University of Huddersfield Repository

Cooper, Christopher D.O., Newman, Joseph A., Aitkenhead, Hazel, Allerston, Charles K. and Gileadi, Opher

Structures of the Ets Protein DNA-binding Domains of Transcription Factors Etv1, Etv4, Etv5, and Fev: Determinants of DNA Binding and Redox Regulation by Disulfide Bond Formation.

Original Citation

Structures of the Ets Protein DNA-binding Domains of Transcription Factors Etv1, Etv4, Etv5, and Fev
DETERMINANTS OF DNA BINDING AND REDOX REGULATION BY DISULFIDE BOND FORMATION*

Christopher D. O. Cooper[1,2], **Joseph A. Newman**[1], **Hazel Aitkenhead, Charles K. Allerston**[3], **and Opher Gileadi**[4]

From the Structural Genomics Consortium, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, United Kingdom

Background: Transcription factors of the Ets family regulate development and cancer.
Results: We show the details of DNA sequence recognition and identify structural changes that drastically affect DNA binding.
Conclusion: The activity of Ets transcription factors is regulated by their oxidation state.
Significance: Ets transcription factors may control the response of cells to oxygen levels.

Ets transcription factors, which share the conserved Ets DNA-binding domain, number nearly 30 members in humans and are particularly involved in developmental processes. Their deregulation following changes in expression, transcriptional activity, or by chromosomal translocation plays a critical role in carcinogenesis. Ets DNA binding, selectivity, and regulation have been extensively studied; however, questions still arise regarding binding specificity outside the core GGA recognition sequence and the mode of action of Ets post-translational modifications. Here, we report the crystal structures of Etv1, Etv4, Etv5, and Fev, alone and in complex with DNA. We identify previously unrecognized features of the protein-DNA interface. Interactions with the DNA backbone account for most of the binding affinity. We describe a highly coordinated network of water molecules acting in base selection upstream of the GGAA core and the structural features that may account for discrimination against methylated cytidine residues. Unexpectedly, all proteins crystallized as disulfide-linked dimers, exhibiting a novel interface (distant to the DNA recognition helix). Homodimers of Etv1, Etv4, and Etv5 could be reduced to monomers, leading to a 40–200-fold increase in DNA binding affinity. Hence, we present the first indication of a redox-dependent regulatory mechanism that may control the activity of this subset of oncogenic Ets transcription factors.

Ets transcription factors, which share the conserved Ets DNA-binding domain, number nearly 30 members in humans and are particularly involved in developmental processes. Their deregulation following changes in expression, transcriptional activity, or by chromosomal translocation plays a critical role in carcinogenesis. Ets DNA binding, selectivity, and regulation have been extensively studied; however, questions still arise regarding binding specificity outside the core GGA recognition sequence and the mode of action of Ets post-translational modifications. Here, we report the crystal structures of Etv1, Etv4, Etv5, and Fev, alone and in complex with DNA. We identify previously unrecognized features of the protein-DNA interface. Interactions with the DNA backbone account for most of the binding affinity. We describe a highly coordinated network of water molecules acting in base selection upstream of the GGAA core and the structural features that may account for discrimination against methylated cytidine residues. Unexpectedly, all proteins crystallized as disulfide-linked dimers, exhibiting a novel interface (distant to the DNA recognition helix). Homodimers of Etv1, Etv4, and Etv5 could be reduced to monomers, leading to a 40–200-fold increase in DNA binding affinity. Hence, we present the first indication of a redox-dependent regulatory mechanism that may control the activity of this subset of oncogenic Ets transcription factors.

The Ets transcription factor family consists of 28 genes in humans (1, 2) containing the evolutionarily conserved 85 amino acid Ets DNA-binding domain (3), originally identified as a viral oncogene (E26 transformation-specific) (4). Ets proteins exhibit ubiquitous or tissue-specific expression (5) and are particularly involved in differentiation processes and the response to signaling pathways (6, 7). Ets proteins may be classified based on the presence of additional structured domains (5). The small PEA3 subfamily is characterized by an N-terminal transactivation domain and consists of three members, Etv1 (Ets translocation variant 1, ER81), Etv4 (PEA3), and Etv5 (ERM) (8).

PEA3 transcription factors have roles in morphogenesis (9) and neuronal differentiation (10, 11). PEA3 members are also oncoproteins (8), whose overexpression correlates with up-regulation of HER2/Neu and with progression in breast tumors (12). *ETV1* is amplified in >40% of melanomas (13), and overexpression following chromosomal translocation of either *ERG* or *Etv1* to the androgen-inducible *TMPRSS2* promoter is present in most prostate tumors (14). Etv1 directly interacts with the androgen receptor (15) and drives the androgen receptor transcriptional response associated with aggressive prostate cancer (16). Etv1 target genes include *hTERT* (17) and matrix metalloproteinases (18) that may mediate cancer development. Furthermore, Etv1 may be subverted by other oncoproteins such as mutated (activated) KIT, which cooperates with elevated Etv1 levels to promote tumorigenesis (19). Etv4 and Etv5 also play roles in morphogenesis, fertility, and oncogenesis (8). Ewing's sarcomas result from chromosomal translocations that generate dominant transforming fusion proteins of the transactivation domain of the EWS protein with the ETS domain of one of the five Ets proteins Etv1, Etv4, Erg, Fli1, and Fev (Fifth Ewing Variant, PET-1) (20, 21). Fev lacks additional domains and has a restricted tissue expression (21, 22). Fev regulates serotonergic neuronal differentiation, being critical for normal anxiety/aggression development (23), although overexpression is associated with serotonin production in small intestine neuroendocrine tumors, stimulating tumor growth (24). Ets

domains are thus a central nexus in tumor development and disease progression. Targeting transcriptional regulation in cancer with drugs is expected to be challenging (25), but Ets inhibitors have been developed. YK-4-279 targets EWS-FLI1, inhibits growth in Ewing sarcoma (26), and also inhibits Erg/Etv1-driven prostate cancer invasion (27). Although this may be useful for generic treatment of Ets-driven cancers, the lack of specificity could cause off-target effects with other Ets proteins, highlighting the need for further high resolution structural and biochemical studies of Ets proteins.

The Ets domain is a variant helix-turn-helix (winged helix) structure (28–31), comprising three α-helices and a four-stranded antiparallel β-sheet. Ets domains bind to the EBS[5] (Ets-binding site) in dsDNA with the α3 helix inserted into the major groove, and invariant arginine and tyrosine residues hydrogen bonded to bases of the invariant 5′-GGA(A/T)-3′ sequence in the EBS (2). Although binding the GGA(A/T) core is a common property of Ets transcription factors, a genome-wide analysis of all Ets family members has established that Ets domains recognize up to nine bases (32) (three upstream and two downstream of the GGA(A/T) core), and they are classified into four distinct classes based on the sequence preferences for these flanking regions. Etv1, Etv4, Etv5, and Fev all belong to the large and diverse class I (containing 14 of 26 of the mammalian Ets transcription factors), which recognize a consensus sequence ACCGGAAGT(G/A). The structural basis for this sequence discrimination is unclear, as there are no direct base contacts outside of the core. Indirect readout mechanisms have been suggested, based on sequence-dependent DNA conformational preferences (33). However, no mechanism for how sequence preferences would influence shape readout has been proposed nor how the different classes of Ets domains achieve their observed sequence specificities outside the GGA within this paradigm.

A number of other factors influence Ets DNA binding, including interaction with other transcription factors to recognize combined operators in a cooperative manner (34), cooperative binding of palindromic sequences by dimerization (35), and the presence of auto-inhibitory regions surrounding the Ets domain (36–38). DNA binding is also subject to post-translational regulation (39) with protein kinase A phosphorylation of Etv1 at Ser-334 (40) and the equivalent Ser-367 in Etv5 (41) repressing DNA binding. Several Ets proteins have been identified as being subject to redox control, with redox-sensitive cysteines present in GA-binding protein α (GABPα) affecting dimerization and DNA binding (42). However, as with the case for phosphorylation, the structural mechanism for this regulation is unclear.

Additional regulation of Ets DNA binding may occur at the DNA level, with several Ets motifs being over-represented in methylated genomic regions (43). Some Ets proteins are known to bind to their cognate sequence *in vitro* and *in vivo* only in a demethylated state, including GABPα (44, 45) and ETS1 (46). Direct methylation of the class I EBS sequence is possible in two positions (C^{m}GGAA/TTCC^{m}G); however, it is not known whether and to what extent this direct methylation affects Ets binding.

Here, we present the crystal structures of the Ets domains of the entire PEA3 subfamily (Etv1, Etv4, and Etv5) together with the structure of the Erg family member (Fev). All proteins share a highly conserved core Ets structure and make extensive interactions with DNA. Comparisons of apo- and DNA-bound forms of Etv1, Etv4, and Fev allow us to identify residue movements and disorder to order transitions that occur upon DNA binding. In the DNA complex structures of Etv5 and Fev (which are determined to 1.9 and 2.0 Å resolution, respectively), we observe a network of coordinated water molecules contributing to sequence recognition upstream to the GGA core. Structural and biochemical analysis of Etv1 clarifies mechanisms of post-translational regulation in Ets proteins, including PKA-mediated phosphorylation at Ser-334 and the abrogation of Ets binding to sites that include a methylated CpG. Furthermore, a new Ets dimerization interface linked by a redox-sensitive disulfide bond is identified, and this dimerization is shown to result in a significant inhibition of DNA binding, potentially linking Ets transcription factors with cellular redox regulatory mechanisms.

Experimental Procedures

Cloning and Site-directed Mutagenesis—Plasmid DNA templates for full-length Etv1, Etv4, Etv5, and Fev were obtained from the Mammalian Gene Collection (I.M.A.G.E. Consortium Clone IDs 30345383, 3854349, 3050350, and 4130242, respectively) (47). Regions corresponding to the core Ets domains were amplified by PCR using Pfx DNA polymerase (Invitrogen, Paisley, UK). PCR products were ligation independent cloning into the pNIC28-Bsa4 expression vector (GenBank^{TM} accession number EF198106, encompassing a tobacco etch virus (TEV)-cleavable (shown by *) N-terminal His$_6$ tag MHHH-HHHSSGVDLGTENLYFQ*SM), as described elsewhere (48). The initial Etv1 construct contained two primer-incorporated mutations (Y329S and P427S) and thus were designated Etv1$^{Y329S-P427S}$. A subsequent construct containing the wild-type sequence is referred to as Etv1. Additional mutants were generated in a full-length Etv1 construct using the megaprimer method (49). The Gly-326–Asn-429 fragments of the Etv1 mutant derivatives were then subcloned for bacterial expression into pNIC28-Bsa4 as described and confirmed by sequencing. Expression plasmids were transformed into BL21 (DE3) Rosetta-R3 (48) or, when indicated, into Rosetta-gami^{TM} 2 (Novagen®).

Recombinant Protein Expression—Recombinant protein expression was induced by the addition of 0.1 mM isopropyl 1-thio-β-D-galactopyranoside to bacterial cultures grown in TB (Terrific Broth) containing 50 μg/ml kanamycin at an OD$_{600}$ of 3.0 at 37 °C in UltraYield baffled flasks (Thomson Instrument Co, Oceanside, CA). Cultures were further incubated at 18 °C overnight. Selenomethionine-derivatized Etv1$^{Y329S-P427S}$ expression was performed at 37 °C in M9 minimal medium, supplemented with 0.4% glucose, 2 mM MgSO$_4$, 0.1 mM CaCl$_2$, and 50

---

[5] The abbreviations used are: EBS, Ets-binding site; 5meC, 5-methylcytosine; TCEP, tris(2-carboxyethyl)phosphine; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; MWCO, molecular weight cutoff; CV, column volume; PDB, Protein Data Bank; SEC, size exclusion chromatography; TEV, tobacco etch virus; GABP, GA-binding protein α.

μg/ml kanamycin. Cells were cultured to an OD_{600} of 0.8 and 25 μg/ml selenomethionine was added, along with leucine, isoleucine, and valine to 50 μg/ml, and lysine, threonine, and phenylalanine to 100 μg/ml. Cultures were further incubated until OD_{600} of 1.2, and protein expression was induced by addition of isopropyl 1-thio-β-D-galactopyranoside and selenomethionine to final concentrations of 0.1 mM and of 75 μg/ml, respectively. Cells were harvested after further overnight incubation at 18 °C and stored at −80 °C.

