Published in final edited form as:

Biochemistry. 2016 July 26; 55(29): 4105–4118. doi:10.1021/acs.biochem.6b00460.

Conformational Dynamics and the Binding of Specific and Nonspecific DNA by the Autoinhibited Transcription Factor Ets-1

Geneviève Desjardins^{†,||}, Mark Okon[†], Barbara J. Graves^{‡,§}, and Lawrence P. McIntosh^{*,†}

[†]Department of Biochemistry and Molecular Biology, Department of Chemistry, and Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada

[‡]Department of Oncological Sciences, University of Utah School of Medicine, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah 84112-5550, United States

[§]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, United States

Abstract

The affinity of the Ets-1 transcription factor for DNA is autoinhibited by an intrinsically disordered serine-rich region (SRR) and a helical inhibitory module (IM) appended to its winged helix–turn–helix ETS domain. Using NMR spectroscopy, we investigated how Ets-1 recognizes specific versus nonspecific DNA, with a focus on the roles of protein dynamics and autoinhibition in these processes. Upon binding either DNA, the two marginally stable N-terminal helices of the IM predominantly unfold, but still sample partially ordered conformations. Also, on the basis of amide chemical shift perturbation mapping, Ets-1 associates with both specific and nonspecific DNA through the same canonical ETS domain interface. These interactions are structurally independent of the SRR, and thus autoinhibition does not impart DNA-binding specificity. However, relative to the pronounced NMR spectroscopic changes in Ets-1 resulting from specific DNA binding, the spectra of the nonspecific DNA complexes showed conformational exchange broadening and lacked several diagnostic amide and indole signals attributable to hydrogen bonding interactions seen in reported X-ray crystallographic structures of this transcription factor with its cognate DNA sequences. Such differences are highlighted by the chemical shift and relaxation properties of several interfacial lysine and arginine side chains. Collectively, these data support a general model in which Ets-1 interacts with nonspecific DNA via dynamic electrostatic interactions, whereas hydrogen bonding drives the formation of well-ordered complexes with specific DNA.

Graphical abstract

*Corresponding Author: mcintosh@chem.ubc.ca. Phone: 604-822-3341.

||Present Address: (G.D.) Zymeworks Inc., Vancouver BC, V6H 3 V9, Canada.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b00460.

Site-size assays, assigned NMR spectra, detailed NMR spectral comparisons, and NMR relaxation data (PDF)

The authors declare no competing financial interest.

A long-standing challenge in the field of gene expression is to understand how sequence-specific transcription factors rapidly search through a vast background of nonspecific genomic DNA, for which they typically have micromolar affinity, in order to bind a small number of specific cognate target sites with sub-nanomolar affinity.^{1,2} It is generally accepted that nonspecific DNA both buffers the concentrations of free transcription factors available for specific DNA binding and aids in the search process via facilitated diffusion.^{3,4} That is, the proteins undergo one-dimensional sliding along nonspecific DNA, combined with three-dimensional jumping (dissociation followed by reassociation) and direct intersegment transfer, to move from one region of DNA to another distal region. The first process enables rapid scanning of uncondensed DNA in a localized manner, whereas the latter allows a more coarse-scale search in order to avoid being trapped in areas of the chromosome that lack accessible cognate sites.

In this study, we explored how Ets-1 recognizes specific versus nonspecific DNA and investigated the roles of protein dynamics and autoinhibition in these processes. Ets-1 is a well-characterized transcription factor belonging to the ETS family of eukaryotic transcription factors. All members of this family contain a conserved winged helix–turn–helix (wHTH) ETS domain that binds DNA with a consensus 5′-GGA(A/T)-3′ sequence.⁵ The activity of Ets-1 is regulated by a variety of signaling events, such as post-translational modifications and partner–protein interactions.⁶ At the nexus of this regulation is the autoinhibition of DNA binding, which involves a helical inhibitory module (IM) that packs along helix H1 of the ETS domain and is distal to the helix(H2)–turn–helix(H3) DNA-recognition interface.

Prior studies have shown that the marginally stable IM of Ets-1 is allosterically disrupted upon DNA binding.⁷ However, the IM alone imparts only ~2-fold inhibition, and an adjacent intrinsically disordered serine-rich region (SRR) is required for the full ~20-fold effect exhibited by native Ets-1.^{8,9} Through transient interactions, the SRR plays coupled steric and allosteric roles by occluding the ETS domain DNA-binding interface and by stabilizing the IM and reducing motions of the ETS domain, respectively.^{8,10} In particular, the DNA recognition helix H3 undergoes fluctuations detectable by amide hydrogen exchange (HX), and these fluctuations dampen with increasing auto-inhibition due to progressive multisite SRR phosphorylation by calmodulin-dependent kinase II (CaMKII) in response to calcium signaling.^{8,9,11} Aside from perturbation of IM helices, the static crystal structures of the Ets-1 ETS domain in its free and specific DNA-bound forms are highly similar.¹² This suggests that any mean conformational changes in the ETS domain are subtle and that flexibility may play an important role in the interactions of Ets-1 with both nonspecific and specific DNA.

Using NMR spectroscopy, we characterized the structural and dynamic features of autoinhibited Ets-1 fragments bound to 12 bp specific and nonspecific oligonucleotides. In this context, specific refers to DNA bearing a consensus ETS recognition motif, whereas nonspecific refers to any unrelated sequence. Upon binding either DNA, the N-terminal sequences forming both inhibitory helices HI-1 and HI-2 of the Ets-1 IM become predominantly unfolded, yet still transiently sample ordered conformations. However, the structural and dynamic changes observed in the backbone and arginine/ lysine side chains of the core ETS domain are much larger with specific DNA than with nonspecific DNA. This is consistent with the general observation that sequence-specific transcription factors bind nonspecific sequences via “loose” dynamic electrostatic interactions, whereas their consensus cognate sites are recognized through high-affinity interactions including hydrogen bonding between side chains and DNA base pairs.¹³ This work thus broadens our understanding of DNA binding autoinhibition and sequence recognition by Ets-1.

MATERIALS AND METHODS

Protein Expression and Purification

The plasmid encoding residues 279–440 of murine Ets-1 (denoted as N279; Figure 1A and Table 1) with a thrombin-cleavable N-terminal His₆-tag was obtained by ligating a synthetic, codon-optimized gene (Blue Heron Biotech, USA) into the pET28a vector via *Bam*HI/*Nco*I restriction sites.¹⁰ The nontagged N301 construct (residues 301–440) was then generated by PCR from this optimized sequence and inserted into pET28a with the same restriction sites. Codon optimization led to only modest improvements in protein yields.

The N279 and N301 constructs were expressed in *Escherichia coli* HMS174 (λ DE3). To produce full protonated protein, the bacteria were grown in 1 L of M9 minimal media at 37 °C to OD₆₀₀ = 0.8, induced with 0.4 mM isopropyl- β -D-thiogalactopyranoside (IPTG), and then grown at 30 °C for an additional 4 h. Kanamycin (40 μ g/mL) was used as the selection antibiotic, and the M9 minimal media was supplemented with 1 g/L (¹⁵N, 99%)-NH₄Cl for uniform ¹⁵N-labeling or 1 g/L (¹⁵N, 99%)-NH₄Cl and 3 g/L (¹³C₆, 99%)-D-glucose for uniform ¹³C/¹⁵N-labeling (Sigma-Aldrich or Cambridge Isotope Laboratories). For 70% partial deuteration, the above M9 media were prepared with 70% D₂O/30% H₂O v/v. For full deuteration, the M9 minimal media was prepared with 99% D₂O (Cambridge Isotope Laboratories) and contained 1 g/L (¹⁵N, 99%)-NH₄Cl for uniform ²H/¹⁵N-labeling and 1 g/L (¹⁵N, 99%)-NH₄Cl and 3 g/L (²H₇/¹³C₆, 99%)-D-glucose for uniform ²H/¹³C/¹⁵N-labeling. Initially, the M9 salts and ¹⁵NH₄Cl were dissolved in a small amount of 99% D₂O, lyophilized to remove residual protons, and then added to the remainder of the media in 1.1 L of 99% D₂O. For both partial and full deuteration, an initial bacterial culture was grown at 37 °C in 5 mL of LB media until OD₆₀₀ = 2, pelleted by centrifugation, resuspended in prewarmed 0.1 L deuterated M9 media, and then grown until OD₆₀₀ = 1. Cells were centrifuged again and resuspended in 1 L of fresh prewarmed deuterated M9 media. The cultures were grown at 37 °C until OD₆₀₀ = 0.3–0.5, induced with 0.4 mM IPTG, and incubated at 30 °C for an additional 8–16 h. Cells were centrifuged and the protein was purified in H₂O to exchange all labile deuterons, including those on the amide nitrogens, to protons.

Harvested cells containing N301 or N279 were resuspended, respectively, in lysis buffer A (50 mM sodium citrate pH 5.3, 50 mM NaCl, 0.5 mM EDTA, and 1 mM TCEP) or B (20 mM sodium phosphate pH 7.8, 500 mM NaCl, 10 mM imidazole, 0.5 mM EDTA, and 2 mM DTT) with protease inhibitor cocktail tablets (200 μ L of 1 tablet resuspended in 1 mL of water; Roche). The cells were lysed by high pressure homogenization and sonication.

