The Biophysical Basis for Phosphorylation-Enhanced DNA-Binding Autoinhibition of the ETS1 Transcription Factor

Cecilia Perez-Borrajero [1,2], Chang Sheng-Huei Lin [3], Mark Okon [2,4], Karlton Scheu [2], Barbara J. Graves [5,6], Michael E.P. Murphy [3] and Lawrence P. McIntosh [1,2,4,7]

*1 - **Genome Sciences and Technology Program,** University of British Columbia, Vancouver, BC, Canada V6T 1Z3*
*2 - **Department of Biochemistry and Molecular Biology,** University of British Columbia, Vancouver, BC, Canada V6T 1Z3*
*3 - **Department of Microbiology and Immunology,** University of British Columbia, Vancouver, BC, Canada V6T 1Z3*
*4 - **Department of Chemistry,** University of British Columbia, Vancouver, BC, Canada V6T 1Z1*
*5 - **Department of Oncological Sciences,** School of Medicine, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112-5550, USA*
*6 - **Howard Hughes Medical Institute,** Chevy Chase, MD 20815-6789, USA*
*7 - **Michael Smith Laboratories,** University of British Columbia, Vancouver, BC, Canada V6T 1Z4*

***Correspondence to Lawrence P. McIntosh:** Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC, Canada V6T 1Z3. mcintosh@chem.ubc.ca*
https://doi.org/10.1016/j.jmb.2018.12.011
Edited by C. Kalodimos

Abstract

The eukaryotic transcription factor ETS1 is regulated by an intrinsically disordered serine-rich region (SRR) that transiently associates with the adjacent ETS domain to inhibit DNA binding. In this study, we further elucidated the physicochemical basis for ETS1 autoinhibition by characterizing the interaction of its ETS domain with a series of synthetic peptides corresponding to the SRR. Binding is driven by the hydrophobic effect and enhanced electrostatically by phosphorylation of serines adjacent to aromatic residues in the amphipathic SRR. Structural characterization of the dynamic peptide/protein complex by NMR spectroscopy and X-ray crystallography revealed multiple modes of binding that lead to autoinhibition by synergistically blocking the DNA-binding interface of the ETS domain and stabilizing an appended helical inhibitory module against allosterically induced unfolding. Consistent with these conclusions, the SRR peptide does not interact with DNA-bound ETS1. In addition, we found that the ETS1 SRR phosphopeptide binds to distantly related PU.1 *in vitro*, indicating that autoinhibition exploits features of the ETS domain that are conserved across this family of transcription factors.

© 2018 Elsevier Ltd. All rights reserved.

Introduction

The majority of eukaryotic transcription factors contain long stretches of intrinsically disordered regions that contribute to signal integration and fine tuning of gene expression [1]. These conformationally flexible regions frequently contain sites of post-translational modifications that enable the calibration of transcription factor activity in response to cellular stimuli [2]. In the ETS family of transcription factors, disordered regions and marginally stable helices flanking the DNA-binding ETS domain have been found to be autoinhibitory for ETS1, ETS2, ETV6, ERG, and ETV1/4/5 [3–7]. A common feature of these systems is the dynamic interaction between the autoinhibitory sequences and the core ETS domains, which may allow for the rapid adjustment of conformational equilibria between active and inactive protein forms.

In the case of ETS1, both steric and allosteric mechanisms contribute to DNA-binding inhibition. An inhibitory module (IM) appended to the core ETS domain of ETS1 reduces its affinity for DNA by ~2-fold [8]. This module is composed of two N-terminal (HI-1 and HI-2) and two C-terminal (H4 and H5) helices interfaced with H1 of the ETS domain (Fig. 1a, b). Importantly, the IM is distal to the DNA-recognition helix H3. In an allosteric response

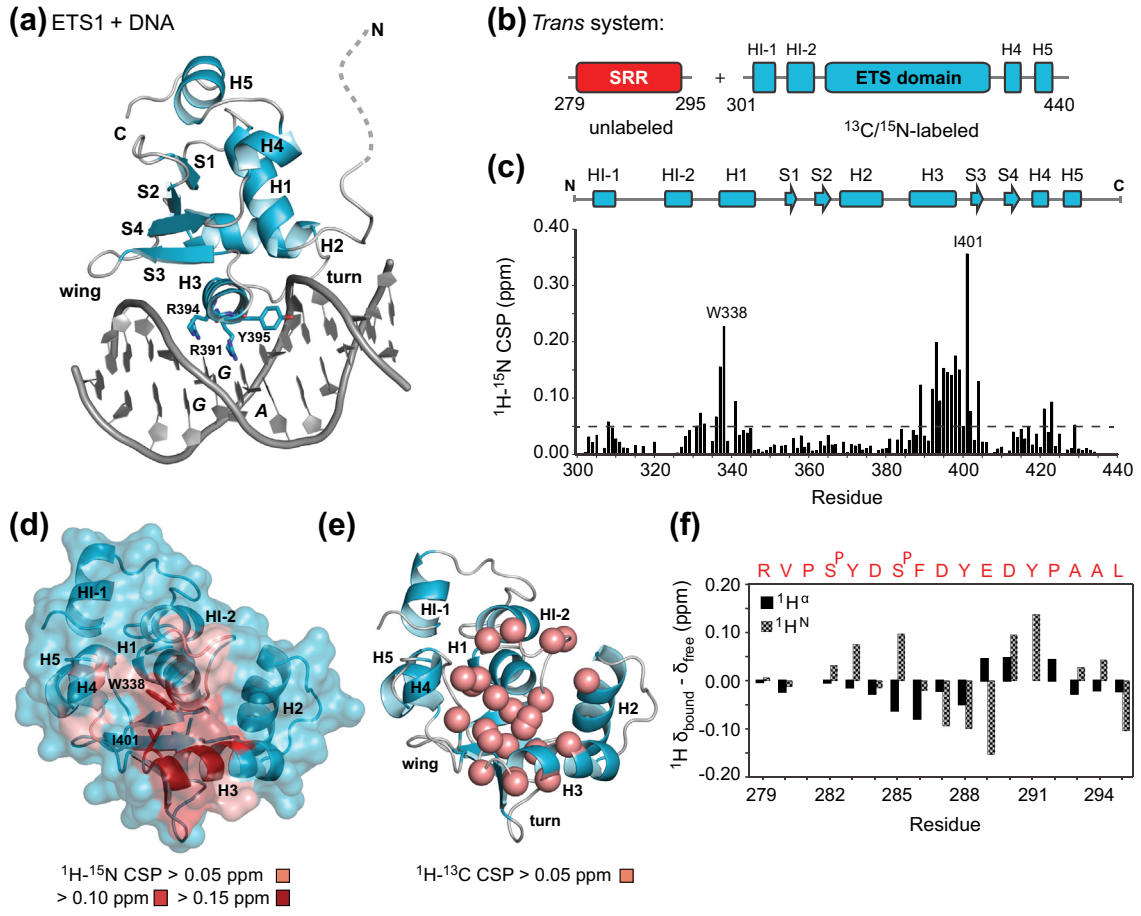

Fig. 1. The SRR interacts with a surface of the ETS domain encompassing the DNA-recognition helix H3 and flanking regions of the IM. (a) The ETS domain binds in the major groove of DNA using a winged helix(H2)–turn–helix(H3) motif. The recognition helix H3 contains a tyrosine (Tyr395) and two arginines (Arg391, Arg394) that make key hydrogen bonds with the core GGA bases of a specific ETS1 binding site. In the DNA-bound state, the inhibitory helices HI-1 and HI-2 are unfolded (adapted from PDB ID: 1MDM). (b) The *trans* system consisted of unlabeled, phosphorylated peptides corresponding to ETS1 SRR residues 279–295 (red) along with $^{15}N/^{13}C$-labeled ETS1$^{301–440}$, containing the core ETS domain and IM with HI-1, HI-2, H3, and H4 interfaced to H1 (cyan). (c, d) Amide $^{1}H–^{15}N$ CSPs of ETS1$^{301–440}$ upon the addition of the WT^{2P} peptide at 5.3-fold molar ratio in 300 mM NaCl (~80% saturation). The largest changes (> 0.05 ppm, dashed line and highlighted in red on the NMR-derived structure of free ETS1, PDB ID: 1R36) clustered around the recognition helix H3 and the loop leading to S3, as well as the N-terminus of H1 and H4 of the IM. (e) The $^{13}C^{\alpha}$, $^{13}C^{\beta}$, methyl, and aromatic moieties of ETS1$^{301–440}$ were also monitored upon the addition of the WT^{2P} peptide to a 1:1 molar ratio in 100 mM NaCl (~85% saturation). Residues with resolved signals showing a $^{1}H–^{13}C$ CSP > 0.05 ppm are identified with pink spheres. (f) The backbone chemical shifts of the WT^{2P} peptide with *trans* X-Pro amides in its free *versus* ETS1$^{301–440}$-bound state are plotted as a function of residue number for $^{1}H^{N}$ (checkered) and $^{1}H^{\alpha}$ (black) nuclei. Residues 283–291 underwent the largest chemical shift changes upon association and thus form the ETS domain-interaction interface of the SRR.

to association with either specific or non-specific DNA, the marginally stable helices HI-1 and HI-2 unfold [9–11]. The energetic penalty of this coupled conformational change attenuates the net affinity of ETS1 for DNA.

ETS1 is further regulated by an intrinsically disordered serine-rich region (SRR), composed of ~60 residues preceding helix HI-1 [12]. Previously, we showed that the SRR both stabilizes the IM and sterically inhibits DNA binding by transiently associating with an interface coarsely mapped to the recognition interface of the ETS domain [8]. This results in a combined ~20-fold basal attenuation of DNA binding. The SRR is phosphorylated by calmodulin-dependent kinase II (CaMKII) in response to calcium signaling [13,14]. An increasing number of phosphoserines causes DNA binding to become progressively weaker (up to ~1000-fold), thus tuning autoinhibition in a “rheostat” or “dimmer switch” manner [12]. Removal of the regulatory layers provided by the IM and the SRR through alternative splicing, mutation as seen with the oncoviral v-Ets, or protein partnerships results in increased affinity for DNA [3].

Recently, we reported that both phosphoserine and flanking aromatic residues in the SRR contribute synergistically to ETS1 autoinhibition [15]. However, the interactions between the disordered SRR and the ETS domain are dynamic, and thus the physicochemical basis of autoinhibition was incompletely defined. Using a combination of NMR spectroscopy and X-ray crystallography, we now present detailed structural models of the ETS domain/SRR complex and discuss the implications toward the regulation of ETS1. In particular, we demonstrate that *trans*-peptides corresponding to the SRR interact in multiple modes with an extended surface of the ETS domain partially overlapping both the DNA-recognition interface and portions of the appended IM. Binding is driven by the hydrophobic effect due to aromatic residues in the amphipathic SRR, combined with global electrostatic contributions provided by adjacent phosphoserines and aspartate/glutamate residues. Consistent with the structural models and a steric mechanism of inhibition, binding of DNA and the SRR peptide by the ETS domain are mutually exclusive. Finally, we found that the ETS1 SRR peptide can associate *in vitro* with the ETS domain of PU.1, a divergent paralog that does not exhibit autoinhibition. This raises the general possibility that proteins with intrinsically disordered sequences similar to the SRR may be capable of regulating, in *trans*, DNA binding by members of the ETS family.

Results

The SRR peptide interacts with a surface overlapping the ETS domain DNA-recognition interface

To characterize the region of ETS1 that associates with the SRR, we used a *trans* system similar to that previously described [15]. This consisted of unlabeled peptide corresponding to a truncated SRR (residues 279–295), along with isotopically labeled ETS1$^{301–440}$ containing the IM and ETS domain (Fig. 1b). We denote the peptide, with blocked termini and phosphorylated serines 282 and 285, as WT^{2P} (Table 1). Although the full SRR encompasses ~60 residues (244–300), a partial SRR with these two CaMKII phospho-acceptor serines recapitulates most of the inhibition seen with wild-type ETS1 and is well-suited for NMR spectroscopic studies [8].

Upon progressive addition of the WT^{2P} peptide to ^{15}N-labeled ETS1$^{301–440}$, many amide signals in its ^{15}N heteronuclear single quantum correlation (HSQC) spectra shifted in the fast exchange regime [16] (Fig. 2a). This indicated relatively weak binding of the peptide under conditions of moderate ionic strength (K_D 140 ± 20 μM in 300 mM NaCl; Table 1). Amide ^{1}H^{N}–^{15}N signals are very sensitive to even subtle structural perturbations. Thus, we also monitored changes in the ^{13}C-HSQC spectra of ETS1$^{301–440}$ upon peptide titration (Fig. 2b–d). Overall, the ^{1}H–^{13}C shift perturbations of aliphatic and aromatic moieties were relatively small, demonstrating that the tertiary structure of the protein was not significantly altered upon binding. ETS1$^{301–440}$ residues exhibiting the largest ^{1}H–^{15}N and ^{1}H–^{13}C chemical shift perturbations (CSPs) overlapped closely and coarsely defined the SRR peptide-interaction interface of the ETS domain as an extended surface partially encompassing the DNA-recognition helix H3, as well as H1 and portions of the IM (Fig. 1c–e). Notably, of the 24 residues exhibiting ^{1}H–^{15}N CSP values > 0.05 ppm, one-third were in helix H3 and over half were hydrophobic (Val, Leu, Ile, Tyr, and Trp).

Using conventional and double-filtered ^{1}H–^{1}H total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY) experiments, we assigned the majority of the proton

Table 1. Probing the sequence dependence of the SRR peptide interactions with ETS1$^{301–440}$

Peptide	Sequence[a]	Length	Charge[b]	K_D (μM)[c]
^{5f}Phe^{2P*}	RVPS^{P}**F**DS^{P}**F**D**F**ED**F**PAALW	18	−6	9 ± 6
Trp^{2P*}	RVPS^{P}**W**DS^{P}**W**D**W**ED**W**PAALW	18	−6	26 ± 8
Leu^{2P*}	RVPS^{P}**L**DS^{P}**L**D**L**ED**L**PAALW	18	−6	43 ± 5
Phe^{2P*}	RVPS^{P}**F**DS^{P}**F**D**F**ED**F**PAALW	18	−6	60 ± 15
WT^{2P*}	RVPS^{P}**Y**DS^{P}**F**D**Y**ED**Y**PAALW	18	−6	130 ± 35
WT^{2P}	RVPS^{P}**Y**DS^{P}**F**D**Y**ED**Y**PAAL	17	−6	140 ± 20
Val^{2P*}	RVPS^{P}**V**DS^{P}**V**D**V**ED**V**PAALW	18	−6	250 ± 50
Ala^{2P*}	RVPS^{P}**A**DS^{P}**A**D**A**ED**A**PAALW	18	−6	930 ± 170
WT^{0P*}	RVPS^{0}**Y**DS^{0}**F**D**Y**ED**Y**PAALW	18	−3	1200 ± 200
Trp^{0P*}	RVPS^{0}**W**DS^{0}**W**D**W**ED**W**PAALW	18	−3	n.d.[d]
Clus^{2P*}	RPDS^{P}DES^{P}D**Y**P**Y**VAL**Y**A**F**W	18	−6	n.d.[d]

[a] Residues in bold correspond to the four aromatic positions substituted in the peptide variants. The absence (0P) or presence (2P) of phosphorylation of Ser282 and Ser285 is indicated. The pentafluoro-phenylalanines in ^{5f}Phe^{2P*} are denoted as F. Clus^{2P}* is a sequence in which the charged and hydrophobic residues have been clustered near the N-terminus and C-terminus, respectively. The * denotes a non-native C-terminal tryptophan residue (W) used for quantitation.

