Published online 4 January 2018 *Nucleic Acids Research, 2018, Vol. 46, No. 5* ***2243–2251***
doi: 10.1093/nar/gkx1297

Phosphorylation of an intrinsically disordered region of Ets1 shifts a multi-modal interaction ensemble to an auto-inhibitory state

Kota Kasahara[1,*,†], Masaaki Shiina[2,*,†], Junichi Higo[3], Kazuhiro Ogata[2] and Haruki Nakamura[3]

[1]College of Life Sciences, Ritsumeikan University, Noji-higashi 1-1-1, Kusatsu, Shiga 525-8577, Japan, [2]Graduate School of Medicine, Yokohama City University, Fuku-ura 3–9, Kanazawa-ku, Yokohama, Kanagawa 236-0004, Japan and [3]Institute for Protein Research, Osaka University, Yamada-oka 3-2, Suita, Osaka 565-0871, Japan

Received August 29, 2017; Revised December 15, 2017; Editorial Decision December 17, 2017; Accepted December 19, 2017

ABSTRACT

Multi-modal interactions are frequently observed in intrinsically disordered regions (IDRs) of proteins upon binding to their partners. In many cases, post-translational modifications in IDRs are accompanied by coupled folding and binding. From both molecular simulations and biochemical experiments with mutational studies, we show that the IDR including a Ser rich region (SRR) of the transcription factor Ets1, just before the DNA-binding core domain, undergoes multi-modal interactions when the SRR is not phosphorylated. In the phosphorylated state, the SRR forms a few specific complex structures with the Ets1 core, covering the recognition helix in the core and drastically reducing the DNA binding affinities as the auto-inhibitory state. The binding kinetics of mutated Ets1 indicates that aromatic residues in the SRR can be substituted with other hydrophobic residues for the interactions with the Ets1 core.

INTRODUCTION

Intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) are abundantly found in cells and nuclei, and play diverse and important roles in biological systems (1–3). A major key role of IDPs/IDRs is to mediate signal transduction controlled by several types of input, i.e. interactions with other molecules and post-translational modifications (PTM) (4–6). In particular, previous high-throughput studies demonstrated that the target sites of kinases are enriched within the IDRs (7).

The characteristic feature of IDPs/IDRs is their disordered structures when they exist alone, but upon binding to their partner molecules they fold into particular structures in many cases. This phenomenon is called 'Coupled folding and binding' (1,2). The folded conformations of IDPs/IDRs are not unique, and different ones are formed depending on the partner molecules, thus providing the promiscuous nature to those IDPs/IDRs as hub proteins (4,8–10). Accordingly, the mechanism of coupled folding and binding is not considered to be simple allostery, and the term 'multistery' may be more suitable for this mechanism (4). In addition, PTMs are frequently observed within IDPs/IDRs, and they seem to regulate the multisteric nature to select their partner proteins by changing their affinities (4).

Recent observations of IDP/IDR binding to one particular partner protein revealed that several or many different complex structures were formed instead of a single specific interaction, in a phenomenon referred to as a fuzzy interaction (11). However, the identification of fuzziness is not straightforward by ordinary methods, which observe only the averaged nature of IDPs/IDRs.

The PTMs of IDRs in eukaryotic transcription factors are crucial for transforming signaling information into gene expression (12). Among them, the transcription factor Ets1 is an illustrative example of a protein that regulates pathways depending on the phosphorylation of IDRs. Ets1 is a regulator of many pathways that play key roles in diseases, including cancer and autoimmunity (13,14). There are two IDRs with phosphorylation sites in Ets1. The first is the N-terminal IDR, including the phosphorylation sites Tyr38 and Ser41. The phosphorylation of these sites regulates the binding of the adjacent structured domain, the PNT (pointed) domain, to a co-activator, CBP/p300 (15). The second IDR is located on the C-terminal side of the PNT domain, and its disordered serine-rich region (SRR) has several Ser residues that can be phosphorylated to regulate the DNA binding affinity of the C-terminal ETS core domain. When the Ser residues, in particular Ser282 and

*To whom correspondence should be addressed. Tel: +81 77 566 1111; Fax: +81 77 561 3729; Email: ktkshr@fc.ritsumei.ac.jp
Correspondence may also be addressed to Masaaki Shiina. Tel: +81 45 787 2590; Fax: +81 45 784 4530; Email: mshiina@med.yokohama-cu.ac.jp
† These authors contributed equally to this work as first authors.

Ser285, in the SRR are phosphorylated, the SRR binds to the Ets1 core domain, inhibiting double-stranded DNA (dsDNA) binding to the Ets1 core. However, when these phosphorylated Ser residues were replaced with Ala residues, the binding affinity of SRR to the Ets1 core domain decreased significantly, and the dsDNA could bind more tightly to Ets1 (16). In 2014, Desjardins *et al*. observed that the phosphorylated SRR interacted with the Ets1 core domain in a fuzzy manner, and that the aromatic residues just after the phosphorylated Ser residues are essential to control the DNA binding. However, the authors did not provide any clear argument about how those aromatic residues associate by the fuzzy interaction (17).

In this report, we aimed to clarify the atomistic details of the interactions between the phosphorylated SRR and the core domain, by both theoretical and experimental methods. We sought to answer the following questions: 'What kind of physicochemical interactions are involved?', 'How fuzzy are they?', 'What conformations are adopted by the IDR?', 'How do the structures differ between the phosphorylated and unmodified IDRs?', and 'How do the aromatic residues next to the phosphoserine residues work?'. We first performed molecular simulations to predict the ensembles of highly probable complex structures of Ets1 with phosphorylated and unmodified SRRs, from analyses of the free-energy landscapes. In addition, experimental kinetic measurements of DNA binding by Ets1 were performed for various mutants with phosphorylated and unmodified SRRs. The interaction schemes of the phosphorylated and unmodified SRRs with the Ets1 core domain and the biological role of the phosphorylation of IDRs are discussed.

