

Articles

Analysis of the 2.0 Å Crystal Structure of the Protein–DNA Complex of the Human PDEF Ets Domain Bound to the Prostate Specific Antigen Regulatory Site^{†,‡}

Yangzhou Wang,^{*,§} Lin Feng, Meriem Said, Sophia Balderman,^{||} Zahra Fayazi,[⊥] Yu Liu,[⊥] Debashis Ghosh, and Andrew M. Gulick

Hauptman Woodward Medical Institute, 73 High Street, Buffalo, New York 14203

Received December 16, 2004; Revised Manuscript Received February 18, 2005

ABSTRACT: PDEF, a prostate epithelial specific transcription factor, is a member of the Ets family of DNA binding proteins. Here we report a 2.0 Å crystal structure of the PDEF Ets domain in complex with a natural, high-affinity DNA binding site in the promoter/enhancer region of the human prostate specific antigen gene. Comparison of the PDEF–DNA complex with other Ets complexes revealed key features that are shared among Ets members, as well as important differences in substrate specification at both the “GGA” core and the flanking regions of the DNA site. The combination of the serine residue at position 308 and the glutamine at position 311 explains the previous observation that the PDEF binds preferentially to a thymine at the +4 position of its binding site. Despite the common essential features that are shared among Ets members, PDEF demonstrates distinct patterns of interactions at different positions of DNA in achieving sequence specific recognition. Collectively, the common and unique interactions with both the DNA bases and the backbone phosphates lead to substrate specificity and individual preference for certain DNA sites.

Interactions between transcription factors and their DNA targets are crucial in virtually all aspects of development and differentiation. Typically, those regulatory proteins contain

a DNA binding domain that is responsible for direct binding to its target site and is sensitive to other regulatory cues (see ref 1 for a review). Mutations that alter the patterns of DNA binding by transcription factors have been associated with diseases in humans, including cancers and other developmental disorders.

The Ets¹ DNA binding domain was first identified as one class of such DNA binding domains that binds to a target site containing a “GGA” core sequence in the middle. It has

[†] Supported by the Dr. Louis Sklarow Memorial Fund (Y.Z.W.) and the Richard W. and Mae Stone Goode Trust (Y.Z.W.) and NIH Grant GM-068440 (A.M.G.). Additional support for the manuscript comes from the Hewlett-Packard Laboratories.

[‡] The coordinates and structural factors have been deposited into the Protein Data Bank as entry 1YO5.

^{*} To whom correspondence should be addressed: Advanced Studies, Hewlett-Packard Laboratories, 1501 Page Mill Rd., Mail Stop 1169, Palo Alto, CA 94304. Telephone: (650) 857-5065. Fax: (650) 857-4146. E-mail: yangzhou.wang@hp.com.

[§] Present address: Hewlett-Packard Laboratories, Palo Alto, CA 94304.

^{||} Present address: State University of New York Upstate Medical University, Syracuse, NY 13210-2375.

[⊥] Present address: State University of New York, Buffalo, NY 14260-1600.

¹ Abbreviations: Ets, E twenty-six avian erythroblastosis virus oncogene; PDEF, prostate-derived Ets factor; PSA, prostate specific antigen; AR, androgen receptor; IPTG, isopropyl β-D-thiogalactopyranoside; DTT, dithiothreitol; GST, glutathione S-transferase; TEV, tobacco etch virus; HPLC, high-performance liquid chromatography; DIG, digoxigenin; PEG, polyethylene glycol; GABPα, GA-binding protein α; rms, root-mean-square; CPK, Corey, Pauling, and Koltun coloring scheme; TEAB, triethylammonium bicarbonate.

been shown that Ets domains of transcription factors usually function in coordination with other protein factors (2, 3). Recent structural studies revealed the complex interactions among certain Ets domains, their DNA, and other regulatory proteins (4–8), indicating that Ets DNA binding domains are subject to regulation and modulation externally in achieving an additional level of specificity.

It has also been shown, however, that “innate” DNA binding specificity exists as a result of the conserved and nonconserved protein residues, as well as the sequence-dependent structural properties of the DNA substrate (9–12). Unique features associated with individual Ets members are sufficient to confer preferences over different DNA binding sites without additional regulatory proteins. Apparently, inherent structural features are able to determine the DNA binding specificities for individual Ets members.

PDEF is a prostate epithelial specific transcription factor that is involved in the regulation of prostate specific antigen (PSA) expression and also acts as a coregulator of the androgen receptor (AR) (13). The full-length 335-amino acid PDEF protein contains a C-terminal Ets domain that binds to the PSA promoter/enhancer region; the Ets domain is composed of 88 residues from position 248 to 335. Characterization of DNA binding sites revealed a consensus sequence typical for an Ets factor that contains the GGA core. The PDEF Ets domain, however, demonstrates a preference for GGAT rather than the GGAA preference that was observed for other characterized Ets factors (13). Comparison of PDEF with other Ets factors indicates a high level of homology at the primary sequence level but also identifies important differences.

As a step toward understanding the structural foundation for PDEF functions and the ability of the PDEF Ets domain to bind GGAT sequences, we determined the structure of the PDEF Ets domain in complex with a native, high-affinity E-site located in the promoter/enhancer region of the human PSA gene. The results reveal common features, as well as interesting differences, when compared to those of other Ets structures. Two residues, serine 308 and glutamine 311, that are unique to PDEF appear to be responsible for the preference of PDEF over GGAT sequences. In addition, PDEF Ets interacts with DNA flanking regions with different amino acid residues. Our data contribute to the identification of the array of interactions that are essential for the interactions between Ets family members and their DNA partners. Our data and that of others suggest a redundant mode for DNA interaction and a complicated “circuitry” in determining the DNA binding specificity by the Ets domain.

EXPERIMENTAL PROCEDURES

Expression, Purification, Crystallization, and Binding Assays. The Ets domain of PDEF (residues 247–335) was PCR amplified and cloned into an N-terminal GST fusion expression vector using *vaccinia* topoisomerase (S. Balderman et al., unpublished results) under the control of a T7 promoter. The expression plasmid, pGST/Topo-PDEF-Ets, was transformed into *Escherichia coli* BL21(λ DE3) and grown in LB medium at 37 °C until the OD₆₀₀ reached \approx 0.4. Cell culture was subsequently induced with 0.4 mM IPTG at 30 °C for 4 h. Cells were harvested and resuspended in buffer A [100 mM Tris (pH 8.0), 500 mM NaCl, 1 mM

EDTA, 0.1% NP-40, and 5 mM DTT] in the presence of the protease inhibitor cocktail (Sigma) and 1 mg/mL lysozyme (Sigma). The cell suspension was subjected to three rounds of treatment with a French press to ensure a complete lysis. Clear lysate that resulted from the centrifugation was filtered with a 0.2 μ m filter and loaded onto a 5 mL GSTrap affinity column (Amersham Biosciences) equilibrated with buffer A. Bound protein was eluted with buffer B (buffer A and 25 mM reduced glutathione). GST–PDEF Ets fusion protein from the GST peak fractions was dialyzed into buffer C [50 mM Tris (pH 8.0), 25 mM NaCl, 0.5 mM EDTA, and 5 mM DTT] in the presence of TEV protease (14, 15) to remove the GST moiety at 4 °C overnight. The resulting PDEF Ets domain contains a GSLDALGS leader sequence that does not appear to affect the DNA binding activities of the Ets domain, as shown in Figure 1C. Ets binding to the E-site is specific since it does not bind to a control DNA sequence without the consensus PDEF binding cassette (data not shown). The dialyzed sample was filtered again through a 0.2 μ m filter and loaded onto a house-packed Source S (Amersham Biosciences) ion-exchange column equilibrated with buffer C and eluted with a 250 to 500 mM NaCl gradient in buffer C. The PDEF peak fractions were pooled, concentrated, and stored at –80 °C until they were needed. GST affinity chromatography was performed on a BIO-CAD HPLC workstation (Applied Biosystems), and the ion-exchange chromatography was performed on an Akta FPLC workstation (Amersham Biosciences). Samples after each purification or digestion were analyzed with SDS–PAGE and concentration and yield determined via a Bradford assay.

The DNA oligonucleotides used for crystallization and gel shift analyses were synthesized at 1 μ mol scale with the trityl protecting group from Alpha DNA (Montreal, PQ). Oligos were adsorbed onto a Dynamax 300 PureDNA21.4 reverse phase column (Varian) at a flow rate of 20 mL/min. The bound trityl group was removed on the column with 0.5% TFA (5). Final purified DNA was eluted with an acetonitrile gradient and pooled and dialyzed into 10 mM TEAB (pH 7.0). Double-stranded DNA substrate for PDEF Ets crystallization was prepared by annealing the sense strand 5'-TAGCAGGATGTGT-3' to the antisense strand 5'-ACA-CATCTGCTA-3'.

