Conformational dynamics and the binding of specific and non-specific DNA by the autoinhibited transcription factor Ets-1

Genevieve Desjardins,[†,ǁ] 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 BC, V6T 1Z3, Canada

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

[§] Howard Hughes Medical Institute, Chevy Chase, MD, 20815, USA

[ǁ] Current address: Zymeworks Inc., Vancouver BC, V6H 3V9, Canada

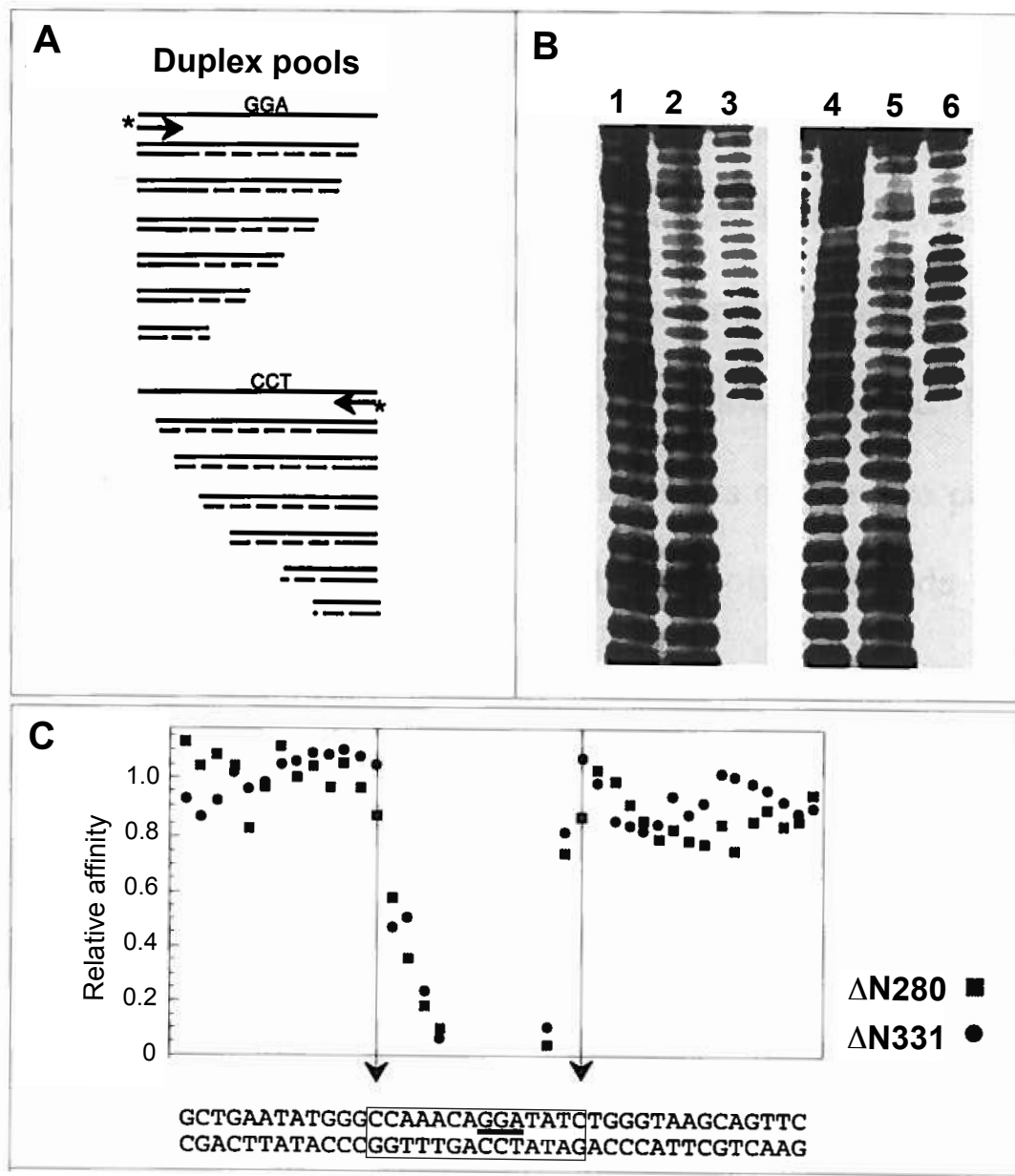[*] Corresponding author e-mail: mcintosh@chem.ubc.ca (phone: 604-822-3341)

Figure S1. Determining the length of an Ets-1 binding site. (A) Schematic showing a site-size selection experiment, where a modified dideoxynucleotide extension procedure was used to generate a pool of truncated DNA duplexes differing in length from 12 to 41 base pairs. The starting ^{32}P-labeled (*) DNA was based on the MSV LTR enhancer,[1] and extension was randomly stopped upon incorporation of a dideoxynucleotide. The probes capable of binding