MATERIALS AND METHODS

System preparation for molecular simulations

The multicanonical molecular dynamics (McMD) simulations were performed with the molecular models composed of the Ets1 core domain and the adjacent IDR including the SRR, Arg279 through Glu441. The structure of the Ets1 core domain, Lys301 through Glu441, was obtained from the NMR solution structure (PDB ID: 1R36, model 1, (18)), and a random coil structure of the IDR from Arg279 to Pro300 with the N-terminal acetyl cap was built with the Modeller software (19), without any template structure. Divalent phosphoric acids were then added on Ser282 and Ser285. This Ets1 model was bathed in a 150 mM NaCl solution, and minimizations with the steepest descent and the conjugate gradient methods were successively applied. Then, a relaxation run was performed for 1.0 ns with 0.5 fs time steps, and the temperature was gradually increased from 10 to 300 K for the first 500 ps of the relaxation. In addition, the positions of the heavy atoms in the proteins were restrained. Successively, a 3.0 ns equilibration was performed with 1.0 fs time steps and the LINCS constraint for bonds with hydrogen atoms (20). We refer to this model as the 'P-state' model. On the basis of this model, the unmodified 'U-state' was modeled, by replacing the pSer residues with Ser residues and equilibrating the system with a 1.0 ns NPT simulation. These preparations were performed with the GROMACS software (21). For the potential parameters, the AMBER-based hybrid force field (22) was applied to Ets1, using methods that were indicated to be effective for simulations of short peptides and IDRs (23), with the parameters reported by Homeyer *et al.* (24) for pSer. The TIP3P model (25) with the ion parameters reported by Joung and Cheatham (26) was utilized as the solvent. The electrostatic potentials were calculated by using the particle mesh Ewald method (27) for the system preparations with GROMACS. The simulation models investigated in this report are summarized in Supplementary Table S1.

Enhanced conformational sampling

By using the P-state and U-state models, the conformational space of the IDR was explored. In order to prevent the denaturation of the Ets1 core domain, the interatomic distances between the Cα atoms in Thr303–Glu441, except for neighboring pairs (≤4 residues in the sequence order), were weakly restrained within the range between –0.5 and +3.0 Å from the initial structure (the restrained potential was applied only when the distance was beyond this range). The same restrictions were applied to both the P- and U-state models.

With these restrictions, first a wide range of conformations of the IDR were obtained by using the canonical simulation at the high temperature condition, 800 K, and taking snapshots every 10 ps. Fifty snapshots were used for the initial structures of 50 runs of independently performed McMD simulations, and the ensemble was obtained by combining all of the trajectories based on the trivial trajectory parallelization theory (28). For the McMD, the virtual-system coupled McMD (V-McMD) method was applied (29), which simulates the behavior of the real-system by coupling with a virtual-system governing the biasing potentials of the real-system. The energy range to be sampled was divided into seven mutually overlapped regions. See reference (29) for details.

In the McMD simulations, the integration time step was 1.2 fs, and covalent bonds with hydrogen atoms were constrained with the SHAKE algorithm (30). The same force field used in the relaxation process was applied. The electrostatic potential was computed by the zero-dipole summation method (31), which is one of the so-called non-Ewald methods (32). The precision of this method and its generalized form, the zero-multipole summation method, have been proven in a variety of systems (33–37). In addition, we recently reported the applicability of this method for the MD studies of the complexes including the Ets1 core domain (38). At the final production run, 9×10^6 steps were computed for each of 50 runs (total 540 ns) in the V-McMD scheme (29). The computation was performed with the myPresto/psygene-G software (39).

Post simulation analysis

In order to visualize the free-energy landscapes (FELs), the conformational spaces were projected onto the two-dimensional (2D) space defined by the principal component analysis (PCA) over the conformational ensembles assembling the P- and U-states. For the PCA analysis, each sampled conformation was expressed as a 119-dimensional real value vector, consisting of the distances between the Cα

atoms of the SRR (Pro281–Asp287; 7 × 6/2 = 21 pairs) and the distances of all pairs of Cα atoms between the SRR (Pro281–Asp287) and the H3 helix (Tyr386–Lys399; 7 × 14 = 98 pairs). In order to evaluate the statistics, the bootstrap analysis was performed by dividing the trajectory into 20 bins. A trajectory was reconstructed by random sampling with the replacement of the 20 bins. The statistics was obtained over 1,000 reconstructed trajectories. The analyses were performed with in-house programs.

Mutagenesis for DNA encoding the Ets1 fragment (residues 276–441)

DNA mutations were introduced by PCR with the primers listed in Supplementary Table S2, or the entire DNA sequence encoding mutated Ets1 (Ser276–Glu441) was synthesized.

Expression and purification of Ets1 fragments

The experimental procedures for the preparation of human Ets1 fragments were described previously (40). Briefly, wild type and mutant Ets1 fragments (amino acids residues 276–441) were bacterially overexpressed and purified using ion-exchange, hydrophobic interaction, and gel-filtration chromatographies. The purified sample was buffer-exchanged into a water solution containing 200 mM NaCl and concentrated by ultrafiltration.

In vitro phosphorylation of Ets1 fragments

The purified Ets1 fragments (100 μg) were phosphorylated with 200 nM CaMKII in buffer solution A (50 mM HEPES, pH 8.0, 10 mM magnesium acetate, 0.5 mM dithiothreitol (DTT), 1 mM ATP and 0.01 mM calmodulin) at 25°C for 3 h, except for the F286G mutant. For the unphosphorylated controls, the Ets1 fragments were incubated in the same manner, in buffer solution A without ATP. In the case of the F286G mutant, the phosphorylation step was performed at 25°C overnight with three additions of CaMKII. The phosphorylation state of each Ets1 fragment was monitored by Phos-tag sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) (Supplementary Figure S1) (41), according to the manufacturer's (Wako) instructions.

Surface plasmon resonance (SPR)

The hemi-biotinylated and non-biotinylated single stranded DNA fragments containing the Runx1 binding site from the *TCRα* enhancer were purchased from FASMAC, and annealed by heat denaturation followed by gradual cooling. The SPR experiment was performed at 25°C, using a BIAcore2000 (GE Healthcare). The CM5 sensorchip (GE Healthcare) was pre-treated with 10 μl of each of the following solutions: 50 mM NaOH, 0.1% HCl, 0.1% SDS and 0.085% H_3PO_4. The sensorchip was then coated with streptavidin by the amine-coupling procedure, in which the following solutions were sequentially applied to the sensorchip: 700 μl of a mixture of N-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC), 200 μl of 0.2 mg/ml streptavidin solution in 1 mM sodium acetate, pH 4.0, and 500 μl of 1 M ethanolamine hydrochloride–NaOH, pH 8.5. To immobilize the DNA fragment via biotin–streptavidin binding to the sensorchip, the biotinylated DNA was injected onto the sensorchip until the resonance units reached 30–50.