Crystals of the protein–DNA complex were prepared by mixing at a DNA:protein molar ratio of 1.2:1 with the PDEF Ets concentration around 0.4 mM. The protein/DNA mixture was dialyzed into a minimal buffer [10 mM Tris (pH 8.0) and 1 mM DTT] (5) and concentrated again to ensure the protein concentration was around 0.4 mM. PDEF Ets and DNA complex crystallization trials were performed by hanging drop vapor diffusion experiments using a 500 μ L reservoir solution of 100 mM citrate (pH 5.0) and 20% PEG 4000 and a mixture of 1 μ L of the complex and 1 μ L of well solution. Crystals were obtained at 22 °C in 3–5 days. The larger crystals are of thick plates with a typical size of \sim 0.2 mm \times 0.4 mm \times 0.6 mm, and the largest crystal can reach more than 1 mm in the longest dimension.

For the biochemical assay of DNA binding by the PDEF Ets domain, gel shift reactions were performed with procedures described previously (16). The protein and DNA were incubated in binding buffer [100 mM HEPES (pH 7.6), 5 mM EDTA, 50 mM (NH₄)₂SO₄, 5 mM DTT, 1% (w/v) Tween 20, and 150 mM KCl], along with 50 ng/ μ L poly-

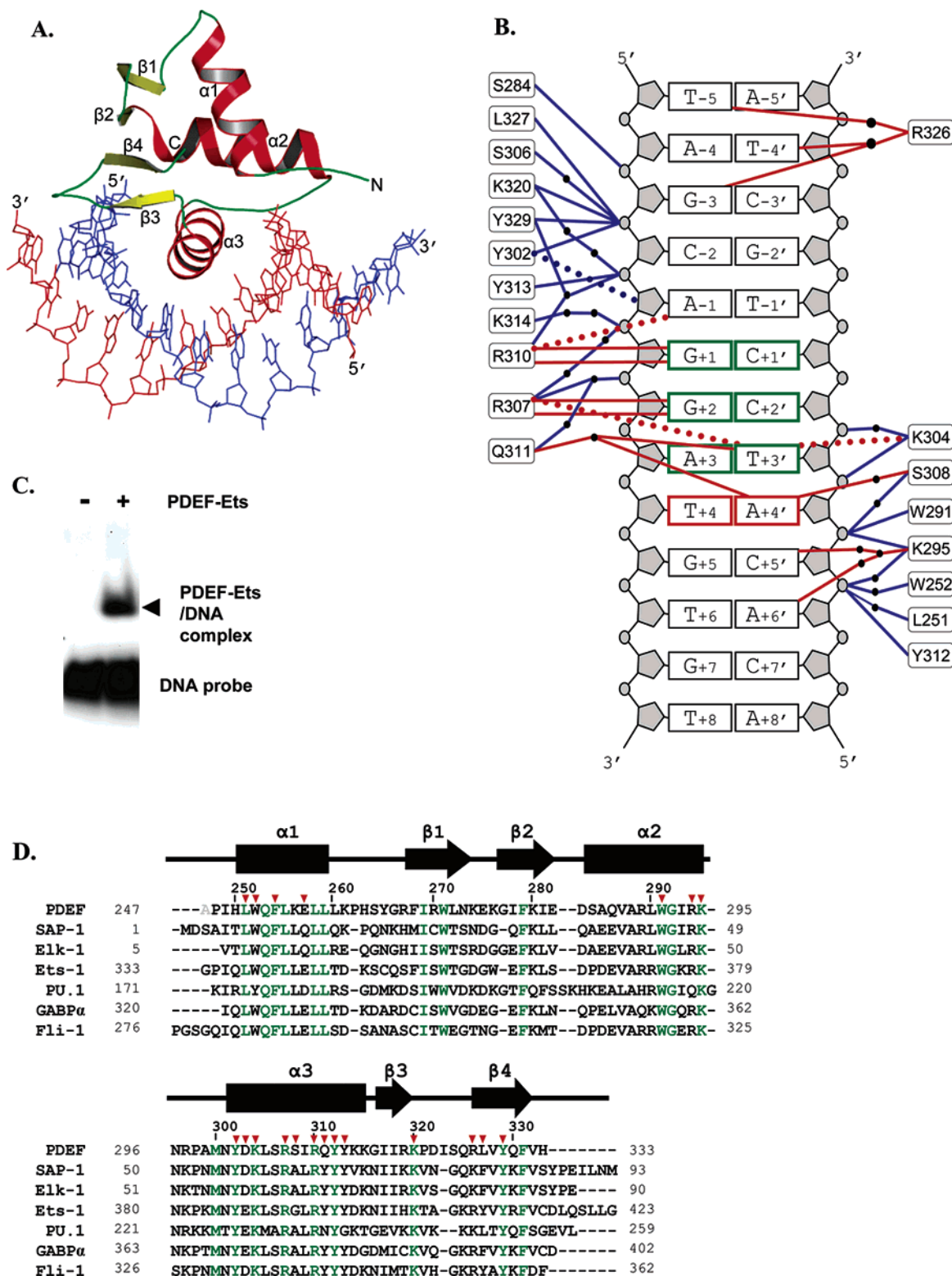


FIGURE 1: PDEF Ets–DNA complex and protein–DNA interactions. (A) Ribbon diagram of the PDEF Ets–PSA enhancer E-site complex. α -Helices are colored red and β -strands yellow. The sense DNA strand is colored blue and the antisense strand red. (B) Schematic diagram of protein–DNA interactions. DNA contacts through hydrogen bonding are depicted as solid lines and van der Waals contacts as dotted lines. DNA base specific contacts are depicted as red lines and backbone contacts as blue lines. Water molecules are represented by black balls. The core GGA region is colored green and the +4 thymine position red. (C) Association of PDEF Ets with the PSA E-site. Gel shift experiments were performed as described in Experimental Procedures. (D) Multisequence alignment of existing Ets structures in the Protein Data Bank. Secondary structural features are indicated above the aligned sequences. Black solid blocks represent α -helices and arrows β -strands. Conserved amino acid residues are represented by green letters. Residues discussed in the text are denoted with red triangles on top of the alignment. The first alanine residue of the PDEF Ets is not part of its natural sequence, rather, a modeled residue due to the disordered side chain of Q247 at this position.

d(I/C) (Roche) and 0.1 $\mu\text{g}/\mu\text{L}$ BSA (New England Biolabs). Approximately 70 fmol of PDEF Ets and 160 fmol of the

labeled probe were used in each gel shift reaction in a final volume of 20 μL . Control reactions with an ~ 125 -fold excess

Table 1: Structural Features of Ets Family Members Bound to DNA^a

protein	DNA used in the complex	overall curvature (deg)	minor groove width (Å)	PDB entry
PDEF	13 bp high-affinity E-site of PSA promoter	16.30	7.46	1YO5
Elk-1 (12)	high-affinity binding site of <i>Drosophila</i> E74 promoter	23.80	7.86	1DUX
Ets1 (5)	high-affinity GGAA complex	19.90	7.56	1K79
	low-affinity GGAG complex	23.32	7.66	1K7A
SAP-1 (11)	high-affinity binding site of the E74 promoter	11.58	7.66	1BC8
	low-affinity binding site of the <i>c-fos</i> promoter	10.54	7.53	1BC7
PU.1 (34)	in vitro binding site from crystallization screening	28.09	7.56	1PUE

^a DNA curvatures were analyzed with CURVE 5.2 (35). As a reference, typical curvatures for B-DNA and A-DNA have been determined to be 4.47° and 36.8°, respectively. The typical minor groove widths for B-DNA and A-DNA have been determined to be 5.90 and 9.84 Å, respectively (11). High and low DNA affinity is categorically described, which is based upon the prominence of supershifted bands of DNA–protein complexes in electrophoretic gel mobility shift analyses.