autoinhibited ΔN280 and uninhibited ΔN331 (lacking the SRR and N-terminal inhibitory helices) were selected from this pool in an electrophoretic gel mobility shift assay. (B) Autoradiography of a denaturing polyacrylamide gel used to size ΔN331-bound and unbound DNA molecules with single nucleotide resolution. Lane 1: the original pool of DNA; lane 2: the unbound fraction; lane 3: the ΔN331-bound fraction. Lanes 4 - 6 are replicates of the first experiment, with the additional step of treating the original pool of DNA with Mung Bean Nuclease to remove single stranded extensions. (C) The plotted relative affinities for each protein were calculated from the band intensities corresponding to duplex oligonucleotides of varying lengths versus that for the longest oligonucleotide. The core GGA is underlined and the optimal 13 - 14 bp binding site is boxed. Uninhibited ΔN331 and autoinhibited ΔN280 exhibited the same DNA length dependence. These experiments were contributed by Qing Ping Xu and reproduced from Gillespie.[2]
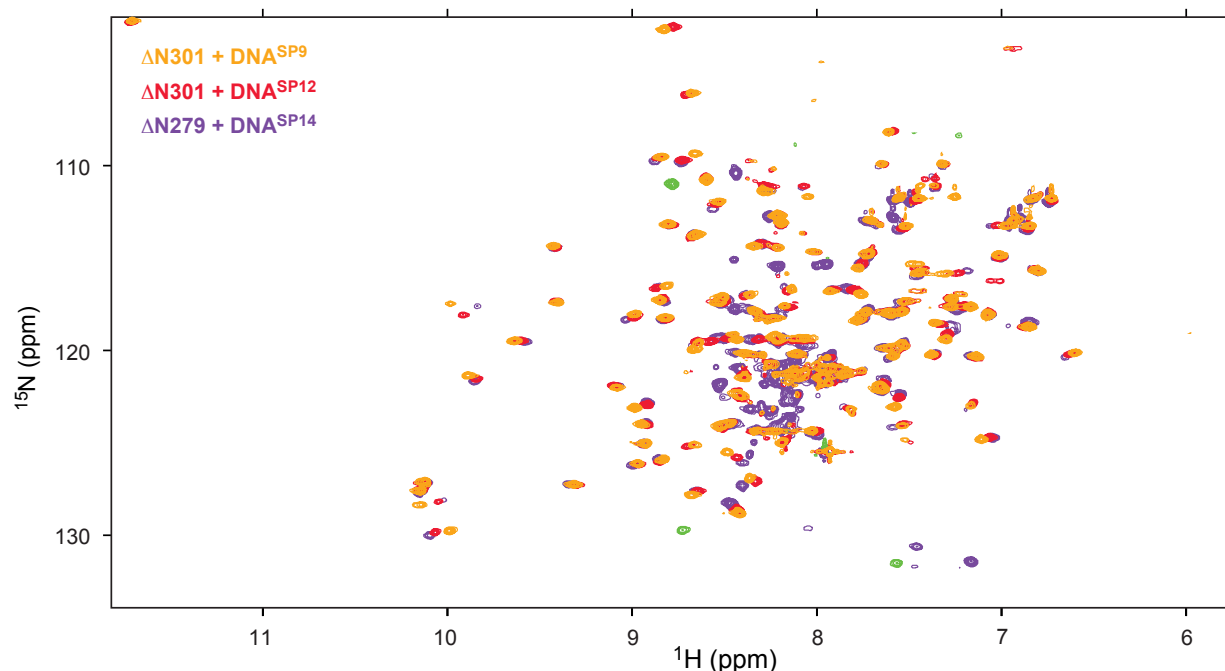
Figure S2. Selection of the oligonucleotide for detailed studies of Ets-1 complexed with specific DNA. Shown are overlaid ^{15}N-HSQC spectra of the following saturated complexes: ΔN301 with DNA^{SP9} (1.5:1 DNA:protein ratio, yellow); ΔN301 with DNA^{SP12} (1.5:1 ratio, red); and ΔN279 with DNA^{SP14} (2:1 ratio, purple/green). Resolved aliased amide and arginine ^{15}N$^{\varepsilon}$H^{+} signals from the ΔN279/DNA^{SP14} complex (green) do not overlap with those from the other complexes due to differing ^{15}N spectral