[b] Approximate net charge of the peptides, with blocked termini, at pH 6.5, assuming −1.5 for pSer, −1 for Glu/Asp, and + 1 for Arg.

[c] Determined from ^{15}N-HSQC monitored titrations in buffer containing 300 mM NaCl.

[d] Not determined (n.d.) as the addition of Trp^{0P*} and Clus^{2P*} peptides caused protein aggregation.

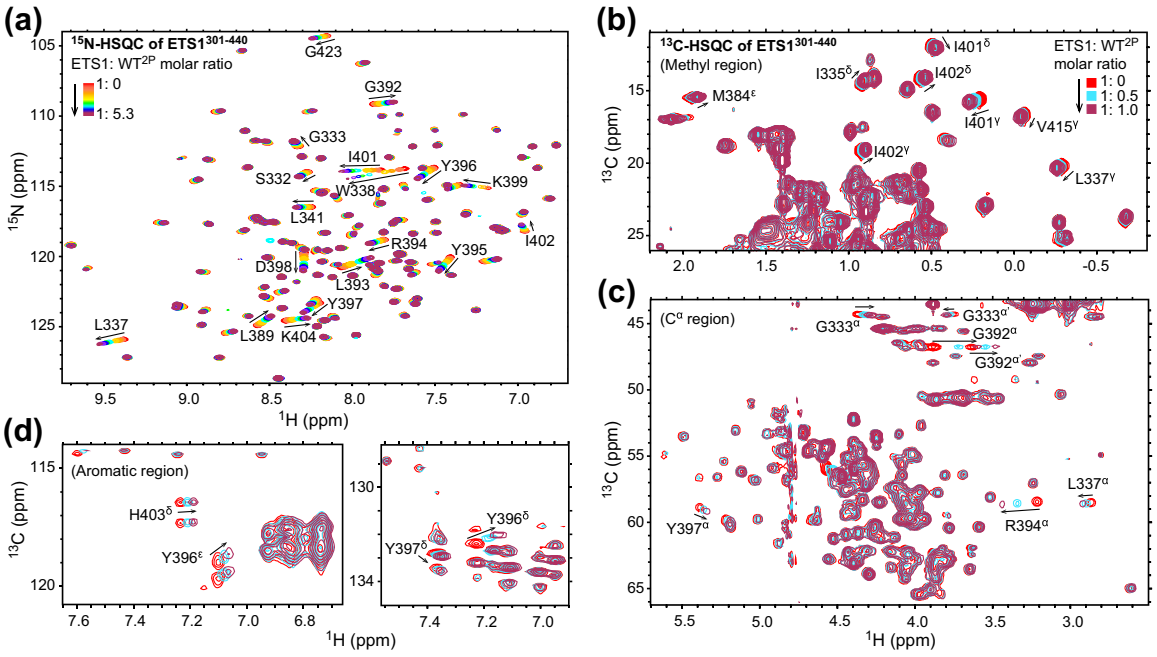

Fig. 2. NMR monitored titrations of ETS1$^{301–440}$ upon the addition of the WT^{2P} peptide. (a) Overlaid ^{15}N-HSQC spectra of ^{15}N-labeled ETS1$^{301–440}$ recorded upon titration with small aliquots of the unlabeled WT^{2P} peptide. The peptide caused many amide $^{1}H^{N}$–^{15}N signals (as well as the indole $^{1}H^{\varepsilon1}$–$^{15}N^{\varepsilon1}$ signals of Trp338 and Trp375; not shown) in the protein to shift in the fast exchange regime from the free (red) to the bound state (maroon). Residues exhibiting particularly large amide CSPs are labeled and correspond to those in the DNA-recognition helix H3 (~392–399) and nearby regions of the ETS domain/IM (see Fig. 1c, d). Under the sample conditions (300 mM NaCl, K_D ~ 140 μM), the protein:peptide molar ratio of 1:5.3 corresponds to ~80% saturation of the ETS1$^{301–440}$/WT^{2P} complex. (b–d) Overlaid ^{13}C-HSQC spectra of ^{15}N/^{13}C-labeled ETS1$^{301–440}$ in the absence (red) and presence of 0.5 (cyan) and 1.0 (maroon) molar equivalents of the unlabeled WT^{2P} peptide. Peaks in the (b) methyl, (c) C$^{\alpha}$, and (d) aromatic regions also shifted in the fast exchange regime. Overall, the peptide caused relatively small spectral perturbations, indicating that the ETS1$^{301–440}$ structure remained essentially unchanged. Residues exhibiting the greatest CSPs are labeled and highlighted in Fig. 1e. Due to spectral overlap, not all perturbed residues were assigned. Under these conditions (100 mM NaCl, K_D ~ 10 μM), the ETS1$^{301–440}$/WT^{2P} complex is ~85% saturated at the final titration point.

chemical shifts of the unlabeled WT^{2P} peptide in its free form and bound to invisible ^{15}N/^{13}C-labeled ETS1$^{301–440}$, respectively (Fig. S1). The chemical shifts of the peptide in these two states were similar, with the largest differences in $^{1}H^{\alpha}$ values relative to those predicted for a random coil polypeptide being <0.15 ppm (not shown). This indicated that the peptide did not undergo folding-upon binding to adopt a predominant well-defined secondary structural conformation. Nevertheless, residues ~283–291 exhibited small yet clear $^{1}H^{\alpha}$ and $^{1}H^{N}$ CSPs, demonstrating that they were involved in association with the ETS domain (Fig. 1f).

Collectively, these NMR spectral data define the intermolecular interfaces between the ETS1 ETS domain/IM and the SRR WT^{2P} peptide. This same interface was previously identified from a comparison of the ^{15}N-HSQC spectra of ETS1$^{301–440}$ *versus* ETS1$^{279–440,2P}$, with the latter species containing the phosphorylated SRR in *cis* [8]. Thus, the *trans*-peptide recapitulates the *cis* interaction of the SRR with the ETS domain/IM. Also, as expected, the WT^{2P} peptide showed similar behavior to that of a longer SRR peptide reported earlier [15]. These introductory results set the stage for the following expanded studies.

Increased SRR hydrophobicity strengthens interactions with the ETS domain

Within the SRR, four aromatic residues (Tyr283, Phe286, Tyr288, and Tyr291) were shown to be crucial for the phosphorylation-enhanced interaction with the ETS domain of ETS1 [15]. Mutation of these to alanines or glycines significantly impaired SRR-mediated DNA-binding autoinhibition in both intra- and intermolecular contexts. Are the tyrosine and phenylalanine residues important due to their aromatic or hydrophobic character? How do they cooperate with the adjacent phosphorylated serine residues? To answer these questions and better understand the nature of the ETS domain/SRR interaction, we expanded the *trans* system to additional peptides that retained the overall pattern of

the wild-type SRR, including phosphorylation sites, but had alternative residues at the four aromatic positions. This generated a series of peptide variants with a range of hydrophobic and aromatic properties (Table 1).

Titrations of the SRR peptide variants into ^{15}N-labeled ETS1$^{301–440}$ were monitored with ^{15}N-HSQC spectra, thereby providing structural insights and allowing the determination of equilibrium dissociation K_D values (Table 1 and Figs. 2a and S2). In all cases, residues in the protein exhibiting large amide CSPs clustered around the same region bound by the WT^{2P} peptide (Fig. 3a). This demonstrated that the general features of the SRR peptides were sufficient to direct binding to a common interface on the ETS domain. Furthermore, increasing the hydrophobic nature of the peptides resulted in tighter binding to ETS1$^{301–440}$ in the order ^{5f}Phe^{2P*} > Trp^{2P*} > Leu^{2P*} > Phe^{2P*} > WT^{2P*} > Val^{2P*} > Ala^{2P*}. Therefore, hydrophobicity, and not strictly aromaticity, is important for the interaction (Fig. 3b).

We also tested a peptide with the same amino acid composition as the WT^{2P*}, but with the negatively charged pSer, Asp, and Glu residues in one contiguous block, followed by the hydrophobic residues. Upon addition of this clustered peptide (Clus^{2P*}), ETS1$^{301–440}$ aggregated, thus precluding the determination of a K_D value (Table 1). Nevertheless, this

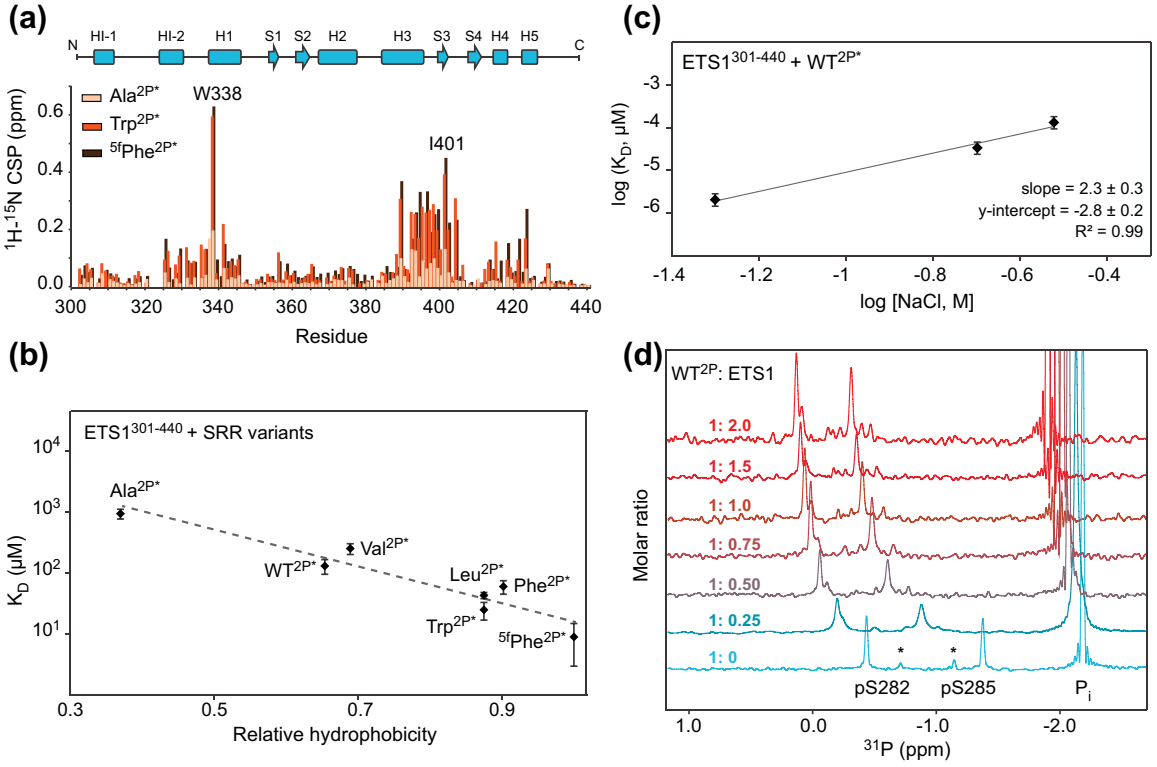

Fig. 3. Hydrophobic and electrostatic interactions drive ETS domain/SRR binding. (a) Addition of phosphorylated peptide variants to ^{15}N-labeled ETS1$^{301–440}$ caused large CSPs in the same region of the protein as affected by the WT^{2P} peptide (300 mM NaCl; compare with Fig. 1c). Shown are ^{1}H–^{15}N CSP values for representative cases of the Ala^{2P*} (beige, 1:14.5 protein:peptide ratio, ~60% saturation), Trp^{2P*} (orange, 1:5, ~95%), and ^{5f}Phe^{2P*} (brown, 1:4.5, ~98%) peptides variants. (b) The NMR-derived K_D values obtained from titrations with the phosphorylated peptides (Table 1) are plotted as a function of hydrophobicity relative to the ^{5F}Phe^{2P*} peptide, assigned a value of 1. The scale is based on side chain hydrophobicity values for the standard amino acids [46], and predicted partition coefficients (octanol/water) of Phe relative to ^{5F}Phe. The dashed line is a visual aid to highlight the trend of decreasing K_D values with increasing hydrophobicity. This overall trend is the same with other hydrophobicity scales (not shown). (c) The interaction between the SRR peptide and the ETS domain weakened with increasing ionic strength. Titrations of the WT^{2P*} peptide into ^{15}N-labeled ETS1$^{301–440}$ were carried out under conditions of differing salt concentrations (0.05 M, 0.2 and 0.3 M NaCl) and fit to obtain K_D values (Table 2). The slope and error estimates were obtained from linear regression analysis. (d) The phosphate groups on the WT^{2P} peptide are involved in the interaction with ETS1$^{301–440}$. The ^{31}P-NMR signals of the pSer282 and pSer285 in the WT^{2P} peptide shifted downfield upon titration with ETS1$^{301–440}$. The * identify signals arising from a minor population of the peptide with a *cis* Val280-Pro281 amide. A slight increase in sample pH from 6.41 to 6.51 during the titration was also observed. This is reflected by the small downfield changes in the signal from buffer inorganic phosphate, P_i. Control pH titrations confirmed that the perturbations of the phosphoserine signals were due to primarily to protein binding, with only minor contributions from the change in sample pH (not shown).

result showed that functional interactions between the ETS domain and the SRR are dependent on the amphipathic sequence pattern, and not simply on the amino acid composition of the phosphopeptide. Similarly, addition of the Trp^{0P*} peptide, which is very hydrophobic and lacks phosphorylation, also caused protein aggregation.