Table 2: Crystallographic Data for Ets Data Sets

resolution (Å)	2.0
space group	<i>P</i> 2 ₁
unit cell dimensions	
<i>a</i> (Å)	36.08
<i>b</i> (Å)	71.48
<i>c</i> (Å)	39.13
β (deg)	113.36
Matthews coefficient (Å ³ /Da)	2.32
<i>R</i> _{merge} (%)	8.2 (23.0) ^a
completeness (%)	94.7 (85.4) ^a
<i>I</i> /σ	17.1
no. of observations	41144
no. of reflections	11755

^a Values for the highest-resolution shell (2.07–2.0 Å) are given in parentheses.

of the unlabeled cold probe were used to chase the shifted band on the gel. Reactions were performed at room temperature for 15 min before quenching, and samples were applied immediately onto a 4% nondenaturing polyacrylamide gel and run at 100 V for ~4 h at 4 °C. DNA probes were end-labeled with digoxigenin-11-ddUTP at the 3' end of the DNA using a DIG Second Generation Gel Shift Kit (Roche).

Crystallographic Data Collection and Structure Determination of the PDEF Ets Domain–DNA Complex. Crystals of the protein–DNA complex were transferred to a series of cryoprotectant solutions containing 20% PEG 4000, 100 mM citrate (pH 5.0), and 8, 16, or 24% ethylene glycol for ~20 min at room temperature. Crystals were then transferred to a stream of cryocooled N₂ gas at –170 °C. Data were collected using a Rigaku RU-H3RHB rotating anode equipped with a Cu anode, Osmic Max-Flux confocal focusing mirrors (MSC), and an R-Axis IV detector (Rigaku). Diffraction was observed to 1.7 Å; however, because of increased mosaicity at the highest resolution, the data were processed and used to 2.0 Å. The data were integrated and scaled using DENZO/SCALEPACK (17). The monoclinic cell dimensions suggested that a single 20 kDa Ets domain–DNA complex was present in the asymmetric unit. Data collection statistics are given in Table 2.

The CCP4 suite of programs (18) was used for determining the structure of the protein–DNA complex. The protein–DNA complex of the human Elk-1 Ets domain (PDB entry 1DUX) (12) was used as a search model for MOLREP (19). Water molecules were removed from 1DUX, and the rotation and translation searches were performed with data to 3.0 Å. A single solution was obtained and examined manually for overlaps between symmetry-related molecules. The complex

Table 3: Refinement Statistics

resolution range (Å)	25.0–2.0
<i>R</i> _{cryst} (%) (overall/highest-resolution shell)	20.3 (25.1) ^a
<i>R</i> _{free} (%) (overall/highest-resolution shell)	23.9 (28.7) ^a
Wilson <i>B</i> -factor (Å ²)	25.6
average <i>B</i> -factor, overall (Å ²)	25.0
average <i>B</i> -factor, protein (Å ²) (main chain, side chain)	23.1, 25.7
average <i>B</i> -factor, solvent (Å ²) (no. of molecules)	28.9 (145)
rms deviation for bond lengths (Å), angles (deg)	0.008, 1.328

^a The highest-resolution shell is from 2.05 to 2.00 Å.

was submitted to a round of refinement with REFMAC (20), in which the *R*-factor and *R*-free values dropped to 35.0 and 42.5%, respectively, using data to 2.2 Å. Continued rounds of manual model building to correct nonconserved side chains and the differences in the DNA sequence followed by refinement with REFMAC improved the overall model quality to the final statistics (Table 3). Figures and illustrations were generated with Pymol and VMD suites of programs (21, 22).

RESULTS

Overview of the PDEF Ets Domain in Complex with DNA.

The crystal structure of the PDEF Ets domain in complex with a 13 bp DNA substrate was determined at a resolution of 2.0 Å using the molecular replacement method with PDB entry 1DUX (human Elk-1 Ets domain) as the search model (12). The DNA sequence was derived from a high-affinity E-site in the human PSA promoter/enhancer regulatory region (13). The electrophoretic mobility shift assay shows that the PDEF Ets domain binds its DNA target well in vitro (Figure 1C). One PDEF Ets–DNA complex is present in the asymmetric unit of the crystal. The final model includes the entire DNA substrate and Ets domain residues 248–333 from the PDEF. The side chain of Q247 is disordered and modeled as an alanine. The Ramachandran plot demonstrates that 89.7% of the residues lie in the most favored region with 10.3% in the additionally allowed region. There are no residues in the generously allowed or disallowed region. The final model contains 145 water molecules, and the refinement statistics are given in Table 2.

An overall structure of the PDEF Ets–DNA complex is shown in Figure 1A. In general, the PDEF Ets domain is very similar to the structures of other Ets factors from different species (reviewed in ref 2). It demonstrates three α-helices (α1–α3) and a four-stranded antiparallel β-sheet (β1–β4). The α3 recognition helix penetrates into the major groove of the DNA substrate and makes contacts with the highly conserved GGA core. This recognition helix also

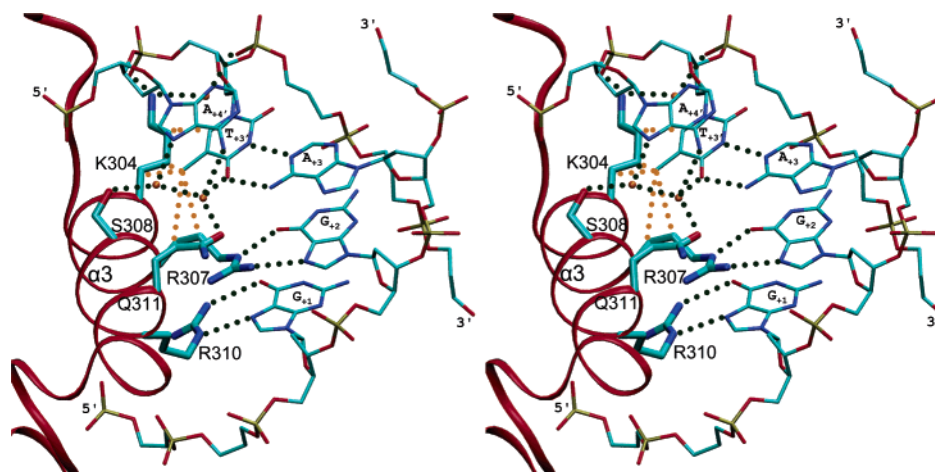


FIGURE 2: Cross-eye stereodigram of the interactions between the residues of the PDEF Ets $\alpha 3$ recognition helix and core region of the E-site. Helix $\alpha 3$ is shown as a red ribbon with side chains of key residues interacting with DNA protruding. Hydrogen bonds are depicted as green dotted lines and van der Waals interactions as orange dotted lines. Water molecules are represented by orange spheres. All cross-eye stereoisomages were generated with VMD (22).

makes a number of contacts with the phosphate backbone of the DNA, as do residues from the $\beta 3$ – $\beta 4$ turn and $\alpha 2$ – $\alpha 3$ turn regions of the protein. In addition, the $\beta 3$ – $\beta 4$ turn and $\alpha 2$ – $\alpha 3$ turn of the PDEF Ets domain colocalize with the ends of the 9 bp Ets–DNA binding site. The DNA substrate in the PDEF Ets complex is curved $\sim 16.3^\circ$ toward the protein and falls within the spectrum of conformations seen in the DNA substrates of other Ets domains (Table 1). With the existing structures, there appears to be no direct correlation between the curvature of DNA and the affinity for an Ets domain. The minor groove, as seen in other Ets–DNA substrates, is also wider than the typical B-form DNA (Table 1).

The DNA substrate within the PDEF crystal makes only a single hydrogen bonding interaction between a phosphate oxygen on C_{-3}' of the antisense strand and the main chain amide nitrogen of K274 of a symmetry-related complex. Base pair stacking has the appearance of being continuous between symmetry-related DNA molecules; however, the DNA molecules are rotated so that they do not form a continuous double helix. In light of the fact that curvature of the DNA has been observed in prior structurally characterized members of the family (Table 1), the observed PDEF DNA curvature is not likely a result from the limited lattice interactions during the crystal packing, but rather is intrinsic to the protein–DNA complex.

DNA Contacts by the PDEF Ets Domain at the GGA Core. The PDEF Ets domain makes a number of contacts with its DNA substrate through bases and the phosphate backbone (Figure 1B). As shown in Figure 2, major interactions with bases occur at the two most conserved arginine residues, R310 and R307, which make key hydrogen bonds with G_{+1} and G_{+2} of the GGA core, respectively. The guanidino side chain of each of the arginine residues makes two hydrogen bonds with the respective guanines. Such interactions are also seen in other published Ets structures (5, 10–12), with the exception of PU.1 which exhibits a single hydrogen bond between the corresponding arginine and each of the two guanine bases in the GGA core (10). In addition, the side chain of R307 also forms van der Waals interactions with the methyl group of the T_{+3}' base.