widths. The spectra with DNA^{SP12} and DNA^{SP14} are most similar, whereas DNA^{SP9} yielded slightly smaller chemical shift perturbations relative to the free protein (Figure 1B). Thus, DNA^{SP12} was chosen as the optimal oligonucleotide that recapitulated the full complement of interactions with Ets-1, while also having a small size for favorable NMR spectroscopic behavior. Note also that the similarity of the dispersed signals of most amides in ΔN301 and ΔN279 indicates that the presence of the SRR does not perturb the structure of the bound complexes. The signals from amides in the unstructured SRR and the unfolded HI-1 and HI-2 are clustered around 8 - 8.5 ppm in the ^{1}H dimension.
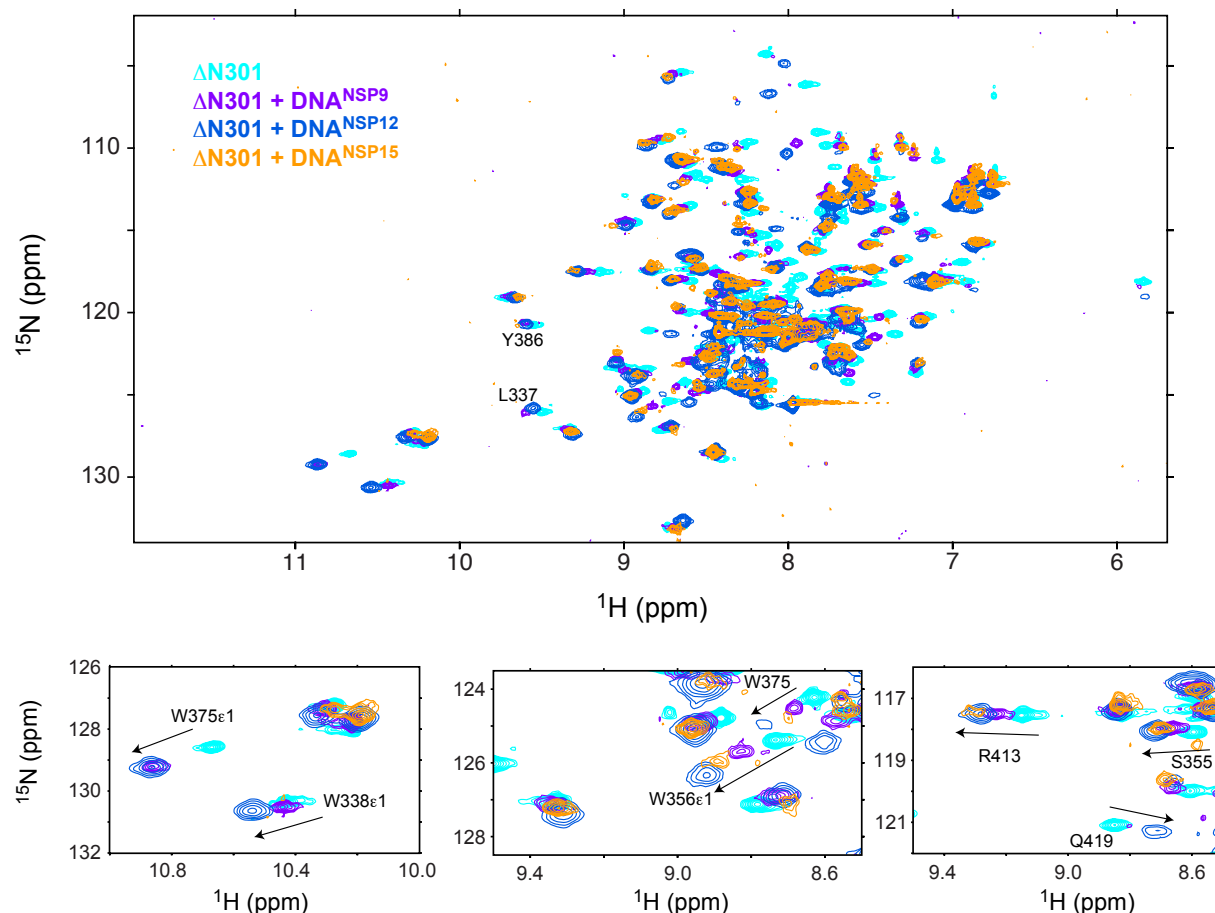
Figure S3. Selection of the oligonucleotide for detailed studies of Ets-1 complexed with non-specific DNA. Shown are ^{15}N-HSQC spectra of ^{15}N-labeled ΔN301 alone (cyan) and in complex with the non-specific oligonucleotides DNA^{NSP9} (3.3:1 DNA:protein ratio, purple), DNA^{NSP12} (5:1, blue), and DNA^{NSP15} (3.8:1, orange) at 28 °C. The three lower panels are expanded views of selected regions. Fewer signals are detected with the DNA^{NSP15} complex than the smaller DNA^{NSP9} or DNA^{NSP12} complexes. For example, signals from the indoles of W338 and W375, and the amides of Q419, L337 and Y386 are broadened beyond detection in the DNA^{NSP15} complex, but are observed with DNA^{NSP9} and DNA^{NSP12}. This is attributed in part to conformational exchange between multiple binding sites on the longer oligonucleotide. In support of this