Electrostatic interactions with phosphoserines are also important for SRR association to the ETS domain

The absence of phosphorylation in the WT^{0P*} peptide resulted in a ~10-fold decrease in binding affinity to $ETS1^{301–440}$ (Table 1). This is consistent with previous studies showing that phosphorylation at Ser282 and Ser285 is important for binding of the SRR peptide, as well as for autoinhibition of ETS1 [8,12,15]. To gain further insights into the nature of these interactions, we carried out NMR-monitored titrations as a function of NaCl concentration. Increasing ionic strength caused binding of the phosphorylated WT^{2P*} peptide to $ETS1^{301–440}$ to weaken significantly (Table 2). This highlighted an electrostatic contribution to their interaction. Based on the slope of a $\log(K_D)$ *versus* log[NaCl] plot, we estimate that the net equivalent of 2 to 3 singly charged counterions is released upon association of this SRR peptide with the ETS domain (Fig. 3c) [17].

In addition to two phosphoserines, the SRR contains three aspartate residues, one glutamate, and one arginine that could contribute electrostatically to binding. To determine whether the phosphate groups are involved in the association with the ETS domain, the changes in the ^{31}P signals of the WT^{2P} peptide upon titration with $ETS1^{301–440}$ were also monitored. These signals were assigned to pSer282 and pSer285 via $^{3}J_{PH}$ scalar couplings with the phosphoserine $^{1}H^{\beta}$ nuclei (Fig. S3). As shown in Fig. 3d, the ^{31}P signals deriving from the phosphoserine residues in both major (*trans* Val280-Pro281) and minor (*cis*) conformers of the peptide shifted downfield upon the addition of the protein, with pSer285 exhibiting the largest CSP.

These ^{31}P-NMR experiments confirmed that the phosphate groups of pSer282 and pSer285 indeed interact with the ETS domain. The observed CSPs could have arisen directly from changes in the magnetic environments of the ^{31}P nuclei, due for example to hydrogen bonding or salt-bridge formation with the protein. However, based on coarse reference ^{31}P-NMR spectra of the free WT^{2P} peptide recorded as a function of pH, under these experimental conditions (pH ~6.5), the phosphoserines exist in acid–base equilibria with roughly equal populations of mono- and di-anionic species (not shown). In addition, it is well established that deprotonation of phosphoserine leads to a downfield ^{31}P chemical shift change [18]. Thus, the observed downfield shifts in the ^{31}P signals of the WT^{2P} peptide upon protein binding (Fig. 3d) may be the result of an electrostatically induced change in this equilibrium to favor the di-anionic species (i.e., a reduction in phosphoserine pK_a value). This is consistent with our structural results, presented below, showing that the phosphoserines localize to a positively charged region of the ETS domain without necessarily forming specific salt-bridges.

Table 2. Increasing ionic strength weakens the $ETS1^{301–440}$/WT^{2P*} interaction

Peptide	Sequence	[NaCl] (mM)	K_D (μM)
WT^{2P*}	RVP**S^{P}**YD**S^{P}**FDYEDYPAALW	50	~2[a]
		200	33 ± 9
		300	130 ± 35

[a] Estimated due to relatively tight binding and deviation from fast exchange on the NMR chemical shift timescale.

NMR-derived model of the WT^{2P}/$ETS1^{301–440}$ complex

In previous studies of inhibited ETS1 fragments, such as $ETS1^{279–440}$, the intramolecular interaction interface of the SRR and the ETS domain/IM was coarsely mapped using CSPs and paramagnetic relaxation enhancement data [8,15]. However, no unambiguous $^{1}H–^{1}H$ NOEs between protons in the SRR and the ETS domain were detected. Along with ^{15}N relaxation and amide hydrogen exchange measurements, this indicated that the SRR only transiently binds the ETS domain to mediate autoinhibition. Over the course of this current study, however, it became apparent that under conditions of reduced ionic strength, the WT^{2P}/$ETS1^{301–440}$ complex formed with low μM affinity (Table 2), allowing more detailed structural studies.

Three-dimensional filtered-edited NOESY experiments were used to selectively observe intermolecular $^{1}H–^{1}H$ NOEs between the $^{15}N/^{13}C$-labeled $ETS1^{301–440}$ and the unlabeled WT^{2P} peptide. The signals obtained were relatively weak, but with long collection times (~3 days), we identified 16 unambiguous and 17 ambiguous intermolecular $^{1}H–^{1}H$ NOEs (Fig. S4 and Table S1), as well as 114 intramolecular peptide–peptide $^{1}H–^{1}H$ NOEs, to use as distance restraints. These new experimental restraints were combined with published distance limits for free $ETS1^{301–440}$ (PDB ID: 1R36) in order to build a low resolution model of the WT^{2P} peptide docked on the ETS domain/IM.

The structural ensemble of the WT^{2P}/$ETS1^{301–440}$ complex is presented in Fig. 4. As expected with a sparse number of assigned intermolecular NOEs, the peptide has high pairwise RMSDs of 3.0 ± 0.7 Å (Fig. 4a). Nevertheless, it is well docked against the protein, adopting a loosely defined extended

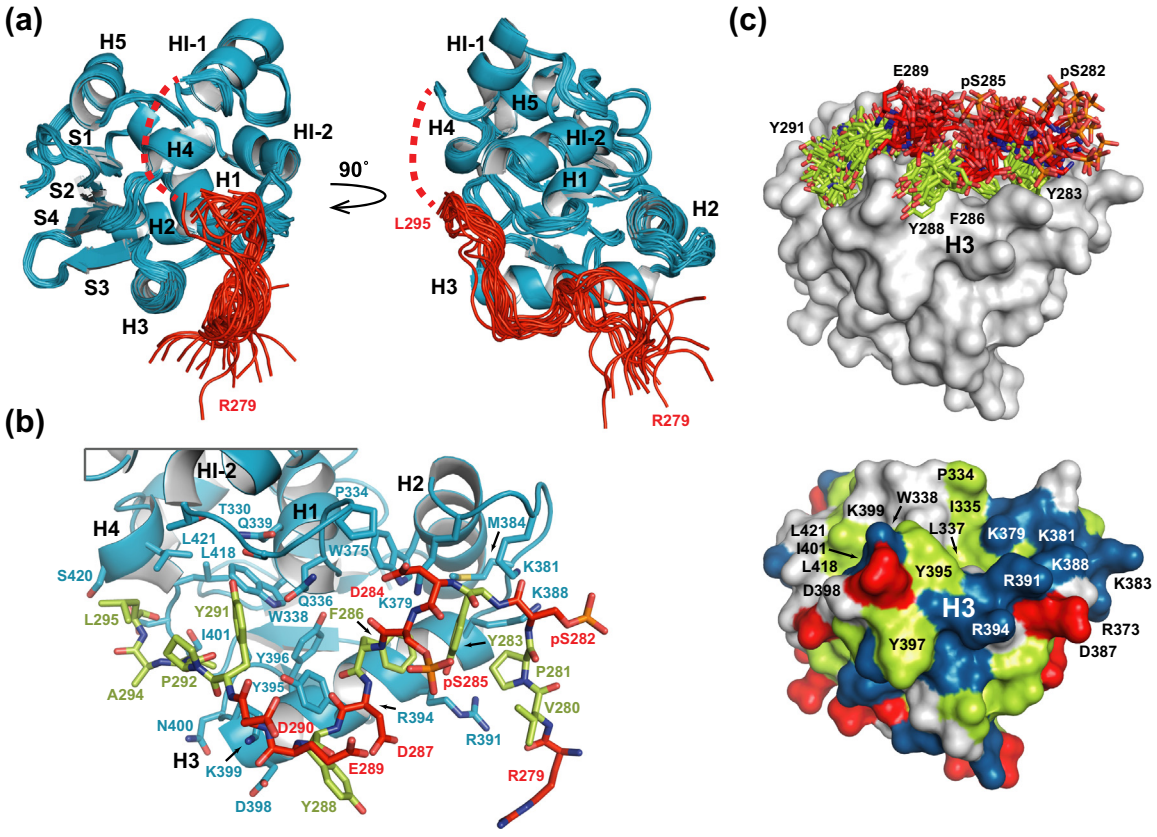

Fig. 4. NMR-derived model of the WT^{2P}/ETS1$^{301–440}$ complex. (a) Based on experimental intermolecular NOE restraints, the WT^{2P} peptide (red) is bound along a surface of ETS1$^{301–440}$ (cyan) extending from the IM to the recognition helix H3. The dashed line represents the connection between the SRR and helix HI-1 of the IM expected in the *cis* context of the native protein. (b) The top-ranked structural model calculated by CYANA is shown in detail, with the carbons of selected ETS1$^{301–440}$ residues in blue and those of the peptide in red (polar) or lime green (hydrophobic/aromatic). Also, oxygen, red; nitrogen, blue; phosphorous, orange; hydrogens, omitted. (c) Physicochemical complementarity exists between the protein and bound peptide. Hydrophobic residues (A, V, I, L, P, C, M, Y, W, F) on the WT^{2P} peptide (top; protein surface in gray) and ETS1$^{301–440}$ (bottom, peptide omitted) are colored in lime green, negatively charged residues (D, E, pS) in red, and positively charged (R, K, H) residues in blue. The very dynamic N-terminal Arg279, Val280, and Pro281 of the peptide have been excluded for clarity.

S-shape conformation spanning from the DNA-recognition helix H3 to the inhibitory helix H4. The interaction surfaces of both the peptide and protein match those defined by CSP mapping (Fig. 1c–f). Importantly, the model positions the C-terminal Leu295 of WT^{2P} in close proximity to the N-terminal Lys301 of ETS1$^{301–440}$, as would be expected when the two molecules form a continuous polypeptide in the *cis* context of the native protein (dashed line in Fig. 4a).

The WT^{2P}/ETS1$^{301–440}$ model exhibits good physicochemical complementarity with the amphipathic nature of the peptide clearly evident (Fig. 4b, c). Although the exact positioning of the side chains is variable, hydrophobic residues on the SRR are buried along a relatively large hydrophobic patch on the surface of ETS1$^{301–440}$. This docking is defined by several intermolecular NOEs between Tyr283, Phe286, Tyr291, and Leu295 of the WT^{2P} peptide with Leu337 (N-terminus of helix H1), Tyr395 (H3), I401 (H3–S3 loop), and Leu421 (H4) of ETS1$^{301–440}$. As a consequence of these experimental restraints, the negatively charged residues of the peptide (pSer282, Asp284, pSer285, Asp287, Glu289, and Asp290) are solvent exposed and localized along a positively charged region of the ETS domain, in and around helix H3, that is composed of Lys379, Lys381, Lys383, Lys388, Arg391, Arg394, and Lys399.

It is worth stressing that, with limited NOE restraints, the peptide conformation could only be defined with low precision. Furthermore, two confidently assigned NOEs between Ile401 of ETS1$^{301–440}$ and Val280 and Phe286 of the SRR could not be satisfied by the calculated structural ensemble (Fig. 4b). This suggests that, in addition to the predominant docking mode of this ensemble, the peptide also bound the protein in minor alternative orientations or alignments yielding distinct NOE patterns, yet averaged chemical

shifts due to rapid interconversion. This is in accordance with previous amide ^{15}N relaxation studies showing that *in cis* [8] and *in trans* [15], the SRR remains mobile on the sub-nsec timescale when bound to the ETS domain. Most importantly, the ensemble fits the observations that global hydrophobic and electrostatic interactions underpin the association of the SRR peptides with ETS1$^{301–440}$. In addition, it provides an explanation for SRR-mediated autoinhibition via both partial steric blockage of the ETS domain DNA-binding interface, as well as stabilization of the IM.

Crystal structure of ETS1$^{301–440}$ in complex with the ^{5f}Phe^{2P*} peptide

To gain complementary higher-resolution structural insights, we co-crystallized ETS1$^{301–440}$ in complex with the SRR variant peptide ^{5f}Phe^{2P*}, which had the highest affinity for the ETS domain. Diffraction-quality crystals with different morphologies, space groups, and cell dimensions were obtained from two conditions varying mainly in pH value. The structures of these complexes were solved to 2.00 Å (pH 7.3) and 2.35 Å (pH 8.5) resolution using molecular replacement (Figs. 5 and S5, Table S2). Both structures consist of ETS1$^{301–440}$ molecules forming domain-swapped dimers and the component monomers superposed with RMSDs $<$0.44 Å over 135 C$^{\alpha}$ atoms. As observed previously for ETS1 crystallized in the absence of DNA (e.g., 1MD0 [19]), domain swapping occurs with helix HI-1 of one monomer aligning end-on against helix H4 of a second ETS1 monomer to effectively form an extended helix. Within the complex, two ^{5f}Phe^{2P*} peptides are sandwiched symmetrically between a dimer of domain-swapped dimers, which have their DNA-recognition helices H3 adjacent and anti-parallel. This results in a 2:4 peptide:protein stoichiometry (Fig. 5a).

The bound peptide has an extended backbone conformation, interrupted by a turn centered at Glu289/Asp290 to give an overall L-shape (Fig. 5b). The N-terminal half of the peptide partakes in extensive contacts with one monomer of the domain-swapped dimer, while the C-terminus is positioned by a small number of interactions between Leu295 and the non-native Trp296 with the other monomer (Fig. 5c). This leads to the burial of ~1500 Å^2 of total accessible surface area (~1100 Å^2 nonpolar) for the protein and peptide [20]. In particular, Trp338 at the start of helix H1 and Tyr396, Tyr395, and Gly392 along helix H3 of the ETS domain lie parallel against the aromatic rings of ^{5f}Phe283, ^{5f}Phe286, ^{5f}Phe288, and ^{5f}Phe291, respectively. Additional interactions at the hydrophobic surface of ETS1$^{301–440}$ involve Leu337 and Ile401, as well as Cys416 and Leu421 (helix H4) of the IM. Two potential salt-bridges to this monomer are formed by pSer282 with Lys399 and Asp290 with Arg391. The remaining negatively charged side chains of the ^{5f}Phe^{2P*} SRR variant (Asp284, pSer285, Asp287, and Glu289) all extend outward to interact with a constellation of lysine and arginine residues along the DNA-binding interface of one monomer of the partner domain-swapped dimer (Fig. 5d).

The crystal structure of the ^{5f}Phe^{2P*}/ETS1$^{301–440}$ complex recapitulates the key features of the NMR-derived model of the WT^{2P}/ETS1$^{301–440}$ complex. Namely, aromatic residues of the amphipathic peptides dock along the same hydrophobic surface of the ETS domain, while charged residues extend outward for flanking electrostatic interactions. However, several notable differences exist. The 2:4 stoichiometry in the crystals stands in contrast to the 1:1 binding stoichiometry observed by NMR spectroscopy under solution conditions where ETS1$^{301–440}$ is monomeric [10]. Of course, a 1:1 stoichiometry also corresponds to the native protein in which the SRR and ETS domain/IM are connected as a single polypeptide chain. In addition, the peptides are oriented in opposite directions such that the N-terminus of ^{5f}Phe^{2P*} is near the IM helices and the start of the protein. In the NMR-derived models, the C-terminal Leu295 of the WT^{2P} peptide is proximal to Gly301, as should occur in the context of the native protein.