Q311 of the PDEF Ets domain interacts with T_{+3}' and A_{+4}' on the antisense strand of DNA through water-mediated hydrogen bonding (Figure 2). The interaction of the glutamine with DNA is different from those of residues at the same position of other Ets structures. When the corresponding position of Q311 is a tyrosine (such as Ets-1, Elk-1, and SAP-1), the side chain is subject to rearrangement. As a result, it can either interact with the sense strand bases $3'$ to the GGA core or be reoriented so that its side chain is pointed in a direction parallel to the DNA axis. When oriented toward the DNA bases, tyrosine at this position has been shown to contact DNA via van der Waals and hydrophobic interactions with DNA bases on the sense strand, and with other amino acid residues (5). In the case of the PU.1 Ets domain, the corresponding position is an asparagine and it interacts with T_{+4}' and A_{+3} through water-mediated hydrogen bonds (10). Q311 of the PDEF interacts with antisense residues T_{+3}' and A_{+4}' only through water-mediated hydrogen bonds.

Despite the fact that bases at the $+3$ and $+4$ positions are all contacted by different Ets domains, it appears that each Ets domain displays different patterns of interaction at the glutamine position. The Q311 position has been postulated to be a site subject to regulation by other protein factors (5). Because of the natural helical rotation of the DNA, bases from either strand become accessible to residues of the $\alpha 3$ helix at the PDEF Q311 position. It is conceivable that such accessibility to bases from both strands permits different ways of base specification by various residues in the Ets domains, both in terms of the contacting bases from different strands and the types of interactions with those bases. Furthermore, such flexibility could provide varying levels of regulation by other protein partners, as shown in the case of the mouse Ets1 (5, 8).

The corresponding position of the PDEF S308 in other Ets domains is typically an alanine or glycine (Figure 1D). PDEF S308 interacts with DNA through a water-mediated hydrogen bond at the A_{+4}' position on the antisense strand (Figure 2). This observation is in agreement with an earlier modeling prediction of such an interaction with serine at this position (11).

The mechanism of base specification at the $+3$ and $+4$ positions is interesting among Ets members. The correspond-

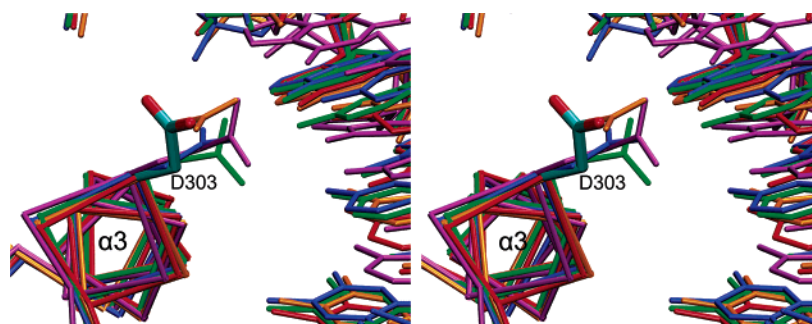


FIGURE 3: Cross-eye stereodigram of superimposition of five Ets–DNA complexes at the D303 position of PDEF. The side chain of D303 of PDEF is shown in CPK color scheme. Color schemes for the coils, other side chains, and DNA are as follows: PDEF Ets in red, Elk-1 (1DUX) in green, Ets-1 (1K79) in orange, SAP-1 (1BC8) in blue, and PU.1 (1PUE) in purple.

ing +4 position in a high-affinity site for another Ets domain, SAP-1, is collectively defined by an alanine (A62) and a tyrosine (Y65), two positions corresponding to those of S308 and Q311 in PDEF. When the low-affinity *c-fos* DNA site is bound by SAP-1, the thymine at the +4 position is contacted by only Y65 of SAP-1. It is possible that the combination of amino acids at the S308 and Q311 positions plays a role in determining a stronger affinity at the +4 position. Another natural variant at position 308 is glycine, and it has been postulated that the presence of a glycine at this position allows more tolerance of different bases at the +4 position after the GGA core (11). Overall, the combination of Q311 and S308 in the PDEF Ets collectively specifies the preference of thymine at the +4 position after the core region.

A highly conserved lysine, K304, makes van der Waals interaction through its C_β , C_γ , and C_δ atoms with T_{+3} (Figure 2). Moreover, it makes a direct hydrogen bond with the phosphate backbone of the T_{+3} position, and a water-mediated hydrogen bond with the backbone of C_{+2} . Those interactions with DNA bases and the backbone are highly conserved among the existing Ets structures and are consistent with the high level of conservation at the lysine position. It is possible that the K304 position is involved in the tethering of DNA along with other residues (see below) to properly orient the DNA molecule.

Previous findings indicate that position 303 (aspartate) adopts different rotamer configurations when interacting with DNA (11, 12). In high-affinity DNA sites, such as the one in the SAP-1–E74 complex, the corresponding aspartate (position 57) makes a water-mediated hydrogen bond with cytosine bases at positions C_{+2} and C_{-1} . In another high-affinity site of the Elk-1–E74 complex, the aspartate (position 58) makes hydrogen bonds with cytosine bases at positions C_{+1} , C_{+2} , C_{-1} , and C_{-2} . With low-affinity sites (1BC7) or with glutamate in place of the aspartate (1K79 and 1PUE), the end atoms of the side chain at this position rotate away so that they no longer point directly toward the DNA and are subsequently unable to make any contacts with the bases. The PDEF Ets domain contains an aspartate at position 303. Despite the presence of a native high-affinity *in vivo* binding site, D303 adopts a rotamer configuration that rotates the end O_δ atoms to a direction that is almost perpendicular to the base pairing plane. Such a configuration does not allow interaction with DNA bases at all (Figure 3). Residues at this position conceivably can be more flexible in interacting with the upstream part of the GGA core since

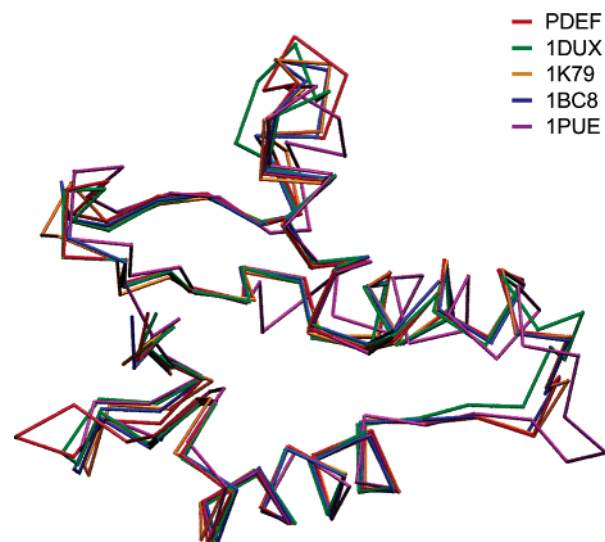


FIGURE 4: Superimposition of the main chains of Ets domains (based upon the C_α positions) from five Ets–DNA binary complexes: PDEF in red, Elk-1 (1DUX) in green, Ets-1 (1K79) in orange, SAP-1 (1BC8) in blue, and PU.1 (1PUE) in purple.

the natural helical turn of DNA allows such accessibility to DNA bases from both the sense and antisense strands, as seen in the cases of Elk-1 and SAP-1. In the case of PDEF, the –1 and –2 positions of the DNA site are open and no contact with DNA is evident (Figure 1B). The presence of the same residue at a structurally aligned position, therefore, does not guarantee the same type of interactions. The lack of D303-mediated DNA contacts is a bit surprising but may help to explain why natural, high-affinity binding sites within the PSA promoter/enhancer regions can accept either adenine or cytosine at the –1 position (13). Overall, it is clear that architectural similarity, as evidenced by the closely matched C_α atoms (Figure 4), does not necessarily correlate with a similar manner of DNA base recognition.