N279 was purified by Ni²⁺-affinity chromatography (5 mL HisTrap HP; GE Healthcare) using an AKTA Prime FPLC. The protein was eluted with an imidazole gradient from 0 to 500 mM in buffer B. Fractions containing N279 were identified by 15% SDS-PAGE, pooled, and dialyzed overnight in cleavage buffer (25 mM Tris pH 7.9, 1 mM EDTA, 10% glycerol, 300 mM NaCl, and 1 mM TCEP) with thrombin (1 U/mg of N279; Novagen). The sample was then subjected to Ni²⁺-affinity chromatography in order to remove the cleaved affinity tag and any uncleaved protein. The protein was concentrated with a 3-kDa molecular mass cutoff (MWCO) Amicon centrifugal filtration device and loaded on a Superdex 75 gel filtration column (GE Healthcare) equilibrated with NMR sample buffer (20 mM MES pH 6.5, 50 mM NaCl, 0.5 mM EDTA, 0.02% NaN₃, and 5 mM DTT). Fractions containing purified N279 were identified by 15% SDS-PAGE, pooled and concentrated to 0.1–0.45 mM for NMR analysis. The N279 protein, with a non-native N-terminal Gly-Ser-His remaining after affinity tag cleavage, has an expected mass of 19 088 Da, and its concentration was determined by absorbance spectroscopy using a predicted $\epsilon_{280} = 39\,880\text{ M}^{-1}\text{ cm}^{-1}$ value.^{14,15}

The N301-containing lysate was loaded on a DEAE column attached in tandem to a strong cation exchange column (SP Sepharose; GE Healthcare) and eluted with an NaCl gradient from 0 to 1 M in buffer A. The fractions containing N301 were pooled and purified further by gel filtration, as described for N279. The N301 protein, with a non-native N-terminal Gly, has an expected mass of 16 300 Da, and its concentration was determined using a predicted $\epsilon_{280} = 35\,410\text{ M}^{-1}\text{ cm}^{-1}$ value.^{14,15}

Duplex Oligonucleotide Preparation

All DNA oligonucleotides were chemically synthesized and purchased from Integrated DNA Technologies (IDT). The corresponding complementary oligonucleotides were resuspended in water, mixed in a 1:1 ratio as determined by absorbance at 260 nm using predicted single strand molar absorptivities ϵ_{260} , and annealed by heating to 95 °C followed by slow cooling (–0.5 °C/min) to favor duplex formation. The duplex DNAs were then purified and exchanged into NMR sample buffer using a Sephadex S75 gel filtration column, followed by concentration with a 3-kDa MWCO Amicon centrifugal filtration device. Final duplex DNA concentrations were determined by absorbance at 260 nm using predicted double strand molar absorptivities ϵ_{260} . The predicted ϵ_{260} values were provided by IDT Technologies¹⁶ and are tabulated in reference.¹⁴

NMR-Monitored DNA Titrations

Small aliquots of the DNA oligonucleotides (initially 0.71–1.12 mM) were titrated into uniformly ¹⁵N-labeled N301 (initially 0.25 mM in 450 μ L) with increasing DNA/protein ratios of 0:1, 0.25:1, 0.5:1, 0.75:1, 1:1, 2:1, 3:1, and varying final ratios depending on the

oligonucleotide. The final DNA/protein molar ratios were 1.5:1 (DNA^{SP9} and DNA^{SP12}), 3.3:1 (DNA^{NSP9}), 3.8:1 (DNA^{NSP15}), and 5:1 (DNA^{NSP12}). In the case of N279, DNA^{SP14} was added progressively to a final 2:1 ratio. The titrations were monitored by recording sensitivity-enhanced ¹⁵N-HSQC spectra¹⁷ at 28 °C with a cryoprobe-equipped 850 MHz Bruker Avance III spectrometer. The number of acquisitions was increased with progressive protein dilution. Amide chemical shift perturbations (CSPs) were calculated as $\delta = \{ (0.2 \delta^N)^2 + (\delta^H)^2 \}^{1/2}$.

NMR Experiments for Spectral Assignments of the DNA/Protein Complexes

NMR spectral assignments of the Ets-1/DNA complexes were obtained using 300 μ L samples of N301 (0.2–0.45 mM) with a DNA/protein ratio of 1.1:1 for DNA^{SP12} and 5:1 for DNA^{NSP12} to ensure saturation. The complexes were in NMR sample buffer (20 mM MES, 50 mM NaCl, 0.5 mM EDTA, 0.02% NaN₃, and 5 mM DTT at pH 6.5). Data were recorded at 10, 13, 28, or 31 °C on cryoprobe-equipped 500, 600 and 850 MHz Bruker Avance III spectrometers. The resulting spectra were processed and analyzed using NMRpipe¹⁸ and Sparky.¹⁹ Signals from backbone and side chain ¹H, ¹³C, and ¹⁵N nuclei were assigned through extensive multidimensional heteronuclear experiments including conventional and TROSY versions of 3D HNCO, HNCA, HN(CA)CO, HN(CO)CA, HNCACB, HN(CO)CACB, C(C)-TOCSY-NH, and ¹⁵N-NOESY-HSQC, and 4D ¹⁵N-HSQC-NOESY-HSQC.^{20,21} Signals from arginine side chain nuclei were assigned with 2D HNCD and H₂CN experiments,²² combined with 2D C(C)-TOCSY-N^H and 3D H(CC)-TOCSY-N^H measurements.²³

Relaxation Measurements

Independent sets of ¹⁵N *T*₁, *T*₂, and heteronuclear ¹H{¹⁵N}-NOE data were collected using a 500 MHz spectrometer with the ¹⁵N transmitter centered on the backbone amides (119 ppm) or the arginine side chains (80 ppm).²⁴ The sample temperature was 31 °C for backbone amides and 13 °C for arginine side chains. Steady-state heteronuclear ¹H{¹⁵N}-NOE spectra were acquired with and without 3 s of ¹H saturation and a total recycle delay of 5 s. Data were processed with NMRpipe¹⁸ using the nlinLS routine to extract peak intensities and then fit as single exponential decays with Matlab (The MathWorks, Inc.). Global tumbling correlation times were calculated with Tensor2²⁵ using amide *R*₁ and *R*₂ values for all residues that had a ¹H{¹⁵N}-NOE ratio greater than 0.6.

RESULTS

Selection of Specific and Nonspecific Oligonucleotides

We first sought to find protein–oligonucleotide combinations that encompassed the salient features of the specific and nonspecific Ets-1/DNA complexes, while optimizing their behavior as required for detailed NMR spectroscopic characterization. Early DNAase I and hydroxyl radical protection studies, combined with ethylation interference assays, indicated that ETS factors bind DNA with a footprint spanning 12–15 bp.²⁶ A site-size selection experiment with a pool of truncated DNAs also demonstrated that maximum binding affinity requires 13–14 bp (Figure S1). However, the intermolecular contacts observed in the X-ray crystallographic structural studies of numerous ETS domain/DNA complexes involve only

~9 bp.^{6,27} Consistent with this observation, the high-affinity sites identified for most ETS factors correspond to a 9 bp consensus 5' ACCGGAAGT3' sequence with the invariant GGA(A/T) core flanked asymmetrically by more variable bases.⁵ Given these differences in defining the ETS domain binding site, we initially examined the interaction of two Ets-1 fragments with a series of duplex oligonucleotides (Figure 1A and Table 1).

The titrations of ¹⁵N-labeled N301 or N279 with 9, 12, and 14 bp specific DNAs corresponding to a consensus ETS recognition sequence²⁶ were monitored via ¹⁵N-HSQC spectra (Figure S2). As expected from the known high affinity of Ets-1 for its cognate sites ($K_D \approx 25$ pM with N301 and 400 pM with N279),⁹ the complexes formed in the slow exchange regime on the chemical shift time scale²⁸ and reached saturation at 1:1 DNA/protein ratios (not shown). Furthermore, binding of the three DNAs induced similar overall chemical shift perturbations (CSPs) in the ¹⁵N-HSQC spectra of the proteins. However, the magnitudes of the CSPs for many amides and indoles were slightly smaller with DNA^{SP9} than with the longer oligonucleotides. Thus, consistent with the site—site selection studies, the 9 bp contacts observed in crystal structures represent the near, but not full, DNA-binding interface of the ETS domain. Consequently, DNA^{SP12} was chosen as the optimal oligonucleotide that recapitulated the entire complement of interactions with Ets-1, while also having a minimal size for favorable NMR spectroscopic behavior.

Titrations of ¹⁵N-labeled N301 with three nonspecific DNAs were also monitored using ¹⁵N-HSQC spectra (Figure S3). In the case of the randomized DNA^{NSP9}, many N301 amide signals shifted progressively with added oligonucleotide, while also exhibiting line broadening at partial saturation (not shown). Although we could not extract an equilibrium dissociation constant from the titration data, binding in this fast-intermediate exchange limit indicated that the K_D value for the N301/ DNA^{NSP9} complex is likely on the order of 10 μ M.²⁸ Similar K_D values in the low micromolar range have been reported for nonspecific DNA binding by ETV6, a related ETS family member.²⁹ In contrast, DNA^{NSP15} interacted in the slow exchange limit, with separate signals from the free and bound states observed at partial saturation (not shown). This behavior suggested that the K_D value for N301/ DNA^{NSP15} is <10 μ M. The higher overall macroscopic affinity of N301 for DNA^{NSP15} versus DNA^{NSP9} most likely results from the need to have more than 9 bp for maximal interactions, combined with the increased number of potential microscopic binding sites in the longer oligonucleotide.³⁰ For example, the 15 bp DNA^{NSP15} could be bound at seven possible 9 bp sites, each of which could be occupied in two orientations.

A comparison of the ¹⁵N-HSQC spectra of the two nonspecific DNA complexes also revealed that many N301 amide ¹H-¹⁵N signals were present with DNA^{NSP9}, yet weak or absent with DNA^{NSP15} (Figure S3). On the basis of the good quality spectrum of N279 with DNA^{SP14}, this is not a simple consequence of the added molecular mass of six base pairs. Therefore, a plausible explanation for the missing signals in the N301/DNA^{NSP15} complex spectrum is conformational exchange broadening due to interconversion between binding sites along the 15 bp DNA, combined with possible interconversion between various oligomeric states. This could occur via diffusion along the DNA or via dissociation/reassociation. Similar length-dependent line broadening has been reported for the lac repressor bound to a series of nonspecific DNAs.³¹ Accordingly, a nonspecific palindromic

sequence DNA^{NSP12} (5' CAAAAA-TTTTGTG3') was designed to both limit the possible adverse effects of conformational averaging and to match the 12 bp size of the specific DNA^{SP12}. Since binding DNA^{NSP12} in either orientation is equivalent, this should help offset the effects of exchange broadening. Although this conjecture was not rigorously tested, the N301/DNA^{NSP12} complex indeed yielded good quality ¹⁵N-HSQC spectra with signals from most amides and indoles visible (Figure S3).

