THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

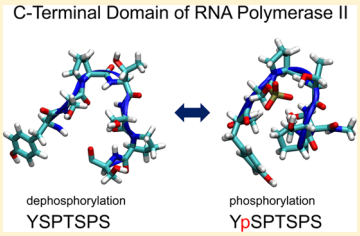
pubs.acs.org/JPCB

Molecular Dynamics Study of the Phosphorylation Effect on the Conformational States of the C-Terminal Domain of RNA Polymerase II

Yasushige Yonezawa*

High Pressure Protein Research Center, Institute of Advanced Technology, Kinki University, 930 Nishimitani, Kinokawa, Wakayama 649-6493, Japan

ABSTRACT: The carboxyl-terminal domain (CTD) of RNA polymerase II in eukaryotes regulates mRNA processing processes by recruiting various regulation factors. A main function of the CTD relies on the heptad consensus sequence (YSPTSPS). The CTD dynamically changes its conformational state to recognize and bind different regulation factors. The dynamical conformation changes are caused by modifications, mainly phosphorylation and dephosphorylation, to the serine residues. In this study, we investigate the conformational states of the unit consensus CTD peptide with various phosphorylation patterns of the serine residues by extended ensemble simulations. The results show that the CTD without phosphorylation has a flexible disordered structure distributed between twisted and extended states, but phosphorylation tends to reduce the conformational space. It was found that phosphorylation induces a β-turn around the phosphorylated serine residue and the cis conformation of the proline residue significantly inhibits the β-turn formation. The β-turn should contribute to specific CTD binding of the different regulation factors by changing the conformation propensity combined with induced fit.



INTRODUCTION

In eukaryotes, RNA polymerase II (Pol II) carries out RNA transcription. The largest subunit of Pol II is Rpb1, which has a long carboxyl-terminal domain (CTD). The CTD is composed of tandem repeat copies of the heptad Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. The consensus sequence is evolutionally well-conserved, providing a dynamic platform to recruit different regulators of the transcription apparatus.[1−7] The heptad is repeated 26−52 times depending on the species. The binding of processing factors is mainly regulated by phosphorylation and dephosphorylation of the residues Ser2 and Ser5 of the CTD. The conformational states of the CTD, called the CTD code, represent a critical regulatory checkpoint for transcription. Efficient capping, splicing, and polyadenylation of mRNA in vivo require the CTD. The CTD is flexibly linked to the Pol II subunit core close to the presumed RNA exit site, and binds to the mRNA-processing factors. Pcf11 is a conserved subunit of the yeast cleavage factor that is required for 3′-RNA processing and transcription termination.[5] Recently, a complex CTD and Pcf11 structure was resolved, which showed a twisted β-turn conformation around the phosphorylated Ser2 and Pro3 residues.[6] A β-spiral model was proposed to explain a CTD conformation state by the authors. However, the molecular basis of the specificity between the CTD and a variety of the binding factors remains unclear. Thus, the conformation states of the CTD due to site-specific modifications (e.g., phosphorylation) are crucial for understanding the CTD code.

In this paper, we investigate the CTD conformational state in solution with several phosphorylation patterns of serine residues using molecular dynamics (MD) simulations. MD simulations are a useful tool for investigating molecular properties of materials, including proteins, DNA, and RNA, and play a complementary role to experiments. However, standard MD simulations of macromolecules using the Boltzmann-weighted ensemble struggle to reach biologically relevant time scales, from microseconds to milliseconds, even using up-to-date modern massive parallel computers, because of the roughness of the free energy and the large energy barriers between stable minima.

To overcome this issue, extended ensemble simulation methods are frequently used. The extended simulations involve replica exchange molecular simulations,[8] the meta-dynamics method,[9] the adaptive-biased umbrella sampling method,[10] and the multi-canonical ensemble method.[11] Multi-canonical molecular dynamics (McMD) treats the temperature as a reaction coordinate, resulting in much faster sampling of the free energy landscape of macromolecules.[12−16]

We have used McMD to study the conformational states of the heptad CTD peptide with explicit water at biological ionic strength under periodic boundary conditions. We simulated and studied various phosphorylation patterns of the CTD peptide: phosphorylation at Ser2, phosphorylation at Ser5, phosphorylation at both Ser2 and Ser5, and the heptad CTD peptide without phosphorylation. The cis conformation of the peptidyl−prolyl bond of the sixth proline and the phosphorylated fifth serine residue of the CTD peptide binding to the transcription factors Nrd1[17] and Ssu72[18] have been exper-

Received: October 14, 2013
Revised: March 7, 2014
Published: March 10, 2014

dx.doi.org/10.1021/jp4101976 | J. Phys. Chem. B 2014, 118, 4471−4478

imentally observed. To investigate the cis conformation effects, we also simulated the CTD peptide with the peptidyl−prolyl bond of the sixth proline set to the cis conformation.

MATERIALS AND METHODS

Preparation of the CTD System. The unit repeat of the CTD heptad peptide Ace-Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7-NMe, whose two termini are methylated to avoid strong electrostatic attractions, was used as the unphosphorylated CTD peptide (u-CTD peptide). On the basis of the u-CTD peptide, we constructed three peptides with the serine residues singly or doubly phosphorylated: the peptide phosphorylated at Ser2 (CTD-pSer2), the peptide phosphorylated at Ser5 (CTD-pSer5), and the peptide doubly phosphorylated at Ser2 and Ser5 (CTD-pSer2/pSer5). For the phosphorylated serine residue, we set the overall charge to −2e. We set the proline residues of the above four peptides to the trans conformation. In addition, we constructed two peptides with Pro6 in the cis conformation: with phosphorylation at Ser5 (CTD-pSer5-cis) and without phosphorylation at Ser5 (u-CTD-cis). We then modeled the simulation system by immersing each of the CTD peptides in the center of a pre-equilibrated 40 × 40 × 40 Å^3 TIP3P water[19] cubic box containing about 2000 water molecules, and removed the overlapping water molecules within 3.0 Å of the CTD peptides. All of the systems were neutralized by adding ions of 150 mM ionic strength. The initial conformation of the CTD peptides was set to the extended form. The center of mass of the CTD peptide was fixed to prevent it from drifting during the simulations.