Protein Purification—For purification of Etv1, Etv4, Etv5, and Fev constructs, ~50 g of cell pellets were thawed and resuspended in buffer A (50 mM HEPES, pH 7.5, 500 mM NaCl, 5% glycerol, 10 mM imidazole, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP)), with the addition of 1× protease inhibitor set VII (Merck, Darmstadt, Germany) and 15 units/ml Benzonase (Merck). Cells were lysed using sonication. Cell debris and nucleic acids were removed by addition of 0.15% polyethyleneimine, pH 7.5, and centrifugation at 40,000 × g for 1 h at 4 °C. Clarified lysates were applied to a 3-ml Ni^{2+}-iminodiacetic acid-immobilized metal ion affinity chromatography gravity flow column (Generon, Maidenhead, UK), washed with 20 column volumes (CV) of buffer A, followed by 20 CV of wash buffer (buffer A with 30 mM imidazole). Fractions were eluted with 5× 2-CV aliquots of buffer A containing 300 mM imidazole and analyzed by SDS-PAGE, and relevant fractions pooled and cleaved with His_6-tagged TEV protease (1:20 mass ratio) overnight at 8 °C. Imidazole was removed by concurrent dialysis during cleavage, using a 3.5-kDa MWCO snakeskin membrane (Thermo Fisher Scientific, Rockford, IL) in buffer B (20 mM HEPES, pH 7.5, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP). TEV protease was removed from dialyzed proteins using Ni-IDA immobilized metal ion affinity chromatography (2-ml CV) and washed with an imidazole gradient in 20 mM steps to 100 mM in buffer B, and cleaved protein was pooled and concentrated with a 3-kDa MWCO centrifugal concentrators (Vivaproducts, Littleton, MA). Final separation was by size exclusion chromatography, using a HiLoad 16/60 Superdex S75 or 200 column equilibrated in buffer B, and run at 1.2 ml/min in buffer B. Protein identity was confirmed by LC/ESI-TOF mass spectrometry. Protein concentrations were calculated from A_{280} (Nanodrop) using the calculated molecular mass and extinction coefficients. The scheme was identical for purification of disulfide-linked proteins but with the omission of TCEP from all buffers.

Crystallization and Structural Determination—The purified Ets domains of Fev and Etv1 were crystallized with double-stranded DNA as follows. Oligonucleotides for co-crystallization were synthesized and used without further purification (Eurofins MWG Operon, Ebersberg, Germany). The oligonucleotides (5′-ACCGGAAGTG-3′) and (5′-CACTTCCGGT-3′), with the Ets core recognition sequence underlined, were annealed by mixing 450 μM of each oligonucleotide in 10 mM Tris, pH 7.5, 50 mM NaCl, heating to 95 °C for 5 min, and cooling to 21 °C slowly in a heating block. Frozen Etv1[Y329S-P427S] or Fev was thawed rapidly, and aggregates were removed by microcentrifugation at 14,000 × g for 10 min at 4 °C. To form DNA complexes, proteins were mixed with annealed oligonucleotides in a 1:1.1 molar ratio in a buffer consisting of 10 mM

HEPES, 166 mM NaCl, 5% glycerol, and 0.5 mM TCEP at protein concentrations of 5.3 and 6 mg/ml, respectively, and incubated on ice for ~30 min. The protein/DNA mixtures were then concentrated by ultrafiltration using a 3000-Da MWCO centrifugal concentrator to an estimated protein concentration of 12.5 and 16 mg/ml for Etv1-DNA and Fev-DNA, respectively. For crystallization of the Etv4 and Etv5 DNA complexes, the same procedure was repeated but using the oligonucleotides (5′-ACCG-GAAGTG-3′) and (5′ACTTCCGGTC3′). In all cases, sitting drop vapor diffusion crystallization trials were set up with a Mosquito (TTP Labtech) crystallization robot. Crystals were obtained in the following conditions: Etv1 (16 mg/ml, 20 °C), 2.5 M sodium formate; Etv-1 DNA (12 mg/ml, 4 °C), 28% PEG smear low molecular weight, 0.2 M sodium chloride, 0.1 M Tris, pH 8.5, 5% glycerol; SeMet-Etv1-DNA (15 mg/ml, 4 °C), 20% PEG 3350, 0.2 M potassium citrate; Etv4 (16 mg/ml, 20 °C), 0.8 M sodium citrate tribasic, 0.1 M cacodylate, pH 6.5; Etv4-DNA (10 mg/ml, 20 °C), 20% PEG 3350, 0.2 M magnesium chloride, 0.1 M BisTris, pH 5.5; Etv5-DNA (8 mg/ml, 20 °C), 40% PEG 300, 0.2 M calcium acetate, 0.1 M cacodylate, pH 6.0; Fev (18 mg/ml, 4 °C), 0.8 M sodium citrate tribasic, 0.1 M cacodylate, pH 6.5; Fev-DNA (16 mg/ml, 20 °C), 25% PEG 3350, 0.1 M BisTris, pH 5.5.

Crystals were cryo-protected by transferring the crystal to a solution of mother liquor supplemented with 25% ethylene glycol and flash-cooling in liquid nitrogen. Datasets were collected for all crystals at Diamond Light source beamlines I04 (Etv1, Etv1-DNA, and Fev-DNA), I02 (Fev), I24 (Etv4-DNA), and I03 (Etv4, Etv5-DNA and selenomethionine Etv1-DNA). Diffraction data were processed with the programs autoPROC (Etv1 and Etv1-DNA) (50), MOSFLM (Fev) (51), and XDS (Etv4, Etv4-DNA, Etv5-DNA and Fev-DNA) (52). The structures of Etv1, Etv4, Etv4-DNA, Etv5-DNA, Fev, and Fev-DNA were solved by molecular replacement using the program PHASER (53). Initial attempts to solve the Etv1-DNA complex by molecular replacement were unsuccessful, and for this reason, the selenomethionine-derivatized protein was produced, and SAD data were collected close to the selenium peak wavelength (0.976 Å). Selenium atom positions were located with the program SHELX (54) and refined using the program SHARP (55). Model building and manipulation were performed in COOT (56), and the structures were refined using autoBUSTER (Global Phasing) (Etv1, Etv1-DNA, and Fev), REFMAC (Etv4) (57), and PHENIX REFINE (Etv4-DNA, Etv5-DNA and Fev-DNA) (58). A summary of the crystallization conditions, data processing, phasing, and refinement statistics for all datasets are shown in Tables 1 and 2.

LC/ESI-TOF Mass Spectrometry of Intact Proteins—30-μl protein samples at 0.02 mg/ml in 0.1% formic acid were injected onto a 4.6 × 50 mm Zorbax 5-μm 300SB-C3 column and resolved by reversed-phase chromatography at 40 °C. The solvent system was 0.1% formic acid in double distilled H_2O (buffer A) and 0.1% formic acid in methanol (buffer B), with 1 min at 5% buffer B and then a linear gradient of 5–95% buffer B over 6 min at 0.5 ml/min. Protein intact mass was determined using an MSD-TOF electrospray ionization orthogonal time-of-flight mass spectrometer (Agilent Technologies, Palo Alto, CA) operated in positive ion mode.

TABLE 1
Data collection and refinement statistics for DNA free structures
Values in parentheses are for the highest resolution shell.

Data collection statistics			
	Etv1	Fev	Etv4
Space group	P1	P 4_3 3 2	P 3_1 2 1
Cell dimensions, *a,b,c* (Å)	33.3, 45.6, 55.4	130.8, 130.8, 130.8	50.8, 50.8, 67.7
Angles α,β,γ (°)	77.9, 84.8, 90.0	90, 90, 90	90, 90, 120
Wavelength (Å)	0.979	0.979	0.979
Resolution (Å)	53.9-1.82 (1.89-1.82)	75-2.64 (2.77-2.64)	26.8-1.05 (1.07-1.05)
R_{merge} (%)[+]	5.4 (33)	8.6 (163)	4.0 (67)
$R_{p.i.m.}$ (%)[*]	6.2 (36)	2.3 (42)	1.6 (34.2)
I/σI	10.4 (2.3)	19.4 (2.0)	21.1 (2.1)
Completeness (%)	97.4 (96.2)	100 (100)	96.5 (67.5)
Multiplicity	2.3 (2.3)	15.9 (16.8)	8.8 (4.7)
Refinement statistics			
	Etv1	Fev	Etv4
Resolution	53.9-1.82 (1.89-1.82)	75-2.64(2.89-2.64)	43 – 1.05 (1.08-1.05)
No. Unique reflections	27819 (2734)	12330 (2730)	43771 (2428)
R_{work}/R_{free} (%)	22.1/25.8 (23.3/31.9)	21.6/22.7 (22.7/23.7)	11.7/13.2 (22.7/22.6)
No. atoms			
Protien	3004	1484	800
Waters	193	9	117
Other	0	0	12
Average B factors (Å²)			
Protein	23.4	87	16
Wilson B	22.0	102	10
R.M.S. deviations			
Bond lengths (Å)	0.01	0.008	0.007
Bond angles (°)	0.93	0.88	0.81
Ramachandran plot			
Favoured (%)	99.5	98.4	100
Allowed (%)	100	99.5	100
PDB ID	4AVP	2YPR	4CO8

[+] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i - I_m|/\Sigma_{hkl}\Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.
[*] $R_{p.i.m.} = \Sigma_{hkl}\sqrt{(1/n - 1)} \Sigma_{i_n = 1}|I_i - I_m|/\Sigma_{hkl}\Sigma_i I_i$.

In Vitro Protein Phosphorylation—Phosphorylation of Etv1 Ser-334 by protein kinase A was performed *in vitro* using 2.5 units of bovine heart PKA catalytic subunit (P2645, Sigma) with 40 µg of Etv1, in modified buffer B (20 mM HEPES, pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM ATP, 10 mM MgCl₂, and 5 mM dithiothreitol). Reactions were performed for 4 h at room temperature and quenched by the addition of 50 mM EDTA, followed by buffer exchange into buffer B with Micro Bio-Spin® P-6 columns (Bio-Rad) and confirmation of phosphorylation by ESI-TOF mass spectrometry of the intact protein.

Electrophoretic Mobility Shift Assays—Equilibrium binding constants (K_d) of Ets constructs were estimated by EMSA. Unless otherwise specified, the dsDNA probe contained the single consensus Ets-binding site used in crystallization (underlined), with oligonucleotides Etv1ALF (5′-ATCTC<u>ACCGGA-AGTGT</u>AGCA-3′) and Etv1ALR (5′-TGCTA<u>CACTTCCGG-TGAGAT</u>-3′). Substrates were prepared by ³²P-end-labeling one oligonucleotide, annealing to the complementary strand in

10 mM Tris-Cl, pH 7.5, 50 mM NaCl, prior to purification with Micro Bio-Spin® P-6 columns. Proteins were generally titrated from 10^{-14} to 10^{-6} M. The DNA probe was used at 2 nM for the initial estimations of K_d (termed K_d^{app1} in Table 3). As the dissociation constants of wild-type Etv1 were estimated to be subnanomolar, we repeated some of the measurements with a DNA probe concentration of 0.1 nM to maintain an excess of protein (results listed as K_d^{app2} in Table 3). As seen in Table 3, the effects of mutations are qualitatively similar at both DNA concentrations. EMSA buffer is composed of 50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 2 mM MgCl₂, 0.01% Tween® 20, and 5% glycerol. Proteins were reduced with 10 mM TCEP prior to assay unless otherwise specified. Reactions were performed for 1 h at room temperature, prior to mixing with loading dye to 0.25%, and resolved by 10% native PAGE at 150 V for 1 h at room temperature. Gels were analyzed using phosphorimaging, and the apparent K_d value was estimated from plots by nonlinear regression with a least squares fit, using a specific one-site

TABLE 2

Data collection and refinement statistics for DNA complexes

Values in parentheses are for the highest resolution shell.

