methyl carbon. In (L), averages and standard deviations of S^2_{axis} from the MD simulations (magenta) and experiments (cyan) by Mittermaier et al.[34] are shown. (M) $S^2_{axis,calc}$ and (N) $S^2_{norm,calc}$ for residues with backbone-nitrogen RMSFs of less than 1 Å are plotted against RMSFs for methyl carbons ($RMSF_C$).

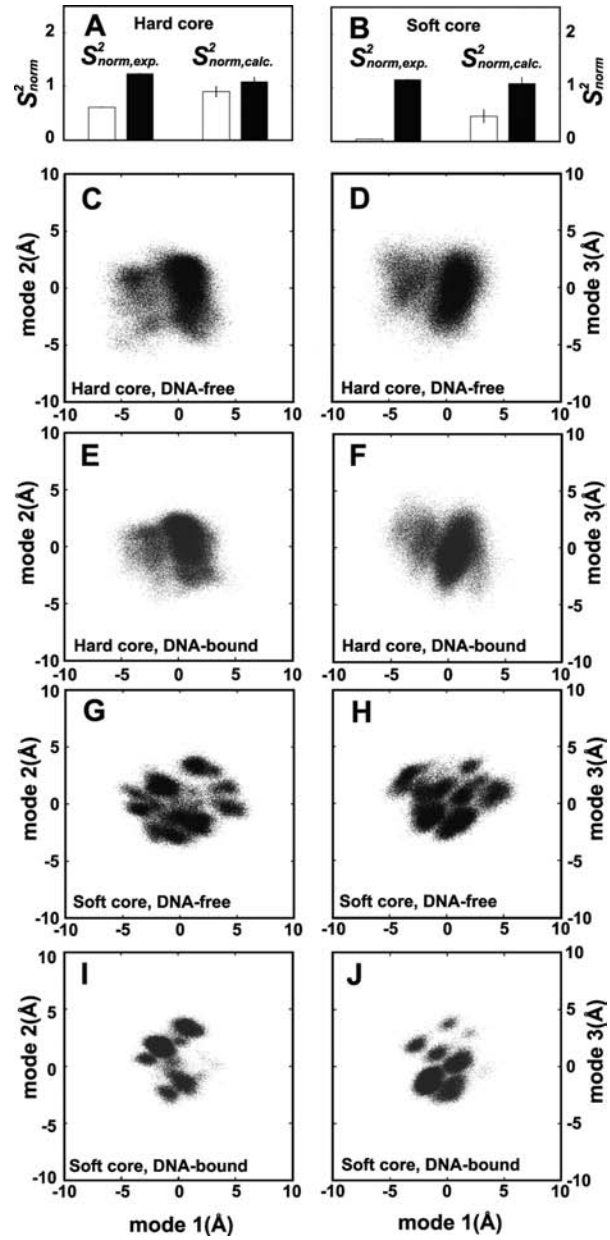

Figure 4. Difference in the side-chain dynamics of the hard and soft hydrophobic cores in the DNA-free and -bound forms. Averages of $S^2_{norm,exp}$ and $S^2_{norm,calc}$ for the (A) hard and (B) soft cores are plotted for the DNA-free (white bar) and -bound (gray bar) forms. PCA results for methyl carbons are shown for the (C−F) hard and (G−J) soft cores. Methyl carbon configurations taken every 1 ps in all trajectories in (C), (D), (G), and (H) for the DNA-free form and (E), (F), (I), and (J) for the DNA-bound form are projected onto the (C, E, G, I) mode 1−mode 2 space and the (D, F, H, J) mode 1−mode 3 space.