We minimized the energy of the system by 300 steps of the steepest descent method followed by 100 steps of the conjugate gradient method. Next, we gradually equilibrated the system from 3 to 310 K over 100 ps using NVT molecular dynamics simulations with the Hoover−Evans method.[20] The SHAKE algorithm[21] was applied to all covalent bonds associated with hydrogen atoms. The Amber force field[22] (parm99) was used to model the atomic interactions, although improved versions of the force field have been proposed.[23] We used the particle mesh Ewald (PME) method[24] for the long-range electrostatic interactions. The parameter α of the PME method was set to 0.35 Å^{-1}, and the real-space cutoff distance was set to 10 Å for all simulations. The mesh size was set to 40 × 40 × 40, thereby ensuring a grid density of 1 Å for the system with sufficient accuracy from the Ewald method. In addition, we used the group-based cutoff scheme. A cutoff of 10 Å was used for the treatment of the short-range van der Waals interactions. We integrated the equations of motion using the velocity Verlet integration algorithm with a time step of 2 fs.

Multi-Canonical Molecular Dynamics Simulations. We used the McMD simulation method[13] to sample the conformational space. The McMD algorithm has been described previously,[15,16] so we will only briefly summarize it here. We generated the McMD ensemble by performing constant-temperature MD simulations at an arbitrarily chosen multi-canonical temperature ($T_{mc} = 1/(k_B\beta_{mc})$) with force scaling:

$$\frac{d\mathbf{p}_i}{dt} = \nu(E)\mathbf{f}_i \tag{1}$$

$$\nu(E) = \frac{\partial \alpha_{mc}(E)}{\partial E} \tag{2}$$

$$\alpha_{mc}(E) = E + \frac{1}{\beta_{mc}} \ln P(E) \tag{3}$$

where β_{mc} is the inverse multi-canonical temperature, k_B is Boltzmann's constant, and $\mathbf{p}_i$ and $\mathbf{f}_i$ are the momentum and the force of the ith atom, respectively. $\alpha_{mc}(E)$ is the multi-canonical energy, and $\nu(E)$ represents the force scaling factor. $\nu(E)$ was determined using the following iterative scheme:

$$\nu^{k+1}(E) = \nu^k(E) + \frac{1}{\beta_{mc}} \frac{\partial \ln P^k(E)}{\partial E} \tag{4}$$

Here, $P^k(E)$ is the probability distribution of the potential energy from the kth iterative run of the McMD. Furthermore, $\nu(E)$ is related to the density of states ($\Omega(E)$) through the following equation:

$$\frac{1}{\Omega(E)} \propto e^{-\beta_{mc}\alpha_{mc}(E)} \tag{5}$$

Once $\nu(E)$ has converged, the system can achieve a random walk on the potential energy landscape. We set the temperature range for the McMD simulation to 270−600 K, because the peptide can take various conformations at these temperatures. We set the multi-canonical temperature to 100 K. We then obtained flat energy distributions covering the temperature range from 270 to 600 K, performed simulations for 1.6 × 10^7 steps, and stored snapshots every 1 ps. All simulations were performed using the *myPresto* MD program.[25]

Potential of Mean Force and Reweighting Method. The potential of mean force (PMF) is defined as

$$\mathrm{PMF} = -k_B T_0 \ln P(E, \beta_0) \tag{6}$$

The potential of mean force at a desired temperature, with the original potential energy from the McMD method, is reproduced using reweighting techniques, as described in the following. A standard multi-canonical reweighting technique to a desired temperature T_0 is given by

$$P(E, \beta_0) = \frac{e^{\beta_{mc}\alpha_{mc}(E)} e^{-\beta_0 E}}{Z_0} P_{mc}(E, \beta_{mc}) \tag{7.1}$$

where

$$Z_0^{-1} = \frac{\int dE \Omega(E) e^{-\beta_{mc}\alpha_{mc}(E)}}{\int dE \Omega(E) e^{-\beta_0 E}} \tag{7.2}$$

In these equations, $P(E, \beta_0)$ represents the probability distribution produced from the potential energy E at $T_0 = 1/(k_B\beta_0)$, where T_0 represents the reweighting (desired) temperature and Z_0 is the partition function. $P_{mc}(E, \beta_{mc})$ stands for the probability distribution from the McMD simulations.

RESULTS AND DISCUSSION

End-to-End Distance. We analyzed the trajectories from the McMD simulations to characterize the conformational state of the CTD peptides. We then plotted the free energy landscapes of the CTD peptides reweighted at 310 K as a function of the distance between the α-carbon atom of the Tyr1 residue and the α-carbon atom of the Ser7 residue of the CTD peptides (end-to-end distance), as shown in Figure 1.

In Figure 1a, CTD peptides whose proline residues are in the trans conformation are shown. The u-CTD peptide has a wide distribution of end-to-end distances ranging from extended to