Data collection statistics					
	Etv1-DNA	**SeMet-Etv1-DNA**	**Etv4-DNA**	**Etv5-DNA**	**Fev-DNA**
Space group	P 4$_3$ 2 2	P 4$_3$ 2 2	C 1 2 1	P 3$_2$ 1 2	C 2 2 2$_1$
Cell dimensions, *a,b,c* (Å)	64.8, 64.8, 129.6	64.7, 64.7, 134.0	176.6, 46.1, 171.1	58.0, 58.0, 84.7	50.2, 64.6, 110.2
Angles α,β,γ (°)	90, 90, 90	90, 90, 90	90, 96.7, 90	90, 90, 120	90, 90, 90
Wavelength (Å)	0.964	0.972	0.968	0.976	0.9795
Resolution (Å)	129-2.9 (3.06-2.9)	134-3.85 (4.06-3.85)	45-2.8 (2.95-2.80)	20-1.95 (2.00-1.95)	110-2.0 (2.05-2.0)
R$_{merge}$ (%)[+]	8 (127)	9.7 (51)	5.0 (46)	7.9 (89.7)	2.7 (40)
R$_{p.i.m.}$ (%)[*]	2.9 (41)	2.3 (10)	3.6 (33.7)	2.8 (33.8)	1.9 (27)
I/σI	17.6 (1.9)	25.5 (9.4)	11.6 (1.8)	16.4 (2.3)	23.6 (3.0)
Completeness (%)	99.2 (100)	99.9 (100)	94.9 (92.2)	99.9 (100)	99.6 (99.8)
Multiplicity	8.9 (9.3)	25.4 (27.7)	2.5 (2.5)	9.7 (9.0)	3.9 (4.0)

Refinement statistics				
	Etv1-DNA	Etv4-DNA	Etv5-DNA	Fev-DNA
Resolution	58-2.9 (3.24-2.9)	40-2.8 (2.89-2.80)	19.5-1.95 (2.0-1.95)	31-2.0 (2.2-2.0)
No. Unique reflections	6594 (1722)	32675 (2746)	12146 (849)	12387 (2895)
R$_{work}$/R$_{free}$ (%)	20.6/22.2 (22.9/24.1)	20.3/24.8 (33.5/38.1)	17.3/20.3 (21.3/25.2)	21.0/25.4 (29.3/33.3)
No. atoms				
Protein	734	6249	793	745
DNA	404	3103	404	404
Waters	0	0	100	36
Average B factors (Å²)				
Protein	124	73	33.6	58
DNA	130	90	35.5	62
Wilson B	127	80	30	39
R.M.S. deviations				
Bond lengths (Å)	0.01	0.003	0.008	0.002
Bond angles (°)	1.35	0.52	1.12	0.70
Ramachandran plot				
Favoured (%)	95.7	94.9	99.0	98.9
Allowed (%)	100	100	100	100
PDB ID	4BNC	4UUV	4UNO	3ZP5

[+] $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i - I_m|/\Sigma_{hkl}\Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.

[*] $R_{p.i.m.} = \Sigma_{hkl}\sqrt{(1/n-1)} \Sigma_{i_n=1}|I_i - I_m|/\Sigma_{hkl}\Sigma_i I_i$.

binding model with Hill slope (Prism, GraphPad, San Diego). Other EMSA substrates are described in the respective figure legends.

Results

DNA-free Structures of the Ets Domains of Etv1, Etv4, and Fev—The full-length sequences of Etv1, Etv4, Etv5, and Fev are predicted to include disordered regions, which often hinder crystallization. We have generated several expression constructs, including the full-length and truncated versions of each protein. Crystals were obtained from constructs that contained the Ets domain, Etv1 (encompassing amino acids 326–429), Etv4 (encompassing amino acids 338–470), and Fev (encompassing amino acids 42–141). The construct of Etv1 included two inadvertent mutations that resulted from primer impurities as follows: Y329S at the N terminus, and P427S at the C terminus; this construct is denoted Etv1$^{Y329S-P427S}$. Our attempts to crystallize the corresponding construct with the wild-

type sequence failed, although the mutated residues are in disordered regions or at the end of the ordered region of the crystal structures of Etv1, so they are unlikely to have any bearing on the structural and functional analyses. We have also attempted to crystallize Etv5 using a similar construct strategy to that defined above (Etv5 construct encompassing amino acids 365–462), but despite obtaining protein of similar purity and yield as the other Etv constructs, we were unable to grow crystals of Etv5 in the absence of DNA.

Crystals of Etv1$^{Y329S-P427S}$ diffracted to 1.82 Å and contained four copies of Etv1 in the asymmetric unit (4AVP, Tables 1 and 2). Electron density is observed for residues Ser-334 to Phe-426 (thus excluding the Y329S and P427S mutations) with most side chains visible. Etv4 crystals diffracted to 1.05 Å and contained a single copy of Etv4 in the asymmetric unit (4CO8, Tables 1 and 2). Fev crystals diffracted to 2.64 Å, with two copies of Fev per asymmetric unit (2YPR, Tables 1 and 2). Ets domains are highly similar in sequence (Fig. 1*A*), with Etv1,

TABLE 3

DNA binding affinity of Ets domain constructs (Etv1 unless specified)

NA indicates undetectable binding or below threshold for regression analysis. K_d^{app1} indicates apparent K_d value measured with a 2 nM DNA probe. K_d^{app2} indicates apparent K_d value measured with a 0.1 nM DNA probe.

Construct (Etv1 unless specified otherwise)	Description/mutation	K_d^{app1} ($\times 10^{-10}$ M)[a]	K_d^{app1}-fold inhibition[b]	K_d^{app2} ($\times 10^{-10}$ M)[a]	K_d^{app2}-fold inhibition[b]
WT	Wild type Ets domain	2.7 ± 0.28	1	5.9 ± 1.5	1
Y329S-P427S	PCR mutant Ets domain	2.5 ± 0.26	0.9	3.1 ± 0.6	0.5
K379A	DNA backbone contact	NA	NA		
D387A	DNA base contact	1.7 ± 0.2	0.6	1.9 ± 0.3	0.3
R391A	DNA base contact	118 ± 14.8	43.7	86.0 ± 16.8	14.6
R394A	DNA base contact	2757 ± 627	1021	ND	ND
Y395F	DNA base contact	2.3 ± 0.5	0.9	13.5 ± 10.5	2.3
Y396F	DNA backbone contact	26390 ± 9349	9774.1		
Y397F	DNA backbone contact	NA	NA		
K404A	DNA backbone contact	NA	NA		
S334E	Phosphoserine mimic	8.9 ± 3.2	3.3	9.4 ± 5.1	1.6
WT (phosphorylated)	PKA-phosphorylated Ser-334	521 ± 455	193		
C416S	Abrogation of disulfides	4.6 ± 0.6	1.7	2.6 ± 0.6	0.4
WT dimer	Disulfide-linked dimeric WT ETV1	195 ± 148	72.2	802 ± 164	136
WT dimer (TCEP reduced)	Dimeric WT ETV1, reduced	11.2 ± 2.6	4.1	2.6 ± 0.4	0.4
WT dimer (9:1 GSH:GSSG)	Dimeric WT ETV1, reduced			171 ± 23	29
WT dimer (1:9 GSH:GSSG)	Dimeric WT ETV1, reduced			272 ± 30	47
Etv4 WT (reduced)	Wild type Ets domain, reduced			3.4 ± 0.66	1
Etv4 WT (oxidized)	Disulfide-linked dimeric WT ETV4			148[c] ± 35	44
Etv5 WT (reduced)	Wild type Ets, reduced			11.6 ± 1.1	1
Etv5 WT (oxidized)	Disulfide-linked dimeric WT ETV5			790 ± 610[c]	68

[a] K_d ± S.E.
[b] Fold increase in K_d value compared with the corresponding wild type.
[c] Estimated, due to binding at regression analysis threshold.

Etv4, and Etv5 all sharing over 90% sequence identities, whereas Fev is slightly more distantly related (~65% identity).

In the structural descriptions that follow, the protein residues are numbered according to their position in Etv1; a list of the corresponding residues in Etv4 and Etv5 and Fev is provided in Table 4.

Etv1, Etv4, and Fev display the typical fold of the Ets domain (Fig. 1, *B* and *C*), which contains three α-helices (α1–3) flanking a four-stranded β-sheet, with an additional C-terminal helix (α4). The structures are highly similar (typical root mean square deviation of 1.0 Å over ~90 residues), with the main difference being a two-amino acid deletion in Fev (Fig. 1*A*), linked to a shift of 30° in the orientation of the C-terminal α4 helix in the Fev structure compared with the other structures (Fig. 1*C*). Superposition of the main chains of other Ets domain structures reveals considerable heterogeneity in the region corresponding to helix α4 (Fig. 1*D*). This region was not expected to be helical from proline-scanning mutagenesis (59) and was not structured in the NMR ensemble of the closest structural homologue of Fev, Fli1 (31).

Crystal Structure of the Etv1, Etv4, Etv5, and Fev DNA Complexes—Both Etv1 and Fev DNA complexes were crystallized using a single 10-base pair DNA duplex containing the class I consensus EBS sequence. Etv4 and Etv5 DNA complexes were crystallized using oligonucleotide duplexes that had nine complementary base pairs (containing the same consensus sequence) with a single overhanging nucleotide on each end, which makes favorable interactions with neighboring DNA molecules (Fig. 2*A*). Etv1-DNA crystals diffracted to 2.9 Å and contained a single copy of Etv1 and a single DNA duplex in the asymmetric unit (4BNC, Table 2). Etv4-DNA crystals diffracted to 2.8 Å resolution and contained eight Etv4 molecules and eight DNA duplexes in the asymmetric unit (4UUV, Table 2), with the only significant differences between the various chains in the asymmetric unit being the chemical environment of the

DNA chains. Etv5-DNA crystals diffracted to 1.95 Å and contained a single copy of Etv5, a single calcium ion (bound close to the N terminus of α4), and a single DNA duplex in the asymmetric unit (4UN0, Tables 1 and 2). Fev-DNA crystals diffracted to 2.64 Å, with two copies of Fev per asymmetric unit (2YPR, Tables 1 and 2). In all the DNA complex crystals, base pair stacking interactions between neighboring molecules allow the DNA to form a pseudo-contiguous helix that, in some cases, runs the entire length of the crystal (Fig. 2*A*).