The opposite alignment of the ^{5f}Phe^{2P*} peptide across the ETS domain in the crystal complexes is inconsistent with the majority of the intermolecular NOEs identified between WT^{2P} and ETS1$^{301–440}$. This is not simply due to differing amino acid compositions since, in solution, WT^{2P} and ^{5f}Phe^{2P*} interact in a similar manner with ETS1$^{301–440}$ as evidenced by common patterns of spectral perturbations (Fig. 3a). However, as discussed above, NMR measurements do suggest multiple binding modes. It is also noteworthy that the peptides have a pseudo-palindromic repeating pattern of aromatic and negatively charged side chains (Table 1). This may enable binding to ETS1$^{301–440}$ in two orientations, with the "reversed" one being favored upon crystallization between a dimer of domain-swapped dimers. Further evidence for multiple binding modes is provided by weak unmodeled electron density observed in the crystal structures of the ^{5f}Phe^{2P*}/ETS1$^{301–440}$ complex that is attributed to partial occupancy of one or more additional peptides along the hydrophobic interfaces of the ETS domains (see Fig. S5 for a detailed explanation).

Comparison of the structures of ETS1 in its free *versus* DNA- or peptide-bound states reveals a particularly interesting difference in the rotamer conformations of Tyr395 (Fig. 5e). In both bound states, Tyr395 extends outward from helix H3 to hydrogen bond with a DNA base or to interact with the peptide through stacking against the ring of ^{5f}Phe288 and hydrogen bonding with Asp290. In contrast, the

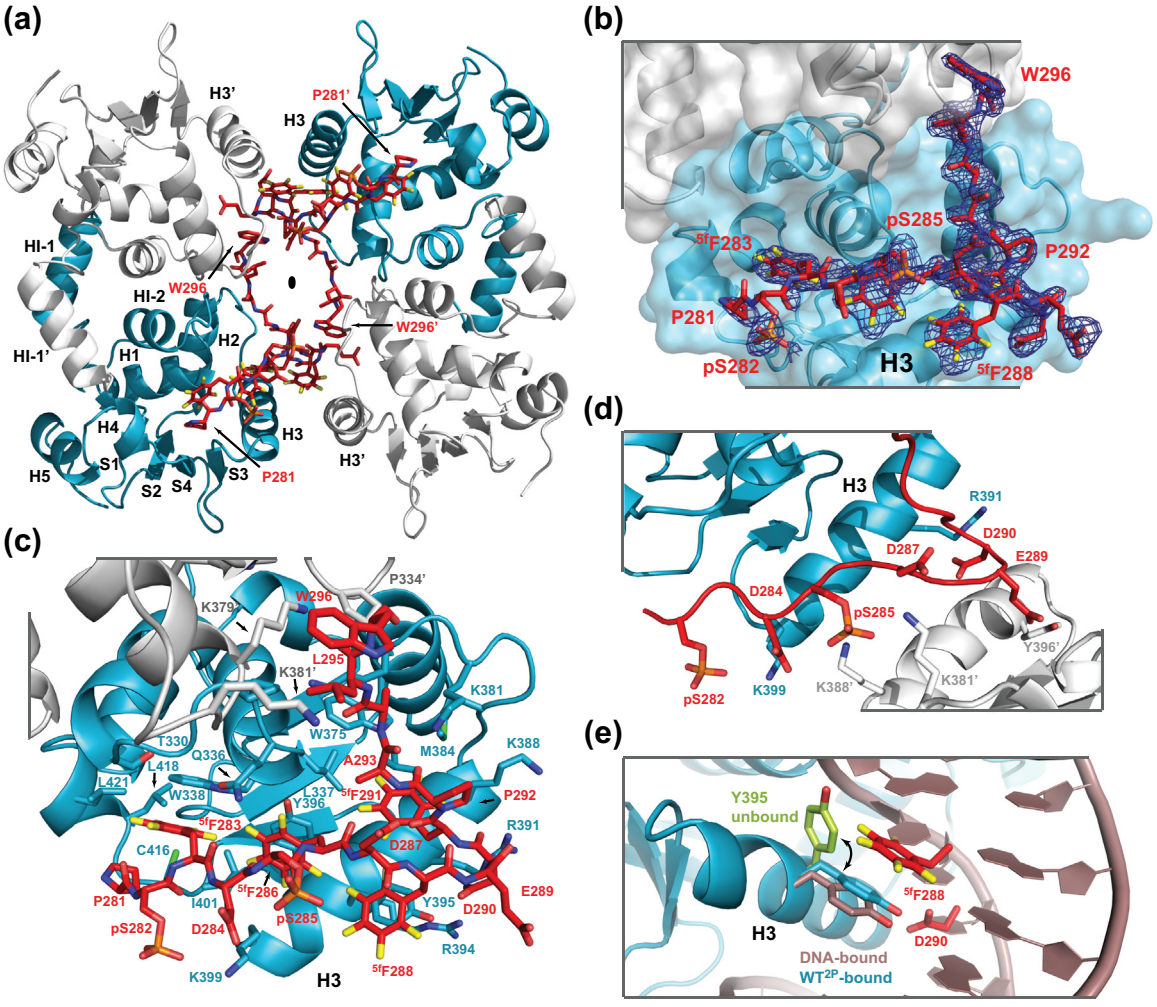

Fig. 5. Crystal structure of the ^{5f}Phe^{2P*}/ETS1$^{301–440}$ complex at 2-Å resolution (pH 7.3). (a) Shown are two domain-swapped ETS dimers related by two-fold crystallographic symmetry indicated by the central black oval. In the domain-swapped dimer, helix H4 of one monomer (cyan) is aligned with helix HI-1 of a second monomer (gray). Two peptides (red) are sandwiched between four ETS1$^{301–440}$ monomers for a 2:4 stoichiometry (oxygen, red; nitrogen, blue; phosphorous, orange; sulfur, green; fluorine, yellow). (b) A single peptide viewed on the molecular surface of an ETS dimer. An omit difference map (dark blue) for ^{5f}Phe^{2P*} was generated by the program phenix.polder and contoured of 1 σ. Density for peptide residues 281–296 is clearly defined, whereas density for the N-terminal Arg279 and Val280 was not observed. The latter are also very dynamic in solution as evidenced by their sharp signals in NMR spectra. (c) Close-up of the primary peptide–protein hydrophobic interface. Residues containing atoms within 5 Å of the peptide are shown in stick format. Notably, ^{5f}Phe283, ^{5f}Phe286, ^{5f}Phe288, and ^{5f}Phe291 in the peptide stack parallel over Trp338, Tyr396, Tyr395, and Gly392 in the ETS domain, respectively. ^{5f}Phe288 and ^{5f}Phe291 are adjacent with a turn centered near Glu289/Asp290. (d) Close-up of the cross dimer interface. While docked against the hydrophobic surface of one protein monomer, negatively charged residues of the peptide, including Asp284, pSer285, Asp287, and Glu289, extend outward to interact with Lys381′, Lys388′, and Tyr396′ of an adjacent symmetry related monomer (gray). (e) The ^{5f}Phe^{2P*} peptide and DNA occupy mutually exclusive positions along the DNA-recognition H3 of the ETS domain. In both bound states, Tyr395 extends outward to interact with a DNA base (in dark mauve, PDB ID: 2NNY) or with the peptide (cyan with only the side chains of ^{5f}Phe288 and Asp290 shown in red). When unbound, Tyr295 lies more parallel to H3 (lime, PDB ID: 1MD0).

tyrosine side chain lies more parallel along H3 in the unbound state. Therefore, Tyr395 undergoes a common conformational change to mediate mutually exclusive binding to either DNA or the inhibitory SRR. Parenthetically, this residue is also involved in the altered specificity of ETS1 for DNA induced by a partnership with the transcription factor Pax5 [21,22].

The SRR peptide does not associate with the DNA-bound ETS domain

Both NMR spectroscopy and X-ray crystallography showed that the SRR peptides bind along the DNA-recognition interface of the ETS domain (Fig. 6a). To test the implications for autoinhibition,

we used NMR spectroscopy to monitor competitive titrations of the WT^{2P} peptide *versus* a high-affinity cognate DNA duplex to ^{15}N-labeled $ETS1^{301–440}$. No changes in the ^{15}N-HSQC spectrum of the protein in a pre-formed 1:1 complex with DNA were observed upon the addition excess WT^{2P} peptide (Fig. 6b). Thus, at these concentrations (~0.16 mM), the peptide did not measurably displace DNA from the ETS domain, nor did it associate with other regions of DNA-bound $ETS1^{301–440}$. Conversely, upon progressive addition of DNA to the preformed $ETS1^{301–440}/WT^{2P}$ complex, amide $^{1}H^{N}–^{15}N$ signals moved in slow exchange to the corresponding chemical shifts of the DNA-bound state (Fig. 6c).

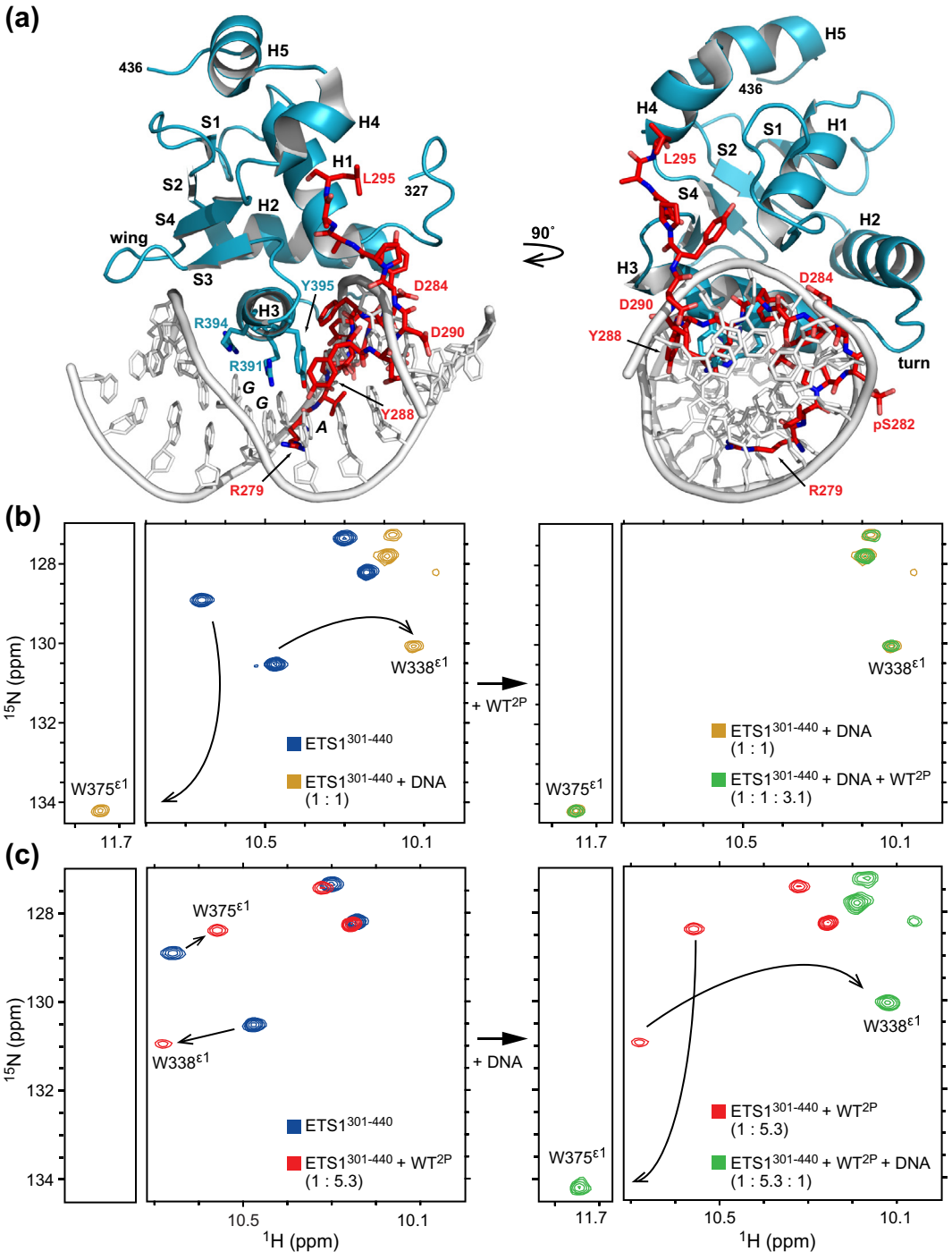

Fig. 6 (*legend on next page*)

This reverse titration demonstrated that DNA binding was favored and precluded interactions with the WT^{2P} peptide. Together, these experiments support a steric mechanism of autoinhibition, in which association of the SRR and DNA by the ETS domain is mutually exclusive. They are also consistent with the significantly higher affinity of $ETS1^{301–440}$ for cognate DNA (nM range [8,23]) *versus* the WT^{2P} peptide (~140 μM in 300 mM NaCl).

The ETS1 SRR peptide can bind distantly related PU.1

PU.1 is a member of the ETS family of transcription factors involved in myeloid and B-cell development [24]. Relative to ETS1, it is the most evolutionarily distant ETS factor [3]. PU.1 also lacks an SRR-like sequence and has not been reported to exhibit DNA-binding autoinhibition [23]. Nevertheless, we used NMR-monitored titrations to test whether the ETS1 WT^{2P} peptide could bind the ETS domain of PU.1 *in vitro*. Addition of this peptide caused amide and indole chemical shift changes in the fast exchange regime, indicating a weak association with $PU.1^{167–272}$ (Fig. S6). However, fewer residues were perturbed than observed with $ETS1^{301–440}$, and the resulting chemical shift changes were generally smaller (Fig. 7a; compare with Fig. 1c). Also, the perturbed residues mapped to two non-contiguous sites on the surface of the PU.1 ETS domain (Fig. 7b). Although estimated in the 0.5-mM range, the presence of multiple binding sites precluded the accurate determination of K_D values between the WT^{2P} peptide and $PU.1^{167–272}$ (Fig. 7c).

One group of $PU.1^{167–272}$ residues that exhibited substantial CSPs upon the addition of the WT^{2P} peptide clustered around the loop between β-strands S3 and S4. This "wing" region of the PU.1 ETS domain, which contains the sequence ^{246}VKKKL^{250}, is positively charged and involved in binding the phosphodiester backbone of DNA [25]. Therefore, the interaction between the negatively charged WT^{2P} and $PU.1^{167–272}$ at this site is likely mediated by electrostatic contacts. The corresponding wing of ETS1 has the divergent sequence ^{405}TAGKRY^{410} and does not contribute to SRR association.