Soft and Hard Hydrophobic Cores. The S^2_{norm} analysis in previous NMR experiments for $PhoB^{DBTD}$ revealed that the side-chain dynamics in hydrophobic cores is significantly altered upon DNA binding.[22] The hydrophobic core of $PhoB^{DBTD}$ can be divided into two cores separated by the central α-helix α1 (Figure 1).[22] The core distant from DNA is mainly composed of the N-terminal β-sheet (β1−4), a face of α1, and the C-terminal half of α3 (blue spheres in Figure 1). Another core located near the DNA is mainly composed of a face of α1, α2, the N-terminal half of α3, and the C-terminal β-sheet (β5−7) (pink spheres in Figure 1). Before DNA binding, side chains in the core close to DNA exhibited a remarkably small S^2_{norm} (Figure 4B) relative to those in the core distant from DNA (Figure 4A), suggesting that the side chains in the core close to DNA are more fluidlike than those in the core distant from DNA. Therefore, the cores close to and distant from DNA are herein called the "soft" and "hard" cores, respectively. The definitions of the cores are the same as those used in the previous NMR study.[22] S^2_{norm} in both hydrophobic cores increased upon DNA

binding (Figure 4A,B), and thus, the side-chain packing appeared to be enhanced by DNA binding. However, the detailed mechanism underlying the changes in side-chain dynamics remains unclear.

The S^2_{norm} values derived from our MD simulations ($S^2_{norm,calc}$) can reproduce the trends observed experimentally for the soft and hard hydrophobic cores (Figure 4A,B). $S^2_{norm,calc}$ was much smaller in the soft core than in the hard core before DNA binding, and then $S^2_{norm,calc}$ in both cores increased upon DNA binding. Because the ensemble of atomic structures in the MD simulations was available, in contrast to the case of the NMR experiments, analyzing the side-chain fluctuations in the structural ensemble allowed us to elucidate the detailed mechanism underlying the changes in side-chain dynamics.

To this end, we performed principal component analysis (PCA) for the methyl carbons in both the hard and soft cores. As shown in Figure 4C−J, projections of the methyl carbon configurations onto the PCA modes demonstrated striking differences between the hard and soft cores. In the hard core, the methyl carbon configurations were simply distributed around one or two maxima. In contrast, the soft core exhibited a number of clusters of methyl carbon configurations. Each of the trajectories of the DNA-free and -bound forms visited multiple clusters. Splitting of the clusters was not due to the fluctuations of the backbones; PCA of the backbones showed only one large cluster for both the hard and soft cores (Figure S3 in the Supporting Information).

In the soft core, some of the clusters existing in the DNA-free state appeared to be disallowed in the DNA-bound state. Upon DNA binding, the populations of some of the clusters were significantly reduced and a few clusters dominated, even though a dynamic equilibrium still existed. To confirm this observation, we performed *K*-means clustering with the total number of clusters ($K = 10$) for methyl carbon configurations of the soft core, including all of the modes instead of projections onto a few modes (Figure S2 in the Supporting Information). Upon DNA binding, the populations of five out of the 10 clusters in the DNA-free state were significantly reduced in the DNA-bound state. In contrast, the populations of the other four clusters were largely enhanced. Even when the number of clusters used as an input for the *K*-means clustering was varied, similar results were obtained. This behavior of side-chain dynamics clearly corresponds to the population-shift mechanism for conformational changes upon ligand binding.

The structural origins of the clusters appearing in the PCA were side-chain rotamers in the soft core. Each cluster could be characterized by a specific combination of side-chain rotamers in the soft core (Table S4 in the Supporting Information). Figure 5A,B shows that PC mode 1 was mainly related to the Leu162 side-chain rotamers a and b, as defined in Table S2 in the Supporting Information. In the DNA-free state, the side chain of Leu162 exhibited interconversions between the two rotamers (Figure 5A). Upon DNA binding, the population of the side-chain conformations of Leu162 shifted toward rotamer a (Figure 5B and Table S3 in the Supporting Information). NMR experiments also showed that the Leu162 side chain equilibrates between the two rotamers (a and b) in the DNA-free state and that rotamer a dominates upon DNA binding.[22] Hence, the experimental observations for the Leu162 side chain were reproduced in our calculation. The dynamical behavior of the Leu162 side chain can be explained by interactions with the adjacent Trp184 (Figure 6B,C). The distance between Leu162 and Trp184 is well-correlated with the rotamers of