In all four DNA complexes, the recognition helix (α3) inserts deep into the major groove and provides multiple contacts with the base pairs of the core GGAA sequence, as well as with the phosphate backbone (Fig. 2, *B* and *C*). Additional contacts to the DNA backbone are provided by the C terminus of α2, the α2-α3 loop, β3, and the β3-β4 loop. The proteins contact the DNA over a span of 9 bp, and the interface accounts for the burying of ~9% (500 Å^2) of the total solvent-accessible surface area. To accommodate these extensive interfaces, the DNA becomes slightly distorted from the canonical B form, exhibiting smooth bending of ~20° toward the protein with significant widening of the major groove (~20.5 Å at widest point) in the vicinity of the recognition helix and narrowing of the minor groove downstream (~9.3 Å at narrowest point) (Fig. 2*D*). The bending of the DNA can be explained by the large number of contacts made to the phosphodiester backbone, many of which are charged in nature possibly introducing bending through asymmetric phosphate neutralization (60). Specifically in Etv1, hydrogen bonds to phosphate oxygens are formed with the side chains of residues Gln-336, Trp-375, Tyr-386, Ser-392, and Tyr-396, whereas salt bridges are formed with Lys-379, Lys-388, Lys-399, and Lys-404 (Fig. 2*C*). Two additional charged residues, Arg-381 and Arg-409, appear to be in position with the potential to interact with the phosphodiester backbone; the former occupies a position between the two DNA strands in the region of the minor groove narrowing, and
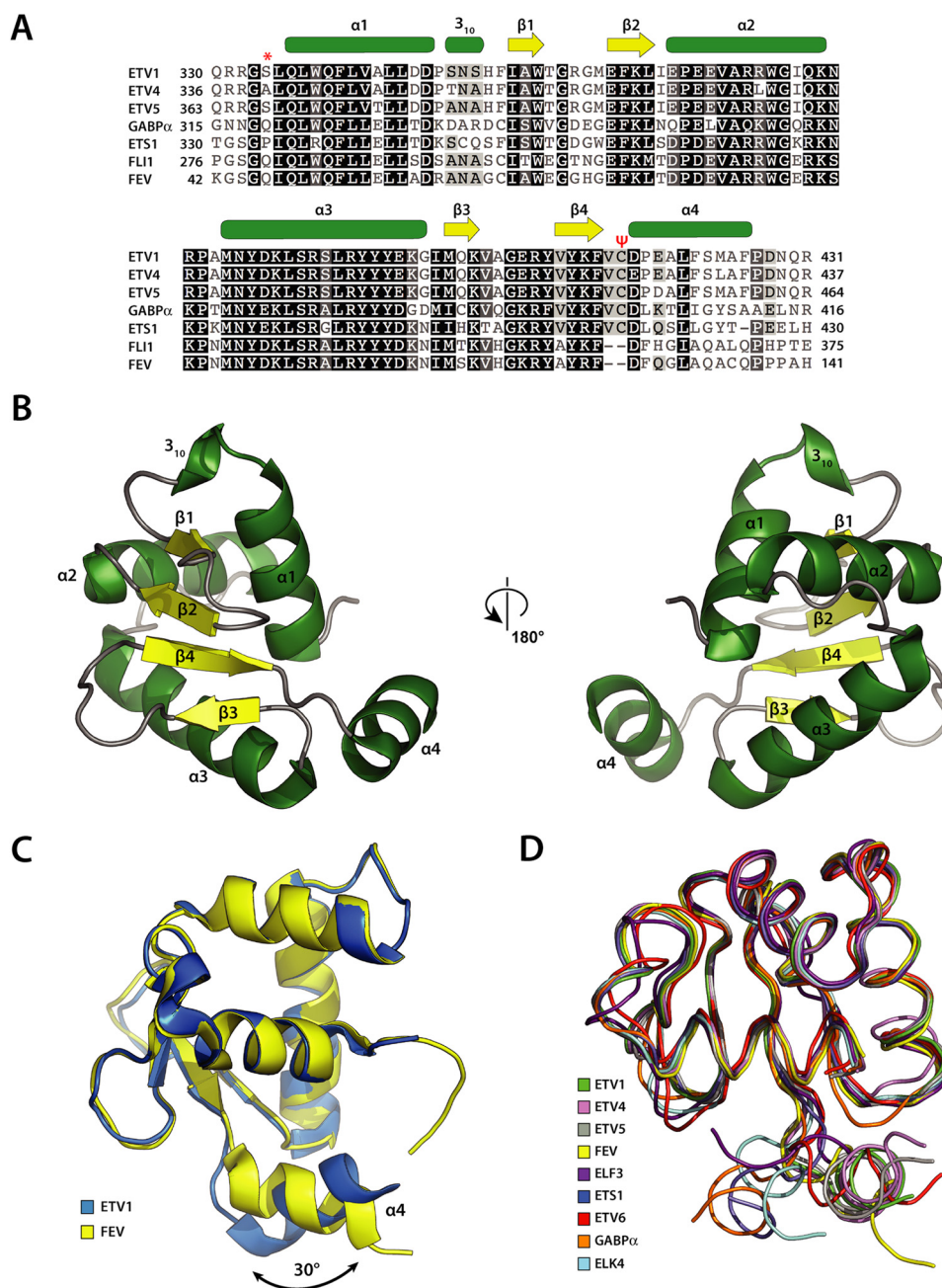
FIGURE 1. **Etv1 and Fev Ets domain topology and sequence alignment.** A, sequence alignment of class I Ets domains. Sequences were Etv1 (AAD29877), Etv4 (AAH16623), Etv5 (CAG33048), GABPα (NP_001184226), ETS1 (CAG47050), FLI1 (AAH10115), and Fev (NP_059991). Secondary structural assignment is for Etv1 (PDB code 4AVP), containing additional flanking residues where appropriate. *Green*, helix (α or 3_{10}); *yellow*, β-strand; *, serine phosphorylated by protein kinase A; ψ, conserved cysteine involved in disulfide bond formation in Etv1, -4, and -5. B, overall cartoon representation of Etv1 Ets domain. *Green*, helix (α or 3_{10}); *yellow*, β-strand. C, superimposition of Etv1 (PDB code 4AVP, *blue*) and Fev (PDB code 2YPR, *yellow*). D, ribbon superimposition showing the variance in the positions of α4. Structures used were Etv1 (PDB code 4AVP), Fev (PDB code 2YPR), ELF3 (PDB code 3JTG), ETS1 (PDB code 3MFK), Etv6 (PDB code 2LF7), GABPα (PDB code 1AWC), and ELK4 (PDB code 1BC8).

the latter could potentially form an additional salt bridge in longer DNA sequences. All of the equivalent interactions are preserved in Etv4, Etv5, and Fev with the exception of Arg-381 and Ser-392, which are replaced by Lys-93 and Ala-104, respectively, in Fev.

Recognition of Specific DNA Substrates by Etv1 and Fev—The Ets family of transcription factors display sequence-specific DNA recognition for nine residues, which we have numbered −3 to +6 around the invariant 5′-GGA-3′ core. This core GGA sequence is directly recognized in all Ets family members by two

arginines and a tyrosine (Arg-391, Arg-394, and Tyr-395 in Etv1; Arg-103, Arg-106, and Tyr-107 in Fev), found invariably on the recognition helix (α3). The two arginine residues are both positioned with their guanidinium groups directly above the O6 and N7 of the two guanine bases and, due to the requirement for two hydrogen bond acceptors, are capable of direct recognition of guanine at positions 1 and 2 (Fig. 3A). Similarly, the tyrosine is in a position to accept a hydrogen bond from the N6 of the adenine residues at position 3 (Fig. 3A); positioning of the complementary thymine methyl group in a nonpolar envi-

TABLE 4
Numbering of equivalent residues in Etv1, -4, and -5 and Fev

Etv1	Etv4	Etv5	Fev
Gln-336	Gln-342	Gln-369	Gln-48
Leu-337	Leu-343	Leu-370	Leu-49
Trp-375	Trp-381	Trp-408	Trp-87
Lys-379	Lys-385	Lys-412	Lys-91
Asp-387	Asp-393	Asp-420	Asp-99
Lys-388	Lys-394	Lys-421	Lys-100
Arg-391	Arg-397	Arg-424	Arg-103
Ser-392	Ser-398	Ser-425	Ala-104
Arg-394	Arg-400	Arg-427	Arg-106
Tyr-395	Tyr-401	Tyr-428	Tyr-107
Tyr-396	Tyr-402	Tyr-429	Tyr-108
Tyr-397	Tyr-403	Tyr-430	Tyr-109
Lys-399	Lys-405	Lys-432	Lys-111
Lys-404	Lys-410	Lys-437	Lys-116
Arg-409	Arg-415	Arg-442	Arg-121
Tyr-410	Tyr-416	Tyr-443	Tyr-122
Cys-416	Cys-422	Cys-449	Phe-F126

ronment created by the aliphatic portion of the side chains of Arg-391 (Fev-Arg-103) and Lys-388 (Fev-Lys-100) may also contribute to recognition at this position. The characteristic positioning of the two arginine side chains is likely to be further stabilized by cation-π stacking interactions, which, in the case of Ets domains, form between the guanidinium group of the arginine and the base at the −1 position, on the same strand as the hydrogen bonded partner. These types of interactions have already been identified in other Ets domain DNA complexes and may contribute to sequence specificity at this site due to the fact that arginine-guanine interactions are energetically the most favorable (61, 62).

The DNA sequences recognized by different subclasses of the Ets family outside the central GGA core are more diverse (32, 33). Etv1, Etv4, Etv5, and Fev all belong to the class I subfamily of Ets domains with a preferred consensus sequence 5′-ACCGGAAGT-3′. Both Etv1 and Fev display an absolute requirement for positions +1 to +3, strong preferences at positions −1, −2, and +4, and somewhat more relaxed sequence selectivity at positions −3, +5, and +6 (32). The means by which this recognition is achieved is not currently well understood, as relatively few direct base contacts outside the core GGA motif can be seen in the current structures in the PDB (30, 63– 65). From analysis of the high resolution structures of the Etv5 and Fev DNA complexes (which are high enough resolution (< 2.0 Å) to reliably locate ordered waters), a number of additional contacts can be seen, which appear to enable direct readout of additional base pairs both upstream and downstream of the GGA core. The two C-G base pairs at positions −2 and −1 lie close to a highly coordinated network of four water molecules conserved in both structures and are coordinated by polar contacts to the side chains of Asp-387 (Fev Asp-99), Ser-390 (Fev Ser-102), Arg-394 (Fev Arg-106), Tyr-412 (Fev Tyr-124), and the phosphate oxygens of the two cytosine nucleotides at positions −2 and −1. These waters are also found in conserved positions in the crystal structures of the DNA complexes of SAP-1 (64) and ELK-1 (65) and form a hydrogen bond to the N4 of the cytosine at position −1 (Fig. 3B, Fig. 4, A−C). A detailed examination of the chemical environment of these water molecules using known hydrogen bonding donors and acceptors as fixed entities (assuming normal protonation states at physiological pH) reveals a single unique
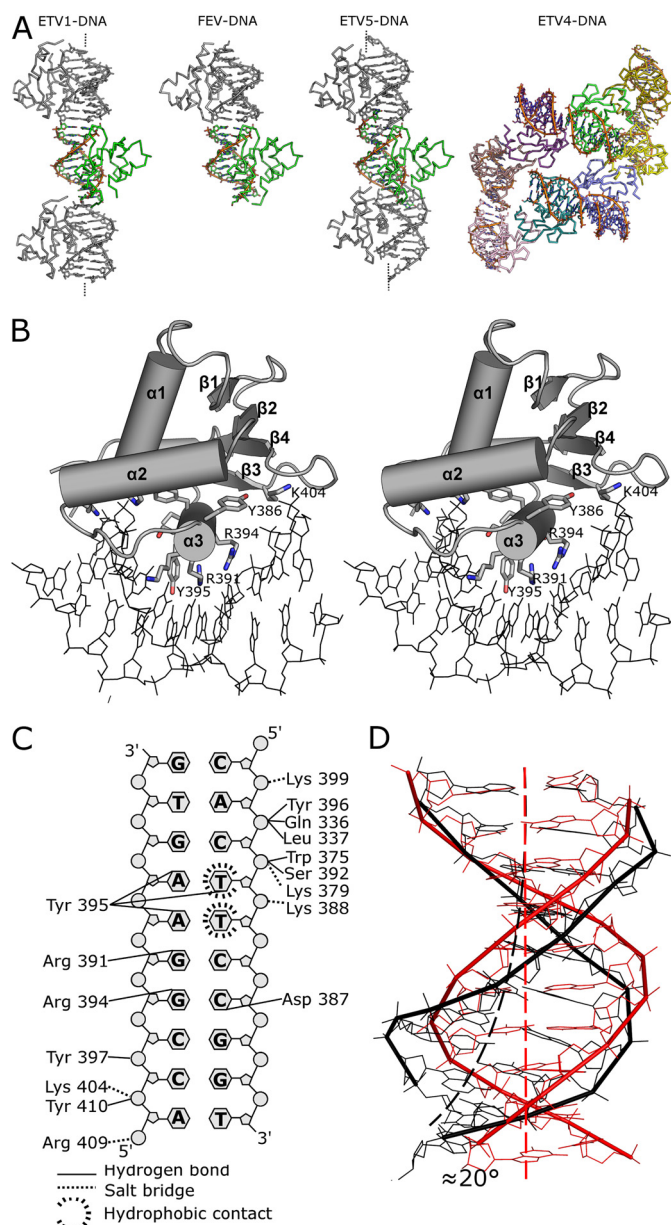
FIGURE 2. **Interactions of Etv1 with DNA.** *A,* contacts formed by DNA complex crystals. The DNA molecules in the DNA complex crystals associate via blunt (Etv1 and Fev) or sticky ends (Etv4 and Etv5) with neighboring molecules in the crystal. The contents of the asymmetric unit are colored by chain with symmetry-related molecules colored *gray*. DNA fibers ending in *dashed lines* form a continuous helix running the entire length of the crystal. *B,* stereo view of the interaction between Etv1 and DNA. Etv1 is shown in cartoon representation with secondary structure elements labeled; residues forming close contacts with the DNA are shown as *sticks,* and the DNA is shown in the *line* representation. The DNA can be seen to undergo a significant widening of the major groove to accommodate extensive interactions with residues from the recognition helix α3. *C,* schematic view of the residues involved in the DNA protein interface of both Etv1 (shown on the *left*) and Fev (shown on the *right*). Hydrogen bonds are depicted as *solid black lines,* and salt bridges are represented as *dashes. D,* comparison of the DNA from the Etv1 DNA complex (*black*) with canonical B form DNA of similar length (*red*). The DNA can be seen to undergo a significant widening of the major groove and bending of ~20° toward the protein interface.