argument, the complex of ΔN301 with the palindromic DNA[NSP12] yielded better quality spectra than with the other DNAs, possibly due to degenerate binding modes. However, the ΔN301 was ^{13}C-labeled and 70% deuterated for studies with DNA[NSP12], the latter of which also leads to sharper NMR signals.

It is difficult to estimate the stoichiometry, and hence molecular mass, of these non-specific complexes due to the presence of multiple binding sites. At initial titration points with DNA:ΔN301 molar ratios < 1, an oligonucleotide molecule may be bound by more than one protein molecule. However, the end-point spectra presented here were recorded with DNA:ΔN301 molar ratios > 3 and thus the proteins appeared to be saturated and most likely bound via the canonical H3 interface to a single oligonucleotide (albeit with a statistical distribution of K_D-weighted stoichiometries). In support of this argument, the ^{15}N-HSQC spectra of the complexes did not change appreciably when the DNA:ΔN301 ratios were increased from ~ 2 to ~ 3 (not shown). However, ^{15}N relaxation studies (Figure S7) showed that the 1:5 ΔN301:DNA[NSP12] complex had an effective 22 nsec correlation time for global tumbling, which is indicative of higher order oligomerization.

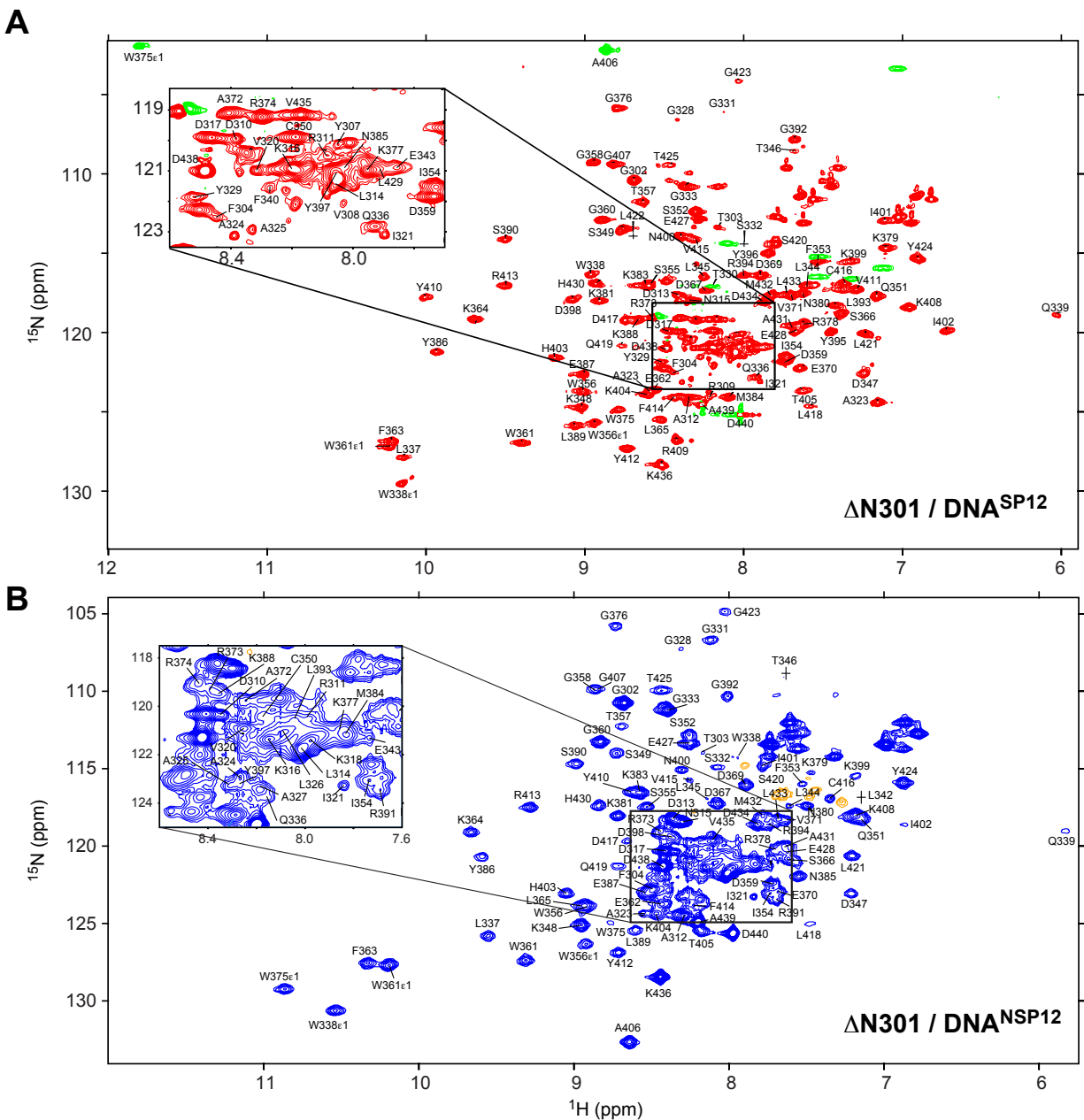**Figure S4. Assigned amide ^{15}N-HSQC spectra of DNA-bound ΔN301.** ^{15}N-HSQC spectra of the saturated (A) specific DNA^{SP12} complex (1.1:1 DNA:protein ratio) and (B) non-specific palindromic DNA^{NSP12} complex (5:1 DNA:protein ratio) recorded on an 850 MHz spectrometer. Aliased backbone amide and indole and arginine sidechains signals are colored green and yellow, respectively. Samples were in 20 mM MES pH 6.5, 50 mM NaCl, 0.5 mM EDTA,