dx.doi.org/10.1021/jp4101976 | J. Phys. Chem. B 2014, 118, 4471−4478

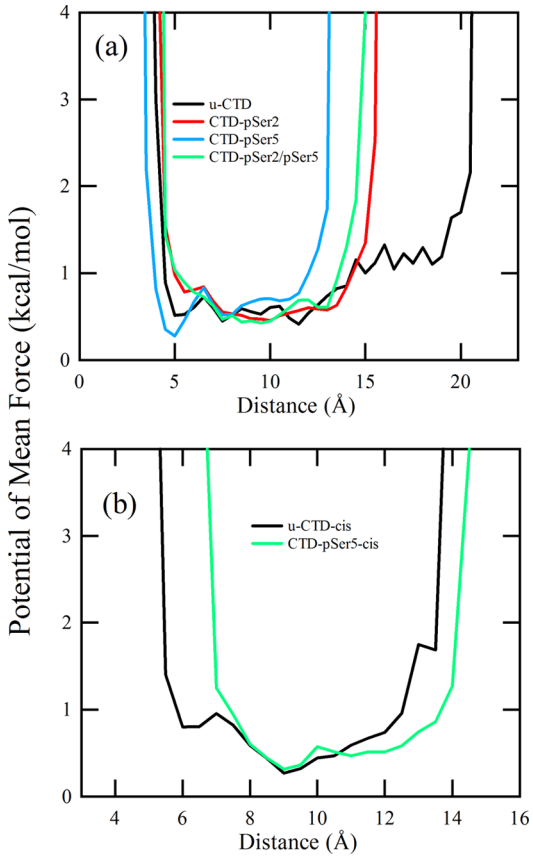


Figure 1. Potential of mean force at 310 K reweighted from the McMD simulations as a function of the distance between the α-carbon atom of the Tyr1 residue and the α-carbon atom of the Ser7 residue of the CTD peptide (end-to-end distance): (a) CTD peptides whose proline residues are in the trans conformation and (b) CTD peptides whose Pro6 residues are in the cis conformation.

twisted conformations, indicating that the u-CTD has a very flexible structure. The end-to-end distance of the CTD-pSer2 peptide has a narrower distribution than that of the u-CTD peptide. Moreover, the end-to-end distance of CTD-pSer5 has a narrower distribution than that of CTD-pSer2. The profiles of the distributions of CTD-pSer2/pSer5 and CTD-pSer2 are very similar. The phosphorylated CTD peptides have a low potential of mean force at small end-to-end distances, and the distributions of the phosphorylated CTD peptides are centered at 9 Å.

The results clearly show that phosphorylation of the serine residues significantly reduces the conformational space of the CTD peptide, resulting in more compact conformational states. Furthermore, phosphorylation tends to induce twisted conformations. It is interesting that the reduction of the conformational space for phosphorylation at Ser5 is greater than that for phosphorylation at Ser2.

In Figure 1b, the end-to-end distances for the CTD peptides whose Pro6 residues are only in the cis conformation are shown. It shows that the u-CTD-cis peptide has smaller end-to-end distances than the u-CTD peptide, because the carboxyl group of Thr4 and the amide group of NMe tend to form hydrogen bonds in the u-CTD-cis peptide. In addition, the end-to-end distance distribution of CTD-pSer5-cis shifts about 2 Å from that of u-CTD-cis.

Potential of Mean Force of the Dihedral Angle Space. The backbone dihedral angles of peptides describe the conformations of peptides. There have been both experimental and theoretical conformational studies of short peptides, and the studies used the dihedral angles to evaluate the conformational space.[26,27] To compare the conformational states of the important residues in the CTD peptides, we evaluated the potential of mean force with respect to dihedral angles in the CTD peptide. We used a pair of dihedral angles: ϕ, C(O)−N(H)−Cα−C(O), and ψ, N(H)−Cα−C(O)−N(H). We then calculated the potential of mean force of the dihedral angles of Ser2, Pro3, Thr4, Ser5, and Pro6 for the different phosphorylation states, because the dihedral angle spaces of the five residues are considered to be important to characterize the conformational state of the CTD peptides. In the following, we show the potential of mean force for the five residues.

The potentials of mean force for Ser2 in the CTD peptides with proline residues in the trans conformation are shown in Figure 2. Herein, we define the free energy local minima as α_R

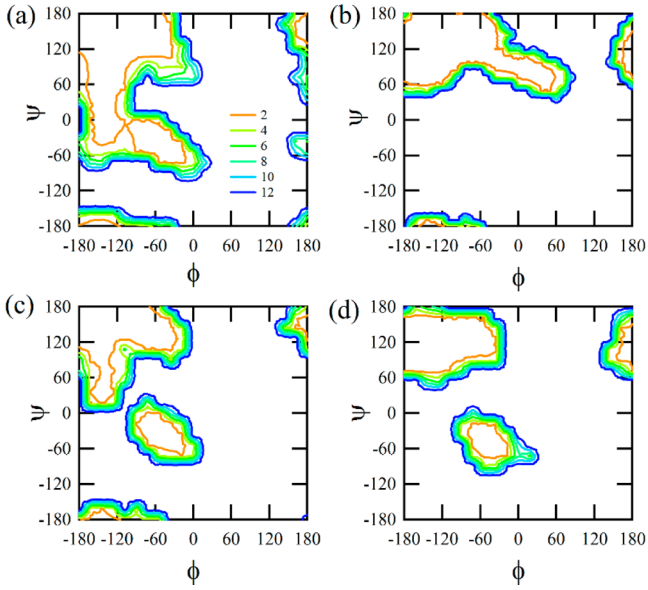


Figure 2. Contour plots of the potential of mean force for the Ser2 dihedral angle space: (a) u-CTD peptide, (b) CTD-pSer2 peptide, (c) CTD-pSer5 peptide, and (d) CTD-pSer2/pSer5 peptide with the proline residues in the trans conformation. The unit of the contours is kcal/mol, and ϕ and ψ are in deg.

(right-handed α-helix region): $-100° < \phi < -30°$ and $-80° < \psi < -5°$, α_L (left-handed α-helix region): $5° < \phi < 75°$ and $25° < \psi < 120°$, βs (region largely involved in β-sheet formation): $-180° < \phi < -50°$ and $80° < \psi < -170°$, P_{II} (region largely associated with the polyproline II-like helix): $-110° < \phi < -50°$ and $120° < \psi < 180°$, and γ (regions forming tight turn): $-100° < \phi < -60°$ and $0° < \psi < 60°$. For the u-CTD peptide, the α_R, βs, and P_{II} regions are observed, which indicates that the residue is very flexible. In the CTD-pSer2 peptide, the α_R region disappeared and the α_L region appeared compared with the u-CTD peptide. Although the potential of mean force of the α_R region is smaller in the CTD-pSer5 peptide, the dihedral angle space profile is similar to that of the u-CTD peptide. It is surprising that an α_R region exists for the CTD-pSer2/pSer5 peptide, while the corresponding region is not seen for the CTD-pSer2 peptide, indicating that phosphorylation at both Ser2 and Ser5 weaken the inhibition effect on the α_R region.