The overlaid ¹⁵N-HSQC spectra of ¹⁵N-labeled N301 in the absence and presence of DNA^{SP12} or DNA^{NSP12} are shown in Figure 1B and Figure S4. These spectra, which enabled the following studies, were assigned using a suite of conventional and TROSY-based heteronuclear correlation experiments, combined with protein deuteration.

HI-1 and HI-2 Unfold upon Specific DNA Binding

Insights into the secondary structure and dynamics of the specific and nonspecific complexes were obtained from main chain nuclei chemical shifts (¹H^N, ¹⁵N, ¹³C^α, ¹³C^β, and ¹³CO). On the basis of these shifts, the algorithm MICS (Motif Identification from Chemical Shifts) was used to determine the likelihood that each residue adopts an α -helix, β -strand, or random coil conformation and to provide a measure of its mobility in terms of a random coil index squared order parameter (RCI-S²).^{32,33} A MICS analysis of the published chemical shifts of unbound N301 (BMRB ID 5991) agrees closely with its secondary structure as determined by NMR spectroscopic and X-ray crystallographic approaches (Figure 2A). The core ETS domain contains three α -helices and four β -strands, and the flanking IM folds as four additional α -helices. Also consistent with previous relaxation studies of the fast time scale motions of N301, the helices and strands are ordered, whereas the termini and intervening loop regions are more flexible.^{9,11} (Helices HI-1 and HI-2 are marginally stable and undergo facile amide HX due to slower time scale motions not reflected by their chemical shifts.) Overall, these results provided a baseline for the analysis of N301 complexed with specific and nonspecific DNAs.

The MISC analysis of N301 bound to specific DNA^{SP12} is shown in Figure 2C. The core ETS domain (residues 335–415) is well folded and ordered, retaining the same secondary structure as in the free protein. The lack of any major conformational change is consistent with the close similarity of all ETS domain structures determined in the absence and presence of specific DNAs by X-ray crystallography. In contrast, the N-terminal portion of the IM changed dramatically with helices HI-1 and HI-2 unfolding as evidenced by high random coil propensities and reduced RSI-S² values. Previous circular dichroism (CD) spectroscopy and partial proteolysis studies also demonstrated the disruption of this region of the IM upon binding specific DNA.⁷ However, fully disordered residues normally yield sharp intense signals with random coil chemical shifts, whereas the amide signals from residues corresponding to HI-1 and HI-2 were weak, and their intensities decreased with increasing spectrometer magnet field strength (Figure S5). Indeed, even with deuteration, resonances from residues spanning 303–310 (HI-1) and 320–329 (HI-2) were anomalously broad in spectra recorded with an 850 MHz NMR spectrometer and could only be assigned using three-dimensional ¹H–¹⁵N–¹³C correlation spectra measured with a 500 MHz instrument. This is diagnostic of conformational exchange in the millisecond–microsecond

time scale regime.²⁸ Partial backbone ordering was also observed through heteronuclear NOE measurements, discussed below. Thus, within the N301/DNA^{SP12} complex, the N-terminal inhibitory residues appear to be predominantly unfolded, but also exchanging with minor ordered conformations. We speculate that the latter are the HI-1 and HI-2 residues refolded, at least in part, as helices.

Helices HI-1 and HI-2 also Unfold upon Nonspecific DNA Binding

The MISC analysis of N301 bound to DNA^{NSP12} is shown in Figure 2B. Similar to the specific complex, the ETS domain retained the well-ordered secondary structural elements characteristic of the free protein. In addition, the N-terminal inhibitory helices HI-1 and HI-2 were predominantly unfolded, with random coil chemical shifts and reduced RCI-S² values. This is also consistent with CD spectroscopy and partial proteolysis studies of Ets-1 bound to nonspecific DNA.⁷ However, as with DNA^{SP12}, the signals from the N-terminal inhibitory residues in the N301/DNA^{NSP12} complex were weak due to conformational exchange broadening (not shown). This precluded the complete assignment of the mainchain ¹H, ¹⁵N, and ¹³C signals from several amino acids within the N-terminal region of the protein. Thus, the inhibitory module is not fully disordered and appears to transiently sample folded conformations when N301 is bound to both specific and nonspecific DNA.

Both Nonspecific and Specific DNA Bind via the Canonical ETS Domain Interface

To gain insights into the structural basis of DNA-binding by Ets-1, assigned ¹⁵N-HSQC spectra of N301 in its free and oligonucleotide-saturated forms were compared (Figure 1B). Using these spectra, the amide CSPs (δ) for N301 due to binding the specific DNA^{SP12} were calculated (Figure 3). Amide chemical shifts are highly sensitive to their local environment, and the CSPs can arise from direct or indirect structural perturbations and/or from the aromatic ring currents and electric fields due to the nearby bases and phosphodiester groups in the DNA. Large perturbations (i.e., $\delta > 0.35$ ppm) were observed for amides at the N-terminus of helix H1, in the turn between helices H2 and H3, along the DNA binding helix H3, and in the loop between strand S3 and S4. When mapped onto the structure of N301, these residues correspond closely to the canonical ETS domain DNA-binding interface of all ETS factors observed by X-ray crystallography (Figure 4). Significant CSPs were also observed for the N-terminal inhibitory residues and amides within the helices H1/H4/H5 onto which they pack. This provides further evidence that helices HI-1 and HI-2 unfold in an allosteric response to DNA binding.

The amide CSPs for N301 due to binding the nonspecific DNA^{NSP12} are also shown in Figures 3 and 4. The largest spectral changes were exhibited by the N-terminal inhibitory region. These closely resemble those observed with DNA^{SP12} and reflect the common unfolding of helices HI-1 and HI-2 that occurs when N301 binds either oligonucleotide. In contrast, rather small CSPs were observed for the remainder of the ETS domain and C-terminal inhibitory region. However, a set of residues, including Trp338, Trp375, Asn385, Gly392, Asp398, Ile401, Ala406, and Tyr410, showed above average CSPs. Additionally, residues including Tyr395, Tyr396, Arg409, and Val411 could not be detected and thus are likely broadened due to conformational exchange. All of these named residues map to the canonical DNA-binding surface of the ETS domain. Furthermore, as seen in the spectra of

Figure 1B and in a CSP projection analysis (Figure S6), the amide $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ signals of many amides in N301 shift in the same direction, but to a lesser magnitude, upon binding DNA^{NSP12} relative to DNA^{SP12}. This indicates that N301 forms generally similar, albeit less well-defined, time-averaged interactions with DNA^{NSP12} than DNA^{SP12}. Such interactions likely involve electrostatic contacts between the positively charged DNA-binding interface of the ETS domain and the negatively charged phosphodiester backbone of DNA. Also, smaller net CSPs could arise due to averaging of potential positive and negative chemical shift changes as the protein rapidly exchanges between binding sites along DNA^{NSP12}.

Upon closer inspection of the NMR data, it is evident that several important contacts formed with DNA^{SP12} were absent in the low affinity N301/DNA^{NSP12} complex (Figures 3 and 4). According to the crystal structures of many Ets-1 complexes, residues in helix H1 (Leu337 and Trp338), helix H2 (Trp375), the turn (Lys379, Lys381 and Met384), helix H3 (Tyr386-Gly392 and Arg394-Tyr397), and the wing (Arg409, Tyr410, and Tyr412) make conserved hydrogen bonding and salt-bridging contacts with specific DNA.^{27,34} Indeed, these, as well as several other surrounding residues, showed significant CSPs upon binding DNA^{SP12}. In contrast, signals from Leu337, Trp375, Met384, Tyr386, Glu387, Ser390, Arg391, and Arg394 were only slightly perturbed in the nonspecific complex. This is especially noteworthy since the amide of Leu337 and the indole of Trp375 form conserved hydrogen bonds with the phosphodiester backbone of DNA.²⁷ As seen in Figure 1B, signals from these two residues shifted downfield substantially upon contacting DNA^{SP12} (diagnostic of hydrogen bond formation), yet showed only small perturbations upon binding DNA^{NSP12}. Overall, we conclude that Ets-1 binds both the nonspecific and specific DNAs via the same general ETS domain interface. However, DNA^{NSP12} is bound through weaker interactions and lacks many of the key contacts necessary for the high affinity binding of DNA^{SP12}.

Backbone Motions in N301 Change upon DNA Binding

The dynamics of N301 in its free and bound states were examined through amide ^{15}N relaxation studies.²⁸ In particular, heteronuclear $^1\text{H}\{^{15}\text{N}\}$ -NOE experiments were used to measure backbone mobility on the sub-nanosecond time scale. As shown in Figure 5, whether free or bound to DNA^{NSP12} or DNA^{SP12}, amides throughout the ETS domain were generally rigid with high NOE values. However, residues in the turn, helix H3, and the wing had slightly lower NOE values in the free protein and the nonspecific complex, indicative of modest flexibility. In contrast, upon binding specific DNA, they become more ordered. This is suggestive of contacts with the bases and phosphodiester backbone that are more persistent than with nonspecific DNA.

Focusing on HI-1 and HI-2, in free N301 the residues forming these helices were also well-ordered on the sub-nanosecond time scale with NOE values comparable to those of the core ETS domain (Figure 5). In contrast, when bound to both nonspecific and specific DNAs, their NOE values dropped substantially, confirming that the two inhibitory helices became predominantly unfolded and conformationally disordered. However, the NOE values were not highly negative (i.e., as seen with the very flexible N- and C-termini), indicating that the fast time scale motions of these residues were partially restricted in both complexes. This is

consistent with the exchange broadening observed for many of these amides in simple ^{15}N -HSQC spectra and suggested that the N-terminal inhibitory residues are in conformational equilibria, perhaps transiently folding as helices.