The Ets1 fragment was serially diluted in filtered and degassed SPR running buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 0.01% Tween20, and 2 mM DTT) from 0 to 32 or 64 nM for unmodified Ets1 (U-state) and from 0 to 800 nM for phosphorylated Ets1 (P-state) and injected onto the sensorchip at the flow rate of 100 μl/min with the 'kinject' method. The obtained sensorgrams were double-subtracted with those from a blank cell (without DNA immobilization) and a buffer-only injection, and the subtracted sensorgrams were globally fitted to a 1:1 Langmuir binding model (a single exponential binding model) with the BIAevaluation 4.1 software (Supplementary Figure S2). The measurements were repeated at least three times for individual samples, and the analyzed data are shown as means with standard deviations.

RESULTS

Canonical ensembles provided by all-atom molecular simulations

We built two molecular models with the Ets1 fragment (residues 279–441), as shown in Figure 1, based on the NMR solution structure (PDB ID: 1R36) with an artificially modeled disordered region (Arg279 through Pro300). The first model, representing the phosphorylated state (P-state model), included the phosphorylated Ser282 and Ser285. In the other model, referred to as the model in the unmodified state (U-state model), both Ser residues were unphosphorylated. After minimization and relaxation processes, the all-atom computations for enhanced conformational sampling of the IDR (from Arg279 to Gly302) were performed with a random walk within the energy space from 300 to 800 K by using the V-McMD method (29), which was developed to create a multicanonical ensemble (42,43) more effectively. These simulations applied weak restraints to the Ets1 core domain (from Thr303 to Asp441), in order to prevent denaturation. The V-McMD simulations generated near-uniform potential energy distributions, and the multi-canonical ensembles (Supplementary Figure S3) yielded the precise canonical ensembles of the P- and U-state models at room temperature (300 K).

A promising hypothesis to explain how the phosphorylated SRR inhibits Ets1–DNA binding involves competitive inhibition by direct interactions between the phosphorylated Ser (pSer) residues and the H3 helix. As Ets1 recognizes a regulatory *cis* element by burying the basic H3 helix into the major groove of DNA, masking the basic residues by negatively charged pSer residues may competitively inhibit the DNA binding. In order to evaluate this hypothesis, we first assessed the probabilities of contacts between the SRR and the H3 helix at 300 K.

Figure 2 shows the probability distributions of the number of contacting residue pairs between the SRR (Ser282 through Asp287) and the H3 helix (Lys388 through Lys399), referred to as N_c. In the U-state, the highest probability was observed at $N_c = 0$ (the dashed line in Figure

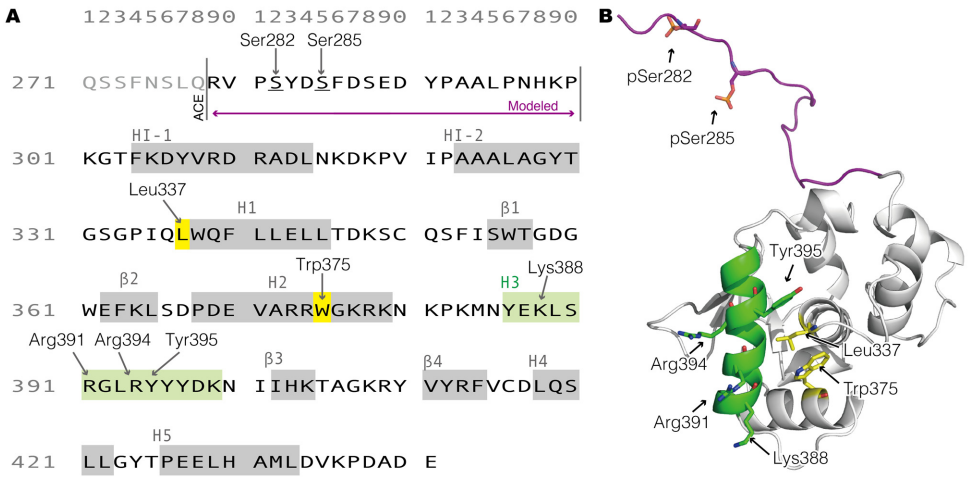

Figure 1. Sequence and structure of Ets1, including the SRR and the core domain. (**A**) Sequence of Ets1. Several residues are pointed out by arrows, and the two Ser residues that are phosphorylated are underlined. The secondary structure elements are shaded and labeled as H and β for α-helix and β-strand, respectively. 'ACE' in front of Arg279 means the acetyl group as the N-terminal cap. The initial structure of the range indicated with the label 'Modeled' was built by the Modeller software, and the structure of the remaining region was taken from the experimental structure (PDB ID: 1R36). (**B**) The initial 3D structure of the simulation model (phosphorylated state). The regions shown in purple and green correspond to the 'Modeled' region including the SRR and H3 helix in panel A, respectively. The residues shown in yellow are part of the hydrophobic core described in the main text.

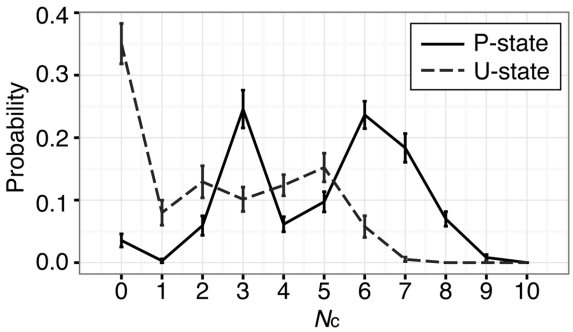

Figure 2. Probability distributions of the number of contacting residue pairs between the SRR and the H3 helix (N_c). The solid and dashed lines indicate the probability distributions of the phosphorylated (P-state) and unmodified (U-state) models, respectively, at 300 K. The error bars indicate the standard deviation estimated by the bootstrap analysis.