The second site of $PU.1^{167–272}$ involved in WT^{2P} peptide binding is akin to the SRR-interaction interface of ETS1. This site was identified based on the CSPs exhibited by the amides of Tyr175 at the start of helix H1, Glu242 following H3, and the indole of Trp215 in H2. The corresponding residues in $ETS1^{301–440}$ are Trp338, Ile401, and Trp375, respectively, all of which exhibit large CSPs upon association with the same peptide. A comparison of the PU.1 and ETS1 ETS domains at this peptide-binding site shows common surface features with a hydrophobic patch flanked by positively charged residues along the DNA-recognition interface (Fig. 7d). Thus, although the autoinhibitory *cis*-SRR is only present in ETS1 and ETS2, it functions by exploiting characteristics of the ETS domain that are conserved even among distantly related family members.

Discussion

Hydrophobic groups in the SRR mediate interactions with the ETS domain

We investigated the mechanisms underlying the phosphorylation-enhanced autoinhibition of ETS1 by characterizing the intermolecular interactions of peptide models of the SRR with the ETS domain and appended IM. In both NMR-derived structural ensembles and two X-ray crystallographic structures of the peptide/protein complexes, Tyr/Phe and ^{5f}Phe residues of the *trans*-SRR, respectively, dock along a complementary hydrophobic patch on the surface of $ETS1^{301–440}$. A series of peptides differing at the wild-type aromatic positions (283, 286, 288, and 291) all bound this same general interface, giving

Fig. 6. Mutually exclusive binding of DNA and the WT^{2P} peptide to the ETS domain. (a) A diagram showing one representative SRR peptide (red) from the NMR-derived model of the $WT^{2P}/ETS1^{301–440}$ complex (Fig. 4) overlaid upon an X-ray crystallographic structure of the ETS1 ETS domain (cyan) complexed with DNA (PDB ID: 2NNY). The SRR and DNA associate with the ETS domain via partially overlapping interfaces. To test the predicted steric mechanism of autoinhibition, ^{15}N-HSQC spectra were used to monitor the competitive interactions of the WT^{2P} peptide and a cognate DNA duplex to ^{15}N-labeled $ETS1^{301–440}$. (b) Left: Starting with free $ETS1^{301–440}$ (blue), DNA was added in small increments to make a 1:1 protein/DNA complex (gold). Consistent with high-affinity binding, the indole $^{1}H^{\varepsilon1}–^{15}N^{\varepsilon1}$ signals of Trp375 and Trp338 shifted in slow exchange from the free to the DNA-bound state [11]. Right: WT^{2P} peptide was then added to a 3.1 molar ratio (green). No further changes were observed indicating that, at these conditions (300 mM NaCl) and concentrations (~0.16 mM), the WT^{2P} peptide does not displace the bound DNA. (c) Left: The reverse experiment with the WT^{2P} peptide first added to a 5.3 molar ratio. As expected for weaker binding, the peaks shifted in fast exchange from the free (blue) to the peptide-bound state (red). Right: Subsequently, the DNA duplex was added in small increments (green). The indole peaks moved in slow exchange from the peptide-bound to the DNA-bound state, indicating displacement of the peptide by the DNA. For clarity, intermediate spectra along the titration series are not shown, and fast and slow exchange behaviors are denoted by the straight and curved arrows, respectively. Corresponding changes in amide $^{1}H^{N}–^{15}N$ signals (top right of each spectra) are not indicated with arrows.

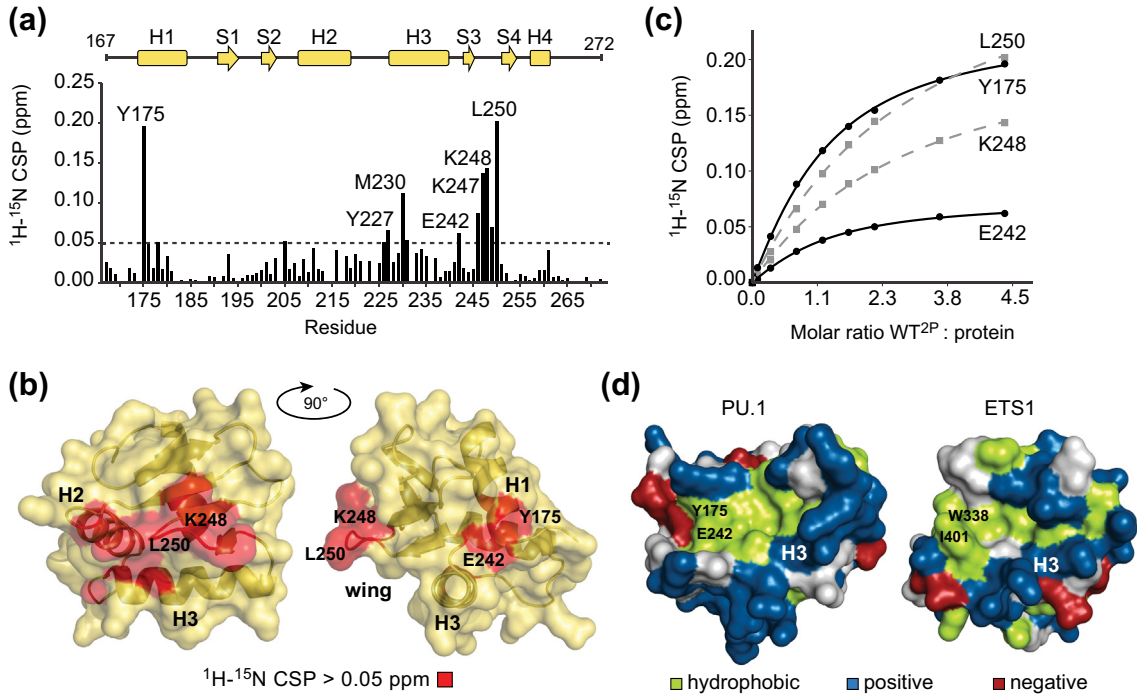

Fig. 7. The ETS1 SRR peptide also interacts with distantly related PU.1. (a) Titration of the ETS1 WT^{2P} peptide into ^{15}N-labeled PU.1$^{167–272}$ caused amide chemical shifts in changes in the fast exchange regime (Fig. S6). Plotted are the resulting ^{1}H–^{15}N CSPs upon the addition of a 4.4-fold molar ratio of WT^{2P} peptide (300 mM NaCl). Blank values correspond to prolines and residues with signals that could not be unambiguously assigned. The top cartoon shows the secondary structural elements of the PU.1 ETS domain. (b) Residues with CSP > 0.05 ppm (dashed horizontal line in panel a) are highlighted in red on the structure of PU.1$^{167–272}$ (PDB ID: 5W3G). The WT^{2P} peptide perturbed two non-contiguous interfaces identified by Lys248 and Leu250 in the wing between strands S3 and S4, and by Tyr175 (helix H1) and Glu242 (following H3). Trp338 and Ile401 of ETS1, which correspond to Tyr175 and Glu242 of PU.1, respectively, also showed large CSPs upon peptide binding. (c) It is plausible that one peptide could span both surfaces of the PU.1 ETS domain. However, amides in these interfaces reported different titration curves, indicative of two independent binding sites with different affinities for the peptide. For example, at the titration end-point, Tyr175 and Glu242 (black solid lines) showed CSP values closer to saturation (and hence higher affinity binding) than did Lys248 and Leu250 (gray dashed lines). (d) The ETS domains of both PU.1 and ETS1 (PDB ID: 1R36) contain a conserved hydrophobic patch flanked by positively charged residues that overlaps the DNA-recognition interface and is perturbed by the WT^{2P} peptide. The surface representations of the core ETS domains are colored according to amino acid properties (green, hydrophobic; blue, positive; red, negative).

similar patterns of CSPs. Importantly, the strength of the interaction correlated with the hydrophobicity of the varied residues, and not necessarily their aromatic character. The observation that the exact sequence is not critical in localizing the peptide, and that multiple variants have relatively high affinity for ETS1$^{301–440}$, is also consistent with the conclusion that the association between the SRR and ETS1 is driven by the hydrophobic effect involving these four residues. However, van der Waals forces and possibly hydrogen bonding with more specific stereochemical constraints likely contribute to association. For example, in the crystal structures, the aromatic rings of ^{5f}Phe283, ^{5f}Phe286, and ^{5f}Phe288 of the peptide stack in a parallel-displaced manner [26] against those of Trp338, Tyr396, and Tyr395, respectively, from the protein. Conversely, since ^{5f}Phe is very hydrophobic yet has reduced aromatic ring electron density, π–cation interactions with lysine or arginine residues do not dictate formation of the high affinity ETS1$^{301–440}$/^{5f}Phe^{2P*} complex [27].

Previously, we reported that aromatic residues in the SRR are critical for ETS1 autoinhibition [15]. This conclusion followed from the observation that ETS1$^{279–440}$ (with the SRR appended in *cis* to the ETS domain/IM) exhibited significantly greater phosphorylation-enhanced inhibition of DNA binding than did variants with the Tyr/Phe at positions 283, 286, 288, and 291 replaced by Val, Ala, or Gly. Unfortunately, a variant with four Leu substitutions could not be tested for DNA binding due to incomplete phosphorylation by CaMKII *in vitro* [28]. Using an extensive set of synthetic phosphopeptides with both aromatic and aliphatic substitutions, we have now found that the affinity of the *trans*-SRR

peptides for the ETS domain/IM correlates with hydrophobicity. For example, although lacking four aromatic residues, Leu^{2P*} bound ETS1$^{301–440}$ with higher affinity than did WT^{2P*}.

This raises the question, why are the Tyr/Phe residues adjacent to the SRR phosphoacceptor serines conserved among ETS1 and closely related ETS2 orthologs [15]? It is plausible that autoinhibition of DNA binding by wild-type ETS1 is dependent upon additional factors not fully recapitulated by the association of model *trans*-SRR peptides to the ETS domain/IM. However, the individual replacement of Tyr283 or Phe286 in the SRR with several large aliphatic amino acids was recently reported to not markedly impact phosphorylation-enhanced autoinhibition of an ETS1 fragment [29]. Thus, the conserved Tyr/Phe may serve additional sequence-specific biological functions, such as mediating phosphorylation of the SRR by CaMKII or dephosphorylation by phosphatases. Also, multiple signaling pathways converge at the SRR to regulate ETS1 function. For example, phosphorylation of Ser282 provides a recognition site for the ubiquitin ligase COP-1, whereas phosphorylation of the adjacent Tyr283 by Src family tyrosine kinases decreases COP-1 binding to prevent ubiquitin-mediated degradation of ETS1 [30]. Finally, the different physicochemical properties of aromatic *versus* aliphatic residues, such as the hydrogen bonding potential of tyrosine, may perhaps influence the solubility of the SRR to reduce potential protein aggregation.

Global electrostatics also contribute to SRR binding

In addition to hydrophobic interactions, binding of the SRR peptides to the ETS domain/IM is dependent upon their complementary electrostatic properties. This conclusion is supported by several lines of evidence, including a diagnostic reduction in affinity with increasing sample ionic strength. The importance of serine phosphorylation, which elevates the predicted net charge of the peptides from approximately −3 to −6, is seen by an 10-fold difference in the K_D values of WT^{2P*} *versus* WT^{0P*} for ETS1$^{301–440}$ (pH 6.5, 300 mM NaCl). Also, NMR-monitored titrations demonstrated that the ^{31}P nuclei of pSer282, and more so pSer285, were perturbed in the presence of ETS1$^{301–440}$. The resulting downfield ^{31}P chemical shift changes are consistent with binding favoring both phosphoserines in their di-anionic forms.

Although the contribution of phosphorylation and charge interactions in ETS1 autoinhibition was previously established [8], an important result of our structural studies is that the binding of the ETS domain/IM by the SRR peptides hinges upon global electrostatics rather than well-defined salt-bridges. In the NMR-derived models of the ETS1$^{301–440}$/WT^{2P} complex, the two phosphoserines, three aspartates, and one glutamate of the SRR are only broadly localized in the general proximity of a continuous electropositive region near the ETS1 DNA-binding interface that is composed of five lysines and two arginines. It must be noted that no intermolecular NOE restraints involving the WT^{2P} charged side chains were confidently detected, and thus their positions in the calculated ensembles are a consequence of satisfying the observed restraints involving adjacent aromatic groups. However, even in the crystal structures of the ETS1/^{5f}Phe^{2P*} complex, only two potential intermolecular salt-bridges (pSer282–Lys399 and Asp290–Arg391) were found at the primary interface. Recently, it was proposed from molecular dynamic simulations that the SRR forms a 3_{10}-helix, positioning pSer282 and pSer285 to salt-bridge with Arg394 and Arg391, respectively, along the recognition helix H3 of the ETS domain [29]. A predominant contribution of such specific electrostatic interactions, along with an induced secondary structural element, is not borne out by our experimental studies.

Synergy of aromatic/hydrophobic residues and phosphoserines

The interaction of the SRR peptides with the ETS domain/IM is dependent upon aromatic/hydrophobic residues as well as phosphoserines. For example, Ala^{2P*} lacking Phe/Tyr and WT^{0P*} lacking phosphorylation both bind ETS1$^{301–440}$ with approximately 10-fold lower affinity than WT^{2P*}. Phosphorylation-enhanced autoinhibition of ETS1 DNA binding is also substantially weakened upon substitution of the four Phe/Tyr in the SRR with Ala or Gly [15]. What is the basis for this synergy? A simple explanation is that complex formation results from the cooperative reinforcement of many individual weak interactions between the amphipathic SRR peptide and the ETS domain/IM. Structurally, burial of SRR aromatic residues against the hydrophobic surface of the ETS domain requires that the interleaved phosphoserine, aspartate, and glutamate residues be solvent exposed and thus favorably positioned along the flanking electropositive protein region. Conversely, positioning of these negatively charged residues for complementary electrostatic interactions with the ETS domain clusters the intervening aromatic groups against the protein's hydrophobic surface.

Previously, we found that aspartate and glutamate are poor mimics of phosphoserine for ETS autoinhibition [15]. This result appears greater than expected solely from the differences in the net charges (−2 for di-anionic pSer *versus* −1 for Asp and Glu) of these side chains. Thus, serine phosphorylation may play a more specific role than simply enhancing electrostatic interactions of the SRR with the ETS domain/IM. For example, synergy could result from an enhancement in the hydrophobic

properties of the SRR aromatic residues as a result of phosphorylation. Phosphate is a salting-out anion in the Hofmeister series, and generally decreases the solubility of non-polar residues and increases protein stability [31,32]. It is plausible that a phosphoserine has similar effects on neighboring hydrophobic groups, possibly through perturbations in local water structure. In support of this hypothesis, it was recently shown that hydrophobic interactions can be strongly increased by the protonation of nearby (nanometer scale) amines to become ammonium ions [33]. The related ammonium cation is also salting-out in the Hofmeister series.