dx.doi.org/10.1021/jp4101976 | J. Phys. Chem. B 2014, 118, 4471−4478

The potentials of mean force for Pro3 of the CTD peptides with proline residues in the trans conformation are shown in Figure 3. In general, proline has restrictions in the $\phi-\psi$ space

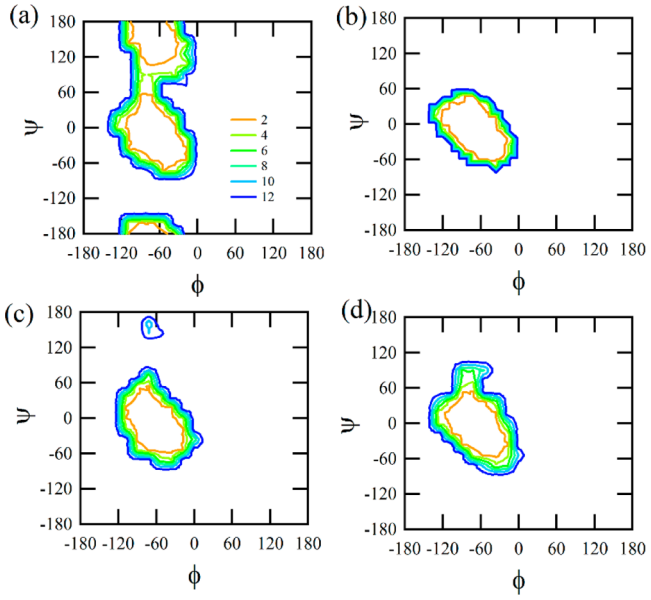


Figure 3. Contour plots of the potential of mean force for the Pro3 dihedral angle space: (a) u-CTD peptide, (b) CTD-pSer2 peptide, (c) CTD-pSer5 peptide, and (d) CTD-pSer2/pSer5 peptide with proline residues in the trans conformation. The unit of the contours is kcal/mol, and ϕ and ψ are in deg.

that arise from the five-membered ring: ϕ is restricted to approximately $-60°$ by the ring, and ψ falls into two groups at about -45 and $+135°$ in the α-helix and β-sheet regions of the Ramachandran plot. In Figure 3a, although the βs and P_{II} regions are observed for the u-CTD peptide, these regions disappeared in the three phosphorylated CTD peptides. Thus, phosphorylation of the serine residues prevents the βs and P_{II} regions of the proline residue due to electrostatic and steric effects. As a result, the Pro3 residue of the CTD peptides with phosphorylation of the serine residues tends to form α-helix rather than β-sheet.

As shown in Figure 4, all of the CTD peptides with respect to the potential of mean force of Thr4 with proline residues in the trans conformation have very similar profiles. All of the threonine residue dihedral angle spaces tend to be in α-helical regions. Therefore, phosphorylation of the serine residues does not significantly affect the dihedral angles of the Thr4 residue.

The potential of mean force for the dihedral angle of Ser5 with proline residues in the trans conformation is shown in Figure 5. For the u-CTD peptide, the α_R, α_L, βs, and P_{II} regions are observed, while the γ region is absent. This indicates that unphosphorylated Ser5 has a very flexible conformation, resulting in disordered structures. For the CTD-pSer2 peptide, the α_R region is observed, while the α_L region is absent. The P_{II} and γ regions are absent in CTD-pSer5. It is interesting that, despite phosphorylation of Ser5 in both cases, the α_R region is observed for CTD-pSer5, while it is not observed for CTD-pSer2/pSer5.

The potential of mean force of Pro6 without phosphorylation (Figure 6a) shows the same profile as Pro3 without phosphorylation (Figure 3a). Moreover, in contrast to the case of Pro3, phosphorylation does not significantly change the

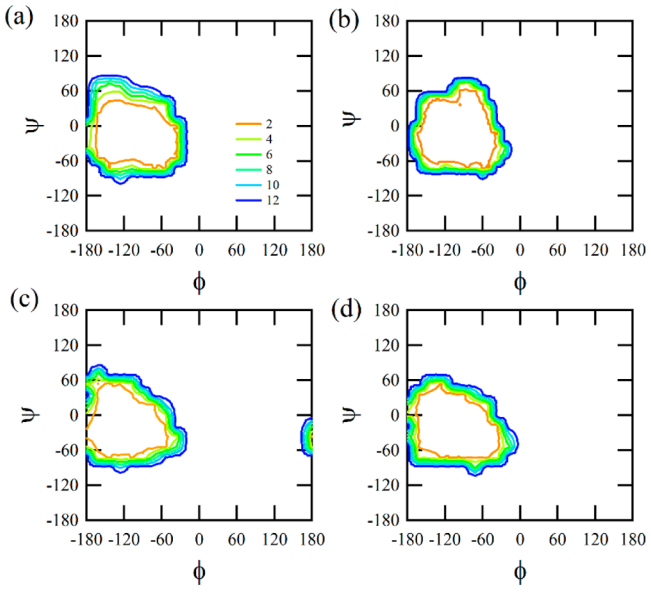


Figure 4. Contour plots of the potential of mean force for the Thr4 dihedral angle space: (a) u-CTD peptide, (b) CTD-pSer2 peptide, (c) CTD-pSer5 peptide, and (d) CTD-pSer2/pSer5 peptide with the proline residues in the trans conformation. The unit of the contour plots is kcal/mol, and ϕ and ψ are in deg.