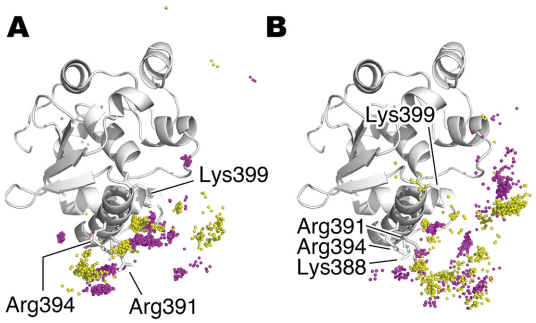

Figure 3. Spatial distribution of Oγ atoms of pSer at the SRR in (**A**) the P-state, and (**B**) that of Ser in the U-states, at 300 K. The purple and yellow dots represent the Oγ atoms of pSer/Ser282 and pSer/Ser285, respectively.

2), which means that a considerable part of the conformational ensemble lacks interactions between the unmodified SRR and the H3 helix. In fact, the Oγ atoms of Ser282 and Ser285 were located at several different positions, not only near the H3 helix but also around other surfaces of the Ets1 core domain, as shown in Figure 3. In Figure 2, the distribution indicated by the dashed line ranges from zero to six, and the summed probability of $N_c \geq 1$ is 0.650 (in the case of the P-state, this probability was 0.965). Thus, the unmodified SRR interacts with the H3 helix with a certain probability.

In contrast, in the P-state, the phosphorylated SRR tended to interact with the H3 helix, and there are two peaks at $N_c = 3$ and $N_c = 6$ (the solid line). This indicates that the phosphorylation of Ser282 and Ser285 facilitates direct contacts with the H3 helix. These interactions are mainly due to salt-bridges between the pSer residues and the positively charged residues in the H3 helix; *i.e.*, Lys388, Arg391, Arg394 and Lys399. The probability distributions of the inter-residue distances show that the salt-bridges, pSer282–Arg394 and pSer285–Arg391, are formed in a stable manner (Supplementary Figure S4). Note that pSer285 interacted more favorably with Arg391 than Arg394, in contrast to the fact that pSer282 interacted with either Arg391 or Arg394 at similar probability levels. As shown in Figure 3, the Oγ atoms of the two pSer residues were, on average, located primarily around the H3 helix, but at several different positions. This result supports the hypothesis that the H3 helix electrostatically attracts the phosphorylated SRR, and this interaction competitively inhibits the H3–DNA binding.

Phosphorylation alters the free-energy landscape of the IDR

The canonical ensembles of the P- and U-states at 300 K are summarized as the free-energy landscapes (FEL), where the distributions of the potential of mean forces (PMFs) are shown depending on the structures and the locations of the IDRs (Figure 4). Here, the axes of the landscape are defined by utilizing the principal component analysis (PCA), on the basis of the Cα–Cα distances among the SRR (from

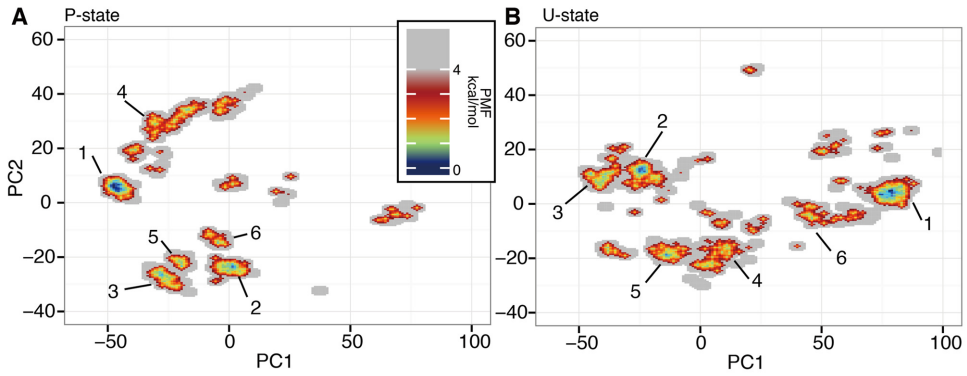

Figure 4. The free energy landscape (FEL) in the canonical ensemble at 300 K of (**A**) the P-state model and (**B**) the U-state model. The horizontal and vertical axes denote the first and second principal component axes, respectively. The definition of the axes is identical for both the P- and U-states. The color graduation from blue to red, and gray, indicates the potential of mean force. The numbers on the map represent the rank of stability of each cluster. The representative structures in each cluster are shown in Supplementary Figure S5.

Pro281 to Asp287) residues, and between the SRR and the H3 helix (from Tyr386 to Lys399), using both the P- and U-state models. The contribution ratios of the first, second, and third principal components (PC1, PC2 and PC3) were 0.545, 0.252 and 0.106, respectively. PC1 reflects the distance between the SRR and the H3 helix (the Pearson correlation coefficient between PC1 and pSer282–Arg394 distance was 0.92). PC2 collectively represents some structural features, rather than a single parameter.

In both the P- and U-state models, the IDR conformations were widely spread in the 2D FEL map, and fell into many clusters (Figure 4A, B, Supplementary Figure S5, and Supplementary Data S1). The two FELs were quite different, suggesting that the phosphorylation of the two serine residues drastically changes the conformational ensemble of the IDR. In fact, the most stable cluster of the P-state model, namely the most popular state, was located at the left-most part of this landscape, and that of the U-state model was located at the right-most part (labeled '1' in Figure 4A and B, respectively), which means that their distances between the SRR and the H3 helix were the nearest and the farthest in the ensemble, respectively.

In the P-state, the most stable conformational cluster 1 had two salt-bridges between the SRR and the H3 helix (pSer282–Arg394 and pSer285–Arg391), as shown in Figure 5, which could competitively inhibit the recognition of DNA by the H3 helix. The SRR and the H3 helix were aligned in the anti-parallel direction. The SRR adopts a compact conformation with a short 3_{10}-helix structure around pSer285. In addition, the aromatic side-chain of Tyr283 stacks with that of Tyr395, and Phe286 forms hydrophobic contacts with Trp375 and Leu337 (Figure 5). These interactions could maintain the SRR at the auto-inhibitory position in a specific manner.