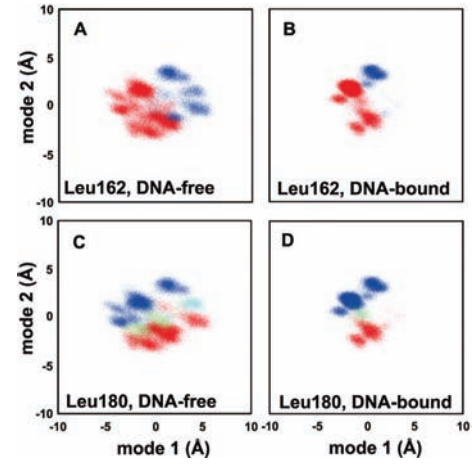

Figure 5. Relationship between side-chain rotamers and PCA projections of methyl carbon configurations in the soft core. In (A) and (B), PCA projections onto the mode 1−mode 2 space are colored by the rotamer type of Leu162: rotamer a (red) and b (blue) (Table S2 in the Supporting Information). In (C) and (D), the same PCA projections are colored by the rotamer type of Leu180: rotamer a (red), b (blue), c (green), and d (cyan).

Leu162 (Figure S5A−C in the Supporting Information): rotamer a prefers the short distance (∼5 Å), while rotamer b favors the long distance (∼6 Å) (Figure S5B). After DNA binding, as a result of positional stabilization of Trp184 by a direct interaction with DNA (Figure 6C), the favorable distance for rotamer a was populated more (Figure S5C), leading to a population shift toward rotamer a in Leu162. The NMR experiments also support the stabilization of Trp184 and adjacent regions upon DNA binding.[22]

Even for each rotamer of Leu162, multiple clusters appeared in the PCA, indicating that the dynamical equilibrium occurring in the soft core of PhoB was not solely due to the rotamers of Leu162 (Table S4 in the Supporting Information). PC mode 2 was associated with side-chain rotamers of Leu180 (Figure 5C,D). The four rotamers a−d (two major and two minor rotamers), were observed for Leu180 in the DNA-free state (Figure 5C). In contrast, rotamer b was mainly selected in the DNA-bound state (Figure 5D and Table S3 in the Supporting Information). The dynamical nature of the Leu180 side chain in the DNA-free state was also found in the NMR experiments.[22] The population shift of the Leu180 side-chain rotamers can be explained by interactions with adjacent Val190. The closest interatomic distance between Leu180 and Val190 is ∼3.5 Å in both the DNA-free and -bound states, indicating that hydrophobic interactions are maintained in both states (Figure S5D−F in the Supporting Information). However, Val190 largely fluctuates in the DNA-free state, and thus, Leu180 changes its rotamers depending on the position of Val190 (Figure 6D). In the DNA-bound state, fluctuations of Val190 are suppressed as a result of interactions with DNA, so the side-chain configurations of Leu180 are stabilized (Figure 6E).

As indicated by Leu162 and Leu180, the population shift in the soft core takes place through conformational stabilization of residues directly interacting with DNA. Figure 6A displays structural mapping of RMSF changes of methyl carbons upon DNA binding. Methyl groups in the interface regions between the protein and DNA exhibited stabilization upon DNA binding. The soft core is surrounded by the interface regions that directly interact with DNA. As for Leu162 and Leu180 described above, Ile199 is also an example of indirect stabilization. In the DNA-