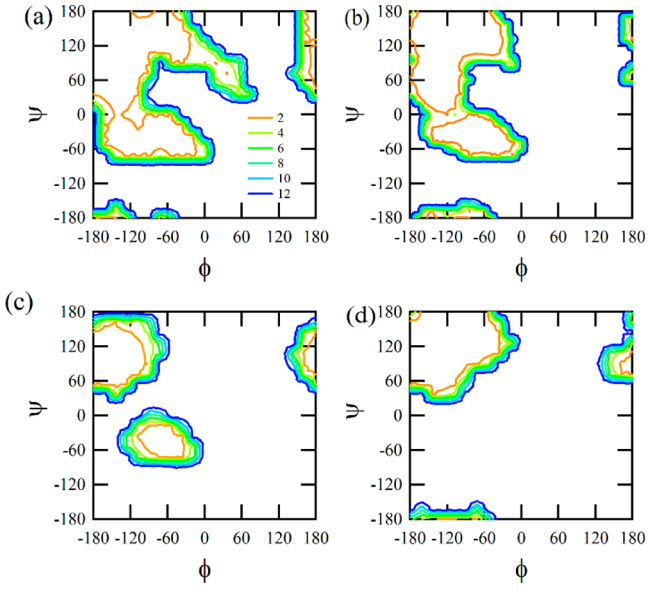


Figure 5. Contour plots of the potential of mean force for the Ser5 dihedral angle space: (a) u-CTD peptide, (b) CTD-pSer2 peptide, (c) CTD-pSer5 peptide, and (d) CTD-pSer2/pSer5 peptide with proline residues in the trans conformation. The unit of the contours is kcal/mol, and ϕ and ψ are in deg.

phosphorylated profiles of CTD-pSer2 compared with the unphosphorylated profile. Furthermore, for CTD-pSer2/pSer5, the βs and P_{II} regions are slightly reduced. The difference between Pro3 and Pro6 should influence the CTD binding specificity.

The potentials of mean force of the dihedral angle space for the u-CTD-cis peptide are shown in Figure 7. In this figure, the α_R and βs regions of Ser2 are clearly separate, indicating strong inhibition for the γ region. Compared with the case of the u-CTD peptide (Figure 3a), the βs and P_{II} regions of Pro3 are

dx.doi.org/10.1021/jp4101976 | J. Phys. Chem. B 2014, 118, 4471−4478

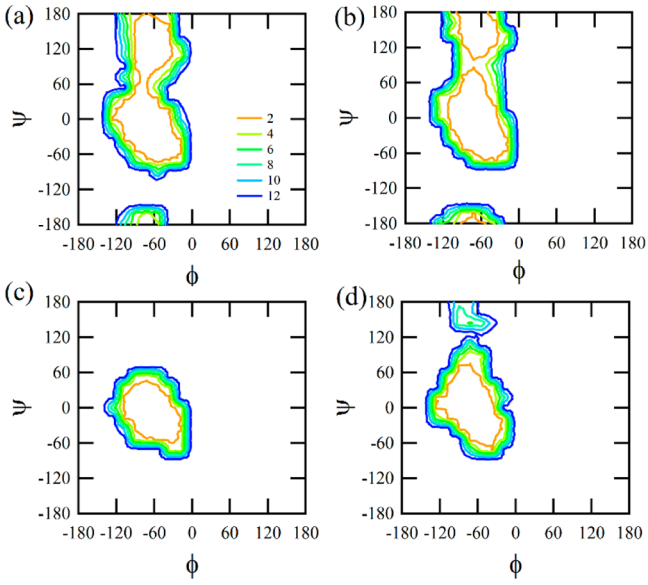


Figure 6. Contour plots of the potential of mean force for the Pro6 dihedral angle space: (a) u-CTD peptide, (b) CTD-pSer2 peptide, (c) CTD-pSer5 peptide, and (d) CTD-pSer2/pSer5 peptide with proline residues in the trans conformation. The unit of the contours is kcal/mol, and ϕ and ψ are in deg.

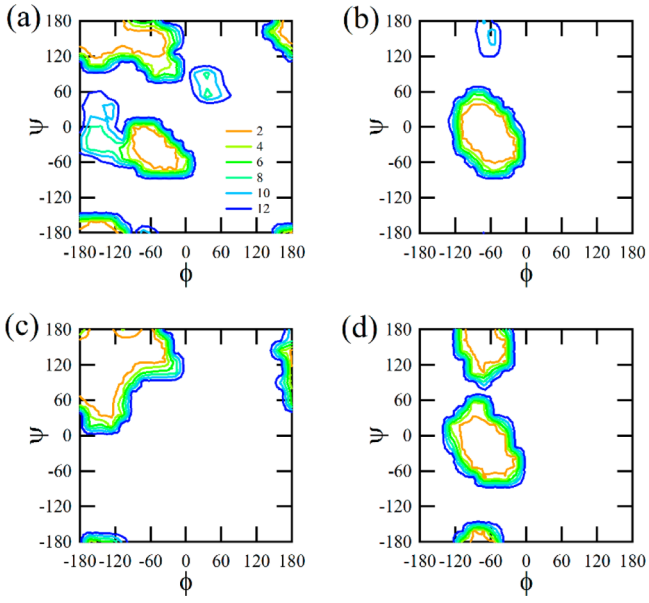


Figure 7. Contour plots of the potential of mean force for the dihedral angle space of the u-CTD-cis peptide for (a) Ser2, (b) Pro3, (c) Ser5, and (d) Pro6. The unit of the contours is kcal/mol, and ϕ and ψ are in deg.

considerably smaller. Moreover, the dihedral angle space of Ser5 is strictly restricted to the βs region, and the profile of the dihedral angle space of Pro6 is similar to that of u-CTD (Figure 6a). Because the profile of Thr4 with Pro6 in the cis conformation shows a very similar trend to that of the CTD peptide with the proline residues in the trans conformation, the data are not shown.