In contrast, in the most probable conformation of the U-state, the SRR is trapped far from the H3 helix and close to the HI2 helix and the HI2–H1 loop, by hydrophobic contacts with Val280, Tyr283, Phe286, Leu295, Ala327, Pro334 and Ile335 (Supplementary Figure S6A). Although the SRR in this conformation does not contact the H3 helix and seems to lack auto-inhibitory effects, the second, third,

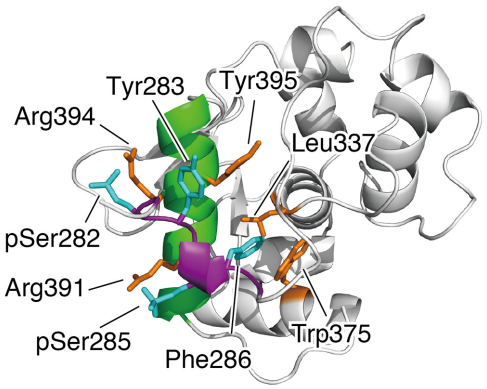

Figure 5. The representative structure in the most stable cluster of the P-state model at 300 K. The backbone of the SRR is colored purple, and that of the H3 helix, which is the DNA binding interface, is shown in green. The side-chains engaged in the important interactions are represented as stick models in cyan and orange (for the SRR and core domain, respectively).

and fourth most stable clusters in the unmodified model form contacts between the SRR and the H3 helix (Supplementary Figure S5B). In the fourth cluster, the SRR has an α-helical structure of the SRR with a couple of salt-bridges between the SRR and the H3 helix: Asp284–Lys388 and Glu289–Arg391 (Supplementary Figure S6B). In contrast to the fact that the SRR tends to form a 3_{10}-helix in the P-state model, it has an α-helical tendency in the U-state model ('G' for 3_{10}-helix and 'H' for α-helix, respectively, in Supplementary Figure S7). These conformational ensembles suggest that the unmodified form also has some auto-inhibitory effects on Ets1–DNA binding, exerted by some sub-stable conformers.

In terms of the interactions between the SRR and the H3 helix, the P-state preferred the anti-parallel and parallel alignments of the SRR and the H3 helix as the first and second most stable conformations, respectively (Supplementary Figures S5A and S8A), due to the salt-bridges between the pSer and Arg residues (Supplementary Figure S4). On the contrary, the U-state favored the perpendicular

arrangement between the SRR and the H3 helix (Supplementary Figures S5B and S8B).

Ets1–DNA binding affinities for the wild-type and mutant proteins

In order to examine the interaction schemes of the P- and U-state models suggested by the all-atom molecular simulations, we performed *in vitro* assays of DNA binding for the Ets1 fragment (residues 276–441) and its various mutants within the SRR (Supplementary Figure S1). From kinetic measurements using the surface plasmon resonance method, the association equilibrium constants of the P- and U-states were observed for the Ets1–DNA (the *TCRα* enhancer) binding. The details of the experimental conditions are described in the Methods section and in our previous paper (44).

The phosphorylation of the SRR of the wild-type (WT) drastically reduced the equilibrium association constant K_A (k_{on}/k_{off}) (Figure 6, Supplementary Figure S9, and Table S3), mainly due to the effect on the on-rate for Ets1–DNA binding (Supplementary Figure S10). When Tyr283 is replaced with Phe, Thr or Gly, the phosphorylation effect was as large as that in the WT, while Ala introduced at position 283 increased the Ets1–DNA binding in both the P- and U-states. In contrast, when Phe286 is replaced with other hydrophobic residues, the phosphorylation distinctively reduced the DNA binding, similar to the WT. While the previous work by Desjardins *et al*. suggested that the aromatic residues next to the pSer have synergetic effects with phosphorylation (17), our results imply that not only aromatic but also hydrophobic side-chains work to form stable interactions between the phosphorylated SRR and the H3 helix. It is interesting that the phosphorylation of the F286Y mutant did not reduce the K_A value. This can be interpreted as the effects of the polar nature of the Tyr side-chain, which can weaken the hydrophobic packing with the core domain (Figure 5). The effects of the phosphorylation of the other mutant proteins are outlined in the Discussion section.

DISCUSSION

Mechanisms of the phosphorylation-dependent auto-inhibition

Our all-atom molecular simulations with the generalized ensemble method provided the canonical ensembles of the phosphorylated state (P-state) and the unmodified state (U-state). There are several structural clusters appearing in the FEL, which correspond to the different binding modes of the IDR to the Ets1 core domain. Although validating our predicted structural ensemble in a quantitative manner is not straightforward, the ensembles successfully explained the observations in the experimental binding assays from a qualitative viewpoint, as discussed below.

The simulated ensemble in the U-state showed a significantly lower frequency of the SRR–H3 helix contacts than that in the P-state (Figure 2), and this result well explains the experimental fact that the k_{on} was significantly reduced by the phosphorylation (Supplementary Figure S10). In most of the P-state clusters, the formation of two salt-bridges between the pSer residues and the H3 helix was frequently observed, masking the positive charges on the H3 helix to competitively inhibit DNA binding. The characteristic feature of the most stable conformation in the P-state is that the SRR adopts a 3_{10}-helix conformation, where the phosphate groups of pSer282 and pSer285 form salt-bridges with the side-chains of Arg394 and Arg391, respectively. As the pitch of an α-helix is 3.6 Å, these two arginine residues on the H3 helix can only form strongly attractive interactions with the two pSer residues when the backbone of residues 282 to 285 forms a 3_{10}-helix structure, as shown in Figure 7. If this SRR adopted an extended structure or an α-helix, then these salt-bridge pairs would not be formed. This is consistent with the fact that the insertions of Ala and Pro (Ins 284A, Ins 285AA and Ins 285P) in the SRR increased the DNA binding and reduced the effect of phosphorylation. Whereas there is no known example of this binding mode, i.e. a doubly phosphorylated SXXS motif forms salt-bridges with the RXXR motif in the PDB (45), numerous interacting pairs of α- and 3_{10}-helices have been reported. Some examples are shown in Supplementary Figure S11, and the details of the analysis using the HOMCOS database (46) are described in Supplementary Text S1.

In addition, the importance of the two aromatic residues adjacent to the pSer for the effects of the phosphorylation suggested by the experiments (See Y283A and F286G in Figure 6) can be interpreted by the observations of hydrophobic contacts observed in the most stable structure of the P-state simulation, i.e. Tyr283 stacking with Tyr395, and Phe286 contacting the hydrophobic core (Leu337 and Trp335; Figure 5). However, it is interesting that the double mutation in the SRR, Y283D/D284Y, significantly decreased the DNA binding of the P-state. This suggests that the stacking effect between Tyr283 and Tyr395 (Figure 5) is not essential. In contrast, the deletion or mutation of the Asp residues in the SRR (Asp284 and Asp287) significantly increased the DNA binding, probably because of the reduction of the electrostatic masking effect for the basic residues of the H3 helix.