arrangement that satisfies all of the interactions in the network (Fig. 4D). The requirement in this network for the water closest to the base at position −1 to be a hydrogen bond acceptor almost perfectly explains the observed sequence preferences at
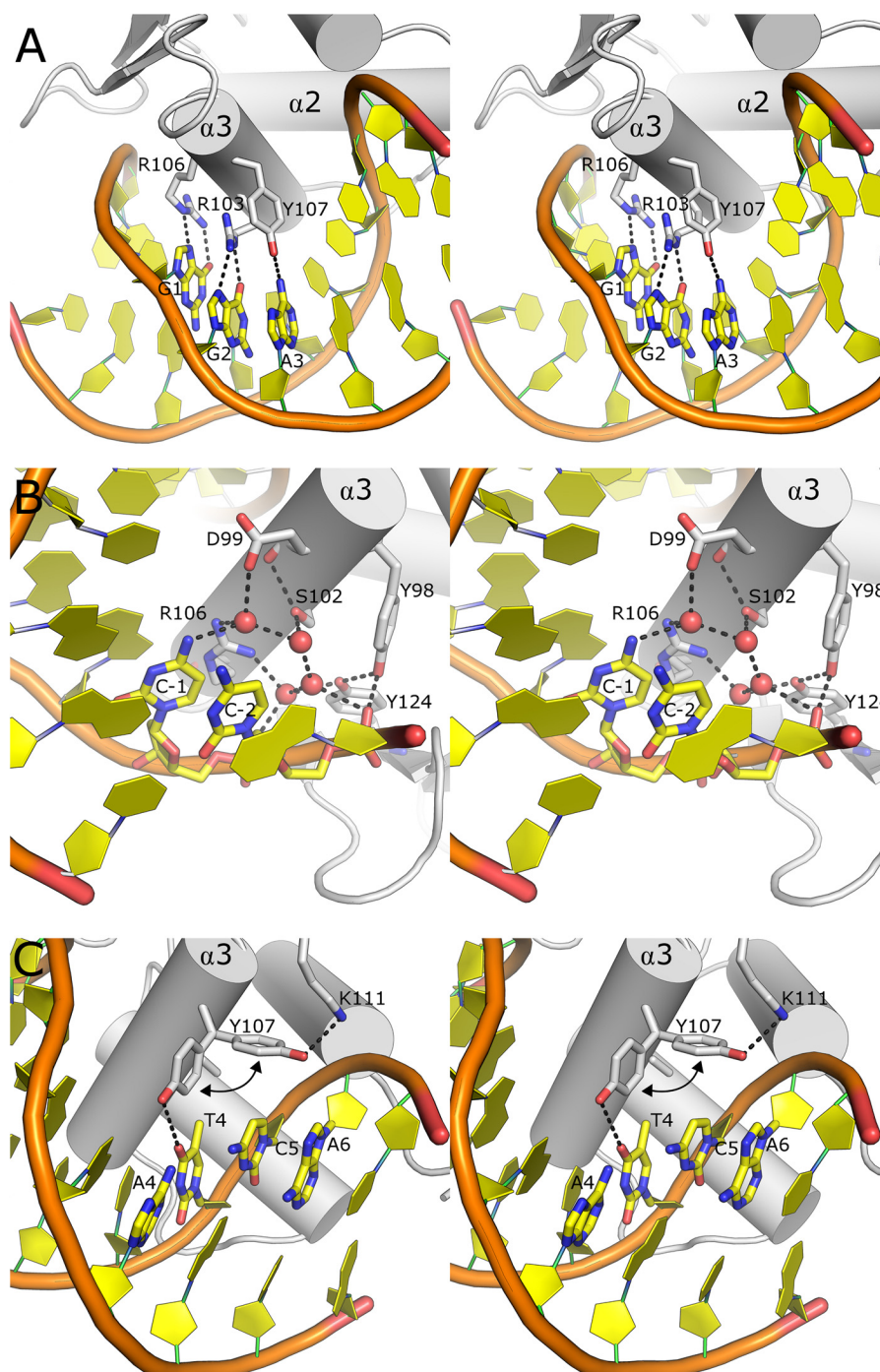
FIGURE 3. **Molecular basis of DNA sequence recognition.** *A,* stereo view of the sequence-specific protein DNA interactions between Fev and DNA that facilitate the recognition of the core GGA motif. Key residues and nucleobases are shown in the *stick* representation, and hydrogen bonds are shown as *dotted lines. B,* recognition of features upstream of the GGA core is achieved via a highly coordinated network of water molecules that are conserved among other class I ETS domains. *C,* recognition of DNA sequence features downstream of the core GGA motif by the dual conformation of Tyr-107, which appears to be able to form both polar and nonpolar interactions with the thymine at position 4, and in its alternative conformation it forms van der Waals interactions with nucleobases at the +4 and +5 positions on the complementary strand from the GGA core.

this position (32), with a strong selection of cytosine (~75% occurrence) and the only other allowed base being adenine (~15% occurrence), which is also able to donate a hydrogen bond, although in this case the hydrogen bonding distance would be significantly longer (≈4.0 Å). The close packing of these water molecules also appears sufficient to explain the selectivity at the −2 position (in which cytosine is the most favored base and thymine is the least frequent), due to the

potential to form favorable van der Waals-type interactions with the nonpolar face of the cytosine and potentially unfavorable interactions or steric clashes with the thymine methyl C7. This mode of recognition also likely occurs at the −1 position; consequently, there would be a similar discrimination against 5-methylcytosine, which could occur in position −1 (which is part of a CpG base step). Interestingly, one high resolution crystal structure in which the positions of these waters is not con-
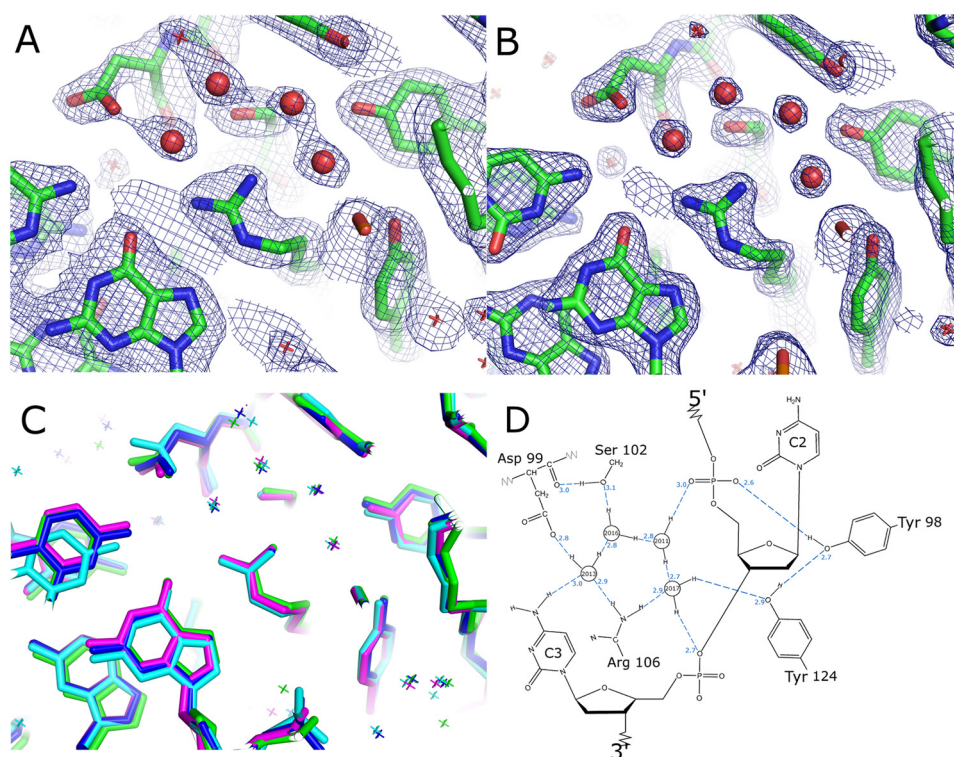
FIGURE 4. **Cluster of conserved water molecules in the protein DNA interface.** *A,* $2F_o - 1F_c$ electron d.5σ. *B,* $2F_o - 1F_c$ electron density map of the Etv5-DNA complex, contoured at 1.5σ. *C,* structural superposition of the Fev (*green*), Etv5 (*blue*), SAP-1 (PDB code 1BC8, *cyan*), and ELK1-DNA (PDB code 1DUX, *pink*) complexes showing the conserved positions of the four water molecules important for recognition of bases at the −1 and −2 positions. *D,* schematic view of the water network with hydrogen atoms shown explicitly and hydrogen bonds and distances shown in *blue*. A single arrangement of hydrogen atoms is sufficient to fulfill all of the required donor-acceptor pairs, and it includes the requirement for a hydrogen bond donor at the −1 position, explaining the observed substrate specificity at this site.

served, the PU.1 DNA complex (66), is a member of the class III subfamily of Ets domains that are specific for EBS sequences containing adenine and guanine in the −1 and −2 positions, respectively (32). The base pairs at positions +4 and +5 downstream from the GGA motif lie close to the side chain of Tyr-395, which has the potential to accept a hydrogen bond from the N6 of the adenine at position +4, and also to form a favorable van der Waals interaction with the corresponding thymine C7 methyl.

In the Etv5 and Fev DNA complex structures, the corresponding residue (Etv5-Tyr-428 and Fev-Tyr-107, respectively) can be seen to occupy two conformations, one in which the two interactions detailed above are conserved, and an additional conformation where the side chain hydroxyl lies close to Etv5-Lys-432/Fev-Lys-111 and provides a large contact surface, with the potential for favorable van der Waals interactions to bases at positions +5 and +6 on the complementary strand (Fig. 3*C*). As discussed below, mutation of the corresponding tyrosine in Etv1 (Tyr-395) to phenylalanine had little effect on DNA binding affinity, suggesting that this van der Waals interaction is the most significant contribution of this residue to binding energy. Although these interactions may not be sufficient to determine the absolute sequence preferences at these positions, the experimentally derived preferences are more relaxed, with a tolerance for thymine at +4, adenine at +5, and cytosine at +6 (32). Both the importance to DNA recognition and the dynamic nature of this residue have already been demonstrated in a comparative structural study of the DNA binding properties of the Elk-1 and Sap-1 Ets domains (65) in which a salt bridge, estab-

lished between the nearby Lys (equivalent to Fev-Lys-111) and a neighboring Asp residue (equivalent to Fev-Asp-110), was thought to stabilize the tyrosine in its alternative position. The finding that both conformations of this residue appear to occur together indicates that the dynamic nature of the side chains of the recognition helix may be even more important for DNA recognition than previously thought.

Order-Disorder Transitions upon DNA Binding—A comparison of the structures of Etv1, Etv4, and Fev in the presence and absence of DNA allows us to identify conformational changes that may occur upon DNA binding. In both cases, the overall structure is very well conserved (root mean square deviation of ~1 Å), with only minor movements of both the N and C termini and the β3-β4 loop, which appear to move away from the DNA in the DNA complexes of both Etv1 and Fev, avoiding potential steric overlaps that can be seen to occur following structural superposition. On an individual residue level, the most striking transitions can be seen to occur in the recognition helix α3, where in Etv1 Asp-387, Arg-391, and Arg-394 can be seen to become ordered upon DNA binding (Fig. 5, *A*, and *B*) and Tyr-395, which can be seen to switch between the two alternative rotamers found to be important for recognition of bases downstream of the GGA motif in the Fev-DNA complex. Similar transitions occur in the Fev structure where Asp-99, Arg-103 and Tyr-107 adopt different rotamers, and the disordered Arg-106 and Lys-111 become ordered upon DNA binding. In Etv4, in addition to three residues Arg-397, Arg-400, and Tyr-401 adopting different rotamers, the C-terminal end of the recog-
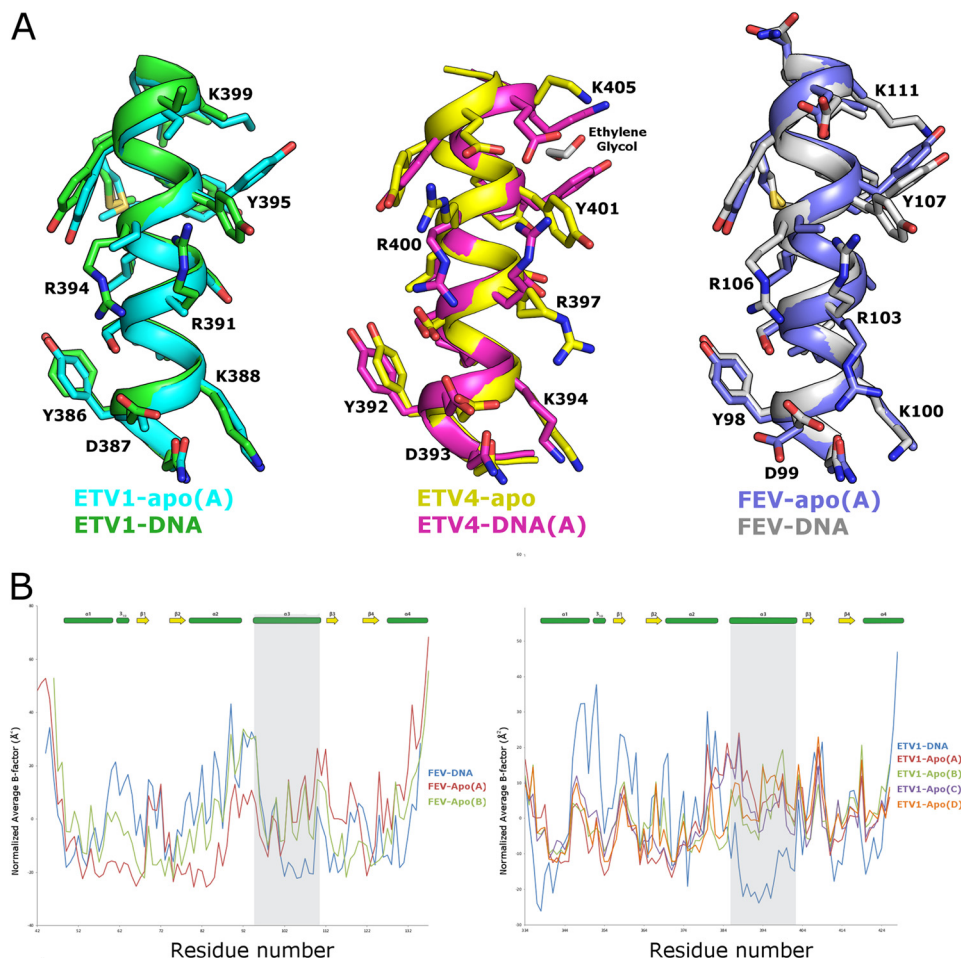
FIGURE 5. **Structural transitions on DNA binding in Fev and Etv1.** *A,* comparison of the apo- and DNA-bound forms of Etv1 (shown on the *left*), Etv4 (*center*), and Fev (shown on the *right*) reveals significant disorder-order transitions and rotamer movements within the recognition helix that accompany DNA binding. *B,* comparison of temperature factors of apo- and DNA-bound forms. The average *B*-factor (normalized so that the mean value is equal to zero) is plotted as a function of residue number for Fev (*left panel*) and Etv1 (*right panel*), with secondary structural elements shown for reference. Residues belonging to the recognition helix α3 are highlighted by a *gray background* and can be seen in both cases to transition from relatively high *B*-factors to among the lowest *B*-factors in the entire structure. These transitions, however, were not seen in a comparison of Etv4 apo- and DNA-bound crystals, probably due to the extensive crystal contacts formed by residues in this helix in the apo-crystals, which consequently diffract to 1.05 Å resolution.