Contacts at the 5' Flanking Region of DNA. In addition to conserved interactions with DNA bases by helix α_3 of the PDEF Ets domain, the Ets domain makes characteristic contacts with phosphate backbones in association with residues from other parts of the protein. The underlined 5' flanking side of the DNA binding site (5'-TAGCAGGAT-GTGT-3') is contacted by a collection of three tyrosines, Y302 and Y313 of helix α_3 and Y329 of strand β_4 . Y302 and Y329 form a network of hydrogen bonds with the phosphate group of C_{-2} (Figure 5A). The phosphate group of A_{-1} is contacted by Y313 through a direct hydrogen bond,

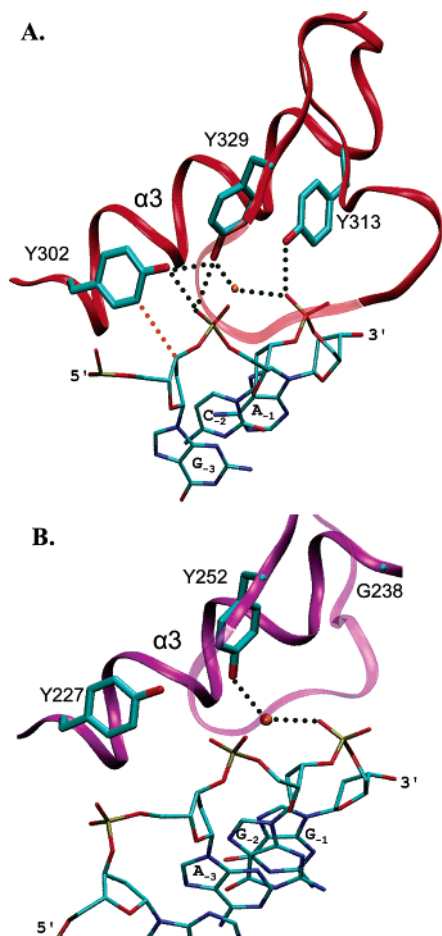


FIGURE 5: Comparison of patterns of DNA backbone interactions of three tyrosines with different Ets factors. The three tyrosines (Y302, Y313, and Y329) of PDEF make backbone contacts at the 5' end of the DNA binding sites that are immediately upstream of the GGA core. Two Ets domains were shown with PDEF (A) in red and PU.1 (IPUE) (B) in purple. Hydrogen bonds are represented by dark green dotted line and van der Waals contacts as orange dotted lines.

and by Y329 through a water-mediated hydrogen bond. Y302 also makes a van der Waals contact with the sugar ring of G₋₃. Comparison of PDEF with the published Ets–DNA complexes shows that the hydrogen bonds with the tyrosines are present in other Ets domains as well. The pattern of interactions is identical in SAP-1, Elk-1, and Ets-1. Those structures all contain the Ets domain with a high-affinity DNA binding site. The PU.1–DNA complex (Figure 5B) demonstrates a different hydrogen bonding pattern compared to the others, although it is not clear whether its binding site is a high-affinity site.

In other cases where an Ets domain is regulated by a protein partner, the corresponding interactions of the three tyrosines are perturbed to various extents. For example, in the GABP α – β –DNA ternary complex where the association of GABP α with DNA is made tighter (4, 23), the tyrosine corresponding to Y329 of PDEF is too far away from the DNA backbone to make a direct hydrogen bond. Nonetheless, at least one or more interactions between the corresponding position of PDEF tyrosines and the DNA backbone are present in all cases. It is conceivable that those tyrosines play some role in properly positioning DNA substrates vis-à-vis their Ets domains.

In the PDEF Ets, the side chain of R326 penetrates into the minor groove of the DNA and makes water-mediated hydrogen bonds with G₋₃, T₋₄, and T₋₅ on both strands (Figure 6A,B). These interactions are absent in other Ets–DNA complexes, although the position is occupied by a positively charged residue with long side chains, typically an arginine or a lysine. In proteins with an arginine, such as the high-affinity complex of human SAP-1, the side chain is positioned over the minor groove and is too far from the DNA to make any base specific interactions. This is also true for the human Elk-1 Ets domain where the corresponding residue is a lysine. It is clear that despite the conservation at this position, the interaction of R326 with DNA bases appears to be a unique feature of the PDEF Ets domain. Of all the reported Ets–DNA complex structures, PDEF is the only member with an arginine side chain penetrating into the minor groove and interacting with DNA bases that are outside the typical 9 bp Ets site.

In addition to the interactions of arginine and tyrosines with DNA at the 5' end, K320 is also a highly conserved residue among all the Ets family members. Its side chain is located above the backbone of the DNA sense strand (Figure 6C), an orientation seen in all Ets–DNA structures. In the PDEF complex, this side chain of lysine makes a direct hydrogen bond to the phosphate group of C₋₂, an interaction seen in some other Ets–DNA complexes. Furthermore, the C₋₂ phosphate backbone is in contact with the amino group of main chain residue L327 (Figure 6C). When the other Ets–DNA structures are examined, K320 and L327 (main chain) appear to form a redundant group to make contacts with the C₋₂ position. In all cases, the C₋₂ backbone is at least contacted by L327, K320, or both.

In every case of the existing Ets–DNA complexes, the side chain of the lysine at the K320 position invariably makes one or more van der Waals contacts with the conserved tyrosine (Y329) located on strand β 4. In PDEF, the C δ and C ϵ atoms of K320 make van der Waals contacts with the C δ and C ϵ atoms of Y329. In other structures, the specific atoms from the lysine that interact with tyrosine may be different, but they all contact one of the C ϵ atoms of tyrosine at the Y329 position. The conservation of such interactions suggests that the K320 position may be involved in the proper positioning of Y329 for it to make its key contacts with the DNA. The direct phosphate backbone contact through hydrogen bonding may be a secondary function of K320 in PDEF. Therefore, it appears that highly conserved residues can function in different contexts through interactions with other residues. Depending upon the circumstance, they display a distinct functional pattern in terms of DNA contacts. This likely contributes to the flexibility of Ets domains in accommodating either different DNA substrates or regulation by other protein factors. Perhaps not surprisingly, multiple natural high-affinity DNA sites that are sufficiently divergent at the 5' flanking region for PDEF exist. Our data and that of others suggest that redundancy in Ets domains for DNA contacts at the 5' end of DNA binding sites exists.

Contacts at the 3' Flanking Region of DNA. Compared to the three tyrosines described above, conservation of interactions of the tyrosine (Y312), tryptophan (W291), and lysine (K295) triad with the DNA backbone at the 3' end of the DNA binding site is more striking. As shown in Figure 7A, Y312, located on the side of helix α 3 opposite from Y313,