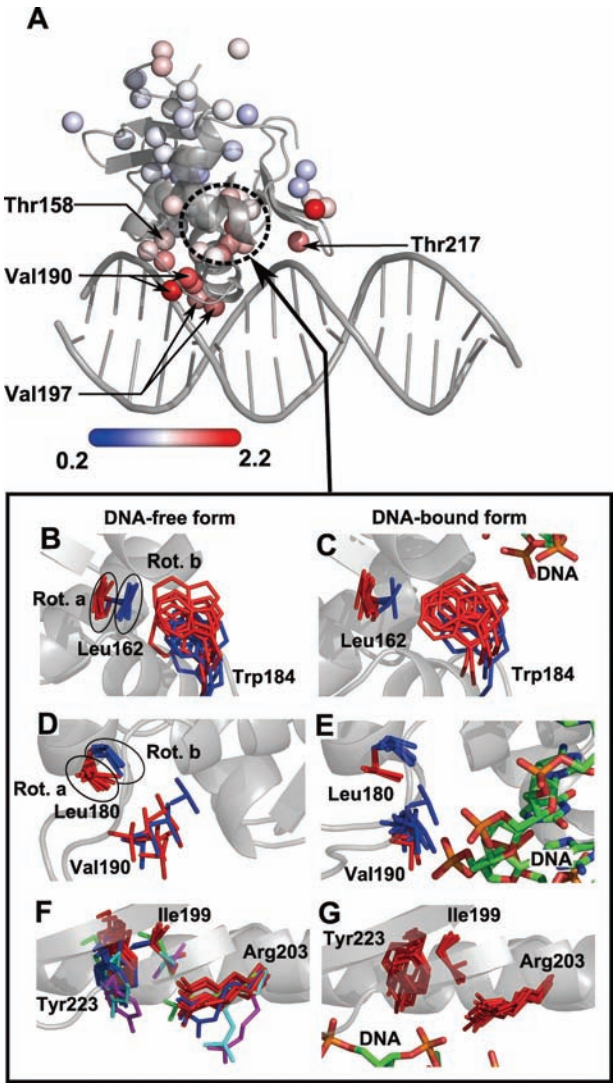

Figure 6. Conformational changes in the side chains of $PhoB^{DBTD}$ upon DNA binding. In (A), the RMSF ratios ($RMSF_{DNA\text{-}free}/RMSF_{DNA\text{-}bound}$) of the methyl carbons are mapped onto the structure. The residues that directly interact with DNA (Thr158, Val190, Val197, and Thr217) are labeled.[28] The soft core is indicated by a black dashed-line circle. Conformations of side chains in the soft core are drawn for (B, C) Leu162, (D, E) Leu180, and (F, G) Ile199. Snapshots were fitted using four atoms (N, Cα, Cβ, C) of the respective residues. In (B) and (C), the side chains of Leu162 and Trp184 are colored by the rotamer type of Leu162: rotamer a (red) and b (blue). In (D) and (E), the side chains of Leu180 and Val190 are colored by the rotamer type of Leu180: rotamer a (red) and b (blue). In (F) and (G), the side chains of Ile199, Arg203, and Tyr223 are colored by the rotamer type of Ile199: rotamer a (red), b (blue), c (green), d (cyan), and e (magenta).

free state, the side chain of Ile199 adopted 7 rotamers (a to g), of which two are major and five are minor (Figure 6F and Tables S2 and S4 in the Supporting Information). In the DNA-bound state, stabilization of the adjacent Arg203 that directly interacts with DNA excluded the minor rotamers of Ile199 that pointed to Arg203 (Figure 6G). Selection between the major rotamers a and b was affected by the position of Tyr223. The side chain of Tyr223 interacting with DNA is packed against rotamer a of Ile199, leading to stabilization of rotamer a and resulting in a population shift toward it upon DNA binding (Figure 6G and Table S3 in the Supporting Information). Stabilization of the Tyr223 side chain upon DNA binding was also observed in the NMR experiments.[22]

Transactivation Loop. After PhoB binds to its cognate DNA, PhoB recruits the σ^{70} subunit of RNA polymerase to trigger transcription initiation. Because free PhoB does not bind to the σ^{70} subunit, some changes in the binding site for the σ^{70} subunit would be expected. In PhoB, the region responsible for binding to σ^{70} is the transactivation loop (158−191) linking α2 and α3 of $PhoB^{DBTD}$. Most of the residues in the transactivation loop are highly flexible in both the DNA-free and -bound forms (Figure 2A). However, the two C-terminal residues of the transactivation loop, Val190 and Glu191, exhibit clear stabilization upon DNA binding as a result of direct interactions with DNA.[28] As described above, stabilization of Val190 upon DNA binding affects the dynamics of the soft core of $PhoB^{DBTD}$ (Figure 6E). Experimentally, mutation of Val190 results in a decreased interaction with RNA polymerase. Interestingly, in addition to Val190, mutation of Trp184, which also affects the dynamics of the soft core (as described above), reduces the interaction with RNA polymerase.[35]