The potentials of mean force of the dihedral angle space for the CTD-pSer5-cis peptide are shown in Figure 8. From the figure, the profiles of Ser2, pSer5, and Pro6 are not affected by phosphorylation at Ser5, and show similar profiles to those for

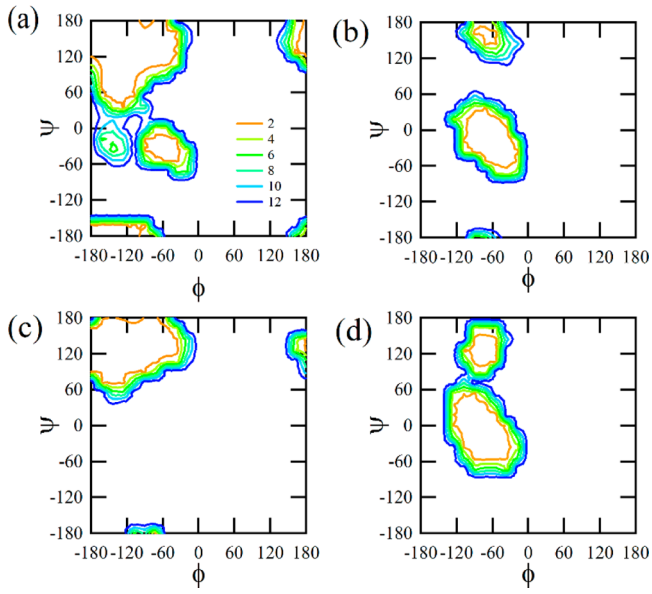


Figure 8. Contour plots of the potential of mean force for the dihedral angle space of the CTD-pSer5-cis peptide for (a) Ser2, (b) Pro3, (c) pSer5, and (d) Pro6. The unit of the contours is kcal/mol, and ϕ and ψ are in deg

the u-CTD-cis peptide, except that a small P_{II} region of Pro3 and a small fraction of the γ region of Ser2 are observed. Because the profile of Thr4 with Pro6 in the cis conformation shows a very similar trend to that of the CTD peptide with the proline residues in the trans conformation, the data are not shown.

β-Turn Conformations. It has been reported that the β-turn conformation of the CTD plays a significant role in recognizing and binding the regulation factors Pcf11.[6,7] β-turn conformations are frequently observed around proline. Here, we evaluate the β-turn conformations around (p)Ser2-Pro3-Thr4 and (p)Ser5-Pro6-Ser7 from the McMD trajectories. From the CTD-Pcf11 complex structure (PDBID: 1SZA), we found that the β-turn along pSer2-Pro3-Thr4 is mainly stabilized by interactions between pSer2 and Thr4 side chains to Pro3. The main chain dihedral angles (ϕ, ψ) of pSer2 and Pro3 are (−143°, 139°) and (−43°, −34°), respectively. This means that to construct the β-stabilizing interactions (hydrogen bonds) both a β-conformation of pSer2 and an α-conformation of Pro3 are required. The results from the potential of mean force of the dihedral angles in Figures 2 and 3, corresponding to Ser2 and Pro3, are in agreement with such a requirement for β-stabilizing interactions. We then defined the β-turn propensity using appropriate simple measures: the distance between the phosphorus atom of pSer2 and the hydroxyl oxygen of Thr4 is used for the β-turn measure of pSer2-Pro3-Thr4 (Figure 9a); the distance between the hydrogen atom of Ser2 and the hydroxyl oxygen of Thr4 is used for Ser2-Pro3-Thr4 (Figure 9b); the distance between the phosphorus atom of pSer5 and the hydroxyl oxygen of Ser7 is used for pSer5-Pro6-Ser7 (Figure 9c); and the distance between the hydrogen atom of Ser5 and the hydroxyl oxygen of Ser7 is used for Ser5-Pro6-Ser7 (Figure 9d). The potential of mean force of the hydrogen bonds as a function of distance for the CTD peptides when the proline residues are in the trans conformation and when Pro6 is in the cis conformation are shown in Figures 10 and 11, respectively.

dx.doi.org/10.1021/jp4101976 | J. Phys. Chem. B 2014, 118, 4471−4478

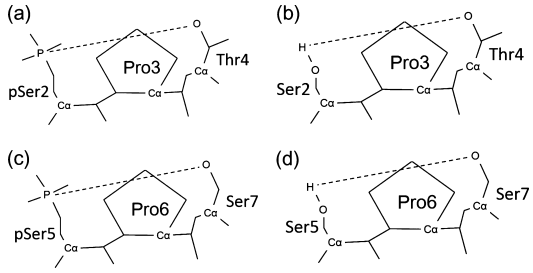


Figure 9. Definitions of the measure of the β-turn stabilized by hydrogen bond interactions in the CTD peptides for (a) pSer2-Pro3-Thr4, (b) Ser2-Pro3-Thr4, (c) pSer5-Pro6-Ser7, and (d) Ser5-Pro6-Ser7.

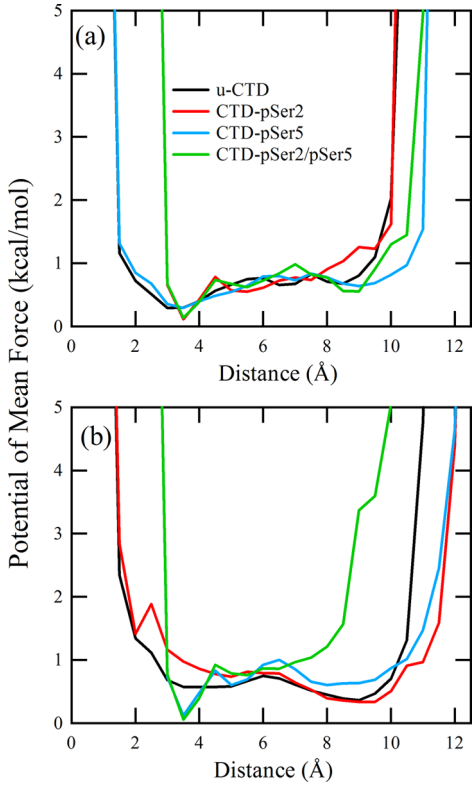


Figure 10. Potential of mean force as a function of the distance for (a) (p)Ser2-Pro3-Thr4 and (b) (p)Ser5-Pro6-Ser7, for CTD peptides where the proline residues are in the trans conformation, which characterizes the hydrogen bond interactions that stabilize the β-turn conformations.

From Figure 10a, for the hydrogen bond interactions around (p)Ser2-Pro3-Thr4, phosphorylation of the Ser2 residue significantly decreases the potential of mean force for the small distance range from 3 to 4.5 Å, resulting in the global minimum of potential of mean force. This minimum does not exist for the u-CTD peptide and reveals that the population of hydrogen bonds has significantly increased. The width of the distribution of CTD-pSer2 is narrower than the distributions of the other peptides.