A "fuzzy" complex

In contrast to intrinsically disordered protein regions that adopt well-defined conformations upon partner binding, earlier NMR chemical shift and ^{15}N relaxation studies revealed that the SRR remains dynamic and without any persistent induced structure when associated with the ETS domain/IM. This holds for both the intermolecular binding of the SRR peptides to ETS1$^{301–440}$ [15] and in the intramolecular context of ETS1$^{279–440}$ or larger ETS1 constructs [8,12].

Further insights into the "fuzzy" nature [34] of the SRR interaction are provided by this investigation. Only a sparse number of relatively weak intermolecular NOE restraints were detected between WT^{2P} and ETS1$^{301–440}$, and thus, the peptide has high RMSDs in the calculated structural ensemble of the complex. Although a lack of restraints can arise in several ways, this is consistent with a dynamic peptide undergoing rapid exchange between multiple energetically similar bound conformations. In X-ray crystal structures, the ^{5f}Phe^{2P*} peptide is ordered yet sandwiched in a 2:4 stoichiometry between two domain-swapped ETS1$^{301–440}$ dimers. Weak electron density attributed to one or more additional bound peptides was also detected, but could not be adequately modeled due to disorder and low occupancy. On one hand, the general features of the structures defined by these two techniques are similar and explain why binding of the amphipathic SRR peptides to the ETS domain/IM is tolerant to amino acid substitutions and dependent upon global hydrophobic and electrostatic interactions. On the other hand, in crystals, the ^{5f}Phe^{2P*} peptides bound with the opposite orientation across the ETS domain than found for the WT^{2P} peptide in the NMR-derived ensemble. However, two confidently assigned intramolecular NOEs not satisfied by the calculated structural ensemble of the WT^{2P}/ETS1$^{301–440}$ complex are roughly consistent with the reversed orientation of ^{5f}Phe^{2P*} in the crystallized complex. This suggests that in solution, the SRR peptides, with a pseudo-palindromic repeating pattern of charged and aromatic side chains, bind the ETS domain/IM in multiple interconverting orientations or alignments. Although the predominant binding mode is that shown in Fig. 4, alternative modes could still preserve a similar overall fuzzy pattern of hydrophobic and electrostatic interactions. Such multi-modal binding is also predicted through molecular dynamics simulations [29].

In native ETS1, the full SRR spans ~60 residues and encompasses three additional CaMKII phosphoacceptor serines. Of these, phosphorylation of Ser251 (in the sequence Asp-Ser-Phe-Glu) also reinforces DNA-binding autoinhibition [12]. Furthermore, the presence of four extra pSer-Phe/Tyr-Asp motifs introduced into the truncated SRR of ETS1$^{279–440}$ essentially abrogated binding to DNA [15]. This repeating motif maintains the amphipathic pattern of alternating aromatic and negatively charged residues needed for association of the SRR peptides to the ETS domain/IM. However, it is unlikely that additional repeats, including that of the full SRR in its native context, can be simultaneously accommodated along the hydrophobic interface of the ETS domain. Rather, multiple interconverting modes of binding may occur, each contributing to the overall free energy of complex formation. Such fuzzy behavior also explains several aspects of ETS1 autoinhibition including the observed graded decrease in DNA affinity with increasing phosphorylation without a strong dependence on the exact sites of SRR modification [12].

Coupled steric and allosteric mechanisms of ETS1 DNA-binding autoinhibition

The insights gained through characterizing the intermolecular interactions of the SRR peptides with ETS1$^{301–440}$ refine our mechanistic understanding of ETS1 autoinhibition. A direct steric contribution is immediately evident from a comparison of the structural models of ETS1 constructs bound to DNA and to the *trans*-peptides (Figs. 6a and 8). Both ligands interact with similar key regions of the ETS domain, including the N-terminus of helix H1 (e.g., Leu337, Trp338), the turn after helix H2 (Lys379, Lys381), the recognition helix H3 (Tyr395 and Tyr396), and the loop to strand S3 (Ile401). Also, NMR-monitored competition studies confirmed that binding of the WT^{2P} SRR peptide and DNA to the ETS domain are mutually exclusive.

In these experiments, the peptide and protein are separated (*trans*), whereas in their native context, they are covalently linked (*cis*) via an additional five dynamic residues (^{296}PNHKP^{300}) [8]. This linkage will at least partially reduce the entropic penalty for the intramolecular association of the SRR with the ETS domain/IM to facilitate autoinhibition [35,36]. To gauge this effect, it is instructive to compare the published amide chemical shift differences between ETS1$^{279–440}$ and ETS1$^{279–440,2P}$ (with the *cis*-SRR)

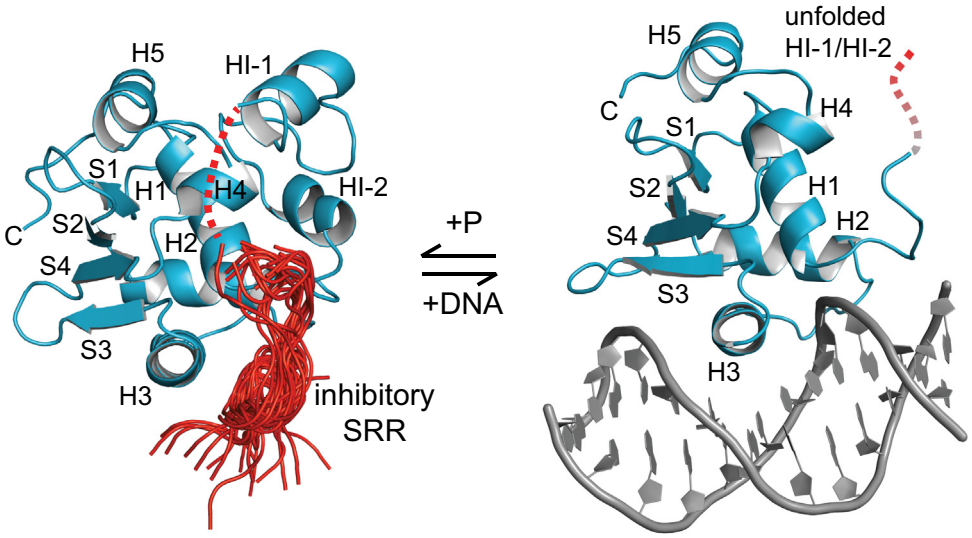

Fig. 8. Coupled steric and allosteric mechanisms of ETS1 DNA-binding autoinhibition. The SRR both sterically blocks the DNA-recognition interface of the ETS domain and stabilizes helices HI-1 and HI-2 of the IM against unfolding that is allosterically linked to DNA binding. The basal interactions of the dynamic SRR with the ETS domain/IM are progressively enhanced by multi-site phosphorylation, thereby providing a graded regulation of autoinhibition [12].

versus ETS1$^{301–440}$ (lacking the SRR) [8] against the extrapolated saturating CSPs for binding of the WT^{0P*} and WT^{2P*} peptides to ETS1$^{301–440}$ measured herein (Fig. S2). From this comparison, we estimate that the ETS domains of ETS1$^{279–440}$ and ETS1$^{279–440,2P}$ are ~40% and ~85% in their SRR-bound states, respectively. Combined with the corresponding K_D values of Table 1, this indicates that the effective concentration [36,37] of the *cis*-SRR is ~1 mM in both its unmodified and phosphorylated forms.

This effective concentration is insufficient to explain the inhibitory effects of the SRR on ETS1 DNA binding solely through a steric blockage mechanism. That is, the K_D value of ETS1$^{301–440}$ for a consensus DNA is ~25 pM and the presence of the unmodified *cis*-SRR (ETS1$^{279–440}$) weakens it by ~20-fold to the basal level of the wild-type protein [8,15]. Assuming a simple competition model [38], the effective concentration of the *cis*-SRR would need to be significantly higher (~25 mM) to account for such a change. More dramatically, phosphorylation of ETS1$^{279–440}$ represses DNA binding by an additional ~100-fold, yet the K_D values for the association of WT^{2P*} *versus* WT^{0P*} to ETS1$^{301–440}$ differ by only 10-fold. Although the experimental conditions (temperature, pH, and ionic strength) differed for these measurements, it is evident from such back-of-the-envelope calculations that ETS1 autoinhibition does not arise only from competitive occlusion of the DNA-binding interface and does not scale linearly with the affinity of *trans*-peptides for the ETS domain/IM. One plausible non-steric contribution toward autoinhibition could result from electrostatic repulsion of DNA by the *cis*-SRR, which has a net negative charge that increases upon phosphorylation (Table 1). However, solvent screening would likely reduce such a long range interaction by the conformationally dynamic SRR.

In addition to steric effects, the SRR is also important for an allosteric pathway leading to autoinhibition. Upon DNA binding, the marginally stable helices HI-1 and HI-2 of the IM unfold [9–11]. The IM is distal to the ETS domain DNA-binding interface, and unfolding may be allosterically linked to subtle conformational changes in the HI-2/H1 loop necessary for the N-terminus of helix H1 to precisely interact with the DNA phosphodiester backbone [10,39,40]. Somewhat surprisingly, despite this apparently large helix-coil structural perturbation, the IM reduces the affinity of the ETS1 ETS domain for DNA by only ~2-fold. The appended SRR is necessary for wild-type levels of autoinhibition [10]. Extensive prior studies demonstrated that the SRR stabilizes the IM and dampens motions in the ETS domain, thereby shifting a conformational equilibrium of the ETS1 from a flexible "active" to a rigid "inactive" state [8,12]. Increasing SRR phosphorylation progressively shifts this equilibrium, thus acting as a "dimmer switch" for DNA binding.

Our current studies also provide a structural explanation for the role of the SRR in this allosteric pathway for autoinhibition. In solution, intermolecular NOEs detected between the C-terminal Ala293/Ala294/Leu295 of WT^{2P*} and Leu418/Leu421 in ETS1$^{301–440}$ place the peptide in close proximity to helix H4 of the IM. In crystal structures, these leucines are near Pro281 and ^{5f}Phe283 of the fluorinated peptide. Thus, with either mode of binding, hydrophobic residues in the SRR directly contact helix H4, which in turn abuts HI-1 (Fig. 8). By thermodynamic

linkage, such favorable interactions will stabilize the IM helical bundle against unfolding. Similarly, docking of the SRR along the hydrophobic surface that extends from the IM to helix H3 will stabilize the ETS domain and restrict its conformational dynamics. It is also noteworthy that disruption of the IM by the mutations Y307P in HI-1 or L429A in H5 impairs the repressive effects of the SRR on DNA binding [13]. A folded IM will aid binding of the adjacent SRR to the ETS domain and thereby contribute to the steric pathway for autoinhibition. Conversely, a disrupted IM, with HI-1 and HI-2 unraveled, will place an additional ~30 disordered residues before the ETS domain. This will reduce the local effective concentration of the SRR for competing with DNA binding [36,37].

In summary, the SRR both sterically blocks DNA binding and stabilizes the IM against allosteric unfolding. By synergistically coupling these two mechanisms, the SRR establishes ETS1 autoinhibition at a functionally relevant and regulatable level.

Conserved ETS domain interface

The SRR interacts with the ETS1 ETS domain via an interface that partially overlaps its DNA-binding surface. The general features of this interface, with an exposed hydrophobic region flanked by positively charged residues, are conserved across the entire ETS family. We speculate that the SRR of ETS1 and closely related ETS2 evolved to exploit this interface for the autoinhibitory regulation of these factors via post-translational modifications and protein partnerships [3]. Similarly, the ETV6 repressor is autoinhibited by an amphipathic α-helix that packs against the corresponding interface of its ETS domain and thereby sterically inhibits DNA-binding [4,41]. It is intriguing that the *trans*-SRR peptide also bound *in vitro* to the ETS domain of distantly related PU.1. In a strict sense, this is not biologically relevant as PU.1 lacks any obvious SRR-like sequence and is not known to be autoinhibited or to interact with ETS1. More generally, however, the observation raises the possibility that other transcriptional regulatory proteins with amphipathic sequences similar to the SRR may be capable of modulating, in *trans*, DNA binding by members of the ETS family.

The same conserved interface was also targeted in a computational screen to identify inhibitors of the prostate cancer-linked ERG [42]. Encouragingly, a small molecule identified by this proof-of-principle approach indeed weakly bound the ETS domain of ERG as predicted and competitively inhibited DNA binding. However, the molecule also bound the ETS domains of PU.1 and ETV4, as expected due to their common surface features. Thus, the specific therapeutic intervention of DNA binding by only a subfamily of ETS factors will also require consideration of their distinguishing, non-conserved properties. For example, a small molecule that stabilizes the folded IM or bound SRR of ETS1 or ETS2, or the inhibitory α-helix of ETV6, may selectively reinforce the autoinhibition of only these selected ETS factors. The structural insights provided herein should help guide the discovery and development of such important potential therapeutics.

Materials and Methods

Expression and purification of ETS1$^{301–440}$

Expression and purification of ETS1$^{301–440}$ (residues 301–440 of ETS1) was performed as previously described [15]. For NMR spectroscopy studies, the final ETS1$^{301–440}$ buffer consisted of 20 mM MES, 50–300 mM NaCl, 5 mM DTT, and 0.5 mM EDTA at pH 6.50. For production of crystals, the protein was further purified using an analytical mono-S cation exchange column (GE Healthcare) equilibrated with 20 mM MES, 50 mM NaCl at pH 6.0 and eluted over 10 column volumes with a gradient to 20 mM MES and 1 M NaCl at pH 6.0. The final sample buffers consisted of 10 mM MES, 50–100 mM NaCl, and 1 mM TCEP at pH 6.50.

Expression and purification of PU.1$^{167–272}$

The cDNA encoding PU.1$^{167–272}$, including the core ETS domain and an appended C-terminal helix H4, was cloned into pET28-MHL [43]. This vector contains an N-terminal His_6-affinity tag followed by a TEV cleavage site. The protein was produced by heterologous expression in freshly transformed *Escherichia coli* BL21 (λDE3) cells grown at 37 °C in minimal M9 media supplemented with 1 g/L of $^{15}NH_4Cl$, 1 × trace metal mix [44], and 35 mg/L kanamycin. At OD_{600} ~0.4, the cells were cooled to 30 °C and expression was induced at OD_{600} ~0.6 using a final concentration of 1 mM IPTG. The cells were harvested ~16 h post-induction, and the pellet was frozen at −80 °C until required. The pellet was resuspended in 40 mL of denaturing binding buffer (20 mM sodium phosphate, 500 mM NaCl, 10 mM imidazole, 4 M guanidinium hydrochloride, pH 7.4) per liter of culture. The cell suspension was lysed by sonication with cooling in an ice-water bath. The lysate was cleared by centrifugation and filtering, then applied to a Ni^{+2}-NTA HisTrap HP column (GE Healthcare). Following a wash with 30 mM imidazole, His-tagged PU.1$^{167–272}$ was eluted in one step with elution buffer (20 mM sodium phosphate, 500 mM NaCl, 1 M imidazole, 4 M guanidinium hydrochloride, pH 7.4). The resulting protein sample was dialyzed against 2 L of refolding buffer (50 mM sodium phosphate, 500 mM NaCl, 0.5 mM EDTA, pH 6.50). Soluble protein was separated by

centrifugation and folded His_6-tagged PU.1$^{167–272}$ was further purified using a Superdex-75 gel filtration chromatography column. The final NMR sample buffer consisted of 20 mM MES, 300 mM NaCl, 5 mM DTT, and 0.5 mM EDTA at pH 6.50.