Recent literature has proposed a dynamic model for addressing how transcription factors selectively recognize specific binding sites among many similar ones available in the genome.[36−38] In this model, allostery plays a key role in recognition and selectivity. Specifically, transcription factor proteins exist as dynamic ensembles of molecules in different conformational states with different DNA-binding properties, and slight differences in DNA sites allosterically affect these protein ensembles at the cofactor binding site. An interesting example of such a protein is the glucocorticoid receptor.[39] The crystal structures of the glucocorticoid receptor bound to three different DNA sites whose sequences are very similar to each other show that DNA binding allosterically leads to a conformational change at the coregulator binding site. Despite the current lack of a PhoB/DNA/RNA polymerase triple complex structure, some changes in the dynamics of the PhoB transactivation loop upon DNA binding were observed as described above, with the soft core exhibiting dynamic changes of the residues adjacent to the transactivation loop upon DNA binding. To extend these observations, detailed investigations of the dynamics in the PhoB/DNA/RNA polymerase complex and/or the PhoB/DNA complex with slightly different DNA sequences are required; these remain the object of future work.

Concluding Remarks

Side-chain conformational changes upon DNA binding in $PhoB^{DBTD}$ were investigated using all-atom MD simulations and analyses of NMR relaxation order parameters. Strong correlations between the simulated and experimental order parameters for both the backbone and side-chain methyl groups indicated that the MD simulations provided an accurate structural interpretation of the order parameters.

The MD simulations demonstrated a striking contrast in the dynamical behavior of the side chains between the soft and hard hydrophobic cores. The side chains in the soft core exhibited

(35) Makino, K.; Amemura, M.; Kawamoto, T.; Kimura, S.; Shinagawa, H.; Nakata, A.; Suzuki, M. *J. Mol. Biol.* **1996**, *259*, 15–26.
(36) Pan, Y.; Tsai, C. J.; Ma, B.; Nussinov, R. *Trends Genet.* **2010**, *26*, 75–83.
(37) Ma, B.; Tsai, C. J.; Pan, Y.; Nussinov, R. *ACS Chem. Biol.* **2010**, *5*, 265–272.
(38) Pan, Y.; Tsai, C. J.; Ma, B.; Nussinov, R. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1118–1120.
(39) Meijsing, S. H.; Pufall, M. A.; So, A. Y.; Bates, D. L.; Chen, L.; Yamamoto, K. R. *Science* **2009**, *324*, 407–410.
(40) DeLano, W. L. *PyMOL*; DeLano Scientific LLC: Palo Alto, CA, 2008.

dynamical conversions among multiple rotamers in the DNA-free state. DNA binding indirectly restricted the side-chain motions in the soft core, resulting in stabilization to a few configurations. This disorder−order transition upon molecular recognition clearly corresponds to the population-shift or pre-existing equilibrium mechanism. To date, the population-shift mechanism upon molecular binding has been recognized for conformational changes in backbones.[4,5] This study demonstrates that the population-shift mechanism operates even in hydrophobic side chains that do not directly interact with DNA. DNA binding directly and indirectly affects dynamics in the domain and thus induces the low-entropy state of the protein. To our knowledge, our calculation is the first to demonstrate this kind of disorder−order transition of the hydrophobic cores upon molecular recognition.

It is noteworthy that the flexible soft core is located close to the molecular recognition site and that the immobile hard core resides far from the recognition site. According to recent literature,[5] ubiquitin also utilizes flexible sites for recognition of partner proteins. In PhoB, in addition to the soft core, the transactivation loop connecting α2 and α3, which is highly flexible even after DNA binding, reportedly interacts with the σ^{70} subunit of RNA polymerase. The flexibility of this loop has been argued to be crucial in the transactivation activity.[27] The disorder−order transition upon molecular recognition generally leads to a loss of entropy and thus has a negative impact on binding affinity. However, because binding of transcription factors to DNA is transient, it is unnecessary for the binding affinity to be too strong. Furthermore, flexibility provides adaptability for molecular recognition. During molecular evolution, flexibility might have advantages for adapting to changes in partners and creating new interactions. The dynamics−function relationship is of great importance in understanding protein function.

Acknowledgment. This work was partly supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) and by Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of MEXT. All of the computations were carried out at the Tsurumi campus of Yokohama City University (Yokohama, Japan).

Supporting Information Available: Correlation between positional fluctuation of methyl carbons and S^2_{axis}; topological effects of side chains on S^2_{axis}; Figures S1−S5; Tables S1−S4. This material is available free of charge via the Internet at http://pubs.acs.org.

JA103218X