0.02% NaN3, 5 mM DTT and 5% D_2O at 28 °C. Several isotopic labelling schemes (Material and Methods) in combination with conventional and TROSY-based experiments allowed the assignment of signals from at least three nuclei for ~ 94 % of the residues in the specific complex and 86 % for the non-specific complex. Most of the missing signals are from amides located in helices HI-1 and H1 and at the DNA-binding interface.
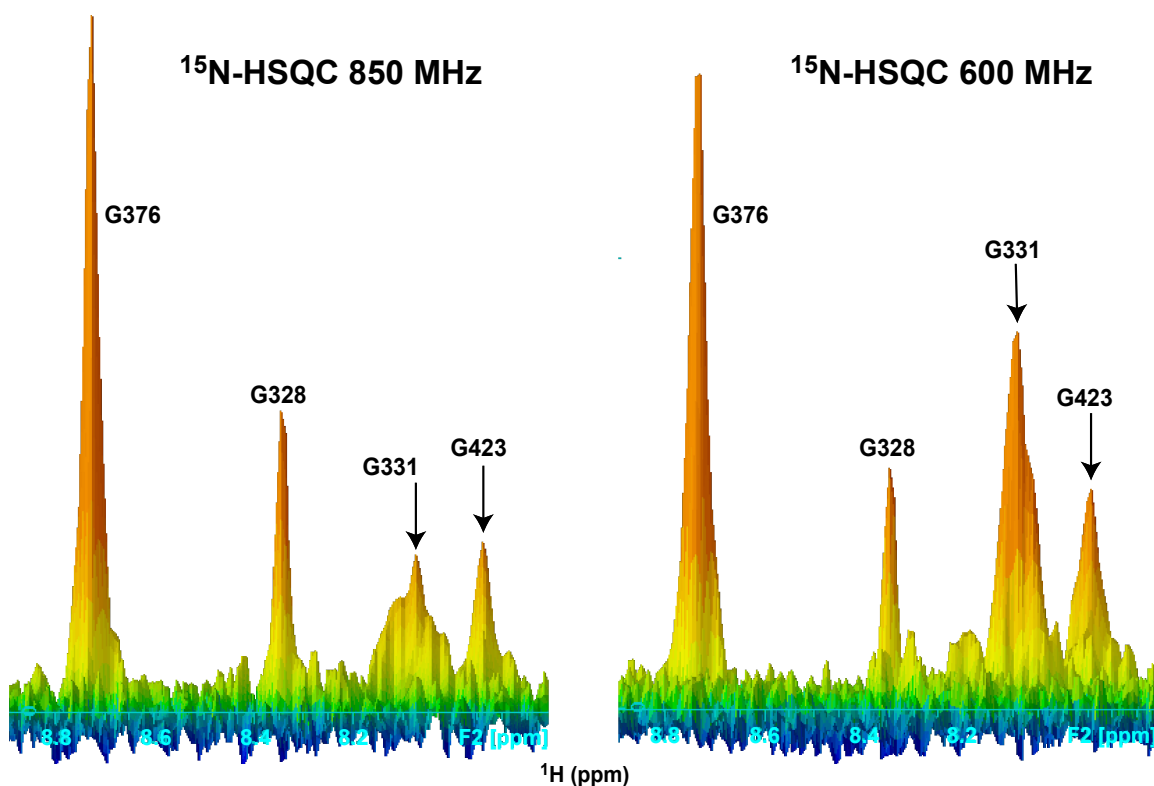
Figure S5. Although predominantly unfolded, residues corresponding to helices HI-1 and HI-2 undergo conformational exchange in both specific and non-specific complexes. Shown are ^{1}H projections through a small ^{15}N region of the ^{15}N-HSQC spectra of the ΔN301/DNA[SP12] complex recorded with 850 and 600 MHz NMR spectrometers. The signals (arrows) corresponding to Gly331 (at the C-terminus of HI-2) and Gly423 (at the interface between helices HI-1 and H4) are weak and their intensities decrease with increasing spectrometer field strength (see also Supplemental Figure S4). This is diagnostic of conformational exchange in the intermediate to fast exchange regime.[3] Similar behavior was observed with non-specific DNA (not shown). Given that ΔN301 is fully saturated with DNA[SP12] or DNA[NSP12], we attribute this behavior to exchange of inhibitory module (HI-1 and HI-2) residues between unfolded and possible folded conformations within the context of the bound complexes.
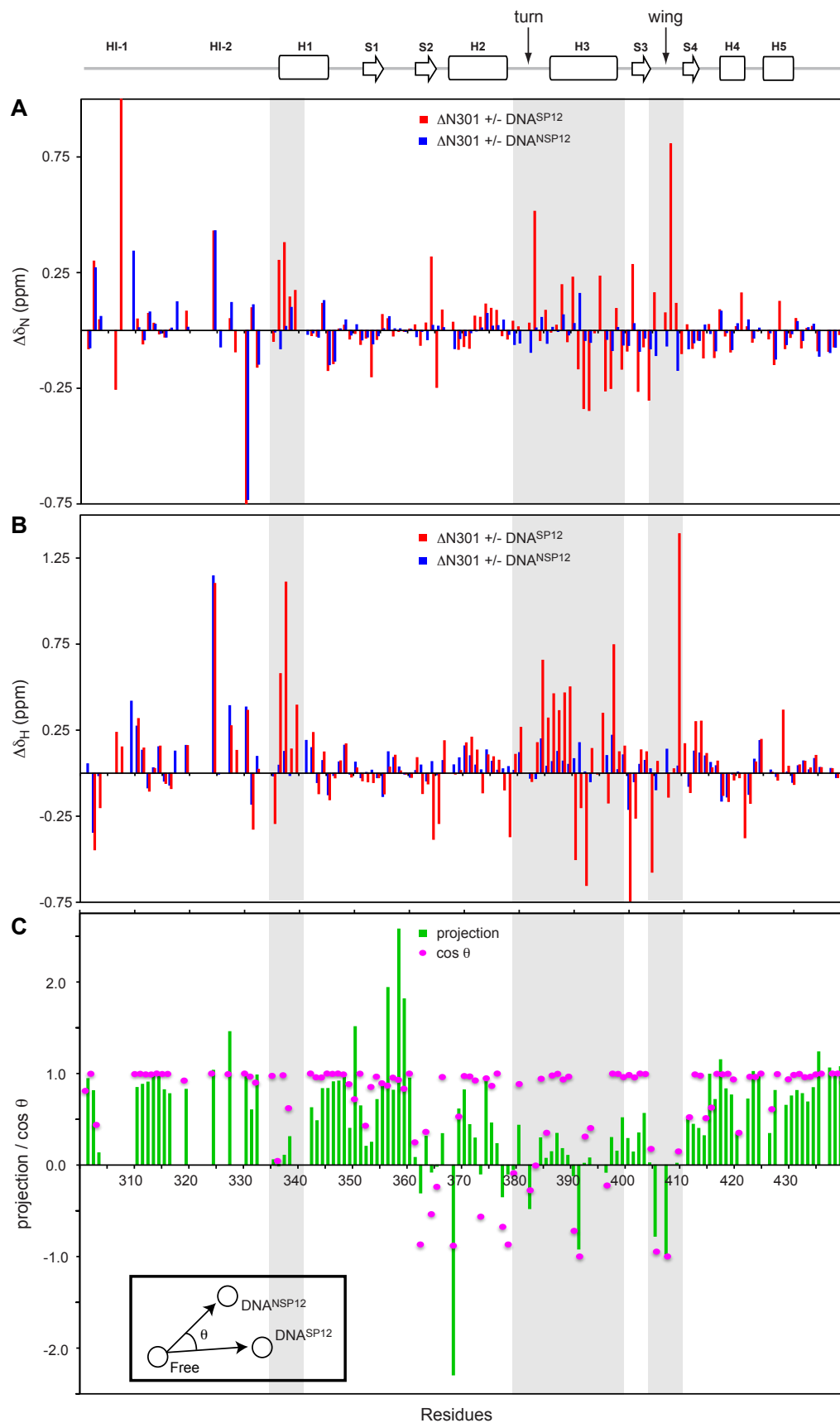
Figure S6. Comparison of the magnitudes and signs of the ΔN301 amide chemical shift perturbations (CSPs) resulting from binding specific versus non-specific DNA. The histograms show the magnitudes and signs of the amide (A) ^{15}N ($\Delta\delta_N$) and (B) ^{1}H^N ($\Delta\delta_H$) chemical shift changes in ΔN301 upon forming the specific DNA^{SP12} (red) and non-specific DNA^{NSP12} (blue) complexes. (C) The relative magnitudes and directions of the CSPs for each residue in the two complexes are also compared by a vector projection analysis (lower inset diagram).[4] The green bars show the projections of the CSP vectors for the non-specific complex along the corresponding vectors for the specific complex. The cosines of the angles θ between the two vectors are given by the magenta dots. Shaded boxes indicate the primary residues involved in DNA binding. Blank values correspond to prolines or unassigned amides.