Specific complex structure of the phosphorylated SRR with the Ets1 core domain

Desjardins *et al*. observed the importance of the two aromatic residues, Tyr283 and Phe286, adjacent to the pSer residues, and concluded that 'synergistic effects' occur between the pSer residues and these aromatic residues, although the interactions have a fuzzy feature (17). However, a model to explain the synergistic effects with the fuzzy interactions was not provided.

Here, we propose the existence of a specific complex structure of the phosphorylated SRR with the H3 helix in the Ets1 core domain, as the most stable structure appearing in the FEL in Figure 4A. The specific structure, shown in Figure 5, is composed of (i) the characteristic salt-bridges between the two pSer residues forming a 3_{10}-helix with the two Arg residues, (ii) the van der Waals interactions between the side-chains of Tyr283 and Tyr395, and (iii) the van der Waals interactions of Phe286, Leu337 and Trp375. These three interactions explain the NMR chemical shift titration experiments by Desjardins *et al*. (17), in which the chemical shifts of Leu337, Trp375, and Tyr395 gradually moved

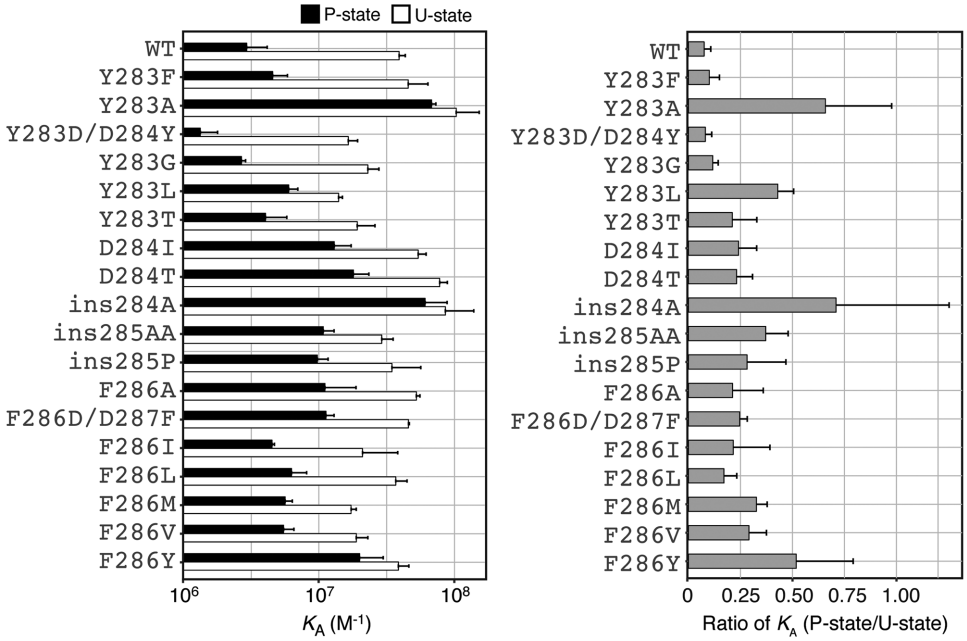

Figure 6. (**A**) K_A values of the wild type and mutants of the Ets1 fragment (residues 276–441) to the *TCRα* enhancer DNA. (**B**) The ratio of K_A values (P-state/U-state).

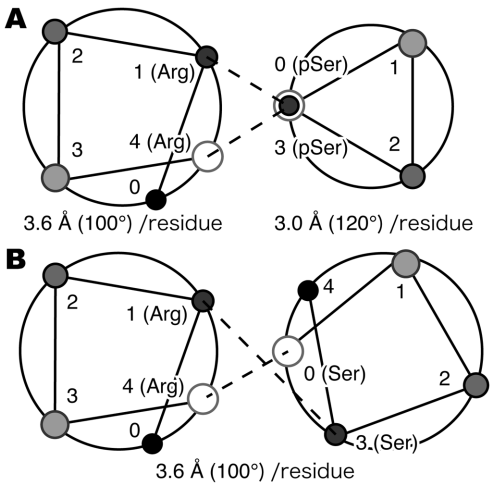

Figure 7. Interaction scheme between the α-helix and the 3_{10}-helix, represented by helical wheel projections. The large circles indicate the helix diameters of the H3 helix and the SRR, viewed from the helical axes. The small circles with the numbers from 0 to 4 represent the positions of residues. The helical pitches of the α- and 3_{10}-helices are 3.6 Å (100°) and 3.0 Å (120°) per residue, respectively. (**A**) When the SRR forms a 3_{10}-helix, the two pSers are located at the same position near the two Arg residues of the H3 helix. (**B**) When the SRR forms an α-helix, the positions of the two pSers are not overlapped and they deviate from the ideal positions to interact with the H3 helix (see text in Discussion section.)

by adding a peptide mimicking the SRR to the Ets1 core domain (residues 301–440, ΔN301). This specific interaction should mask the H3 helix and prevent DNA binding. In the FEL, several sub-stable complex structures are also observed, forming similar salt-bridges between the SRR and the H3 helix with large N_c values and masking the DNA binding surface. Namely, even in the P-state, not only the most stable complex structure but also the sub-stable structures should exist to some extent, with both specific and transient properties.

In contrast, when the two Ser residues, Ser282 and 285, are not phosphorylated, the structure of the FEL is completely different from that in the above phosphorylated state, and weak interactions between the SRR and the H3 helix were observed with small N_c values. Namely, the interactions are multi-modal, and their masking effect is weak. Whereas the populations of the most stable clusters were almost the same between the P- and U-states, i.e. 32.84% and 32.21%, respectively (Supplementary Figure S5), the most stable cluster of the U-state includes very diverse structures that are not clearly distinguished in the 2D FEL, but are shown in the distributions of the secondary structure elements in Supplementary Figure S7D. Thus, it is very different from the most stable structure in the P-state, where the majority forms the specific conformation with the 3_{10}-helical SRR (Supplementary Figure S7B).