Amide longitudinal (R_1) and transverse (R_2) relaxation measurements were also carried out on the free and bound forms of N301 (Figure S7). On the basis of a model-free analysis²⁵ of the relaxation behavior of well-ordered residues in the ETS domain, the effective isotropic rotational correlation time for the global tumbling of the specific N301/DNA^{SP12} complex was 12.3 ns. This is a small increase relative to the value of 9.7 ns reported previously for free N301 under similar conditions¹¹ and indicates that DNA^{SP12} was indeed bound as a well-defined 1:1 complex. In contrast, the nonspecific complex exhibited an effective correlation time of 22 ns. This value, which implies almost 2-fold slower tumbling for the nonspecific versus specific complexes, reflects the substantially slower average longitudinal (R_1) and faster transverse (R_2) relaxation rate constants measured for the ETS domain of N301/DNA^{NSP12}. Although conformational exchange broadening due to interconversion between binding sites along DNA^{NSP12} will increase the effective R_2 values, the slower R_1 relaxation indicates that the N301/DNA^{NSP12} sample has an average molecular mass higher than that of a 1:1 complex. This could result from two or more molecules of N301 bound on one molecule of DNA^{NSP12} (despite recording data with a 5:1 DNA/protein ratio to favor the 1:1 complex). Alternatively, DNA^{NSP12} could interact weakly with N301 via additional noncanonical interfaces to form a distribution of higher order oligomers with more than one DNA molecule bound on each protein molecule.

The SRR Does Not Change the Structural Interactions of Ets-1 with Specific DNA

Similar NMR spectroscopic studies were performed with the N279 construct in order to assess if the presence of the partially truncated SRR inhibitory region alters the behavior of Ets-1 when bound to DNA (Figures S1 and S8). With the exception of added signals having random coil $^1\text{H}^{\text{N}}$ chemical shifts from the disordered SRR residues, the ^{15}N -HSQC spectra of the N301/DNA^{SP12} and N279/DNA^{SP14} complexes superimposed very closely. Furthermore, chemical shift and heteronuclear $^1\text{H}\{^{15}\text{N}\}$ -NOE measurements confirmed that helices HI-1 and HI-2 were predominantly unfolded when N279 is DNA bound. Both complexes also exhibited similar global tumbling times, with a modest increase due to the added mass of the SRR for N279/DNA^{SP12} (14.8 ns) relative to N301/DNA^{SP12} (12.3 ns). Thus, although reinforcing autoinhibition, the presence of the SRR does not measurably perturb the structure of the ETS domain once bound to specific DNA.

Arginine and Lysine Side Chains at the DNA-Binding Interface

In crystal structures, the recognition helix H3 of Ets-1 binds along the major groove of specific DNA such that two invariant arginine side chains (Arg391 and Arg394) hydrogen bond to the two guanine bases in the 5'-GGA(A/T)-3' motif.³⁵ Several lysines and one additional arginine (Arg409) also contribute to the interface as judged by X-ray crystallographic studies. To further characterize the difference between binding specific and nonspecific DNA, we used NMR spectroscopy to investigate these cationic side chains.

The lysine side chains of N301 in its free, as well as DNA^{SP12} and DNA^{NSP12} bound, forms yielded significantly different ¹⁵N-HSQC spectra (Figure 6). Because of rapid HX, no lysine ¹⁵N^εH₃⁺ signals were observed for the isolated protein (pH 6.5 and 13 °C). Upon binding to nonspecific DNA, one weak signal appeared. This indicates that at least one side chain aminium moiety becomes protected from HX, likely due to burial at the interface and hydrogen bonding with the phosphodiester backbone of the oligonucleotide or possibly with an acceptor atom on the Hoogsteen edge of a base pair. Upon binding to specific DNA, signals from at least five lysines become visible, albeit with variable intensities. Unfortunately, we were unable to assign these signals to specific residues. However, the crystal structure of N301 with DNA sequence containing a 5'-GGA(A/T)-3' core (1K79.pdb) shows that five lysine N^εH₃⁺ groups (Lys379, 381, 388, 399, and 404) interact directly with phosphodiester oxygens.

In contrast, eight strong signals from the more slowly exchanging arginine ¹⁵N^εH⁺ pairs were observable in the ¹⁵N-HSQC spectra of the free protein (Figure 6A). The protein has nine arginines, of which NMR signals for seven could be assigned confidently. In the presence of DNA^{SP12}, an entirely different pattern of nine sharp arginine ¹⁵N^εH⁺ signals was observed (Figure 6C). In particular, the ¹H^ε resonances of Arg391 and Arg394 were shifted downfield considerably, as expected due to the formation of hydrogen bonds with the 5'-GGA-3' guanines. In addition, three arginine ¹⁵N^ηH₂⁺ pairs yielded dispersed nondegenerate signals, which is also indicative of well-defined DNA interactions (Figure S9).

Upon binding the nonspecific DNA^{NSP12}, several arginine ¹⁵N^εH⁺ and ¹⁵N^ηH₂⁺ signals were broadened and exhibited chemical shifts clearly different from those of either the free and DNA^{SP12} bound forms of N301 (Figure 6B and Figure S9). Unfortunately, line broadening hampered the detection of scalar correlations to the adjacent ¹³CH₂ groups, and thus we could only assign the ¹⁵N^εH⁺ signals from five arginines (Arg309, 311, 373, 374, and 413) based on a comparison with the spectra of the free and specific DNA complex. Consistent with their similar chemical shifts in all forms of N301, these five arginines are not involved in DNA interactions. Although not identified in the N301/DNA^{NSP12} complex, Arg391 and Arg394 in the recognition helix H3 clearly do not exhibit the chemical shift dispersion induced in the N301/DNA^{SP12} complex. Overall, these NMR spectra suggest that nonspecific DNA complexes involve dynamic electrostatic interactions between the lysine and arginine side chains with the phosphodiester backbone, whereas specific complexes exhibit well-defined base hydrogen bonding.

Arginine Guanidinium Groups Have Different Relaxation Properties That Correlate with Their DNA Interactions

We also used ¹⁵N relaxation to investigate the dynamic properties of the guanidinium ¹⁵N^ε nuclei in the N301 complexes (Figure 7). A wide range of arginine motions can be observed. In the free protein, the solvent exposed Arg391 and Arg394 in the DNA-recognition helix H3, as well as Arg309 and Arg311 in the inhibitory helix HI-1, are mobile on the sub-nanosecond time scale. This is evident by heteronuclear NOE values < 0 and relatively slow transverse *R*₂ relaxation. Although the latter two side chains appear to be

forming salt-bridges with adjacent aspartates (Asp306, Asp313, and/or Asp317) in some, but not all, Ets-1 crystal structure, these NMR measurements indicate they are conformationally flexible in solution. Conversely, Arg373 and Arg374 in helix H2 and Arg413 in strand S4 show more restricted mobility with heteronuclear NOE values > 0 . Thus, the salt-bridges to Glu370 in helix H2 and to Asp359 or Glu362 near strand S2, respectively, seen in most Ets-1 crystal structure appear to persist in solution. Similarly, the side chain of Arg378 at the C-terminal end of helix H2 is likely restricted by an interaction with Glu343 in helix H1. Although this arginine was only characterized in the N301/DNA^{SP12} complex, it is not involved in DNA binding and thus should show the same dynamic behavior in the free protein.

Upon binding of specific DNA^{SP12}, the mobile side chains of Arg391 and Arg394 become well-ordered as shown by heteronuclear NOE values > 0.7 and fast transverse R_2 relaxation. This is expected with their intimate roles in the recognition of the core guanine bases and their essentially invariant positions in all Ets-1/DNA complexes studied to date by X-ray crystallography. In contrast, the side chain of Arg409 is relatively mobile in the N301/DNA^{SP12} complex. Thus, despite appearing close to or hydrogen bonding with the phosphodiester backbone of DNA in these crystal structures, any electrostatic interactions involving Arg409 are relatively weak and not persistent in solution.

As noted above, only five arginine $^{15}\text{N}^{\text{eH}}$ signals could be assigned in the ^{15}N -HSQC spectrum of the N301/DNA^{NSP12} complex. These residues are not involved in DNA binding and exhibit similar chemical shifts and ^{15}N relaxation behavior in both the specific and nonspecific complexes. The conspicuous absence of ^{15}N -HSQC signals from Arg391 and Arg394 in the N301/DNA^{NSP12} complex suggests that they are exchange broadened due to interconversion between possible binding modes along the nonspecific DNA.

DISCUSSION

We investigated the mechanisms by which Ets-1 binds specific and nonspecific DNAs, as well as the interplay of autoinhibition with these processes. With either oligonucleotide, the inhibitory module is disrupted, and the same canonical interface is used for DNA binding. However, the resulting structural and dynamic changes in the backbone and arginine/lysine side chains of the ETS domain are much larger with specific DNA than with nonspecific DNA. This is consistent with the general observation that sequence-specific transcription factors bind nonspecific sequences via dynamic electrostatic interactions, whereas their cognate sites are recognized through high-affinity interactions including hydrogen bonding between protein side chains and DNA base pairs.

HI-1 and HI-2 Predominantly Unfold upon DNA-Binding

Using NMR spectroscopy, we provided new insights into the mechanism of Ets-1 autoinhibition. Upon binding both nonspecific and specific oligonucleotides, residues throughout the N-terminal IM of N301 exhibited similar, large amide CSPs. Analysis of main chain chemical shifts with the MICS algorithm indicated that residues forming helices HI-1 and HI-2 in the free protein become more flexible and adopt random coil conformations in the DNA^{SP12} and DNA^{NSP12} complexes. The unfolding of this portion of

the IM is supported by substantial decreases in the amide heteronuclear $^1\text{H}\{^{15}\text{N}\}$ -NOE and chemical shift-derived RSI-S² values of these residues. However, the ^{15}N -HSQC signals and NOE values from these residues were not as sharp or low, respectively, as expected for a completely disordered polypeptide. This indicates that the sub-nanosecond time scale motions of the N-terminal inhibitory sequence are partially dampened. Furthermore, the NMR signals from these amides broadened with increasing magnetic field strength, which is also diagnostic of conformational exchange on the millisecond-microsecond time scale. We suggest that this reflects the transient refolding of helices HI-1 and HI-2 within the DNA-bound complexes.