SRR peptides

Peptides corresponding to the SRR of ETS1 (residues 279 to 295; Table 1), were purchased from ABI Scientific at 95% purity. The peptides were modified with N-terminal acetylation and C-terminal amidation to avoid charged termini. The majority of the peptides contained a C-terminal non-native tryptophan residue to facilitate quantification by UV-absorbance spectroscopy using predicted molar absorptivities [45]. The presence of this residue did not significantly change the dissociation constant relative to the strictly wild-type sequence (Table 1), although it slightly increased the CSPs observed in NMR-monitored titrations with ETS1$^{301–440}$ (not shown). For titrations with this protein, the lyophilized peptides were resuspended in NMR buffer (20 mM MES, 50 mM NaCl, 5 mM DTT, 0.5 mM EDTA, pH 6.50). The pH was re-adjusted with NaOH, and the samples were dialyzed against 2 L of NMR buffer at the desired NaCl concentrations (Table 2) using Float-A-Lyzer (Spectrum Labs) dialysis devices with MWCO of 100–500 Da. To ensure buffer matching, both the peptide and the protein samples were dialyzed for 48 h at 4 °C in the same container. For studies on the free SRR peptides, the samples were dialyzed against 2 L of 20 mM sodium phosphate and 50 mM NaCl at pH 6.50 at 4 °C. For crystallography, the $^{5f}Phe^{2P*}$ peptide was dialyzed against 2 L of 10 mM MES, 100 mM NaCl, and 1 mM TCEP at pH 6.50. For ^{31}P-NMR studies, the WT^{2P} peptide was dialyzed against 10 mM MES, and 100 mM NaCl at pH 6.50.

SRR peptide hydrophobicity scale determination

The relative hydrophobic character of each peptide was estimated by considering the additive effect at the four substituted sites and using the scale of Monera *et al.* [46] for the standard amino acids at pH 7.0. These were 41 for Ala, 63 for Tyr, 76.4 for Val, 97 for Trp and Leu, and 100 for Phe. Combined hydrophobicity values of 164, 290, 306, 388, and 400, respectively, were obtained by adding the contributions of each amino acid. These were then normalized to 1 using the predicted octanol–water partition coefficients of Fmoc-L-phenylalanine (PubChem CID: 978331) and Fmoc-pentafluoro-L-phenylalanine (PubChem CID: 7020337). The final hydrophobicity values obtained for the peptide variants were 0.37 (Ala^{2P*}), 0.65 (WT^{2P*}), 0.69 (Val^{2P*}), 0.88 (Leu^{2P*}, Trp^{2P*}), 0.90 (Phe^{2P*}), and 1 ($^{5f}Phe^{2P*}$).

DNA oligonucleotides

The complementary oligonucleotides corresponding to a specific ETS1 binding site [11] spanning 12 base pairs, 5′-CAGCC**GGAA**GTG-3′ and 5′-CAC**TTCC**GGCTG-3′, were purchased from Integrated DNA Technologies. The oligonucleotides were resuspended in NMR sample buffer (20 mM MES, 300 mM NaCl, 5 mM DTT, 0.5 mM EDTA, pH 6.50), mixed in a 1:1 molar ratio based on quantitation by UV-absorbance spectroscopy with predicted molar absorptivities (http://biophysics.idtdna.com/UVSpectrum.html), heated to 95 °C, and slowly cooled to allow duplex DNA annealing. The sample was then passed through a Superdex-75 gel filtration column (GE Healthcare) in NMR sample buffer to remove impurities and single stranded DNA. The purest fractions were concentrated to ~1.3 mM and used for NMR-monitored titrations with 0.16 mM ^{15}N-labeled ETS1$^{301–440}$ in the presence or absence of the WT^{2P} peptide.

NMR spectroscopy

NMR experiments were performed at 28 °C using cryoprobe-equipped Bruker Avance III 600 or 850 MHz spectrometers, unless noted otherwise. The spectra were processed using NMRPipe [47] and NMRFAM-Sparky [48]. The previously published chemical shifts of free ETS1$^{301–440}$ [10] were used as starting points to assign signals of the peptide-bound protein via ^{15}N- and ^{13}C-HSQC monitored titrations. The shifts and intramolecular NOEs of the unlabeled WT^{2P} peptide bound to $^{15}N/^{13}C$-labeled protein (1:1 molar ratio WT^{2P}:ETS1$^{301–440}$, 100 mM NaCl, ~85% saturation) were assigned using a combination of double-filtered 2D $^{1}H–^{1}H$ NOESY (150 ms mixing time) and 2D $^{1}H–^{1}H$ TOCSY spectra [49]. Intermolecular NOEs between the unlabeled peptide and labeled protein (220 μM 1:1 WT^{2P}:ETS1$^{301–440}$ complex, 50 mM NaCl) were measured with a filtered-edited three dimensional $^{1}H–^{15}N/^{13}C$-NOESY spectrum (150 ms mixing time) [50]. This experiment detects NOEs between a ^{1}H directly bonded to a $^{12}C/^{14}N$ (present in the unlabeled WT^{2P} peptide) and a ^{1}H directly bonded to $^{15}N/^{13}C$-labeled ETS1$^{301–440}$, with the latter resolved by the shifts of the bonded $^{15}N/^{13}C$ nucleus. Intramolecular protein NOEs, such as those arising from buried sulfur-bonded $^{1}H^{\gamma}$ of Cys350 and neighboring labeled protons, were excluded by collecting control experiments in the absence of the peptide. Due to spectral overlap, particularly involving the relatively large number of aromatic residues present in both the protein and the peptide (14 and 4, respectively), some NOE signals were either unassigned or assigned ambiguously. The resulting NOE restraints were converted to upper distance limits for structure calculations with CYANA [51].

The chemical shifts of the unbound WT^{2P} peptide (285 μM) were assigned using a combination of two-dimensional natural abundance ^{13}C-HSQC, ^{13}C-HMBC, ^{1}H–^{1}H NOESY, and ^{1}H–^{1}H TOCSY experiments. One-dimensional ^{31}P-NMR and two-dimensional ^{31}P-HSQC spectra of the free WT^{2P} peptide were collected with a Bruker Avance III 500 MHz spectrometer to obtain assignments of the pSer282 and pSer285 ^{31}P signals. In subsequent titrations monitored by ^{31}P-NMR, both peptide and protein were initially 480 μM.

CYANA calculations to model the WT^{2P}/$ETS1^{301–440}$ complex

The model of the WT^{2P} peptide docked on $ETS1^{301–440}$ was generated with CYANA (v. 3.97) [51]. The input data consisted of distance restraints and sequence files. The latter contained 17 residues corresponding to the WT SRR peptide (279–295 with *trans* X-Pro amides and unmodified termini), a linker of 13 "ghost" residues, and 140 residues of the ETS1 ETS domain/IM (301–440). The ghost residues were required for structural calculations of the complex using CYANA, but did not provide any restraints to bias the positioning of the peptide on the ETS domain. The CYANA library was also modified to include phosphoserines [52]. Rather than starting with fixed input coordinates, the ETS domain/IM structure was generated for the complex via a standard CYANA protocol using 3268 upper and lower distance limits that were measured previously to calculate the structural ensemble of unbound $ETS1^{301–440}$ (available from PDB ID: 1R36 and Biological Magnetic Resonance Data Bank entry number 5991) [10]. The tendency of the WT^{2P}/$ETS1^{301–440}$ complex to slowly aggregate realistically precluded recording the numerous experiments necessary to fully assign all NMR signals of the bound protein and to confidently measure sufficient interprotein NOEs to independently determine its structure with acceptable precision. Also, based on small ^{1}H–^{13}C CSPs (Figs. 1 and 2), the tertiary structure of $ETS1^{301–440}$ in solution did not change substantially upon peptide binding. In support of this conclusion, the coordinates of the ETS domains in the $^{5f}Phe^{2P*}$/$ETS1^{301–440}$ complexes determined herein superimposed closely upon those of the unbound domain swapped protein (PDB ID: 1MD0) with main-chain RMSDs <0.4 Å.

A total of 17 ambiguous and 16 unambiguous intermolecular NOE restraints detected between unlabeled WT^{2P} and ^{15}N/^{13}C-labeled $ETS1^{301–440}$ were used in the form of upper distance limits to dock the peptide on the protein (Table S1). Due to the dynamic nature of the complex, the relative intensities of NOEs in the filtered-edited 3D ^{1}H–^{15}N/^{13}C NOESY spectrum were not considered and all protein–peptide restraints were set to an upper distance limit of 6 Å. Also included were 114 intramolecular peptide–peptide NOE restraints, derived from the double-filtered ^{1}H–^{1}H NOESY spectra of the complex and assigned automatically with the CYANA noeassign function.

The top-ranking (lowest-energy) 18 calculated structures were used as the final ensemble model of the protein–peptide complex. No additional energy minimization or refinement with explicit solvent was applied. Only six peptide–protein and/or four interpeptide distance restraints were violated by a mean of 0.2 Å in six or more members of this ensemble. Across the ensemble, the WT^{2P} peptide had 83%, 12%, and 5% of residues with (φ,ψ) dihedral angles in the allowed, generously allowed, and disallowed regions, respectively, of the Ramachandran plot [53]. When the 18 members of the ensemble were superimposed based on the helices and strands of $ETS1^{301–440}$, the corresponding mainchain atoms of residues 282–295 of the WT^{2P} peptide had pairwise RMSDs of 3.0 ± 0.7 Å.

CSP analysis and dissociation constant determination

Labeled $ETS1^{301–440}$ and $PU.1^{167–272}$, and unlabeled SRR peptide samples were prepared in 20 mM MES, 300 mM NaCl (unless specified), 5 mM DTT, and 0.5 mM EDTA at pH 6.50. The protein and peptide samples were typically concentrated to ~250 μM and ~2 mM, respectively. The protein and peptide concentrations were calculated by measuring UV-absorbance at 280 nm under native conditions and using the following predicted molar absorptivities [45]: $\varepsilon = 35{,}410\ M^{-1}\ cm^{-1}$ ($ETS1^{301–440}$), $\varepsilon = 4470\ M^{-1}\ cm^{-1}$ (WT^{2P}), $\varepsilon = 9970$ (WT^{2P*}, WT^{0P*}), and $\varepsilon = 5500\ M^{-1}\ cm^{-1}$ ($^{5f}Phe^{2P*}$, Val^{2P*}, Leu^{2P*}, Ala^{2P*}, and Phe^{2P*}). The peptides were added in small increments to $ETS1^{301–440}$ and ^{15}N-HSQC and/or ^{13}C-HSQC spectra were recorded at each titration point. Chemical shifts changed co-linearly with increasing peptide concentration and, for the most part, occurred in fast exchange regime on the NMR timescale. The ^{1}H–^{15}N and ^{1}H–^{13}C CSPs were calculated according to Eqs. (1) and (2), respectively [54].

$$^{1}H–^{15}N\ CSP = \left[(0.14\Delta\delta_N)^2 + (\Delta\delta_H)^2\right]^{1/2} \quad (1)$$

$$^{1}H–^{13}C\ CSP = \left[(0.3\Delta\delta_C)^2 + (\Delta\delta_H)^2\right]^{1/2} \quad (2)$$

$\Delta\delta_N$, $\Delta\delta_C$, and $\Delta\delta_H$ are the changes in ^{15}N, ^{13}C, and ^{1}H chemical shifts (ppm), respectively. For each titration, 10 residues exhibiting the largest ^{1}H–^{15}N CSPs and in fast exchange were fit to a

1:1 binding isotherm (Eq. (3)) using GraphPad Prism in order to obtain the equilibrium dissociation constant (K_D values).

$$\Delta\delta_i = \Delta\delta_{sat}\left(\left([P]_{T,i} + [SRR]_{T,i} + K_D\right) - \sqrt{\left([P]_{T,i} + [SRR]_{T,i} + K_D\right)^2 - 4[P]_{T,i}[SRR]_{T,i}}\right)\Big/\left(2[P]_{T,i}\right) \quad (3)$$

$[P]_{T,i}$ and $[SRR]_{T,i}$ are the total concentrations of labeled protein and unlabeled SRR peptides adjusted for dilution effects, respectively, at each point *i*, and $\Delta\delta_{sat}$ is the extrapolated CSP at saturation. The 10 fit K_D values for each titration were averaged and the mean values with standard deviations, as an indication of precision, are reported in Table 1.

Crystallization and structure determination of the $^{5f}Phe^{2P*}$/ETS1$^{301–440}$ complex

Purified ETS1$^{301–440}$ (0.45 mM) was mixed with the $^{5f}Phe^{2P*}$ peptide at a 1:1.2 molar ratio in 10 mM MES, 75–100 mM NaCl, and 1 mM TCEP at pH 6.50. Crystals of the complex grew within 3 days at 4 °C using the sitting drop vapor diffusion method for which 2 μL of complex solution were mixed with 2 μL reservoir solution containing 100 mM HEPES, 0.2 M Li_2SO_4, and 16%–26% PEG 3350 at pH 7.1–8.9. Two distinct crystal morphologies were initially observed: rhomboid (space group $P3_221$ at pH 7.1–8.5) and needle-like (space group $P4_32_12$ at pH 8.5–8.9). With optimization, diffraction quality rhomboid crystals were obtained from 100 mM MES (pH 7.3), 24% PEG 3350, 0.2 M Li_2SO_4, and 4.5% ethylene glycol. Needle morphology crystals from 100 mM MES (pH 8.5), 18% PEG 3350, and 0.2 M Li_2SO_4 were also suitable for data collection. Cryo-protection was achieved by soaking the crystals with a gradient of PEG 3350 at 20%, 25%, and 30% to 35%, while maintaining the concentrations of the remaining components constant. The crystals were flash-cooled in liquid nitrogen and stored for ~5 days prior to data collection.