nition helix switches from a 3_{10} hydrogen bonding pattern (which is the form adopted in all the other ETS structures) to a more canonical α-helical hydrogen bonding pattern, although in this case the difference may be due to the fact an ethylene glycol molecule is bound in the vicinity (Fig. 5*A*). The dynamic nature of the residues within the recognition helix is likely to be an important factor for Ets domain recognition, and it indicates possible means whereby Ets domains may scan along the DNA duplex in search for high affinity sites following binding to non-specific sites, in a similar manner to that observed for the *lac* repressor DNA complex (67).

Biochemical Analysis of Etv1 DNA Binding—To assess the contribution of individual residues to DNA binding in the solution we have, using Etv1 as a model system, mutated several residues individually to alanine or (in the case of tyrosine residues) to phenylalanine. Isolated Ets domains containing each of the mutations were compared in EMSAs to wild-type Etv1 (Fig. 6*A* and Table 3). We also compared the Etv-1 construct used for crystallization, which contains two primer-induced mutations, against wild-type Etv1 and found no significant difference in DNA binding (Fig. 6*B*). Somewhat surprisingly, mutants in res-
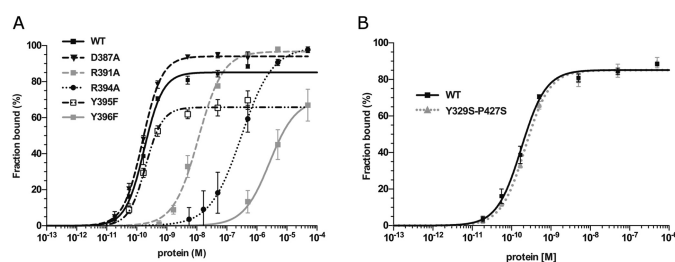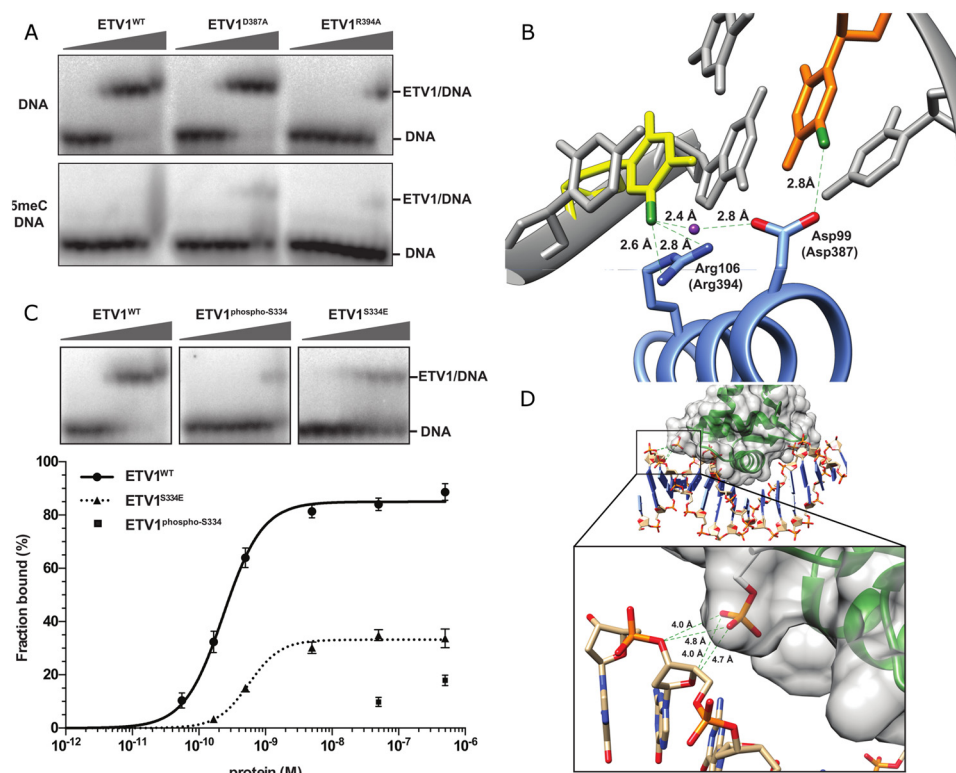

FIGURE 6. **Mutational analysis of Etv1 DNA binding.** *A,* DNA binding isotherms of mutated Etv1 proteins derived from EMSAs using the standard DNA probe at 2 nM. *Error bars* are plotted as ± S.E. of replicates. *B,* comparison of the DNA binding to WT and the primer-induced mutant Etv1 (Y329S/P427S) used for crystallization.

idues involved in salt bridges to the DNA backbone (Lys-379 and Lys-404) exhibited no detectable DNA binding. Similarly, mutants of the tyrosine residues forming hydrogen bonds to the DNA backbone (Tyr-396 and Tyr-397) to either phenylalanine or to alanine also resulted in significant (>1000-fold) reduction in binding affinity to DNA.

FIGURE 7. **Effects of methylation and phosphorylation on Etv1 DNA binding.** *A,* representative EMSA analysis of three Etv1 proteins (wild type, D387A, and R394A) titrated against normal (*upper panel*) and 5meC-modified DNA (Etv1ALF-5meC/Etv1ALR-5meC, *lower panel*). Gradients represent increasing protein concentrations, and unbound DNA substrate and complexes are marked. *B,* model of potential interactions with 5-methylated cytosine, based on the Fev-dsDNA structure (PDB code 3ZP5) with methyl groups on the methylated cytosines shown in *green,* and a conserved water molecule shown in *purple. C,* representative EMSA analysis of wild type (*left panel*), S334E phosphorylation mimic (*right*), and PKA-phosphorylated Etv1 (*center panel*) titrated against normal DNA substrate. The DNA binding isotherms derived from these data are plotted below. *D,* model of potential clashes between phosphorylated Etv1 Ser-334 in the Etv1-dsDNA structure (PDB code 4BNC) superimposed onto a longer dsDNA from PU.1-dsDNA (PDB code 1PUE). Distances between potentially clashing atoms are shown as *green dashed lines.*

Mutations in individual residues involved in nucleobase interactions affected DNA binding to a lesser extent. Mutants of the invariant Arg-391 and Arg-394 bound DNA 50- and 1300-fold less than the wild type, respectively. Interestingly, mutation of Tyr-395 to phenylalanine did not affect the K_d values, but maximal DNA binding at saturation was only 60% of the total probe; this may indicate a rapid dissociation (high k_{off}) of the complexes. Uniquely among the residues tested, a D387A mutation had no effect on DNA binding; it is possible that this residue contributes little to the free energy of binding to the cognate DNA sequence but may be crucial for selectivity (see below). Overall, the mutational analysis demonstrates that the residues interacting with the DNA backbone in the crystal structure are the major contributors to the binding affinity in solution, although the interactions with the bases add a significant but lower energetic contribution.

*Biochemical Analysis of Etv1 Binding to Methylated DNA—*To test whether CpG methylation directly affects Etv1 binding, we modified the 20-bp oligonucleotides used in the EMSA experiments by substituting 5meC at position −1 of the consensus (CGGAA) along with the cytosine complementary to G+1, generating a fully methylated CpG motif (C^mGGAA/ TTCC^mG). EMSA experiments showed that the Etv1 protein failed to bind the methylated DNA (Fig. 7*A*). Modeling of 5meC at these positions in the Etv1-DNA and Fev-DNA structures (Fig. 7*B*) showed Arg-394/106 (Etv1/Fev) was close to the 5meC

at −1 of the CGGAA consensus strand, and Asp-387/99 was close to 5meC at both −1 of the CGGAA strand and +1 on the complementary strand. Methylation of cytosines at these positions could create a steric clash with Arg-394/106 causing loss of hydrogen bonds involved in the conserved water network. EMSA analyses of Etv1^D387A and Etv1^R394A mutants showed that replacement of either of these charged side chains with the smaller and hydrophobic alanine recovered some binding to methylated DNA (3 and 2% of total DNA bound, respectively, at the highest protein concentration) compared with the wild-type protein (Fig. 7*A*). It may be significant that Asp-387 (or glutamate) is almost invariant in the ETS recognition helix, with only class III proteins displaying a substitution to glutamine (32). As class I Ets proteins predominantly select cytosine at the −1 position of the consensus, although class III members typically have adenine (32), it is possible that Asp-387 (or Glu) plays a universal role not only in the preference for cytosine at −1 but also to the discrimination against CpG-methylated sites.