Previously, measurements with CD spectroscopy indicated a loss of helical structure when N280 bound both specific and nonspecific DNA.⁷ This was attributed to the unfolding of HI-1 because of the increased trypsin and chymotrypsin susceptibility of residues within or adjacent to this helix. However, additional studies with other proteases suggested both helices might unfold.³⁶ Subsequent X-ray crystallographic studies of Ets-1 fragments in a ternary complex with Pax-5 on the mb-1 promoter DNA also showed helix HI-1 to be disordered, whereas HI-2 remained folded (1MDM.pdb).³⁵ In contrast, the two inhibitory helices were either folded or unfolded in the crystal structure of a comparable Ets-2/DNA complex (4BQA.pdb).³⁷ Both Ets-1 inhibitory helices were also displaced by the ETS-interacting domain (EID) of Runx1 in a ternary complex of the two partner transcription factors bound to a composite DNA site (4L0Z.pdb).³⁸ In another quaternary complex involving Runx1 lacking the EID, HI-1 is unfolded, whereas HI-2 remains helical but packs against the Ets-1 ETS domain in an alternative register that is incompatible with the native fold of the IM (3WTS.pdb).³⁹ Furthermore, amide HX studies of DNA-free Ets-1 fragments demonstrated that helices HI-1 and HI-2 are marginally stable and protected from exchange by only ~15- and ~75-fold, respectively, relative to a reference random coil polypeptide.^{9,11}

Collectively, this extensive body of data indicates that the energetic difference between the folded and unfolded states of HI-1 and HI-2 in both free and DNA-bound Ets-1 is small. In the absence of DNA, the two helices are predominantly folded, but readily sample unfolded states that are detectable via HX measurements. Conversely, in solution, DNA-binding tips the balance toward their unfolded conformations. However, magnetic field-dependent line broadening suggests that minor folded states still exist transiently. Folding of the inhibitory helices HI-1 and HI-2 is also favored for Ets-1/DNA complexes in certain crystalline environments.

In the absence of the stabilizing SRR, the IM only imparts ~2-fold autoinhibition of specific DNA binding by Ets-1.^{8,9} This modest effect is also in line with the delicate balance between the folded and unfolded forms of the N-terminal inhibitory sequence. Furthermore, a detailed comparison of the structures of various Ets-1 fragments in their free and bound states reveals no obvious differences beyond a small displacement of the flexible loop between helices HI-2 and H1. This displacement appears necessary to avoid a steric clash with the DNA backbone and to enable a critical dipole-enhanced hydrogen bond between the amide of Leu337 at the start of helix H1 and a phosphodiester oxygen.^{11,40} The lack of any clear allosteric pathway linking DNA binding and unfolding of the distal IM suggests that only subtle perturbations are required to tip this energetic balance.

Autoinhibition Does Not Impart DNA-Binding Specificity

The effects of autoinhibition on the affinity of Ets-1 for nonspecific DNA have not been quantitatively determined. However, given the similar disruption of the IM as seen with specific DNA, it is likely that autoinhibition also increases the K_D value of N301 for nonspecific DNA by ~2-fold. Furthermore, since the SRR reinforces autoinhibition by stabilizing the IM and by sterically blocking the DNA-binding interface of the ETS domain, it is reasonable to suggest that full length Ets-1 is inhibited comparably (~20-fold) for binding both specific and nonspecific DNAs.¹⁰ It is also noteworthy that inhibited and uninhibited fragments of Ets-1 showed the same DNA-binding sequence specificity profiles in high-throughput screening assays.⁵ This is consistent with our observations that the SRR does not change the size of the Ets-1 binding site, nor does it perturb the structure of the DNA-bound ETS domain. Similar behavior has been reported for ETV6, a related family member that is autoinhibited due to steric blockage of its ETS domain by an appended C-terminal helix. The unfolding of this helix reduces the affinity of ETV6 for specific and nonspecific DNA by ~50-fold and ~5-fold, as measured using EMSA and ITC, respectively.^{29,41} Collectively, these results indicate that autoinhibition *per se* does not contribute substantially to the specificity (i.e., relative affinity) of these two ETS factors for their cognate sites versus background DNA within the nucleus. Rather, it enables regulation via routes including protein partnerships and post-translational modifications.⁶

Ets-1 Binds Specific and Nonspecific DNA through the Canonical ETS Domain Interface

Autoinhibited Ets-1 binds 5'-GGA(A/T)-3' containing DNA sites with K_D values on the order of 0.4 nM.^{8,9} In contrast, the K_D value for 9 bp nonspecific DNA based on NMR titrations can be estimated as ~10 μ M. This represents a 10⁴- to 10⁵-fold difference in affinity that must arise through structural and dynamic complementarity of the Ets-1 ETS domain with the nucleotide bases (direct base readout), phosphodiester backbone (with sequence-dependent conformations; indirect shape readout) and hydrating waters of its cognate sites.^{42,43} Although seemingly large, this affinity difference is offset by the vast excess of nonspecific versus specific DNA sites within the eukaryotic genome. Also, consensus, near-consensus, and nonconsensus sequences will be bound with a continuum of thermodynamic affinities.^{44,45}

Despite this large difference in affinity, our studies show that N301 binds specific and nonspecific DNA through the same general interface found in all ETS domain/DNA structures. As expected, upon binding the three specific oligonucleotides tested herein, large CSPs are observed for N301 amides in helix H1, throughout the turn preceding H3, the recognition helix H3, the following "wing" and strands S3 and S4. These perturbations likely arise from many sources, including ring current and electric field effects from the closely juxtaposed DNA, subtle changes in the structure and dynamics of the ETS domain, and most importantly, the formation of key inter-molecular hydrogen bonds (e.g., the amide of Leu337 in helix H1 with the phosphodiester backbone, and the side chains of Arg391 and Arg394 in helix H3 with the invariant 5'-GGA(A/T)-3' guanines). In addition, the unfolding of helices HI-1 and HI-2 leads to pronounced CSPs in the IM module and its interface with the ETS domain.

In the case of binding nonspecific DNA, N301 amides including Trp338 (H1), Trp375 (H2), Asn385 (H3), Asp398 (H3 to S1 loop), Ile401 (H3 to S1 loop), and Tyr410 (turn) also show above average CSPs. However, the magnitudes of these CSPs are smaller than observed with DNA^{SP12}, indicating weaker time-averaged interactions than in the specific complex. Furthermore, the amide signals of Tyr395 (H3), Tyr396 (H3), R409 (turn), and Val411 (turn) are broadened beyond detection. Such exchange broadening becomes more pronounced when N301 is saturated with longer oligonucleotides (DNA^{NSP15} versus DNA^{NSP9}) and is somewhat alleviated with the palindromic DNA^{NSP12}. Similarly, the indole ¹⁵N^ε1H signals of Trp338 and Trp375 are broadened beyond detection with the longer DNA^{NSP15}, but present with the palindromic and smaller nonspecific sequences. As seen with the lac repressor, this likely reflects the millisecond–microsecond interconversion of N301 between different possible binding positions and orientations along the nonspecific oligonucleotide.³¹ However, slow exchange could also lead to apparent NMR linebroadening if the protein is partitioned between many microstates that yield slightly different amide chemical shifts. Such interconversion could occur via rotationally coupled sliding along the DNA or from dissociation/ reassociation events with different orientations and locations on the DNA.⁴ Relaxation studies also suggest that the N301/ DNA^{NSP12} complex adopts slower tumbling oligomeric states. This stands in contrast to the specific N301/DNA^{SP12} complex, which exhibits one well-defined 1:1 conformation. Although the stoichiometries of the nonspecific complexes were not determined, if excess DNA^{NSP12} was interacting weakly with N301 via additional noncanonical interfaces, then this could provide a mechanism for the direct intersegmental transfer of Ets-1 between DNA strands in the genome.

It is also noteworthy that the DNA-binding interface of Ets-1 is conformationally dynamic. This is seen experimentally through both NMR relaxation dispersion measurements,⁸ which detect millisecond–microsecond time scale motions, and HX studies, which detect local unfolding transitions allowing the exchange of amide hydrogens with those of solvent water.^{9,11} Very similar behavior is exhibited by ETV6⁴⁶ and is likely a common property of most, if not all, ETS domains. We speculate that this flexibility contributes to the ability of Ets-1 to bind both specific and nonspecific DNAs via the same canonical interface. Furthermore, the flexibility of the ETS domain is dampened through transient, phosphorylation-enhanced interactions with the adjacent SRR, thereby linking DNA-binding autoinhibition with CaMKII-dependent cellular signaling.⁸

Arginine and Lysine Side Chains Reflect Differential Interactions of Ets-1 with Specific versus Nonspecific DNA

X-ray crystallographic studies have revealed that two or three arginine and five lysine side chains are present at the DNA-binding interface of Ets-1. The contributions of these residues in the binding of specific versus nonspecific oligonucleotides were assessed through NMR spectroscopy. In the free protein, all lysine ¹⁵N^εH₃⁺ groups underwent rapid HX with water and were not detectable. However, upon addition of DNA^{NSP12}, one weak signal was observed, whereas upon addition of the 5′-GGAA-3′ containing DNA^{SP} sequence, five distinct lysine signals appeared. Because of their intrinsically slower HX, ¹⁵N^εH⁺ signals from nearly all of the nine arginines of N301 were detected and assigned in its free and

specifically bound states. These guanidinium groups exhibited a wide range of relaxation behaviors that are generally consistent with expectations based on crystallographic studies of several Ets-1/DNA complexes, as well as NMR spectroscopic studies on other DNA-bound transcription factors such as the HoxD9 homeodomain⁴⁷ and the zinc-finger protein Egr-1.⁴⁸ For example, solvent exposed Arg309 and Arg311 are mobile on the sub-nanosecond time scale, whereas salt-bridged Arg373, Arg374, and Arg413 have partially restricted motions with or without DNA.

Most strikingly, the invariant Arg391 and Arg394 in the recognition helix H3 are dynamic in the free protein and become well-ordered upon forming critical hydrogen bonds with the 5'-GGA(A/T)-3' guanines. In contrast, upon binding nonspecific DNA, all arginine side chain signals are exchange broadened. Again, this is consistent with conformational averaging between multiple binding modes. Importantly, the absence of any downfield shifted signals indicated that Arg391 and Arg394 of helix H3 were not forming persistent hydrogen bonds with the DNA bases. Indeed, unsatisfied hydrogen bonding with these arginines would destabilize many possible complexes with nonspecific DNAs and thereby contribute to the discrimination of Ets1 for its cognate sites.⁴⁹ Overall, the distinct spectral patterns shown by the lysine and arginine side chain reflects the electrostatic versus well-defined hydrogen bonding interactions of Ets-1 with nonspecific versus specific DNAs, respectively.