Residues associated with helices HI-1, HI-2 and those in the helices onto which they pack (H1/H4/H5), exhibit projection values ~ 1 and cos(θ) ~ 1. This pattern indicates that their corresponding amide ^{15}N and ^{1}H^N signals shift in the same direction, with a similar magnitude, upon binding DNA^{NSP12} relative to DNA^{SP12}. This is further evidence that HI-1 and HI-2 unfold upon binding either DNA oligonucleotide. As seen in Figure 1B (assignments in Figure S4), a second subset of peaks shows amide ^{15}N and ^{1}H signals that shift in the same direction (cos(θ) ~ 1), but with a magnitude < 1, upon binding non-specific versus specific DNA. Many of these residues are located in helix H1 (e.g., Leu337), the wHTH turn, and the N-terminal portion of the recognition helix H3 (e.g., Tyr386 and Ser390), which is on the edge of the DNA binding motif. This indicates that ΔN301 forms generally similar, albeit less well-defined, time-averaged interactions with DNA^{NSP12} that 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, rather than base-specific hydrogen bonds. In

addition, 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}. As also seen in Figure 1B, a third subset of amides, such as Gly392, Arg394, Tyr397 and Ala406 located in helices H2, H3 and strand S3, show distinctly different patterns of chemical shift changes (including changing in opposite directions, $\cos(\theta) \sim -1$, and hence yielding negative projection values). This indicates that these regions of ΔN301 interact differently with specific and non-specific DNA. These differences likely reflect the formation of base-specific hydrogen-bonding contacts with the $^{5'}$GGA(A/T)$^{3'}$ motif in DNA^{SP12} that are absent in the looser non-specific DNA complex. Collectively, these analyses support the conclusion that Ets-1 binds non-specific and specific DNA by a similar canonical interface of the ETS domain.
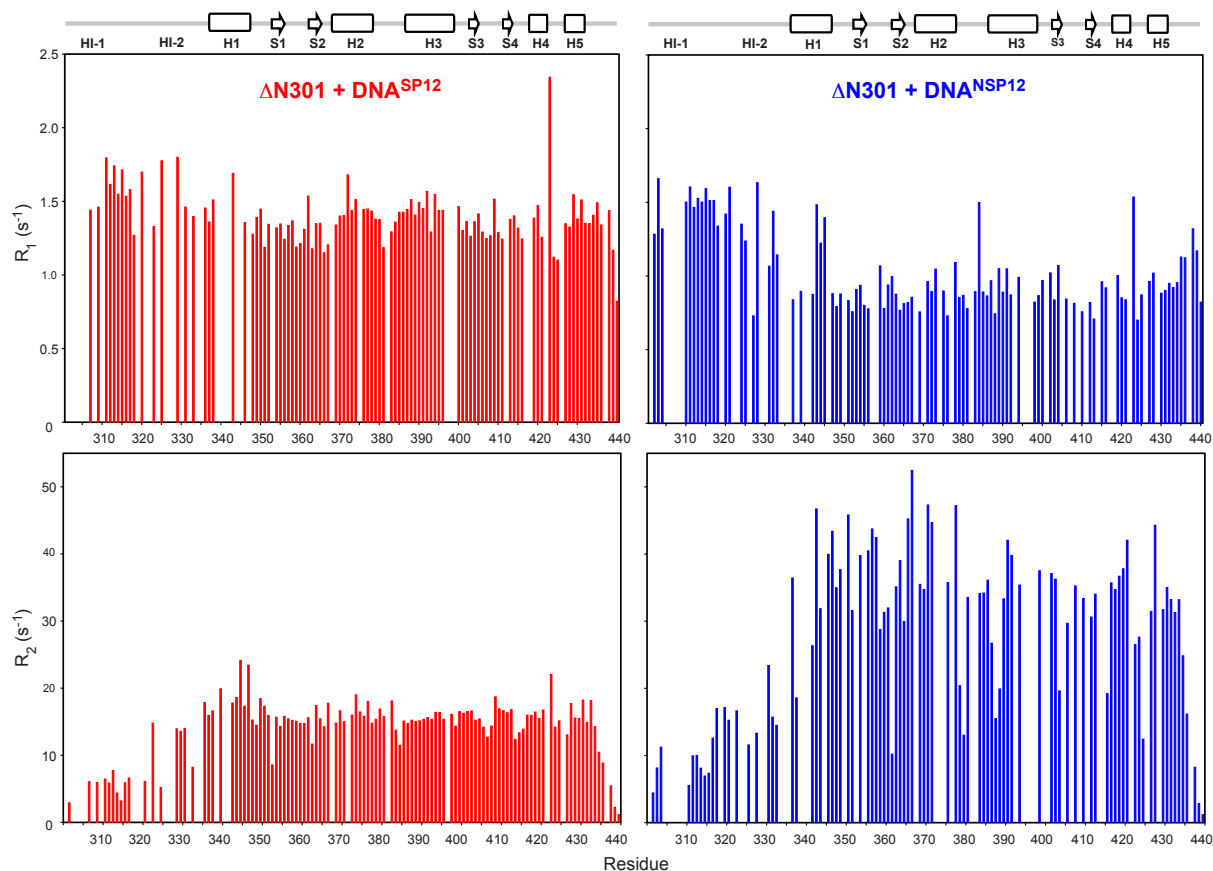
Figure S7. Amide longitudinal (R_1) and transverse (R_2) relaxation of ΔN301 bound to specific and non-specific DNA. Amide ^{15}N R_1 (upper) and R_2 (lower) relaxation rate constants for ΔN301 in the saturated specific DNA^{SP12} complex (1.1:1 DNA:protein ratio; left, red) and non-specific DNA^{NSP12} complex (5:1 DNA:protein ratio; right, blue). Samples were in 20 mM MES pH 6.5, 50 mM NaCl, 0.5 mM EDTA, 0.02% NaN$_3$, 5 mM DTT and 5% D$_2$O and the data were recorded at 31 °C using a 500 MHz spectrometer. Decreasing R_1 and increasing R_2 values reflect slower motion of the $^1H^N$-^{15}N bond vector.[3,5] and analysis of these data for well-ordered ETS domain residues yielded global tumbling correlation times of 12.3 nsec and 22.0 nsec for the DNA^{SP12} and DNA^{NSP12} complexes, respectively. Fitting to an anisotropic rotational diffusion tensors with modeled structure of the Ets-1/DNA complex did not change these results. The