X-ray diffraction data sets were collected at the Canadian Light Source on beamline 08B1–1 or the Stanford Synchrotron Radiation Lightsource on beamline 9–2. Each native data set was collected from a single crystal at 100 K using a wavelength of 0.98 Å. Diffraction data were indexed, integrated and scaled with the iMosflm [55] and Aimless [56] programs. The structures of the complexes were solved by molecular replacement with Phaser-MR [57] using the published coordinates of an ETS1 domain-swapped dimer (PDB ID: 1MD0). The initial models served as starting points for refinement with phenix.refine [58,59] and manual model building with Coot [60]. The $^{5f}Phe^{2P*}$ peptide was manually built into omit difference maps using the positions of aromatic rings as initial reference points. In the $^{5f}Phe^{2P*}$/ETS1$^{301–440}$ (pH 7.3) structure, residue 301 at the N-terminus (chain A) and 439–440 at the C-terminus (chains A and B) of the protein were not included in the model. Also, Lys383, Glu427, and Tyr397 were modeled as Ala. For the $^{5f}Phe^{2P*}$/ETS1$^{301–440}$ (pH 8.5) structure, residues 301 at the N-terminus (chains A-D) and 437–440 at the C-terminus (chains A–D) were not modeled. Arg409 (chain A), Lys364 (chain C), and Arg394 (chain D) were modeled as Ala. Data collection and refinement statistics are presented in Table S2.

Accession numbers

Atomic coordinates and structure factors for the peptide–protein crystal structures have been deposited with the Protein Data bank under accession numbers 6DA1 (pH 7.3) and 6DAT (pH 8.5).

Acknowledgments

We thank our UBC colleagues Geneviève Desjardins, Jörg Gsponer, Jennifer Bui, and Florian Heinkel for fruitful discussions; Anson Chan and Jason Grigg for technical advice; and the laboratory of Natalie Strynadka for access to their X-ray crystallography instrumentation.

Funding

This work was funded by the Canadian Cancer Society Research Institute (2011-700772 to L.P.M.) and the Canadian Institutes of Health Research (CIHR; MOP-136834 to L.P.M. and MOP-49597 to M.E.P.M.). C.P.B. was supported through the Genome Sciences and Technology Program by a Natural Sciences and Engineering Research Council (NSERC) of Canada CREATE grant; C.S. L., through the International Tuition Award (UBC, Canada) and the Council of Indigenous Peoples (Taiwan); and K.S., through a CIHR Studentship. Instrument funding was provided by the Canada Foundation for Innovation (CFI), the British Columbia Knowledge Development Fund, the UBC Blusson Fund, and the Michael Smith Foundation for Health Research. Research described in this paper was performed using beamline 08B1-1 at the Canadian Light Source, which is supported by CFI, NSERC, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and CIHR. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory,

is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is funded by the DOE Office of Biological and Environmental Research and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393).

Declarations of Interest: None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2018.12.011.

Received 28 October 2018;
Received in revised form 14 December 2018;
Accepted 19 December 2018
Available online 28 December 2018

Keywords:
intrinsically disordered regions;
fuzzy complexes;
protein structure and dynamics;
NMR spectroscopy;
X-ray crystallography

Abbreviations used:
^{5f}Phe, 2,3,4,5,6-pentafluoro-L-phenylalanine; CaMKII, calmodulin-dependent kinase II; CSP, chemical shift perturbation; HSQC, heteronuclear single quantum correlation; IM, inhibitory module; NOESY, nuclear Overhauser effect spectroscopy; SRR, serine-rich region; TOCSY, total correlation spectroscopy.

References

[1] J. Liu, N.B. Perumal, C.J. Oldfield, E.W. Su, V.N. Uversky, A.K. Dunker, Intrinsic disorder in transcription factors, Biochemistry 45 (2006) 6873–6888.
[2] A. Bah, J.D. Forman-Kay, Modulation of intrinsically disordered protein function by post-translational modifications, J. Biol. Chem. 291 (2016) 6696–6705.
[3] P.C. Hollenhorst, L.P. McIntosh, B.J. Graves, Genomic and biochemical insights into the specificity of ETS transcription factors, Annu. Rev. Biochem. 80 (2011) 437–471.
[4] H.J. Coyne, S. De, M. Okon, S.M. Green, N. Bhachech, B.J. Graves, L.P. McIntosh, Autoinhibition of ETV6 (TEL) DNA binding: appended helices sterically block the ETS domain, J. Mol. Biol. 421 (2012) 67–84.
[5] M.C. Regan, P.S. Horanyi, E.E. Pryor Jr., J.L. Sarver, D.S. Cafiso, J.H. Bushweller, Structural and dynamic studies of the transcription factor ERG reveal DNA binding is allosterically autoinhibited, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 13374–13379.
[6] J.A. Newman, C.D. Cooper, H. Aitkenhead, O. Gileadi, Structural insights into the autoregulation and cooperativity of the human transcription factor ETS-2, J. Biol. Chem. 290 (2015) 8539–8549.
[7] S.L. Currie, D.K.W. Lau, J.J. Doane, F.G. Whitby, M. Okon, L.P. McIntosh, B.J. Graves, Structured and disordered regions cooperatively mediate DNA-binding autoinhibition of ETS factors ETV1, ETV4 and ETV5, Nucleic Acids Res. 45 (2017) 2223–2241.
[8] G.M. Lee, M.A. Pufall, C.A. Meeker, H.S. Kang, B.J. Graves, L.P. McIntosh, The affinity of Ets-1 for DNA is modulated by phosphorylation through transient interactions of an unstructured region, J. Mol. Biol. 382 (2008) 1014–1030.
[9] J.M. Petersen, J.J. Skalicky, L.W. Donaldson, L.P. McIntosh, T. Alber, B.J. Graves, Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix, Science 269 (1995) 1866–1869.
[10] G.M. Lee, L.W. Donaldson, M.A. Pufall, H.S. Kang, I. Pot, B.J. Graves, L.P. McIntosh, The structural and dynamic basis of Ets-1 DNA binding autoinhibition, J. Biol. Chem. 280 (2005) 7088–7099.
[11] G. Desjardins, M. Okon, B.J. Graves, L.P. McIntosh, Conformational dynamics and the binding of specific and nonspecific DNA by the autoinhibited transcription factor Ets-1, Biochemistry 55 (2016) 4105–4118.
[12] M.A. Pufall, G.M. Lee, M.L. Nelson, H.S. Kang, A. Velyvis, L.E. Kay, L.P. McIntosh, B.J. Graves, Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region, Science 309 (2005) 142–145.
[13] D.O. Cowley, B.J. Graves, Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition, Genes Dev. 14 (2000) 366–376.
[14] H. Liu, T. Grundstrom, Calcium regulation of GM-CSF by calmodulin-dependent kinase ii phosphorylation of Ets1, Mol. Biol. Cell 13 (2002) 4497–4507.
[15] G. Desjardins, C.A. Meeker, N. Bhachech, S.L. Currie, M. Okon, B.J. Graves, L.P. McIntosh, 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 (2014) 11019–11024.
[16] I.R. Kleckner, M.P. Foster, An introduction to NMR-based approaches for measuring protein dynamics, Biochim. Biophys. Acta 1814 (2011) 942–968.
[17] P.L. Privalov, A.I. Dragan, C. Crane-Robinson, Interpreting protein/DNA interactions: distinguishing specific from nonspecific and electrostatic from non-electrostatic components, Nucleic Acids Res. 39 (2011) 2483–2491.
[18] E.A. Bienkiewicz, K.J. Lumb, Random-coil chemical shifts of phosphorylated amino acids, J. Biomol. NMR 15 (1999) 203–206.
[19] C.W. Garvie, M.A. Pufall, B.J. Graves, C. Wolberger, Structural analysis of the autoinhibition of Ets-1 and its role in protein partnerships, J. Biol. Chem. 277 (2002) 45529–45536.
[20] L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R.F. Boyko, B.D. Sykes, et al., VADAR: a web server for quantitative evaluation of protein structure quality, Nucleic Acids Res. 31 (2003) 3316–3319.
[21] C.W. Garvie, J. Hagman, C. Wolberger, Structural studies of Ets-1/Pax5 complex formation on DNA, Mol. Cell 8 (2001) 1267–1276.
[22] M.A. Pufall, B.J. Graves, Ets-1 flips for new partner Pax-5, Structure 10 (2002) 11–14.
[23] S. Wang, M.H. Linde, M. Munde, V.D. Carvalho, W.D. Wilson, G.M. Poon, Mechanistic heterogeneity in site recognition by the structurally homologous DNA-binding domains of the ETS

family transcription factors ETS-1 and PU.1, J. Biol. Chem. 289 (2014) 21605–21616.
[24] E.W. Scott, M.C. Simon, J. Anastasi, H. Singh, Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages, Science 265 (1994) 1573–1577.
[25] R. Kodandapani, F. Pio, C.Z. Ni, G. Piccialli, M. Klemsz, S. McKercher, R.A. Makim, K.R. Ely, A new pattern for helix-turn-helix recognition revealed by the PE.1 ETS-domain-DNA complex, Nature 380 (1996) 456–460.
[26] G.B. McGaughey, M. Gagne, A.K. Rappe, Pi-stacking interactions. Alive and well in proteins, J. Biol. Chem. 273 (1998) 15458–15463.
[27] M.R. Davis, D.A. Dougherty, Cation-pi interactions: computational analyses of the aromatic box motif and the fluorination strategy for experimental evaluation, Phys. Chem. Chem. Phys. 17 (2015) 29262–29270.
[28] G. Desjardins, Structural characterization of DNA binding and autoinhibition by the ETS1 transcription factor, (PhD thesis) University of British Columbia, 2015.
[29] K. Kasahara, M. Shiina, J. Higo, K. Ogata, H. Nakamura, Phosphorylation of an intrinsically disordered region of Ets1 shifts a multi-modal interaction ensemble to an auto-inhibitory state, Nucleic Acids Res. 46 (2018) 2243–2251.
[30] G. Lu, Q. Zhang, Y. Huang, J. Song, R. Tomaino, T. Ehrenberger, E. Lim, W. Liu, R.T. Bronson, M. Bowden, J. Brock, I.E. Krop, D.A. Dillon, S.P. Gygi, G.B. Mills, A.L. Richardson, S. Signoretti, M.B. Yaffe, W.G. Kaelin Jr., Phosphorylation of Ets1 by Src family kinases prevents its recognition by the COP1 tumor suppressor, Cancer Cell 26 (2014) 222–234.
[31] R.L. Baldwin, How Hofmeister ion interactions affect protein stability, Biophys. J. 71 (1996) 2056–2063.
[32] P.H. von Hippel, T. Schleich, Ion effects on the solution structure of biological macromolecules, Acc. Chem. Res. 2 (1968) 257–265.
[33] C.D. Ma, C. Wang, C. Acevedo-Velez, S.H. Gellman, N.L. Abbott, Modulation of hydrophobic interactions by proximally immobilized ions, Nature 517 (2015) 347–350.
[34] M. Fuxreiter, Fuzziness: linking regulation to protein dynamics, Mol. BioSyst. 8 (2012) 168–177.
[35] M.I. Page, W.P. Jencks, Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect, Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 1678–1683.
[36] V.M. Krishnamurthy, V. Semetey, P.J. Bracher, N. Shen, G.M. Whitesides, Dependence of effective molarity on linker length for an intramolecular protein-ligand system, J. Am. Chem. Soc. 129 (2007) 1312–1320.
[37] D.P. Goldenberg, Computational simulation of the statistical properties of unfolded proteins, J. Mol. Biol. 326 (2003) 1615–1633.
[38] J. Kuriyan, B. Konforti, D. Wemmer, The Molecules of Life: Physical and Chemical Principles, Garland Science, New York, 2013.
[39] H. Kamberaj, A. van der Vaart, Correlated motions and interactions at the onset of the DNA-induced partial unfolding of Ets-1, Biophys. J. 96 (2009) 1307–1317.
[40] K. Kasahara, M. Shiina, I. Fukuda, K. Ogata, H. Nakamura, Molecular mechanisms of cooperative binding of transcription factors Runx1-CBFbeta-Ets1 on the TCRalpha gene enhancer, PLoS One 12 (2017), e0172654.
[41] S. De, M. Okon, B.J. Graves, L.P. McIntosh, Autoinhibition of ETV6 DNA binding is established by the stability of its inhibitory helix, J. Mol. Biol. 428 (2016) 1515–1530.
[42] M.S. Butler, M. Roshan-Moniri, M. Hsing, D. Lau, A. Kim, P. Yen, M. Mroczek, M. Nouri, S. Lien, P. Axerio-Cilles, K. Dalal, C. Yau, F. Ghaidi, T. Yamazaki, S. Lawn, M.E. Gleave, C.Y. Gregory-Evans, L.P. McIntosh, M.E. Cox, R.S. Rennie, A. Cherkasov, Discovery and characterization of small molecules targeting the DNA-binding ETS domain of ERG in prostate cancer, Oncotarget 8 (2017) 42438–42454.
[43] D.K.W. Lau, Understanding ETS transcription factors: from ordered domains to disordered sequences, (PhD thesis) University Of British Columbia, 2017.
[44] F.W. Studier, Protein production by auto-induction in high-density shaking cultures, Protein Expr. Purif. 41 (2005) 207–234.
[45] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPasy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31 (2003) 3784–3788.
[46] O.D. Monera, T.J. Sereda, N.E. Zhou, C.M. Kay, R.S. Hodges, Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix, J. Pept. Sci. 1 (1995) 319–329.
[47] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRpipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.
[48] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-Sparky: enhanced software for biomolecular NMR spectroscopy, Bioinformatics 31 (2015) 1325–1327.
[49] J. Iwahara, J.M. Wojciak, R.T. Clubb, Improved NMRspectra of a protein-DNA complex through rational mutagenesis and the application of a sensitivity optimized isotope-filtered noesy experiment, J. Biomol. NMR 19 (2001) 231–241.
[50] C. Zwahlen, P. Legault, S.J.F. Vincent, J. Greenblatt, R. Konrat, L.E. Kay, Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage λ N-peptide/boxB RNA complex, J. Am. Chem. Soc. 119 (1997) 6711–6721.
[51] P. Guntert, Automated NMR structure calculation with CYANA, Methods Mol. Biol. 278 (2004) 353–378.
[52] J.W. Craft Jr., G.B. Legge, An AMBER/DYANA/MolMol phosphorylated amino acid library set and incorporation into NMR structure calculations, J. Biomol. NMR 33 (2005) 15–24.
[53] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, ProCheck—a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[54] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, Prog. Nucl. Magn. Reson. Spectrosc. 73 (2013) 1–16.
[55] T.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G. Leslie, iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 271–281.
[56] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242.
[57] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.

[58] P.V. Afonine, R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, T.C. Terwilliger, A. Urzhumtsev, P.H. Zwart, P.D. Adams, Towards automated crystallographic structure refinement with Phenix.Refine, Acta Crystallogr. D Biol. Crystallogr. 68 (2012) 352–367.

[59] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Gross-Kunstieve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: A comprehensive python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221.

[60] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of COOT, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501.