DNA Binding and Transcriptional Regulation

Although great progress has been made, it remains a challenge to understand the specificity of DNA binding by transcription factors within a cellular context and the relationships to biological function. High-throughput binding studies have shown that transcription factors often recognize multiple distinctly different sequence motifs.^{44,45,50} Furthermore, genome-wide mapping efforts revealed that transcription factors with seemingly well-characterized sequence specificities bind more sites without than with clear consensus motifs, albeit with lower relative occupancies and shorter residence times.^{43,51} This is entirely expected from a thermodynamic perspective and the mass action effects of the local in vivo concentrations of transcription factors and their DNA binding sites.^{52,53} More interestingly, only a small fraction of potential consensus sites with a genome are detectably bound by a given transcription factor, and not all bound sites are functional for transcriptional regulation.^{52,53} Accordingly, numerous additional parameters, including chromatin structure, DNA accessibility, cooperative and competitive protein partnerships, and DNA/protein post-translational modifications, must modulate the partitioning of transcription factors between their unbound, nonspecifically bound, specifically bound, and functional forms.^{43,52}

The vast majority of biophysical studies of transcription factors have focused on their specific DNA complexes. In contrast, only a limited number of nonspecific complexes have been investigated by X-ray crystallography and NMR spectroscopy. These include bacterial repressors (lac and cro),^{13,31,54} steroid receptors,^{55,56} homeodomains (HoxD9 and MAT α 2),^{47,57-59} a HMG-box protein,⁶⁰ a POU domain (Oct-1),^{61,62} a paired domain (Pax5),⁶³ the ETV6 ETS domain,^{29,41} the Max basic region-helix-loop-helix-leucine zipper,⁴⁹ and zinc fingers (ZNF217 and Egr-1).^{64,65} Characterization of these systems has

demonstrated that nonspecific DNA binding is driven largely by electrostatic interactions between positively charged amino acid side chains and the negatively charged phosphodiester backbone.⁵⁴ This is offset by energetic penalty of any unfavorable side chain–base pair contacts, including unsatisfied hydrogen bonds.⁴⁹ In addition, as shown by seminal studies on the lac repressor, the backbone and side chains of proteins in nonspecific complexes are significantly more dynamic and hydrated than with specific complexes.¹³ This behavior enables sliding with rotation⁶⁶ of a protein along the approximately isopotential electrostatic surface of the helical DNA. However, across the continuum of nonconsensus sites, the sequence-dependent shape of DNA can lead to preferential interactions.^{43,51,67} Upon recognition of a specific consensus sequence, both the protein and DNA undergo changes in conformation and hydration to form a stable, long-lived hydrogen-bonded complex.^{68–70}

Collectively, one-dimensional sliding combined with three-dimensional hopping and intersegmental transfer ensures that transcription factors find quickly their specific target sites in the cell and thereby respond efficiently to signaling events.^{3,4} Although derived from in vitro biophysical studies, the general features of this facilitated diffusion target search model are supported by elegant time-resolved single molecule microscopy studies in live cells.^{71–74} However, as expected from the heterogeneous environment of DNA within the nucleus, the search process is complex and often without a simple distinction between specific and nonspecific interactions. For example, recent studies suggest that transcription factor search dynamics are biased by the organization of enhancers into localized three-dimensional clusters.⁷⁵ Also, neighboring regions of the genome may have evolved to contain sequence or shape features that help guide transcription factors to their cognate sites.⁷⁶

Consistent with this general model, our studies show that Ets-1 associates with nonspecific and specific DNAs via the same canonical ETS domain interface. However, binding is driven by distinct mechanisms. In particular, the dynamic nature of nonspecific DNA complexes and importance of well-ordered hydrogen bonds in specific DNA binding are clearly evidenced by the NMR chemical shifts and relaxation properties of several interfacial lysine and arginine side chains. A defining characteristic of several ETS factors is the autoinhibition of DNA binding. Although providing routes for biological regulation, autoinhibition acts primarily on the unbound forms of the ETS family members ETV6 and Ets-1 and thus alone does not appear to contribute significantly to the discrimination of specific versus nonspecific DNA sequences.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This study was funded by the Canadian Cancer Society Research Institute (2011-700772 to L.P.M.), the Canadian Institutes for Health Research (CIHR MOP-136834 to L.P.M.), and the National Institutes of Health (R01GM38663 to B.J.G.). Instrument support was provided by the Canada Foundation for Innovation, the British Columbia

Knowledge Development Fund, the UBC Blusson Fund, and the Michael Smith Foundation for Health Research. G.D. held a CIHR Frederick Banting Graduate Scholarship and a UBC Four Year Fellowship. Funding from the Huntsman Cancer Institute/Huntsman Cancer Foundation and the Howard Hughes Medical Institute is also acknowledged.

ABBREVIATIONS

BMRB	Biological Magnetic Resonance Bank
CSP	chemical shift perturbation
HISQC	heteronuclear in-phase single quantum coherence
HSQC	heteronuclear single quantum correlation
HX	hydrogen exchange
IM	inhibitory module
K_D	equilibrium dissociation constant
MWCO	molecular weight cutoff
NOE	nuclear Overhauser effect
R₁	longitudinal relaxation rate constant
R₂	transverse relaxation rate constant
RCI-S²	random coil index squared order parameter
SRR	serine-rich region
wHTH	winged helix–turn–helix
N279	residues 279–440 of murine Ets-1
N301	residues 301–440 of murine Ets-1

References

1. Hippel PHV, Revzin A, Gross CA, Wang AC. Nonspecific DNA binding of genome regulating proteins as a biological-control mechanism: 1. The lac operon - equilibrium aspects. *Proc. Natl. Acad. Sci. U. S. A.* 1974; 71:4808–4812. [PubMed: 4612528]
2. Lin S, Riggs AD. The general affinity of lac repressor for *E. coli* DNA: implications for gene regulation in procaryotes and eucaryotes. *Cell.* 1975; 4:107–111. [PubMed: 1092468]
3. von Hippel PH, Berg OG. Facilitated target location in biological systems. *J. Biol. Chem.* 1989; 264:675–678. [PubMed: 2642903]
4. Marius Clore GM. Exploring translocation of proteins on DNA by NMR. *J. Biomol. NMR.* 2011; 51:209–219. [PubMed: 21847629]
5. Wei GH, Badis G, Berger MF, Kivioja T, Palin K, Enge M, Bonke M, Jolma A, Varjosalo M, Gehrke AR, Yan J, Talukder S, Turunen M, Taipale M, Stunnenberg HG, Ukkonen E, Hughes TR, Bulyk ML, Taipale J. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.* 2010; 29:2147–2160. [PubMed: 20517297]
6. Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu. Rev. Biochem.* 2011; 80:437–471. [PubMed: 21548782]

7. Petersen JM, Skalicky JJ, Donaldson LW, McIntosh LP, Alber T, Graves BJ. Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science*. 1995; 269:1866–1869. [PubMed: 7569926]
8. Pufall MA, Lee GM, Nelson ML, Kang HS, Velyvis A, Kay LE, McIntosh LP, Graves BJ. Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science*. 2005; 309:142–145. [PubMed: 15994560]
9. Lee GM, Pufall MA, Meeker CA, Kang HS, Graves BJ, McIntosh LP. The affinity of Ets-1 for DNA is modulated by phosphorylation through transient interactions of an unstructured region. *J. Mol. Biol.* 2008; 382:1014–1030. [PubMed: 18692067]
10. Desjardins G, Meeker CA, Bhachech N, Currie SL, Okon M, Graves BJ, McIntosh LP. 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.* 2014; 111:11019–11024. [PubMed: 25024220]
11. Lee GM, Donaldson LW, Pufall MA, Kang HS, Pot I, Graves BJ, McIntosh LP. The structural and dynamic basis of Ets-1 DNA binding autoinhibition. *J. Biol. Chem.* 2005; 280:7088–7099. [PubMed: 15591056]
12. Garvie CW, Pufall MA, Graves BJ, Wolberger C. Structural analysis of the autoinhibition of Ets-1 and its role in protein partnerships. *J. Biol. Chem.* 2002; 277:45529–45536. [PubMed: 12221090]
13. Kalodimos CG, Boelens R, Kaptein R. Toward an integrated model of protein-DNA recognition as inferred from NMR studies on the lac repressor system. *Chem. Rev.* 2004; 104:3567–3586. [PubMed: 15303828]
14. Desjardins, G. Ph.D. thesis. University British Columbia; 2015. Structural characterization of DNA binding and autoinhibition by the Ets-1 transcription factor.
15. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 2012; 40:W597–603. [PubMed: 22661580]
16. Tataurov AV, You Y, Owczarzy R. Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys. Chem.* 2008; 133:66–70. [PubMed: 18201813]
17. Kay LE. Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution. *Prog. Biophys. Mol. Biol.* 1995; 63:277–299. [PubMed: 8599031]
18. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 1995; 6:277–293. [PubMed: 8520220]
19. Goddard, TD., Kneller, DG. Sparky 3. University of California; San Francisco:
20. Gardner KH, Kay LE. The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* 1998; 27:357–406. [PubMed: 9646872]
21. Yang D, Kay LE. Improved $^1\text{H}^{\text{N}}$ -detected triple resonance TROSY-based experiments. *J. Biomol. NMR.* 1999; 13:3–10. [PubMed: 21080259]
22. Andre I, Linse S, Mulder FA. Residue-specific pKa determination of lysine and arginine side chains by indirect ^{15}N and ^{13}C NMR spectroscopy: application to apo calmodulin. *J. Am. Chem. Soc.* 2007; 129:15805–15813. [PubMed: 18044888]
23. Rao NS, Legault P, Muhandiram DR, Greenblatt J, Battiste JL, Williamson JR, Kay LE. NMR pulse schemes for the sequential assignment of arginine side-chain H epsilon protons. *J. Magn. Reson., Ser. B.* 1996; 113:272–276. [PubMed: 8995846]
24. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry.* 1994; 33:5984–6003. [PubMed: 7514039]
25. Dosset P, Hus JC, Blackledge M, Marion D. Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR.* 2000; 16:23–28. [PubMed: 10718609]