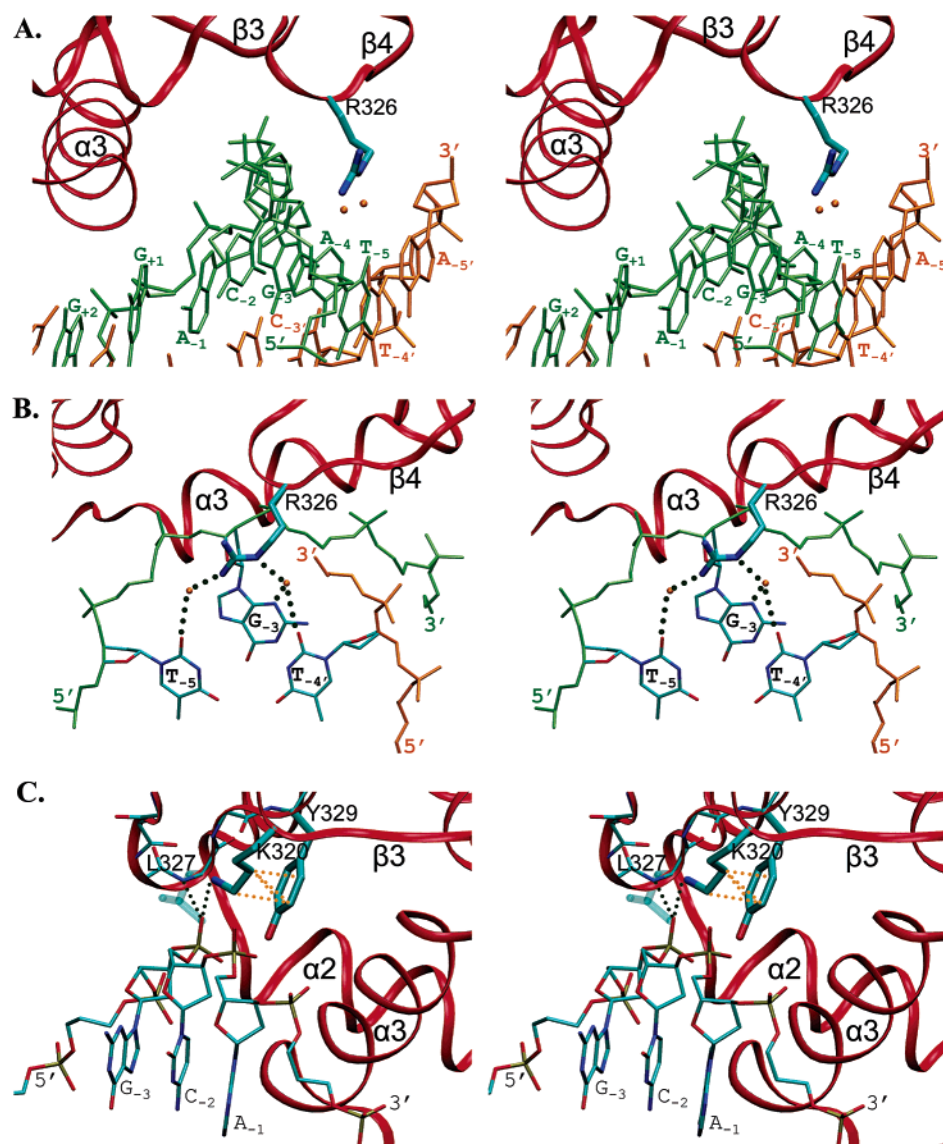


FIGURE 6: Cross-eye stereodiagnostics of interactions of the PDEF Ets at the 5' region of the E-site. (A and B) Two views of base contacts mediated by the PDEF arginine at position 326. In panel A, the R326 side chain is shown to insert into the minor groove of DNA, and in panel B, hydrogen bonds with bases from both the sense and antisense DNA strands are shown. Water molecules are represented by orange spheres. (C) Interactions of the PDEF K320 side chain and the L327 main chain with the backbone of the sense DNA strand and interaction of K320 with Y329. Hydrogen bonds are represented by dark green dotted lines and van der Waals contacts as orange dotted lines.

makes a direct hydrogen bond to the phosphate group of $C_{+5'}$. It also interacts with the main chain amino group of the tryptophan at position 252. In addition to Y312, the main chain amino group of L251 directly interacts with the phosphate group of $C_{+5'}$ through a hydrogen bond. Both L251 and W252 are located at the beginning of helix $\alpha 1$. The phosphate group of the next base, $A_{+4'}$, is contacted by W291 and K295, which are both located on helix $\alpha 2$. In addition, van der Waals contacts are present between W291 and K295.

The pattern of hydrogen bonding made by Y312, W291, and K295, as well as the two main chain atoms of W252 and L251, is present in all the existing structures (a few examples are shown in Figure 7B–E). The van der Waals interactions between W291 and K295 are also consistently present in the reported Ets structures. Those residues appear to function in coordination, and their DNA contacts are not perturbed by the presence of an additional regulatory protein.

Not surprisingly, all three residues (Y312, W291, and K295) are strictly conserved among the Ets domains, whereas W252 and L251 are slightly less conserved, presumably because the side chains do not participate directly in the pattern described above. Those interactions potentially represent essential contacts that must be engaged by the Ets domain as a prerequisite for sequence specific contacts.

In addition to participating in the network of interactions with DNA backbones described above, K295 also interacts with specific DNA bases through water-mediated hydrogen bonds (Figure 8A). Structural alignment of five Ets family members indicates that, despite deviations in the conformation of the loop between helices $\alpha 2$ and $\alpha 3$, the side chain amino groups of these lysine residues superimpose with an rms deviation of 0.07 Å in the proximity for DNA contacts (Figure 8B). However, the pattern of interaction at the PDEF K295 position with $C_{+5'}$ and $A_{+6'}$ is only seen in the Elk-1 Ets domain (1DUX) and not in other Ets domains, despite

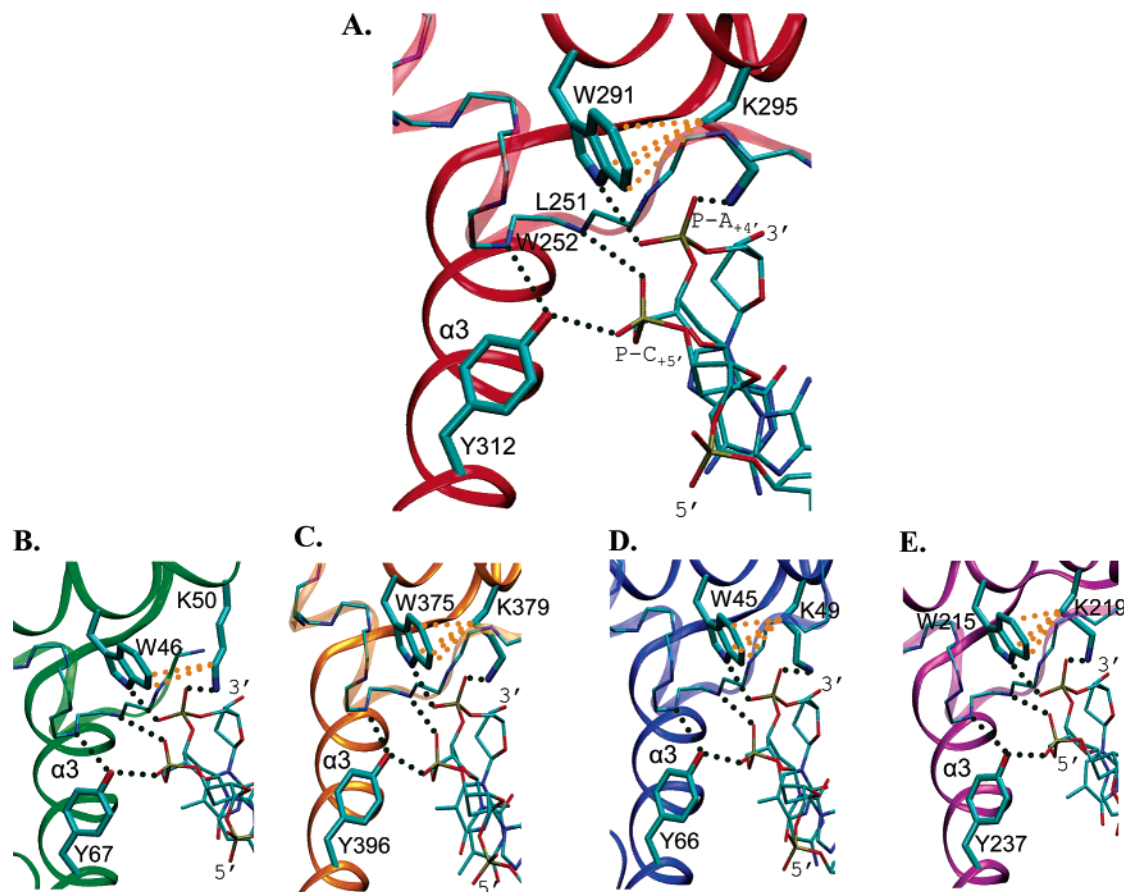


FIGURE 7: Comparison of DNA backbone interactions of Y312, K295, W291, and the main chain amino groups of L251 and W252 of PDEF with other Ets factors. Residues shown here make DNA backbone contacts at the 3' end of the E-site. Color schemes are as follows: (A) PDEF Ets in red with phosphate groups of bases in contact with the PDEF marked, (B) Elk-1 (1DUX) in green, (C) Ets-1 (1K79) in orange, (D) SAP-1 (1BC8) in blue, and (E) PU.1 (1PUE) in purple. Hydrogen bonds are represented by dark green dotted lines and van der Waals contacts as orange dotted lines.

the fact that the end nitrogen atoms, N_{ξ} , are well aligned, especially in the case of Ets-1 (1K79) (Figure 8B). Thus, other than the conserved hydrogen bonding with the phosphate backbone (Figure 7), the interactions of residues between the 295 position and the DNA bases appear to be different for individual Ets members.

Additional Interactions of Interest. In addition to the various essential and redundant patterns of interactions with DNA seen in different Ets members, redundancies were also seen in regions of the protein that are not directly involved in DNA recognition. In PDEF, interactions of the $\alpha 1$ and $\alpha 2$ helices are mediated by a number of highly conserved residues, some of which are strictly conserved among Ets members. Two hydrogen bonds formed between E257 and R294 in PDEF Ets appear to be important in maintaining the spatial conformation between helices $\alpha 1$ and $\alpha 2$ (Figure 9). Additional van der Waals interactions are also present between the conserved residues of helices $\alpha 1$ and $\alpha 2$, and contribute to the proper structural orientation of the helices in PDEF Ets (Figure 9).

Given the presence of such a regular pattern of interactions between the E257 and R294 residues of PDEF Ets, it is likely to be important structurally and/or functionally. However, comparison of PDEF Ets with other Ets structures indicates that not all the Ets domains contain the pair of interactions as seen between E257 and R294 in PDEF. Furthermore, other than the conserved van der Waals interactions between the

phenylalanine at position 254 and the tryptophan at position 291, the rest of the interactions between the two helices often vary to different extents. But when the protein backbones are superimposed, they are aligned well at these regions, suggesting that the interactions between the two helices may be redundant and can reach equivalent orientations for helices with varying numbers of interactions. Such redundancy in Ets domains may determine the intrinsic flexibility that is adaptable to differential properties of target DNA.