Thus, the phosphorylation, as the post-translational modification, switches the biological signal, as suggested by Van Roey *et al.* (5,6). Without phosphorylation, the recognition helix is weakly masked by the SRR, with interactions in a multi-modal manner. The phosphorylation of the two Ser residues creates more specific and strong interactions between the recognition helix and the SRR, and the backbone of the SRR folds into a 3_{10}-helix.

Recently, the phosphorylation of Ser282 in the SRR of Ets1 was shown to provide a binding site for COP-1, a ubiquitin ligase component. This report also demonstrated that Tyr283 is a phosphorylation target of Src family tyrosine kinases, which are known to promote tumor growth, invasion, and metastasis. Interestingly, the phosphorylation

of Tyr283 decreased the COP-1 binding and prevented the ubiquitin-mediated degradation of phosphorylated Ets1, leading to the enhancement of Ets1 activity (47). Since Ets1 directly regulates the genes involved in tumor metastasis and invasion, such as integrins and MMP9 (48), the authors suggested that the accumulation of Ets1 contributes to the oncogenic activity of Src family kinases.

Tyr283 is located near the induced 3_{10}-helix involved in the specific interaction between the phosphorylated SRR and the core domain (Figure 5). Phosphorylation of Tyr283 may reduce the interaction and enhance the DNA binding activity of Ets1, even if the SRR is phosphorylated. This may explain, at least partially, the reported enhancement of Ets1 activity even in the P-state, and also may provide a possible reason why Tyr283 is highly conserved even though the other hydrophobic residues can largely assume the role of Tyr283 in the P-state (Figure 6).

We showed that the phosphorylation of the Ser residues in the IDR trigger the shift from multi-modal to more specific interactions with the DNA recognition helix H3. Additional PTMs, such as the phosphorylation of Tyr283 by Src family kinases, may counteract the 'induced' specific interaction between the SRR and the Ets1 core domain.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The MD simulations were performed on the TSUBAME2.5 supercomputer at the Tokyo Institute of Technology, provided through the HPCI System Research Project (Project IDs: hp140032, hp150015, hp170020 and hp170025). The numerical computations for post-simulation analyses were performed using the supercomputer system provided by the National Institute of Genetics, Research Organization of Information and Systems, Japan. This work was performed in part under the Cooperative Research Program of Institute for Protein Research, Osaka University, CR-17–05. We thank Takeshi Kawabata for analyses of helix–helix interactions in the PDB by using HOMCOS and fruitful discussions.

FUNDING

Japan Society for the Promotion of Science KAKENHI, Grants-in-Aid for Scientific Research on Innovative Areas [JP24118005, JP24118008], Grant-in-Aid for Exploratory Research [JP16K14711], Grant-in-Aid for Scientific Research (B) [JP16H03293], Grant-in-Aid for Young Scientists (B) [JP16K18526]. Japan Science and Technology Agency, Creation of Innovation Centers for Advanced Interdisciplinary Research Areas Program in the Project for Developing Innovation Systems. Japan Agency for Medical Research and Development, The Development of Core Technologies for Innovative Drug Development based upon IT; Funding for open access charge: The Institute of Science and Engineering, Ritsumeikan University, Support of the Submission of Academic Papers.
Conflict of interest statement. None declared.

REFERENCES

1. Wright,P.E. and Dyson,H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, **293**, 321–331.
2. Wright,P.E. and Dyson,H.J. (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell. Biol.*, **16**, 18–29.
3. Babu,M.M. (2016) The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem. Soc. Trans.*, **44**, 1185–1200.
4. Tompa,P. (2014) Multisteric regulation by structural disorder in modular signaling proteins: an extension of the concept of allostery. *Chem. Rev.*, **114**, 6715–6732.
5. Van Roey,K., Gibson,T.J. and Davey,N.E. (2012) Motif switches: decision-making in cell regulation. *Curr. Opin. Struct. Biol.*, **22**, 378–385.
6. Van Roey,K., Dinkel,H., Weatheritt,R.J., Gibson,T.J. and Davey,N.E. (2013) The switches.ELM Resource: a compendium of conditional regulatory interaction interfaces. *Sci. Signal.*, **6**, rs7–rs7.
7. Meggio,F. and Pinna,L.A. (2003) One-thousand-and-one substrates of protein kinase CK2? *FASEB J.*, **17**, 349–368.
8. Patil,A. and Nakamura,H. (2006) Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.*, **580**, 2041–2045.
9. Dunker,A.K., Silman,I., Uversky,V.N. and Sussman,J.L. (2008) Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.*, **18**, 756–764.
10. Patil,A., Kinoshita,K. and Nakamura,H. (2010) Hub promiscuity in protein-protein interaction networks. *Int. J. Mol. Sci.*, **11**, 1930–1943.
11. Tompa,P. and Fuxreiter,M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.*, **33**, 2–8.
12. Gardner,K.H. (2005) Can you hear me now? regulating transcriptional activators by phosphorylation. *Science's STKE*, **2005**, pe44.
13. Dittmer,J. (2003) The biology of the Ets1 proto-oncogene. *Mol. Cancer*, **2**, 29.
14. Garrett-Sinha,L.A. (2013) Review of Ets1 structure, function, and roles in immunity. *Cell. Mol. Life Sci.*, **70**, 3375–3390.
15. Bui,J.M. and Gsponer,J. (2014) Phosphorylation of an intrinsically disordered segment in Ets1 shifts conformational sampling toward binding-competent substates. *Structure*, **22**, 1196–1203.
16. Pufall,M.A. (2005) Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science*, **309**, 142–145.
17. Desjardins,G., Meeker,C.A., Bhachech,N., Currie,S.L., Okon,M., Graves,B.J. and McIntosh,L.P. (2014) Synergy of aromatic residues and phosphoserines within the intrinsically disordered DNA-binding inhibitory elements of the Ets-1 transcription factor. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 11019–11024.
18. Lee,G.M. (2005) The structural and dynamic basis of Ets-1 DNA binding autoinhibition. *J. Biol. Chem.*, **280**, 7088–7099.
19. Eswar,N., Webb,B., Marti-Renom,M.A., Madhusudhan,M.S., Eramian,D., Shen,M.-Y., Pieper,U. and Sali,A. (2006) Comparative protein structure modeling using modeller. *Curr. Protoc. Bioinformatics*, **5.6**, 1–30.
20. Hess,B., Bekker,H., Berendsen,H.J.C. and Fraaije,J.G.E.M. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, **18**, 1463–1472.
21. Pronk,S., Páll,S., Schulz,R., Larsson,P., Bjelkmar,P., Apostolov,R., Shirts,M.R., Smith,J.C., Kasson,P.M., van der Spoel,D. *et al.* (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, **29**, 845–854.
22. Kamiya,N., Watanabe,Y.S., Ono,S. and Higo,J. (2005) AMBER-based hybrid force field for conformational sampling of polypeptides. *Chem. Phys. Lett.*, **401**, 312–317.
23. Higo,J., Nishimura,Y. and Nakamura,H. (2011) A free-energy landscape for coupled folding and binding of an intrinsically disordered protein in explicit solvent from detailed all-atom computations. *J. Am. Chem. Soc.*, **133**, 10448–10458.
24. Homeyer,N., Horn,A.H.C., Lanig,H. and Sticht,H. (2005) AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.*, **12**, 281–289.

25. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
26. Joung,I.S. and Cheatham,T.E. (2008) Determination of alkali and halide monovalent Ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B*, **112**, 9020–9041.
27. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
28. Ikebe,J., Umezawa,K., Kamiya,N., Sugihara,T., Yonezawa,Y., Takano,Y., Nakamura,H. and Higo,J. (2010) Theory for trivial trajectory parallelization of multicanonical molecular dynamics and application to a polypeptide in water. *J. Comput. Chem.*, **32**, 1286–1297.
29. Higo,J., Umezawa,K. and Nakamura,H. (2013) A virtual-system coupled multicanonical molecular dynamics simulation: Principles and applications to free-energy landscape of protein–protein interaction with an all-atom model in explicit solvent. *J. Chem. Phys.*, **138**, 184106.
30. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327–341.
31. Fukuda,I., Yonezawa,Y. and Nakamura,H. (2011) Molecular dynamics scheme for precise estimation of electrostatic interaction via zero-dipole summation principle. *J. Chem. Phys.*, **134**, 164107.
32. Fukuda,I. and Nakamura,H. (2012) Non-Ewald methods: theory and applications to molecular systems. *Biophys. Rev.*, **4**, 161–170.
33. Fukuda,I., Kamiya,N., Yonezawa,Y. and Nakamura,H. (2012) Simple and accurate scheme to compute electrostatic interaction: zero-dipole summation technique for molecular system and application to bulk water. *J. Chem. Phys.*, **137**, 054314.
34. Arakawa,T., Kamiya,N., Nakamura,H. and Fukuda,I. (2013) Molecular dynamics simulations of double-stranded DNA in an explicit solvent model with the zero-dipole summation method. *PLoS ONE*, **8**, e76606.
35. Fukuda,I. (2013) Zero-multipole summation method for efficiently estimating electrostatic interactions in molecular system. *J. Chem. Phys.*, **139**, 174107.
36. Fukuda,I., Kamiya,N. and Nakamura,H. (2014) The zero-multipole summation method for estimating electrostatic interactions in molecular dynamics: analysis of the accuracy and application to liquid systems. *J. Chem. Phys.*, **140**, 194307.
37. Wang,H., Nakamura,H. and Fukuda,I. (2016) A critical appraisal of the zero-multipole method: structural, thermodynamic, dielectric, and dynamical properties of a water system. *J. Chem. Phys.*, **144**, 114503.
38. Kasahara,K., Shiina,M., Fukuda,I., Ogata,K. and Nakamura,H. (2017) Molecular mechanisms of cooperative binding of transcription factors Runx1–CBFβ–Ets1 on the TCRα gene enhancer. *PLoS ONE*, **12**, e0172654.
39. Mashimo,T., Fukunishi,Y., Kamiya,N., Takano,Y., Fukuda,I. and Nakamura,H. (2013) Molecular dynamics simulations accelerated by GPU for biological macromolecules with a non-Ewald scheme for electrostatic interactions. *J. Chem. Theory Comput.*, **9**, 5599–5609.
40. Shiina,M., Hamada,K., Inoue-Bungo,T., Shimamura,M., Baba,S., Sato,K. and Ogata,K. (2014) Crystallization of the Ets1–Runx1–CBFβ–DNA complex formed on the TCRα gene enhancer. *Acta Crystallogr. F Struct. Biol. Commun.*, **70**, 1380–1384.
41. Kinoshita,E., Kinoshita-Kikuta,E., Takiyama,K. and Koike,T. (2006) Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics*, **5**, 749–757.
42. Nakajima,N., Nakamura,H. and Kidera,A. (1997) Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. *J. Phys. Chem. B*, **101**, 817–824.
43. Higo,J., Ikebe,J., Kamiya,N. and Nakamura,H. (2012) Enhanced and effective conformational sampling of protein molecular systems for their free energy landscapes. *Biophys. Rev.*, **4**, 27–44.
44. Shiina,M., Hamada,K., Inoue-Bungo,T., Shimamura,M., Uchiyama,A., Baba,S., Sato,K., Yamamoto,M. and Ogata,K. (2015) A novel allosteric mechanism on protein-DNA interactions underlying the phosphorylation-dependent regulation of Ets1 target gene expressions. *J. Mol. Biol.*, **427**, 1655–1669.
45. Berman,H., Henrick,K., Nakamura,H. and Markley,J.L. (2007) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.*, **35**, D301–D303.
46. Kawabata,T. (2016) HOMCOS: an updated server to search and model complex 3D structures. *J. Struct. Funct. Genomics*, **17**, 83–99.
47. Lu,G., Zhang,Q., Huang,Y., Song,J., Tomaino,R., Ehrenberger,T., Lim,E., Liu,W., Bronson,R.T., Bowden,M. *et al.* (2014) Phosphorylation of ETS1 by Src family kinases prevents its recognition by the COP1 tumor suppressor. *Cancer Cell*, **26**, 222–234.
48. Hahne,J.C., Okuducu,A.F., Kaminski,A., Florin,A., Soncin,F. and Wernert,N. (2005) Ets-1 expression promotes epithelial cell transformation by inducing migration, invasion and anchorage-independent growth. *Oncogene*, **24**, 5384–5388.