correlation times demonstrate that ΔN301 forms a 1:1 complex with DNA^{SP12} and higher order oligomers with DNA^{NSP12}. Consistent with the ^{1}H{^{15}N}-NOE data of Figure 5, the faster R_1 and slower R_2 relaxation of amides within the N-terminal IM confirm that helices HI-1 and HI-2 are unfolded in the presence of either DNA.
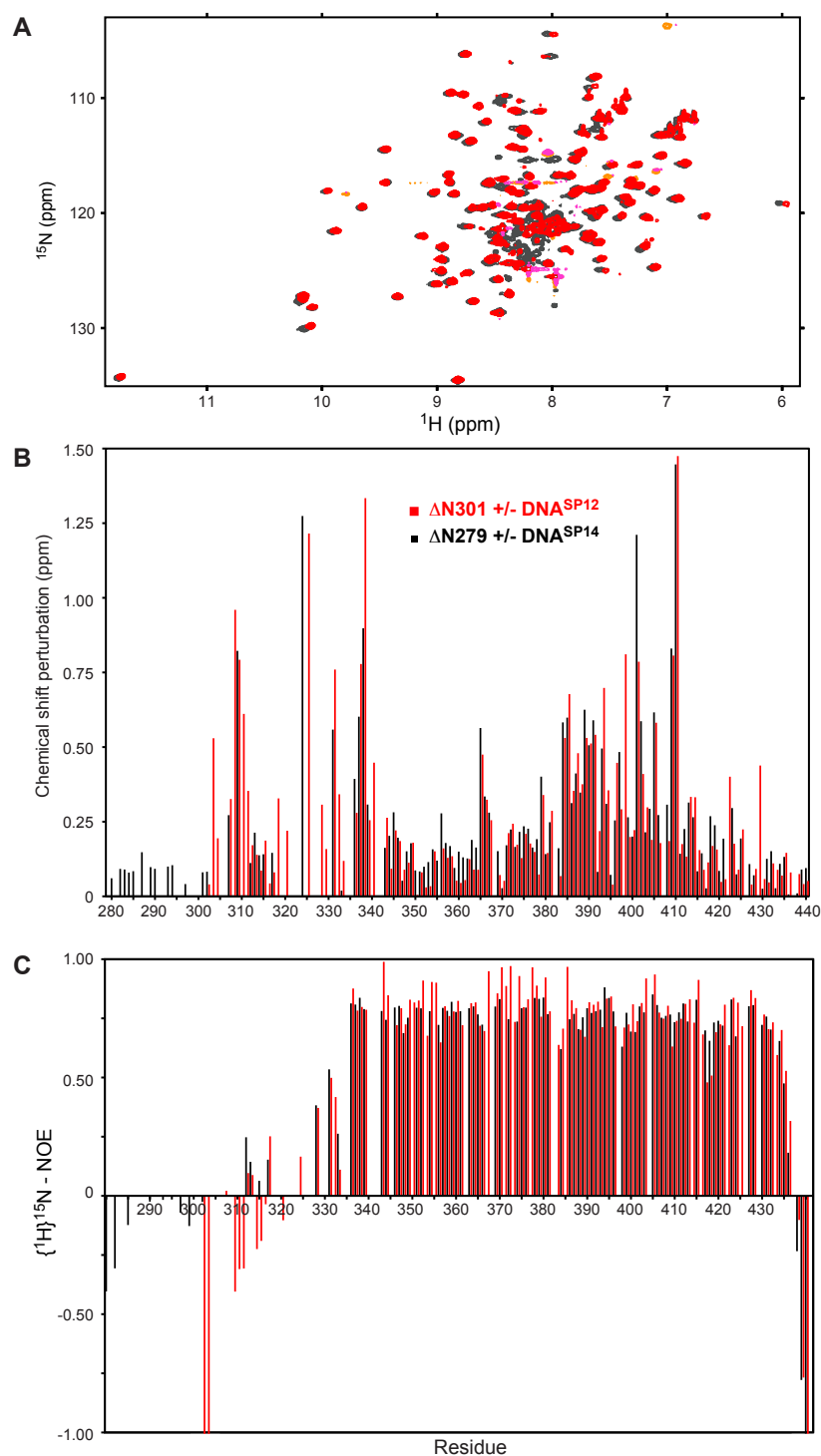
Figure S8. The presence of the SRR does not change the DNA-binding interface or dynamics of Ets-1 bound to specific DNA. (A) Overlaid ¹⁵N-HSQC spectra of the saturated