DISCUSSION

Overall, it is apparent from our analysis that different Ets domains demonstrate differences in DNA sequence specification, despite their overall structural similarities and a shared mechanism of recognition at the core and the flanking regions of their DNA sites. The combination of different amino acids at key locations in individual Ets domains results in different specificity and affinity for DNA sites for each member. For PDEF, S308 and Q311 are able to collectively determine the preference of thymine at the +4 position immediately downstream of the GGA core, which is required for the high-affinity binding *in vivo* (13). Like other Ets domains, PDEF Ets does not make contacts with every nonconserved DNA base of its recognition sequence. It appears that the sequence specificity is “inferred” from the combination of DNA base contacts with the GGA core, a

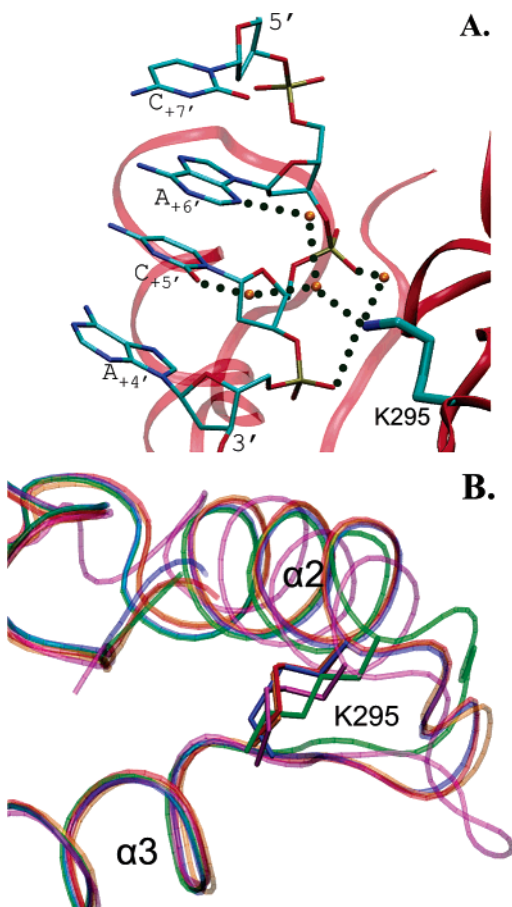


FIGURE 8: Details of DNA interactions with the PDEF lysine residue at position 295. (A) Base and backbone contacts by K295. Hydrogen bonds are represented by dark green dotted lines. (B) Superimposition of the $\alpha 2$ and $\alpha 3$ loop regions of five Ets factors. Main chains are shown in transparent coils, and the side chains of the lysines corresponding to K295 of PDEF are depicted as solid lines. Color schemes are as follows: PDEF in red, Elk-1 (1DUX) in green, Ets-1 (1K79) in orange, SAP-1 (1BC8) in blue, and PU.1 (1PUE) in purple.

few other base contacts flanking the core, and a number of backbone contacts.

Although a strong pattern of interactions with DNA backbone conferring binding specificity is not obvious, we speculate that interactions with the backbone may play a role in sequence-dependent DNA bending. Protein-induced bending of DNA substrate is seen in virtually all Ets complexes and is also present in other types of DNA binding proteins (24). Studies of DNA bending by asymmetric substitution of methyl phosphonate linkages showed that an approximately 20° bending results when three negatively charged phosphate groups on each side of one minor groove are neutralized (25, 26). Additional evidence of DNA bending by attachment of primary amines to position 5 of pyrimidine residues demonstrates a lysine-like effect on the bending of DNA (27, 28). In that regard, it is not surprising that the two highly conserved lysine residues (K295 and K320) both adopt a similar rotamer conformation and make consistent interactions with the phosphate backbone. Furthermore, molecular modeling studies of DNA binding by the *E. coli* CAP protein showed a dependence of bending on the DNA sequence (29). From those studies, it appears that the sequence composition of DNA can affect the degree of its natural curvature and can further influence the degree of bending elicited by external forces.

There is a great variation of curvature of DNA substrates for the existing Ets structures, ranging from $\sim 10^\circ$ in SAP-1 to $\sim 30^\circ$ in Ets-1 (Table 1). PDEF Ets induces the DNA to bend $\sim 16.3^\circ$. Results from SAP-1 and Ets-1 suggest that the degree of bending of different DNA sites by the same Ets domain appears to be more comparable and consistent in range. Furthermore, individual Ets members induce different levels of DNA bending, as seen in bending of the same E74 site by Elk-1 and SAP-1 (11, 12). There is, however, no apparent association of the binding affinity with the degree of DNA curvature. Since it has been shown that DNA curvature is at least in part determined by its inherent sequence composition, it is possible that the combination of DNA backbone contacts, subtle conformational and side chain variations of individual Ets domain, and the DNA conformational flexibility collectively contribute to the preference for a particular target sequence. In light of the

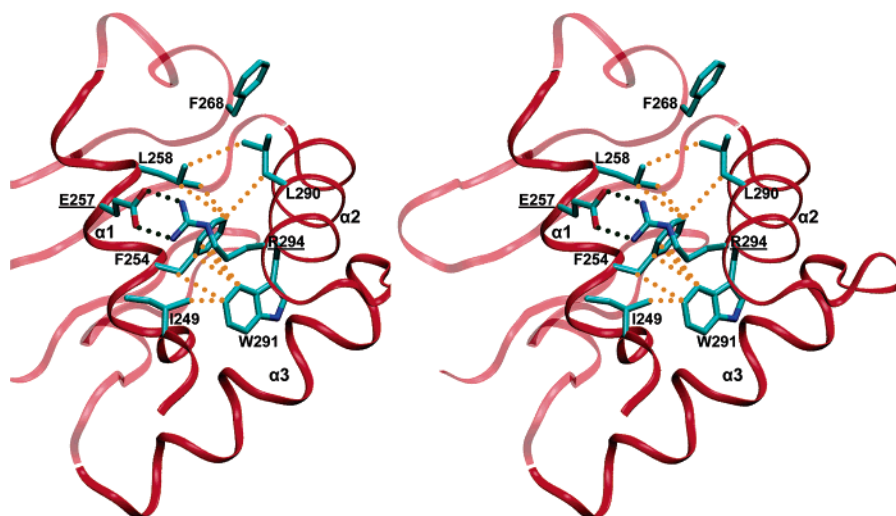


FIGURE 9: Cross-eye stereodialogram of interactions between residues of helices $\alpha 1$ and $\alpha 2$ of the PDEF Ets domain. Side chains of residues that are involved in the interactions between helices $\alpha 1$ and $\alpha 2$ are shown. Hydrogen bonds are represented as dark green dotted lines and van der Waals interactions as orange dotted lines. The phenylalanine at position 268 does not appear to contribute to the interactions between the two helices in PDEF. However, residues at the corresponding position in other Ets domains are involved in the interactions with residues along the $\alpha 1$ – $\alpha 2$ interface.

fact that not all DNA bases are contacted by the protein, it is quite possible that backbone contacts may play an indirect role in the determination of DNA sequence specificity. However, due to the lack of systematic and consistent biochemical analysis in quantifying the level of DNA binding affinity, it is difficult, at present, to establish potential correlations between the DNA sequence and structural variations with binding affinity. Further efforts are necessary to explore such a possibility.

On the basis of our analysis of the PDEF Ets domain in comparison with other Ets structures, it appears that two events play a role in DNA sequence specificity determination. First, PDEF Ets may engage an array of highly conserved interactions to distinguish the minimal and necessary sequential and structural requirements of its DNA substrates. Those key interactions involve both the specific base and backbone contacts. This may lead to the topological change of DNA to an extent that is dependent upon sequences outside the GGA core, as well as upon the ability of the individual protein to induce changes such as bending. On top of the key “signature” interactions, individual Ets family members may engage additional specific interactions, both to the base and to the backbone, which may be further dependent upon the DNA sequence composition. We make no assumption about the order of the two events, and it is not clear from current data whether they happen sequentially or simultaneously. However, the combination of those necessary and sufficient interactions collectively leads to substrate specificity and individual preferences for certain DNA sites.