In Figure 10b, for (p)Ser5-Pro6-Ser7, phosphorylation of Ser5 decreases the potential of mean force for small distances from 3 to 4.5 Å, resulting in the global minimum of potential of mean force. The minimum was not observed when Ser5 is unphosphorylated.

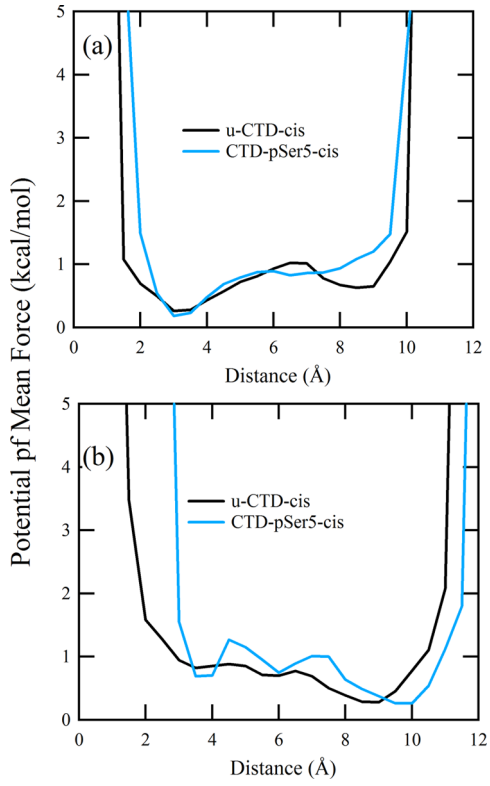


Figure 11. Potential of mean force as a function of the distance for (a) Ser2-Pro3-Thr4 and (b) (p)Ser5-Pro6-Ser7, for CTD peptides with Pro6 in the cis conformation, which characterizes the hydrogen bond interactions that stabilize the β-turn conformations.

From Figure 11a, the profiles of the hydrogen bonding measure with and without phosphorylation at Ser5 are very similar, indicating that the phosphorylation does not significantly affect the hydrogen bonding. In Figure 11b, the global minimum of the potential of mean force shifts to between 8 and 10 Å, indicating that the systems prefer not to form a hydrogen bond. This clearly demonstrates that the cis conformation at Pro6 structurally prevents close contact between the side chains of (p)Ser5 and Ser7.

Phosphorylation Effect on the CTD Mechanism. Noble and co-workers[7] used far-UV circular dichroism (CD) and NMR spectroscopy to study a variety of heptad CTD peptides of different lengths with and without phosphorylation of the serine residues. They concluded from the CD experiments that all the peptides have the same conformation regardless of the phosphorylation site, and the conformation seems to be largely disordered. Moreover, they did not observe stable hydrogen bonds at the phosphorylation sites, which are necessary to form a β-turn, from the NMR measurements. This observation is in agreement with our simulation results, although our results probably provide further details of the conformational states. Our results show that all of the CTD peptides act as flexible disordered peptides with large conformational spaces, and phosphorylation of the serine residues clearly reduces the conformation space. It seems that CD measurement might not be sensitive enough to detect the reduction of the conformational space, probably because of the limitations of CD measurements. Indeed, from only CD measurements, it is generally difficult to determine the accurate conformation of peptides. Stable hydrogen bonds between the phosphorylated serine residues and adjacent proline residues were not observed

dx.doi.org/10.1021/jp4101976 | J. Phys. Chem. B 2014, 118, 4471−4478

in our simulations. However, a considerable increase in the potential of mean force at small end-to-end particle distances, corresponding to hydrogen bonding, was clearly observed in our simulations. Owing to the hydrogen bonds being weak, it might not be possible to determine the reduction of the conformational space by the hydrogen bond and interactions in the NMR spectra. We found that the phosphorylation effects on the dihedral angle space between Pro3 and Pro6 were different. Moreover, we found that the cis conformation at Pro6 significantly inhibits the hydrogen bond around Pro6, indicating that proline with the cis conformation destabilizes the β-turn conformation. This is likely due to Pcf11 indirectly recognizing the phosphate group related to Pro3, while Nrd1 directly recognizes the phosphate group related to Pro6, where a peptidyl−prolyl bond of pSer5-Pro6 in the cis state is formed.[17] Moreover, the effect of double phosphorylation of the CTD is unclear from the current results.

Although it is clear from the current analysis that phosphorylation of the serine residues significantly modifies or restricts the conformational space of the CTD peptide, the CTD peptides still seem to have flexible disordered conformations. When the phosphorylated CTD peptide approaches the regulation factors, the CTD peptides with restricted conformations would recognize and bind the regulation factors more sensitively than the unphosphorylated CTD peptide. After recognition, because of the interaction, conformational change of the CTD should occur to fit the binding site (induced fit). Accordingly, we propose that the CTD uses reduction of conformational states by phosphorylation (the population shift) combined with an induced fit to recognize and bind the regulation factors.

CONCLUSION

We have studied the effect of phosphorylation of the Ser2 and Ser5 residues on the conformational states of the CTD peptide using McMD simulations. We prepared the following systems with explicit solvent under periodic boundary conditions: unit repeat CTD peptide without phosphorylation, the CTD peptide phosphorylated at Ser2, the CTD peptide phosphorylated at Ser5, the CTD peptide phosphorylated at both Ser2 and Ser5, and the CTD peptide with Pro6 in the cis conformation with and without phosphorylation at Ser5. Analysis of the end-to-end distances of the CTD peptides shows that phosphorylation at Ser2 and Ser5 gives more compact conformations than the CTD without phosphorylation.

The potential of mean force with respect to the dihedral angle space showed that phosphorylation of the serine residues significantly reduces the α-helical conformational space, resulting in β-rich conformations. The hydrogen bonds between the side chain of Ser2 and the hydroxyl oxygen of Thr4 and between the side chain of Ser5 and the hydroxyl oxygen of Ser7 are strengthened by phosphorylation of the serine residues. The strengthened hydrogen bonds stabilize the β-turn near the proline residues in the trans conformation, while Pro6 in the cis conformation significantly destabilizes the β-turn. The increased β-turn population should contribute to recognition of the regulation factors binding to the CTD.