(red/pink) ΔN301/DNA^{SP12} (1.1:1 DNA:protein ratio) and (black/orange) ΔN279/DNA^{SP12} (2:1 DNA:protein ratio) complexes. Aliased signals are in pink or orange, respectively. (B) Chemical shift perturbations observed upon binding specific DNA by ΔN279 and ΔN301. (C) Heteronuclear ^{1}H{^{15}N}-NOE observed in both complexes at 31 °C using a 500 MHz spectrometer. Histogram bars with values below -1.0 have been clipped. Samples were in 20 mM MES pH 6.5, 50 mM NaCl, 0.5 mM EDTA, 0.02% NaN3, 5 mM DTT and 5% D$_2$O. With the exception of the additional signals from the disordered SRR residues (with random coil ^{1}H^{N} chemical shifts near 8 - 8.5 ppm), the two complexes show very similar spectra, CSPs, and relaxation behavior.
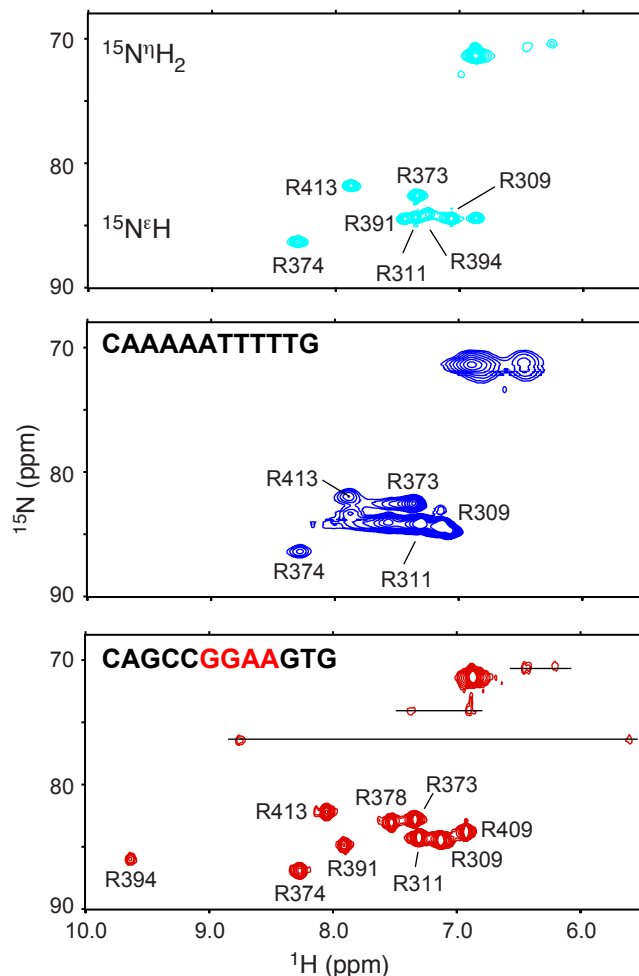
Figure S9. Characterizing the interactions of ΔN301 arginine sidechains with DNA. Selected ^{15}N-HSQC spectral regions showing ^{15}N$^{\varepsilon}$H^{+} (~ 85 ppm) and ^{15}N$^{\eta}$H$_2^{+}$ (~ 70 ppm) signals from the arginine guanidinium moieties of ΔN301 when (A) free, (B) bound to non-specific DNA^{NSP12}, and (C) bound to specific DNA^{SP12} (pH 6.5 and 13 °C). Due HX and/or rotation about the N$^{\varepsilon}$-C$^{\zeta}$ and C$^{\zeta}$-N$^{\eta}$ partial double bonds,[6] only broad ^{15}N$^{\eta}$H$_2^{+}$ signals are seen with free and DNA^{NSP12}-bound ΔN301. In contrast, when bound to DNA^{SP12}, at least three arginines yielded sharper, dispersed ^{15}N$^{\eta}$H$_2^{+}$ signals that are non-degenerate (connected by horizontal lines) due to restricted rotation within the guanidinium moiety. This likely arises from hydrogen bonding of the arginine sidechains with DNA^{NSP12}.

Supporting Information References

[1] Nye, J. A., Petersen, J. M., Gunther, C. V., Jonsen, M. D., and Graves, B. J. (1992) Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev. 6*, 975-990.

[2] Gillespie, M. E. (1998) Structural and genetic investigation of DNA binding by the murine Ets-1 Ets domain, a winged HTH transcription factor. *PhD thesis, Univ. Utah.*

[3] Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta 1814*, 942-968.

[4] Selvaratnam, R., Mazhab-Jafari, M. T., Das, R., and Melacini, G. (2012) The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PLoS One 7*, e48707.

[5] Dosset, P., Hus, J. C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR 16*, 23-28.

[6] Henry, G. D., and Sykes, B. D. (1995) Determination of the rotational dynamics and pH-dependence of the hydrogen exchange rates of the arginine guanidino group using NMR spectroscopy. *J. Biomol. NMR 6*, 59-66.