*Direct Regulation of DNA Binding by PKA Phosphorylation of the Etv1 Ets Domain—*Ets proteins are regulated by phosphorylation (68); for example, the protein kinases RSK1 and PKA phosphorylate Etv1 at multiple positions *in vivo* (40). PKA phosphorylation of Etv1 Ser-334 (at the N-terminal edge of the Ets domain) inhibits *in vitro* DNA binding of the full-length protein but increases transcriptional transactivation potential, as does
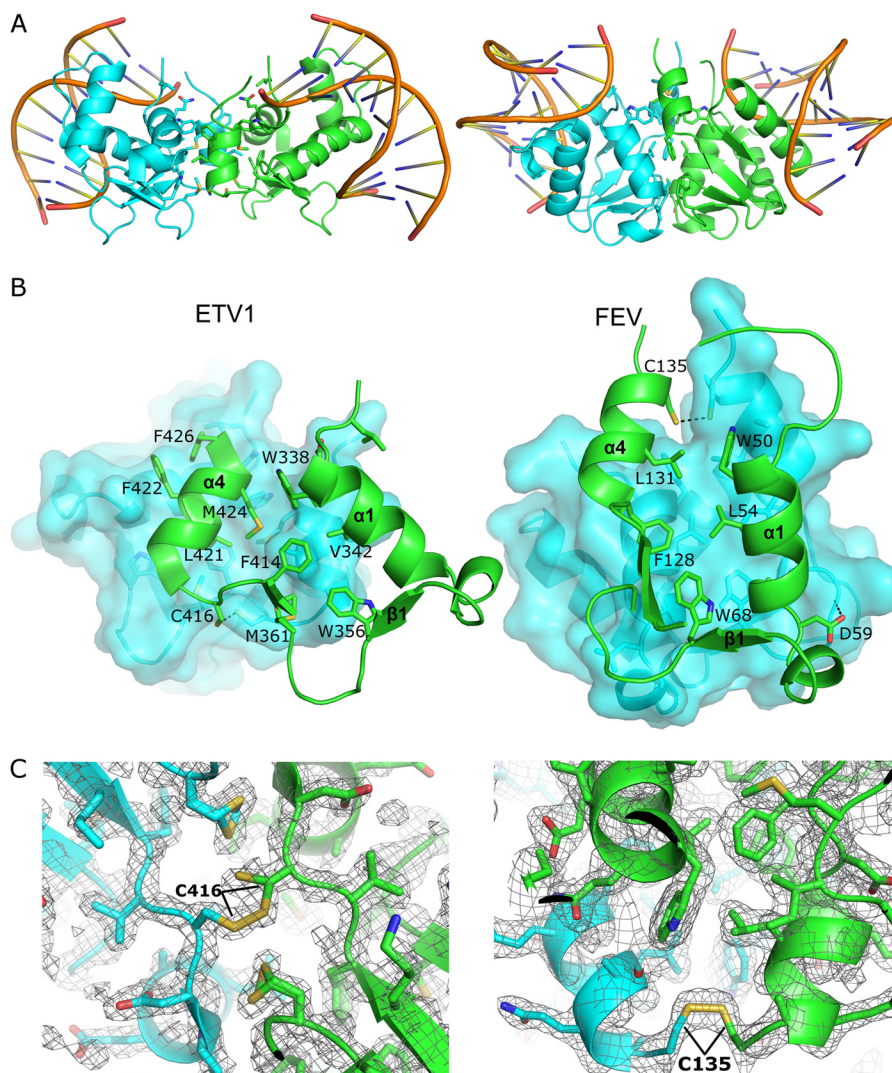
FIGURE 8. **Overview of the dimerization interfaces.** *A,* view of the potential dimers formed in Etv1 (shown on the *left*) and Fev (*right*) DNA complexes with the 2-fold axis vertical, showing the DNA in cartoon representation. The structures of the Etv4 and Etv5 dimers are essentially identical to Etv1. *B,* close view of the secondary structural elements and residues that compose the dimer interfaces for Etv1 (*left*) and Fev (*right*) with polar contacts depicted as *dashed lines. C,* $2F_o - 1F_c$ electron density maps, for Etv1 (*left*) and Fev (*right*), both contoured at 1.0 σ in the vicinity of the intermolecular disulfide bond. Two conformations can be seen for Cys-416 in the Etv1 structure (also observed in the high resolution structures of Etv4 and Etv5), although it is not clear whether this is induced by radiation damage or an intrinsic property of this residue.

phosphorylation of the equivalent Ser-367 in Etv5 (41). In contrast, Etv4 lacks a serine at the equivalent position (*asterisk,* Fig. 1*A*). To assess whether Ser-334 phosphorylation of the Ets domain affects DNA binding directly, we phosphorylated Etv1 with PKA on Ser-334 *in vitro* (Fig. 7, *C* and *D*). In parallel, we produced a phosphomimetic mutant of Ser-334 to glutamate. EMSA experiments showed that phosphorylation of Ser-334 reduced the binding affinity by at least 200-fold (Fig. 7*C*). This confirms that Ser-334 phosphorylation can directly interfere with DNA binding of the Etv1 Ets domain, independent of other regions of Etv1 or of other proteins (40). Interestingly, the S334E mutation caused only a modest reduction in the binding constant, but overall probe binding levels were 35% at apparently saturating DNA concentration. This may indicate a kinetic instability of DNA binding by the mutant protein. A phosphoserine residue is bulkier and carries more negative charge than glutamate, suggesting that the strong effect of the phosphate is due to charge repulsion. Modeling of a phosphor-

ylated Ser-334 in the Etv1-dsDNA structure (PDB code 4BNC) superimposed onto a longer dsDNA (Fig. 7*D*) shows a possible repulsive interaction between the phosphate on the protein and the DNA backbone, which may underlie the decrease in DNA binding.

Etv1/4/5 and Fev Are Disulfide-linked Homodimers—Analysis of the contacts between adjacent molecules in the Etv1, Etv4, and Etv5 crystals reveals a conserved potential dimer interface that is distant from the DNA binding surface and is present in all the DNA-free and DNA-bound crystal forms (Fig. 8*A*). This interface is centered around contributions from the N- and C-terminal helices, α1 and α4, with additional contacts involving residues from the β1-β2 loop (Fig. 8*B*). The contact area for this interface is ∼800 Å², includes contributions from 20 residues, and accounts for the burying of 15% of the monomer accessible surface area upon complex formation. The majority of the interface is nonpolar in nature with extensive hydrophobic interactions. Similarly, an analysis of the crystal contacts in
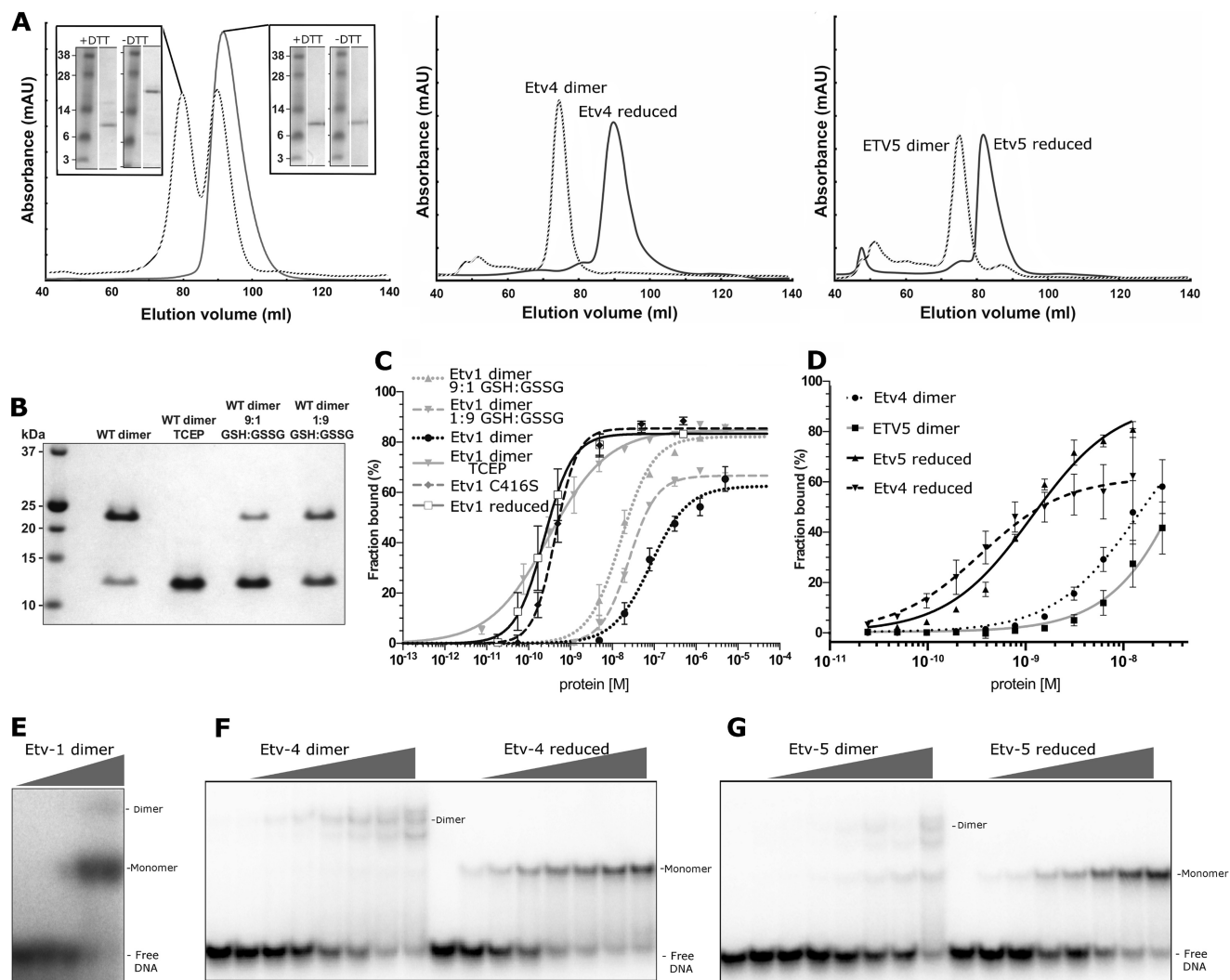
FIGURE 9. **Redox-dependent regulation of Etv1, Etv4, and Etv5.** *A,* purification of Etv1 (*left panel*), Etv4 (*center panel*), and Etv5 (*right panel*) under reducing (*solid lines*) or nonreducing conditions (*dashed lines*). The *inset in the left-hand panel* shows the SDS-PAGE analysis of the size exclusion fractions denatured in the presence (*upper right*) or absence (*lower right*) of 5 mM DTT. *B,* nonreducing SDS-PAGE of an SEC fraction of Etv1 eluting as a dimer (WT dimer) or treated prior to gel loading under various reducing conditions: 10 mM TCEP, and different ratios of reduced (GSH) and oxidized (GSSG) glutathione. *C,* DNA binding isotherms of different oligomeric/redox states of Etv1 from EMSA analysis. The proteins used were as follows: Etv1 purified under reducing conditions (monomeric); Etv1-C416S mutant (monomeric); Etv1 dimer purified under oxidizing conditions; Etv1 dimer treated with 10 mM TCEP; Etv1 dimer treated with reduced glutathione: and Etv1 dimer treated with oxidized glutathione. *D,* DNA binding isotherms of Etv4 and Etv5 from EMSA analysis. The proteins used were as follows: Etv4 purified under oxidized conditions (dimeric); Etv5 purified under oxidized conditions (dimeric); Etv4 dimer treated with 10 mM TCEP; and Etv5 dimer treated with 10 mM TCEP. DNA probe concentration was 0.2 nM. *Error bars* are plotted as ± S.E. *E,* EMSA of dimeric Etv1 showing a slower mobility Etv1-DNA species that is presumed to represent the dimeric species. *F,* EMSA of dimeric and reduced Etv5 showing bands representing monomeric and dimeric DNA complexes. *G,* EMSA of dimeric and reduced Etv5 showing bands representing monomeric and dimeric DNA complexes.

both the Fev and Fev-DNA crystals reveals an interface that includes the same components (α1 and α4 and the β1-β2 loop) as that of the other three Etv proteins. However, the dimer interfaces in Etv1/4/5 and in Fev are not equivalent; rather, they differ by a relative rotation of ~90°, reflecting the different orientations of the α4 helix (Fig. 1C). The contact area for the Fev dimerization interface is ~700 Å², spans 21 residues, and accounts for 12.5% of the monomer accessible surface area. Similar to Etv1, the interface is dominated by contacts from hydrophobic residues, a single hydrogen bond (Asp-59–His-72), and an intermolecular disulfide bond between Cys-135 and its symmetry mate (Fig. 8, *B* and *C*). All the proteins were purified and crystallized in the presence of reducing agent (1 mM TCEP), although the reducing agent may have become oxidized during crystallization in an oxygen-containing environment.

Intermolecular dimerization has been documented previously in Ets factors (69), with ETS1 forming head-to-head homodimers required for cooperative binding on palindromic and other sites (70). ELK1 dimerization appears to regulate ELK1 cellular stability (71), with a homotypic interface utilizing the α1 helix, similar to that of Etv1/Fev, but with a different positioning and orientation (65). The presence of an intermolecular disulfide bond formed between similar C-terminal regions Etv1, Etv4, Etv5, and Fev is particularly significant, as they are homologous to ETS1 that can form disulfide-linked dimers *in vivo* (72). Although present in the crystal structures, disulfide-linked Ets domain dimers were not initially observed in solution (Fig. 9A). When Etv1, Etv4, or Etv5 was expressed in Rosetta-gami™ 2 *Escherichia coli* (73) and purified under nonreducing conditions, a peak consistent with the dimer

molecular weight was observed in size exclusion chromatography (Fig. 9A). SDS-PAGE in the presence or absence of DTT suggests the higher molecular weight peak contained a reducing agent-sensitive dimer, which was further confirmed by mass spectrometry. We have been unable to test whether Fev also forms disulfide-linked homodimers, as our Fev constructs did not produce soluble protein in the Rosetta-gami™ 2 *E. coli* cells.