26. Nye JA, Petersen JM, Gunther CV, Jonsen MD, Graves BJ. Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.* 1992; 6:975–990. [PubMed: 1592264]
27. Grishin AV, Alekseevsky AV, Spirin SA, Kariagina AS. [Conserved structural features of ETS domain-DNA complexes]. *Mol. Biol. (Moscow, Russ. Fed., Engl. Ed.)*. 2009; 43:612–619.
28. Kleckner IR, Foster MP. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta, Proteins Proteomics*. 2011; 1814:942–968.
29. De S, Chan AC, Coyne HJ 3rd, Bhachech N, Hermsdorf U, Okon M, Murphy ME, Graves BJ, McIntosh LP. Steric mechanism of auto-inhibitory regulation of specific and nonspecific DNA binding by the ETS transcriptional repressor ETV6. *J. Mol. Biol.* 2014; 426:1390–1406. [PubMed: 24333486]
30. Chaires JB. Analysis and interpretation of ligand-DNA binding isotherms. *Methods Enzymol.* 2001; 340:3–22. [PubMed: 11494856]
31. Loth K, Gnida M, Romanuka J, Kaptein R, Boelens R. Sliding and target location of DNA-binding proteins: an NMR view of the lac repressor system. *J. Biomol. NMR*. 2013; 56:41–49. [PubMed: 23568265]
32. Shen Y, Bax A. Identification of helix capping and β -turn motifs from NMR chemical shifts. *J. Biomol. NMR*. 2012; 52:211–232. [PubMed: 22314702]
33. Berjanskii M, Wishart DS. NMR: prediction of protein flexibility. *Nat. Protoc.* 2006; 1:683–688. [PubMed: 17406296]
34. Garvie CW, Wolberger C. Recognition of specific DNA sequences. *Mol. Cell.* 2001; 8:937–946. [PubMed: 11741530]
35. Garvie CW, Hagman J, Wolberger C. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol. Cell.* 2001; 8:1267–1276. [PubMed: 11779502]
36. Jonsen, MD. Ph.D. thesis. University of Utah; 1999. Auto-inhibition of DNA binding in the transcription factor Ets-1: A mechanistic and structural study.
37. Newman JA, Cooper CD, Aitkenhead H, Gileadi O. Structural Insights into the Auto-regulation and Cooperativity of the Human Transcription Factor Ets-2. *J. Biol. Chem.* 2015; 290:8539–8549. [PubMed: 25670864]
38. Shrivastava T, Mino K, Babayeva ND, Baranovskaya OI, Rizzino A, Tahirov TH. Structural basis of Ets1 activation by Runx1. *Leukemia*. 2014; 28:2040–2048. [PubMed: 24646888]
39. Shiina M, Hamada K, Inoue-Bungo T, Shimamura M, Uchiyama A, Baba S, Sato K, Yamamoto M, Ogata K. A novel allosteric mechanism on protein-DNA interactions underlying the phosphorylation-dependent regulation of Ets1 target gene expressions. *J. Mol. Biol.* 2015; 427:1655–1669. [PubMed: 25083921]
40. Wang H, McIntosh LP, Graves BJ. Inhibitory module of Ets-1 allosterically regulates DNA binding through a dipole-facilitated phosphate contact. *J. Biol. Chem.* 2002; 277:2225–2233. [PubMed: 11689571]
41. De S, Okon M, Graves BJ, McIntosh LP. Autoinhibition of ETV6 DNA binding is established by the stability of its inhibitory helix. *J. Mol. Biol.* 2016; 428:1515–1530. [PubMed: 26920109]
42. Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* 2010; 79:233–269. [PubMed: 20334529]
43. Slattery M, Zhou T, Yang L, Dantas Machado AC, Gordan R, Rohs R. Absence of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.* 2014; 39:381–399. [PubMed: 25129887]
44. Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X, Kuznetsov H, Wang CF, Coburn D, Newburger DE, Morris Q, Hughes TR, Bulky ML. Diversity and complexity in DNA recognition by transcription factors. *Science*. 2009; 324:1720–1723. [PubMed: 19443739]
45. Jolma A, Yan J, Whittington T, Toivonen J, Nitta KR, Rastas P, Morgunova E, Enge M, Taipale M, Wei G, Palin K, Vaquerizas JM, Vincentelli R, Luscombe NM, Hughes TR, Lemaire P, Ukkonen E, Kivioja T, Taipale J. DNA-binding specificities of human transcription factors. *Cell*. 2013; 152:327–339. [PubMed: 23332764]

46. Coyne HJ 3rd, De S, Okon M, Green SM, Bhachech N, Graves BJ, McIntosh LP. Autoinhibition of ETV6 (TEL) DNA binding: appended helices sterically block the ETS domain. *J. Mol. Biol.* 2012; 421:67–84. [PubMed: 22584210]
47. Iwahara J, Zweckstetter M, Clore GM. NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:15062–15067. [PubMed: 17008406]
48. Esadze A, Chen C, Zandarashvili L, Roy S, Pettit BM, Iwahara J. Changes in conformational dynamics of basic side chains upon protein–DNA association. *Nucleic Acids Res.* 2016 gkw531.
49. Sauve S, Naud JF, Lavigne P. The mechanism of discrimination between cognate and non-specific DNA by dimeric b/ HLH/LZ transcription factors. *J. Mol. Biol.* 2007; 365:1163–1175. [PubMed: 17109882]
50. Rhee HS, Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell.* 2011; 147:1408–1419. [PubMed: 22153082]
51. Zentner GE, Kasinathan S, Xin B, Rohs R, Henikoff S. ChEC-seq kinetics discriminates transcription factor binding sites by DNA sequence and shape in vivo. *Nat. Commun.* 2015; 6:8733. [PubMed: 26490019]
52. Biggin MD. Animal transcription networks as highly connected, quantitative continua. *Dev. Cell.* 2011; 21:611–626. [PubMed: 22014521]
53. Fisher WW, Li JJ, Hammonds AS, Brown JB, Pfeiffer BD, Weiszmam R, MacArthur S, Thomas S, Stamatoyannopoulos JA, Eisen MB, Bickel PJ, Biggin MD, Celniker SE. DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 2012; 109:21330–21335. [PubMed: 23236164]
54. Kalodimos CG, Biris N, Bonvin AM, Levandoski MM, Guennuegues M, Boelens R, Kaptein R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science.* 2004; 305:386–389. [PubMed: 15256668]
55. Gewirth DT, Sigler PB. The basis for half-site specificity explored through a non-cognate steroid receptor-DNA complex. *Nat. Struct. Biol.* 1995; 2:386–394. [PubMed: 7664096]
56. Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature.* 1991; 352:497–505. [PubMed: 1865905]
57. Anderson KM, Esadze A, Manoharan M, Bruschweiler R, Gorenstein DG, Iwahara J. Direct observation of the ion-pair dynamics at a protein-DNA interface by NMR spectroscopy. *J. Am. Chem. Soc.* 2013; 135:3613–3619. [PubMed: 23406569]
58. Aishima J, Wolberger C. Insights into nonspecific binding of homeodomains from a structure of MAT α 2 bound to DNA. *Proteins: Struct., Funct., Genet.* 2003; 51:544–551. [PubMed: 12784213]
59. Iwahara J, Clore GM. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature.* 2006; 440:1227–1230. [PubMed: 16642002]
60. Iwahara J, Schwieters CD, Clore GM. Characterization of nonspecific protein-DNA interactions by ¹H paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* 2004; 126:12800–12808. [PubMed: 15469275]
61. Takayama Y, Clore GM. Intra- and intermolecular translocation of the bi-domain transcription factor Oct1 characterized by liquid crystal and paramagnetic NMR. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:E169–176. [PubMed: 21555551]
62. Takayama Y, Clore GM. Interplay between minor and major groove-binding transcription factors Sox2 and Oct1 in translocation on DNA studied by paramagnetic and diamagnetic NMR. *J. Biol. Chem.* 2012; 287:14349–14363. [PubMed: 22396547]
63. Perez-Borraero C, Okon M, McIntosh LP. Structural and dynamics studies of Pax5 reveal asymmetry in stability and DNA binding by the Paired domain. *J. Mol. Biol.* 2016; 428:2372. [PubMed: 27067111]
64. Zandarashvili L, Vuzman D, Esadze A, Takayama Y, Sahu D, Levy Y, Iwahara J. Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1. *Proc. Natl. Acad. Sci. U. S. A.* 2012; 109:E1724–1732. [PubMed: 22675124]
65. Nunez N, Clifton MM, Funnell AP, Artuz C, Hallal S, Quinlan KG, Font J, Vandevenne M, Setiyaputra S, Pearson RC, Mackay JP, Crossley M. The multi-zinc finger protein ZNF217