It is also important to note that factors contributing to the specificity are not necessarily confined to the helix at the DNA protein interface per se; rather, they may be scattered at the other parts of the protein (30, 31). Our data and that of others have demonstrated the distinct interactive patterns and redundant capabilities within the Ets family members, not only at the DNA and protein interface but also at regions that do not directly participate in DNA binding. Since regions such as helices $\alpha 1$ and $\alpha 2$ shown in Figure 9 may play a role in maintaining the overall topology or flexibility of the protein, it is important to identify potential interdependencies among those regions of the protein with those that function directly at the DNA site. Although our efforts and those of others have contributed to the understanding of the mechanism of base specification at the local DNA interface, identification of interactions at the interface is only one of the steps in revealing the underlying mechanisms of sequence specific DNA recognition. Perhaps methods that take into consideration the important residues and the interdependencies of those residues at both the local DNA interface and the whole domain level can be an important step toward the determination of principles for Ets specificity.

Apparently, as a DNA binding unit that serves to bring other functional components to their DNA target, the Ets domains, as well as other DNA binding domains, are expected to contain additional structural features that allow them to interact with other regulatory protein modules. Furthermore, regulatory proteins may directly alter the conformation of certain residues that interact with DNA, thereby further altering the DNA binding specificity (5). It is, therefore, important to classify “socket” residues within the Ets domain that are subject to external regulation, and

their influence over specific interactions with DNA. Although there are systematic efforts to characterize the spatial relationships of DNA binding domains (32, 33), a more global approach that takes into consideration conformational changes in DNA and the interdependencies among regions of the whole domain, as well as the regulatory proteins, may be necessary to shed more light on the structural foundations of sequence specific DNA binding. Collectively, the ability to differentiate determinants inherent to the Ets domains for DNA interactions from those for external regulatory proteins will contribute to our understanding of the nature of sequence specific DNA binding by transcription factors.

ACKNOWLEDGMENT

We remember and thank the late Carol Yaborough for her technical assistance. We thank Dr. Ashwani Sood for the PDEF cDNA clone. We thank Richard Carter, Peter Markstein, and Jeremy Bruenn for the helpful comments.

REFERENCES

- Garvie, C. W., and Wolberger, C. (2001) Recognition of specific DNA sequences, *Mol. Cell* 8, 937–946.
- Sharrocks, A. D. (2001) The ETS-domain transcription factor family, *Nat. Rev. Mol. Cell Biol.* 2, 827–837.
- Li, R., Pei, H., and Watson, D. K. (2000) Regulation of Ets function by protein–protein interactions, *Oncogene* 19, 6514–6523.
- Batchelor, A. H., Piper, D. E., de la Brousse, F. C., McKnight, S. L., and Wolberger, C. (1998) The structure of GABP α/β : An ETS domain-ankyrin repeat heterodimer bound to DNA, *Science* 279, 1037–1041.
- Garvie, C. W., Hagman, J., and Wolberger, C. (2001) Structural studies of Ets-1/Pax5 complex formation on DNA, *Mol. Cell* 8, 1267–1276.
- Hassler, M., and Richmond, T. J. (2001) The B-box dominates SAP-1-SRF interactions in the structure of the ternary complex, *EMBO J.* 20, 3018–3028.
- Mo, Y., Ho, W., Johnston, K., and Marmorstein, R. (2001) Crystal structure of a ternary SAP-1/SRF/c-fos SRE DNA complex, *J. Mol. Biol.* 314, 495–506.
- Garvie, C. W., Pufall, M. A., Graves, B. J., and Wolberger, C. (2002) Structural analysis of the autoinhibition of Ets-1 and its role in protein partnerships, *J. Biol. Chem.* 277, 45529–45536.
- Verger, A., and Duterrque-Coquillaud, M. (2002) When Ets transcription factors meet their partners, *BioEssays* 24, 362–370.
- Kodandapani, R., Pio, F., Ni, C. Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R. A., and Ely, K. R. (1996) A new pattern for helix–turn–helix recognition revealed by the PU.1 ETS-domain–DNA complex, *Nature* 380, 456–460.
- Mo, Y., Vaessen, B., Johnston, K., and Marmorstein, R. (1998) Structures of SAP-1 bound to DNA targets from the E74 and c-fos promoters: Insights into DNA sequence discrimination by Ets proteins, *Mol. Cell* 2, 201–212.
- Mo, Y., Vaessen, B., Johnston, K., and Marmorstein, R. (2000) Structure of the elk-1–DNA complex reveals how DNA-distal residues affect ETS domain recognition of DNA, *Nat. Struct. Biol.* 7, 292–297.
- Oettgen, P., Finger, E., Sun, Z., Akbarali, Y., Thamrongsak, U., Boltax, J., Grall, F., Dube, A., Weiss, A., Brown, L., Quinn, G., Kas, K., Endress, G., Kunsch, C., and Libermann, T. A. (2000) PDEF, a novel prostate epithelium-specific Ets transcription factor, interacts with the androgen receptor and activates prostate-specific antigen gene expression, *J. Biol. Chem.* 275, 1216–1225.
- Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D., and Waugh, D. S. (2001) Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency, *Protein Eng.* 14, 993–1000.

15. Kapust, R. B., and Waugh, D. S. (2000) Controlled intracellular processing of fusion proteins by TEV protease, *Protein Expression Purif.* 19, 312–318.
16. Libermann, T. A., and Baltimore, D. (1993) Pi, a pre-B-cell-specific enhancer element in the immunoglobulin heavy-chain enhancer, *Mol. Cell. Biol.* 13, 5957–5969.
17. Otwinowski, Z., and Minor, M. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol.* 276, 307–326.
18. Collaborative Computational Project No. 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D* 50, 760–763.
19. Vagin, A., and Teplyakov, A. (1997) MOLREP: An automated program for molecular replacement, *J. Appl. Crystallogr.* 30, 1022–1025.
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D* 53, 240–255.
21. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
22. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics, *J. Mol. Graphics* 14, 33–38, 27–28.
23. Thompson, C. C., Brown, T. A., and McKnight, S. L. (1991) Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex, *Science* 253, 762–768.
24. Shakked, Z., Guzikevich-Guerstein, G., Frolow, F., Rabinovich, D., Joachimiak, A., and Sigler, P. B. (1994) Determinants of repressor/operator recognition from the structure of the Trp operator binding site, *Nature* 368, 469–473.
25. Strauss, J. K., and Maher, L. J., III (1994) DNA bending by asymmetric phosphate neutralization, *Science* 266, 1829–1834.
26. Williams, L. D., and Maher, L. J., III (2000) Electrostatic mechanisms of DNA deformation, *Annu. Rev. Biophys. Biomol. Struct.* 29, 497–521.
27. Strauss, J. K., Roberts, C., Nelson, M. G., Switzer, C., and Maher, L. J., III (1996) DNA bending by hexamethylene-tethered ammonium ions, *Proc. Natl. Acad. Sci. U.S.A.* 93, 9515–9520.
28. Strauss, J. K., Prakash, T. P., Roberts, C., Switzer, C., and Maher, L. J. (1996) DNA bending by a phantom protein, *Chem. Biol.* 3, 671–678.
29. Gurlie, R., Duong, T. H., and Zakrzewska, K. (1999) The role of DNA-protein salt bridges in molecular recognition: A model study, *Biopolymers* 49, 313–327.
30. Shore, P., Whitmarsh, A. J., Bhaskaran, R., Davis, R. J., Waltho, J. P., and Sharrocks, A. D. (1996) Determinants of DNA-binding specificity of ETS-domain transcription factors, *Mol. Cell. Biol.* 16, 3338–3349.
31. Shore, P., and Sharrocks, A. D. (1995) The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities, *Nucleic Acids Res.* 23, 4698–4706.
32. Pabo, C. O., and Nekludova, L. (2000) Geometric analysis and comparison of protein-DNA interfaces: Why is there no simple code for recognition? *J. Mol. Biol.* 301, 597–624.
33. Mirny, L. A., and Gelfand, M. S. (2002) Structural analysis of conserved base pairs in protein-DNA complexes, *Nucleic Acids Res.* 30, 1704–1711.
34. Pio, F., Ni, C. Z., Mitchell, R. S., Knight, J., McKercher, S., Klemsz, M., Lombardo, A., Maki, R. A., and Ely, K. R. (1995) Co-crystallization of an ETS domain (PU.1) in complex with DNA. Engineering the length of both protein and oligonucleotide, *J. Biol. Chem.* 270, 24258–24263.
35. Lavery, R., and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: Conventions and principles, *J. Biomol. Struct. Dyn.* 6, 655–667.

BI047352T