AUTHOR INFORMATION

Corresponding Author

*Phone: +81-736-77-0345(2209). E-mail: yonezawa@waka.kindai.ac.jp.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The author thanks Prof. Nakamura for his suggestions and help. Moreover, the author thanks the reviewers for helpful comments and suggestions. This work was supported by Grants-in-Aid for Scientific Research on Innovative Areas “Transcription Cycle”, JSPS KAKENHI Grant Number 25118521.

REFERENCES

(1) Jasnovidova, O.; Stefl, R. The CTD Code of RNA Polymerase II: a Structural View. *Wiley Interdiscip. Rev.: RNA* **2013**, *4*, 1−16.

(2) Egloff, S.; Dienstbier, M.; Murphy, S. Updating the RNA Polymerase CTD Code: Adding Gene-specific Layers. *Trends Genet.* **2012**, *28* (7), 333−341.

(3) Zhang, M.; Gill, G. N.; Zhang, Y. Bio-molecular Architects: a Scaffold provided by the C-Terminal Domain of Eukaryotic RNA Polymerase II. *Nano Rev.* **2010**, *1−11*, 1.

(4) Phatnani, H. P.; Greenleaf, A. L. Phosphorylation and Functions of the RNA Polymerase II CTD. *Genes Dev.* **2006**, *20*, 2922−2936.

(5) Barilla, D.; Lee, B. A.; Proudfoot, N. J. Cleavage/polyadenylation Factor IA Associates with the Carboxyl-Terminal Domain of RNA Polymerase II in Saccharomyces Cervisiae. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 445−450.

(6) Meinhart, A.; Cramer, P. Recognition of RNA Polymerase II Carboxyl-Terminal Domain by 3′-RNA-processing Factors. *Nature* **2004**, *430*, 223−226.

(7) Noble, C. G.; Hollingworth, D.; Martin, S. R.; Ennis-Adeniran, V.; Smerdon, S. J.; Kelly, G.; Taylor, I. A.; Ramos, A. Key Features of the Interaction between Pcf11 CID and RNA Polymerase II CTD. *Nat. Struct. Mol. Biol.* **2005**, *12* (2), 114−151.

(8) Hukushima, K.; Nemoto, K. Exchange Monte Carlo Method and Application to Spin Glass Simulations. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604−1608.

(9) Laio, A.; Parrinello, M. Escaping Free-energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (20), 12562−12566.

(10) Babin, V.; Roland, C.; Darden, T. A.; Sagui, C. The Free Energy Landscape of Small Peptide as obtained from Metadynamics with Umbrella Sampling Corrections. *J. Chem. Phys.* **2006**, *125*, 204909.

(11) Berg, B.; Neuhaus, T. Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions. *Phys. Rev. Lett.* **1991**, *68* (1), 9−12.

(12) Ulrich, H. E. H.; Okamoto, Y.; Eisenmenger, F. Molecular Dynamics, Langevin and Hybrid Monte Carlo Simulations in a Multicanonical Ensemble. *Chem. Phys. Lett.* **1996**, *259* (3−4), 321−330.

(13) Nakajima, N.; Nakamura, H.; Kidera, A. Multicanonical Ensemble generated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *J. Phys. Chem. B* **1997**, *101*, 817−824.

(14) Hansmann, U. H. E.; Okamoto, Y. Prediction of Peptide Conformation by Multicanonical Algorithm: New Approach to the Multiple-Minima Problem. *J. Comput. Chem.* **1993**, *14* (11), 1333−1338.

(15) Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H. Free Energy Landscape of Peptides by Enhanced Conformational Sampling. *J. Mol. Biol.* **2000**, *296*, 197−216.

(16) Yonezawa, Y.; Shimoyama, H.; Nakamura, H. Multicanonical Molecular Dynamics Simulations combined with Metadynamics for the Free Energy Landscape of a Biomolecular System with High Energy Barriers. *Chem. Phys. Lett.* **2011**, *501*, 598−602.

(17) Kubicek, K.; Cerna, H.; Holub, P.; Pasulka, J.; Hrossova, D.; Loehr, F.; Hofr, C.; Vanacova, S.; Stefl, R. Serine Phosphorylation and Proline Isomerization in RNA II CTD Control Recuruitment of Nrd1. *Genes Dev.* **2012**, *26*, 1891−1896.

(18) Xiang, K.; Nagaike, T.; Xiang, S.; Kilic, T.; Beh, M. M.; Manley, J. L.; Tong, L. Crystal Structure of the Human Symplekin-Ssu72-CTD Phosphopeptide Complex. *Nature* **2010**, *467*, 729−733.
(19) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926.
(20) Evans, D. J.; Hoover, W. G.; Failor, B. H.; Moran, B.; Ladd, A. J. C Nonequilibrium Molecular Dynamics via Gauss's Principle of Least Constraint. *Phys. Rev. A* **1983**, *28*, 1016−1046.
(21) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(22) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(23) Lindorff-Larsen, K.; et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950−1958. Best, R. B.; Sancho, D.; Mittal, J. Residue-Specific α-Helix Propensities from Molecular Simulation. *Biophys. J.* **2012**, *102*, 1462−1467.
(24) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(25) Yonezawa, Y. A Long-range Electrostatic Potential based on the Wolf method Charge-Neutral Condition. *J. Chem. Phys.* **2012**, *136*, 244103.
(26) Christoph, F. W.; Weisshaar, J. C. Conformational Analysis of Alanine Dipeptide from Dipolar Couplings in a Water-Based Liquid Crystal. *J. Phys. Chem. B* **2003**, *107*, 3265−3277.
(27) Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A. Conformational Study of the Alanine Dipeptide at the MP2 and DFT Levels. *J. Phys. Chem. A* **2002**, *106*, 3213−3218.