- contacts DNA through a two-finger domain. *J. Biol. Chem.* 2011; 286:38190–38201. [PubMed: 21908891]
66. Blainey PC, Luo G, Kou SC, Mangel WF, Verdine GL, Bagchi B, Xie XS. Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.* 2009; 16:1224–1229. [PubMed: 19898474]
67. Zhou TY, Shen N, Yang L, Abe N, Horton J, Mann RS, Bussemaker HJ, Gordan R, Rohs R. Quantitative modeling of transcription factor binding specificities using DNA shape. *Proc. Natl. Acad. Sci. U. S. A.* 2015; 112:4654–4659. [PubMed: 25775564]
68. Marcovitz A, Levy Y. Frustration in protein-DNA binding influences conformational switching and target search kinetics. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:17957–17962. [PubMed: 22003125]
69. Sanchez IE, Ferreira DU, Dellarole M, de Prat-Gay G. Experimental snapshots of a protein-DNA binding landscape. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:7751–7756. [PubMed: 20375284]
70. Zhou HX. Rapid search for specific sites on DNA through conformational switch of nonspecifically bound proteins. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:8651–8656. [PubMed: 21543711]
71. Hammar P, Leroy P, Mahmutovic A, Marklund EG, Berg OG, Elf J. The lac repressor displays facilitated diffusion in living cells. *Science.* 2011; 336:1595–1598.
72. Gebhardt JC, Suter DM, Roy R, Zhao ZW, Chapman AR, Basu S, Maniatis T, Xie XS. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat. Methods.* 2013; 10:421–426. [PubMed: 23524394]
73. Chen J, Zhang Z, Li L, Chen BC, Revyakin A, Hajj B, Legant W, Dahan M, Lionnet T, Betzig E, Tjian R, Liu Z. Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell.* 2014; 156:1274–1285. [PubMed: 24630727]
74. Normanno D, Boudarene L, Dugast-Darzacq C, Chen JJ, Richter C, Proux F, Benichou O, Voituriez R, Darzacq X, Dahan M. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat. Commun.* 2015; 6:7357. [PubMed: 26151127]
75. Liu Z, Legant WR, Chen BC, Li L, Grimm JB, Lavis LD, Betzig E, Tjian R. 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *eLife.* 2014; 3:e04236. [PubMed: 25537195]
76. Dror I, Golan T, Levy C, Rohs R, Mandel-Gutfreund Y. A widespread role of the motif environment in transcription factor binding across diverse protein families. *Genome Res.* 2015; 25:1268–1280. [PubMed: 26160164]
77. Esadze A, Zandarashvili L, Iwahara J. Effective strategy to assign H-1-N-15 heteronuclear correlation NMR signals from lysine side-chain NH_3^+ groups of proteins at low temperature. *J. Biomol. NMR.* 2014; 60:23–27. [PubMed: 25129623]

Figure 1.

Modular organization of Ets-1 and the ¹⁵N-HSQC spectra of N301 bound to nonspecific and specific DNA. (A) Ets-1 contains a PNT domain, a serine-rich region (SRR), an inhibitory module (IM), and an ETS domain. (B) The amide region of the ¹⁵N-HSQC spectra of ¹⁵N-labeled N301 alone (cyan) and in complex with palindromic nonspecific DNA^{NSP12} (5:1 DNA/protein ratio; blue) or with specific DNA^{SP12} (1.1:1 DNA/protein ratio; red). Selected peaks are assigned and an * indicates an aliased signal with an actual ¹⁵N chemical shift >134 ppm. The peaks with ¹H^N shifts clustered near 8–8.5 ppm arise from the unfolded HI-1 and HI-2 residues in the DNA complexes. Full spectral assignments are provided in Figure S4.

Figure 2.

Inhibitory helices HI-1 and HI-2 unfold upon binding nonspecific and specific DNA. Shown are chemical shift-based predictions of the secondary structural elements and backbone dynamics (RCI-S² values) of free and DNA-bound $\Delta N301$. The propensities of forming α -helices (red histogram bars), β -strands (blue), or random coil (green) conformations, according to the MICS algorithm,^{32,33} are given for each residue of $\Delta N301$ in (A) its free form and when bound to (B) nonspecific $\text{DNA}^{\text{NSP12}}$ and (C) specific DNA^{SP12} . Although MICS provides normalized scores for all three conformations, only the most probable is presented for each residue. Also shown as black lines are the RCI-S² order parameters;

values decreasing from 1 to 0 indicate increased predicted flexibility. The cartoon above (A) depicts the secondary structural elements of unbound N301 determined by NMR spectroscopy (1R36.pdb; red rectangles; α -helices, blue arrows; β -strands), and that above (B) reflects the MICS analysis of the DNA-bound proteins with HI-1 and HI-2 unfolded. Although conserved, the short helical turn between H1 and S1 is not explicitly named in the standard nomenclature used for ETS domains.

Figure 3.

Amide chemical shift perturbations (CSPs) resulting from binding of specific and nonspecific DNA to N301. Amide CSPs of N301 upon binding DNA^{SP12} (final 1.07:1 DNA/protein ratio) are shown in red and $\text{DNA}^{\text{NSP12}}$ (5:1) in blue. The top cartoon indicates the secondary structure of the bound protein (rectangles, α -helices; arrows, β -strands), including the turn and wing involved in DNA binding by the winged helix–turn–helix motif. Blank values correspond to prolines and unassigned amides in the free and/or bound forms of the protein. These CSPs are analyzed by a projection analysis in Figure S6 and mapped onto the structure of free N301 in Figure 4.

Figure 4.

Ets-1 binds specific and nonspecific DNA by the same canonical ETS domain interface. CSPs due to binding of (A) specific DNA^{SP12} and (B) nonspecific DNA^{NSP12} are mapped onto the structure of free N301 (1R36.pdb; left and right panels are rotated by 180° about the vertical axis). The different CSP ($\Delta\delta$) color codes are indicated, and gray identifies prolines or amides for which $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ assignments were not obtained in the free and/or bound forms of the protein. CSPs in the IM arise as helices HI-1 and HI-2 are unfolded in both bound complexes. In addition, large CSPs with DNA^{SP12} map to the well-characterized DNA-binding interface of the ETS domain that is centered on the recognition helix H3. In

contrast, smaller CSPs in this interface result from DNA^{NSP12} binding. Also found in this interface are several residues (purple) whose signals were absent in the ¹⁵N-HSQC spectra of the DNA^{NSP12} complex due to exchange broadening, yet were assigned in the spectra of the free and DNA^{SP12} bound forms. With an estimated $K_D \approx 10 \mu\text{M}$, N301 should be >99% bound in the presence of a 5-fold molar ratio of DNA^{NSP}, and thus the smaller CSPs relative to the specific complex do not reflect incomplete saturation.

Figure 5.

Amide heteronuclear ${}^{15}\text{N}$ -NOE measurements indicate that helices HI-1 and HI-2 of the inhibitory module are unfolded, but not conformationally unrestricted, in the Ets-1 DNA complexes. ${}^1\text{H}\{^{15}\text{N}\}$ -NOE values of ΔN301 in its free (black) form and when bound to DNA^{SP12} (red, upper) and $\text{DNA}^{\text{NSP12}}$ (blue, lower). DNA/protein ratios are as in Figure 3. Values near ~ 0.8 indicate well-ordered amides, whereas decreasing values indicate increasing flexibility on the subnsec time scale. Missing residues correspond to prolines, and those with weak, overlapping, or unassigned signals. Histogram bars with values below -1.0 have been clipped.

Figure 6.

Characterizing the interactions of lysine and arginine side chains with DNA. Selected ^{15}N -HSQC spectral regions showing signals from the lysine $^{15}\text{N}\epsilon\text{H}_3^+$ (left) and the arginine $^{15}\text{N}\epsilon\text{H}^+$ (right) signals of N301 when (A) free, (B) bound to nonspecific DNA^{NSP12}, and (C) bound to specific DNA^{SP12} (pH 6.5 and 13 °C). Because of rapid HX with the water, unprotected lysines are not seen in the absence of DNA. In contrast to the five lysine $^{15}\text{N}\epsilon\text{H}_3^+$ signals observed with the specific complex, only one appears with the nonspecific complex. Use of lysine-optimized HSQC pulse sequences⁷⁷ did not reveal any additional signals. Although the slower exchanging arginines are seen in all three spectra, the $^{15}\text{N}\epsilon\text{H}^+$ signals are exchange broadened in the presence of DNA^{NSP12} and exhibited substantial chemical shift changes when bound to DNA^{SP12}. Spectra of the guanidinium $^{15}\text{N}\eta\text{H}_2^+$ moieties are provided in Figure S9.

Figure 7.

Dynamic properties of the arginine side chains in free and DNA-bound N301.

Arginine $^{15}\text{N}^{\epsilon}$ longitudinal R_1 and transverse R_2 relaxation rate constants and heteronuclear $\{^1\text{H}\}^{15}\text{N}$ -NOE values for N301 in its free form (black) and when bound to nonspecific DNA^{NSP12} (blue) and specific DNA^{SP12} (red). Also presented is a cartoon representation of free N301 (1R36.pdb) with arginines in stick format, docked onto DNA (1MDM.pdb) with HI-1 and HI-2 still folded. Increasing mobility on the sub-nanosecond time scale leads to lower heteronuclear NOE values and slower transverse R_2 relaxation. Of note, the guanidinium groups of Arg391 and Arg394 in the recognition helix H3 are highly flexible in

the free protein, yet well-ordered in the N301/DNA^{SP12} complex. This is certainly due to hydrogen bonding to guanidine bases in the core 5' GGA(A/T)3' motif. The spectra were recorded at 13 °C and pH 6.5 using a 500 MHz spectrometer. Missing data correspond to unassigned or overlapping signals.

Table 1**Ets-1 Constructs and DNA Sequences**

name	sequence	comment
N279	Ets-1 ²⁷⁹⁻⁴⁴⁰	ETS domain with IM and SRR
N301	Ets-1 ³⁰¹⁻⁴⁴⁰	ETS domain with IM
DNA ^{SP9}	5'GCC GGA AGT ^{3'} 3'CGG CCTT CA ^{5'}	9 bp specific DNA ^a
DNA ^{SP12}	5'CAGCC GGA AGTG ^{3'} 3'GTCGG CCTT CAC ^{5'}	12 bp specific DNA ^a
DNA ^{SP14}	5'CCAAGCC GGA AGTG ^{3'} 3'GGTTCGG CCTT CAC ^{5'}	14 bp specific DNA ^a
DNA ^{NSP9}	5'GCACTCTAG ^{3'} 3'CGTGAGATC ^{5'}	9 bp nonspecific DNA ^b
DNA ^{NSP12}	5'CAAAAATTTTG ^{3'} 3'GTTTTTAAAAAC ^{5'}	12 bp nonspecific DNA (palindromic)
DNA ^{NSP15}	5'GATGCAGTGTAGTCG ^{3'} 3'CTACGTCACATCAGC ^{5'}	15 bp nonspecific DNA ^b

^aOn the basis of data in ref 26 and the consensus 5'GGA(A/T)3' motif is underlined in bold.

^bManually randomized to avoid any GGA sequences, to minimize potential hairpin or polymer formation, and to reduce end-fraying with terminal GC base pairs.