Disulfide Bond Formation and Dimerization Inhibit DNA Binding—To discover the functional consequences of disulfide-dependent dimerization, we compared the DNA binding affinities of purified monomeric and dimeric Etv1 using EMSA. The reduced monomeric Etv1 protein bound the DNA probe with an affinity up to 200-fold higher than the oxidized dimeric form (Table 3; Fig. 9, *A* and *C*); a C416S mutant of Etv1 bound DNA with an affinity similar to the reduced wild-type Etv1 (0.26 nM). Similarly, dimeric Etv4 and Etv5 also bind DNA with affinities that are 2 orders of magnitude lower than the reduced monomeric forms (Fig. 9, *D, F,* and *G* and Table 3; note that the K_d values of the dimeric forms are rough estimates, as saturation could not be reached). A graded response to physiological redox potentials could be seen when Etv1 dimers were treated with different ratios of reduced and oxidized glutathione; the extent of DNA binding roughly correlates with increasing the reducing agent potential and dimer dissociation (Fig. 9, *B* and *C*). We conclude that Etv1, Etv4, and Etv5 binding to DNA is reversibly inhibited by the formation of a disulfide-linked dimer.

DNA complexes of dimeric Etv4 and Etv5 migrate at a lower mobility than complexes of the monomeric proteins (Fig. 9, *F* and *G*), indicating the presence of a dimer of Etv4/5 with one or two bound probe molecules. EMSA using the dimeric form of Etv1 results in a band with mobility of a monomer (Fig. 9E); we assume this reflects the small amount of monomeric protein that is present in the high molecular weight SEC fractions (Fig. 9B, *left lane*); this monomeric subpopulation could not be removed by repeated chromatography. However, at the highest concentrations of dimeric Etv1 and DNA, a weak slower migrating band could be observed with a mobility similar to that of the Etv4 and Etv5 dimeric complexes. The paradoxical observation of the dimer in complex with DNA in the various crystal structures may be accounted for by high concentrations of both protein and DNA (~1 mM) during crystallization, which may be well in excess of the dissociation constant. We also note that both DNA-bound and -unbound forms of the proteins crystallized in the disulfide-bonded form, perhaps indicating that the monomeric form is considerably more flexible and difficult to crystallize.

Discussion

Although the interaction between Ets domains and DNA has been the subject of numerous previous structural and biochemical studies (32, 74), the question of how Ets domains are able to achieve DNA sequence specificity beyond the GGA consensus motif has remained open. Furthermore, although a variety of post-translational modifications regulate Ets transcription factors (39, 40, 75–77), the mechanisms involved have yet to be unambiguously clarified (8). In particular, it is not known

whether modification exerts a direct effect on Ets protein-DNA binding, protein-protein interactions, or stability.

Our structural analysis of the Etv1, Etv4, Etv5, and Fev Ets domains both in the presence and absence of DNA has allowed us to identify additional features of the protein-DNA interface. Mutation of individual residues supports predictions from the crystal structure for binding the GGA core for most residues tested. Remarkably, mutating any of the residues of Etv1, which interact directly with the DNA backbone (Lys-379, Tyr-396, Tyr-397, and Lys-404), led to complete loss of binding. The extent to which these substitutions affect the binding affinity indicates that these residues might play a crucial role in opening the DNA to allow the $\alpha3$ recognition helix to access the widened major groove or anchor the Ets domain once recognition has occurred. Further interface features include the dynamic nature of a conserved tyrosine (Tyr-395 in Etv1), which may allow for recognition for up to three bases downstream from the GGA core. In addition, a conserved cluster of coordinated water molecules supports a structural basis for sequence recognition of two bases upstream of the GGA core, including selectivity against thymine as well as 5-methylcytosine bases at position -1 of the consensus motif. Methylated bases may create an energetically unfavorable environment for conserved aspartate and arginine residues, preventing Ets proteins from binding when transcription is to be repressed by CpG methylation (43), as observed previously for Ets proteins (44–46). Although difficult to rule out the role of indirect recognition through sequence dependence of the DNA bending observed in our structures, the additional direct and water-mediated interactions seen in the Etv1 and Fev protein-DNA crystal structures appear to be sufficient to explain almost entirely the observed sequence specificities.

Etv1 is post-translationally regulated by Rsk1 and PKA phosphorylation *in vivo* (40). We were able to reproduce the inhibition of DNA binding by specific phosphorylation of Ser-334, a known PKA target site, in the isolated Ets domain *in vitro*. This demonstrates a direct effect of post-translational regulation on an Ets domain, and our Etv1-DNA structure suggests that an increase in negative charge around Ser-334 following phosphorylation may directly abrogate DNA binding following electrostatic clashes.

Crystal structures of all four Ets domains exhibit disulfide-linked homodimers. Dimerization of Etv1, -4, and -5 is mediated by the homologous cysteine residues (Cys-416, Cys-422, and Cys-449, respectively). Fev dimers are linked by a different, nonconserved cysteine (Cys-135).

We found that disulfide-mediated dimerization strongly inhibited Etv1, Etv4, and Etv5 DNA binding *in vitro*, which was directly reversible with increasing redox potential. Our crystal structures do not provide an explanation for this inhibition; the DNA in the crystal is bound to the dimeric protein, and the dimerization interface is sufficiently distant from the DNA binding interface to eliminate the possibility of steric effects with the substrates used (20-mer with the consensus sequence centrally located). Examination of the protein DNA interfaces in these crystals and comparison with other active Etv DNA complexes reveal very similar interfaces, with similar contacts made to the phosphodiester backbone and the major sequence-
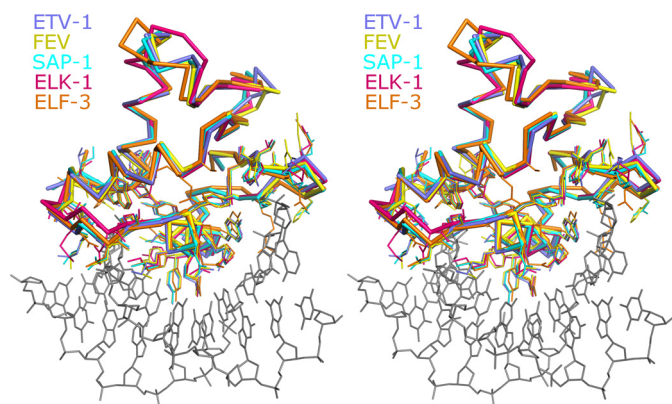
FIGURE 10. **Comparison of the overall features of DNA binding of Etv1 and Fev with other high affinity Ets DNA complexes.** Stereo view of a structural superposition of Etv1 (*blue*) and Fev (*yellow*) (which are in the inhibited form) with three other noninhibited Ets DNA complexes SAP-1 (*cyan*), ELK-1 (*pink*), and ELF-3 (*orange*). A single DNA molecule (taken from the Etv1-DNA complex) is shown in the *stick format* for reference.

specific DNA interactions provided by the recognition helix being totally conserved (Fig. 10). Thus, we believe the interfaces in our dimeric crystal structures to be representative of the higher affinity monomer DNA interface. This leads us to suggest that that the most likely explanation for the inhibition caused by dimerization is an allosteric mechanism of inhibition, which acts through changes in protein flexibility or dynamics. The mechanism may be similar to the autoinhibitory mechanism of ETS1, where packing of a 4-helical inhibitory module also distal to the DNA-binding face allosterically modulates subtle structural changes that inhibit DNA binding (78). It is also possible that the constraints of crystallization favor a protein conformation that is less stable in solution. However, structural studies of allosteric regulation do not always reveal a clear-cut mechanism, and there are cases where allosteric regulation acts through changes in protein flexibility or dynamics (79–81).

Redox control of transcription factors is a recognized regulatory mechanism (80, 82–84). Examples include intermolecular disulfide bridges in bZIP proteins, such as AP-1 (85), and cysteines outside the DNA-binding motif mediating redox-sensitive dimerization in plant homeodomains (86). Ets proteins have been implicated in redox signaling *in vivo*, such as the oxidative inactivation of GABPα by the Hippo pathway (87). The mechanism of redox-dependent DNA binding inhibition in Ets factors has remained unclear, as many contain multiple cysteine residues. Here, we show that a single redox-sensitive cysteine is sufficient to confer inhibition of DNA binding by dimerization.

Notably, cysteines at positions equivalent to Etv1-Cys-416 exist in addition only in GABPα, ETS1, and ETS2 out of all human Ets domain proteins. Although dimerization and disulfide formation seem to have regulatory influence on proteins, including GABPα (88) and ETS1, the equivalent cysteines in crystal structures of ETS1, GABPα, and ETS2 (89, 90) do not seem to be involved in intermolecular disulfide bridges. This indicates that the precise mode of redox regulation seen for Etv1/4/5 may be restricted to a small subset of the Ets protein family.

Etv1, -4, and -5 are strongly implicated in cancer (8). Our data suggest the possibility that redox-mediated dimerization could link Etv1/4/5 factors to the response of cancer cells to their microenvironment. If proven, chemoprevention by targeting such redox-sensitive transcription factors could potentially deliver novel therapeutic strategies (91).

References

1. Sharrocks, A. D. (2001) The ETS-domain transcription factor family. *Nat. Rev. Mol. Cell Biol.* **2,** 827–837
2. Hollenhorst, P. C., McIntosh, L. P., and Graves, B. J. (2011) Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu. Rev. Biochem.* **80,** 437–471
3. Degnan, B. M., Degnan, S. M., Naganuma, T., and Morse, D. E. (1993) The ets multigene family is conserved throughout the Metazoa. *Nucleic Acids Res.* **21,** 3479–3484
4. Leprince, D., Gegonne, A., Coll, J., de Taisne, C., Schneeberger, A., Lagrou, C., and Stehelin, D. (1983) A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306,** 395–397
5. Oikawa, T., and Yamada, T. (2003) Molecular biology of the Ets family of transcription factors. *Gene* **303,** 11–34
6. Lelièvre, E., Lionneton, F., Soncin, F., and Vandenbunder, B. (2001) The Ets family contains transcriptional activators and repressors involved in angiogenesis. *Int. J. Biochem. Cell Biol.* **33,** 391–407
7. Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E. A., Iwasaki-Arai, J., Mizuno, S., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., Fenyus, M. L., Elf, S., Chan, S., Kastner, P., Huettner, C. S., *et al.* (2005) Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106,** 1590–1600
8. Oh, S., Shin, S., and Janknecht, R. (2012) Etv1, 4 and 5: an oncogenic subfamily of ETS transcription factors. *Biochim. Biophys. Acta* **1826,** 1–12
9. Chotteau-Lelievre, A., Montesano, R., Soriano, J., Soulie, P., Desbiens, X., and de Launoit, Y. (2003) PEA3 transcription factors are expressed in tissues undergoing branching morphogenesis and promote formation of duct-like structures by mammary epithelial cells *in vitro*. *Dev. Biol.* **259,** 241–257
10. Abe, H., Okazawa, M., and Nakanishi, S. (2011) The Etv1/Er81 transcription factor orchestrates activity-dependent gene regulation in the terminal maturation program of cerebellar granule cells. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 12497–12502
11. Sedý, J., Tseng, S., Walro, J. M., Grim, M., and Kucera, J. (2006) ETS transcription factor ER81 is required for the Pacinian corpuscle development. *Dev. Dyn.* **235,** 1081–1089
12. Wang, Y., Wang, L., Chen, Y., Li, L., Yang, X., Li, B., Song, S., Yang, L., Hao, Y., and Yang, J. (2011) ER81 expression in breast cancers and hyperplasia. *Patholog. Res. Int.* **2011,** 980513
13. Jané-Valbuena, J., Widlund, H. R., Perner, S., Johnson, L. A., Dibner, A. C., Lin, W. M., Baker, A. C., Nazarian, R. M., Vijayendran, K. G., Sellers, W. R., Hahn, W. C., Duncan, L. M., Rubin, M. A., Fisher, D. E., and Garraway, L. A. (2010) An oncogenic role for Etv1 in melanoma. *Cancer Res.* **70,** 2075–2084
14. Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R.,

Sun, X. W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., Lee, C., Montie, J. E., Shah, R. B., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310,** 644–648

15. Shin, S., Kim, T. D., Jin, F., van Deursen, J. M., Dehm, S. M., Tindall, D. J., Grande, J. P., Munz, J. M., Vasmatzis, G., and Janknecht, R. (2009) Induction of prostatic intraepithelial neoplasia and modulation of androgen receptor by ETS variant 1/ETS-related protein 81. *Cancer Res.* **69,** 8102–8110

16. Baena, E., Shao, Z., Linn, D. E., Glass, K., Hamblen, M. J., Fujiwara, Y., Kim, J., Nguyen, M., Zhang, X., Godinho, F. J., Bronson, R. T., Mucci, L. A., Loda, M., Yuan, G. C., Orkin, S. H., and Li, Z. (2013) Etv1 directs androgen metabolism and confers aggressive prostate cancer in targeted mice and patients. *Genes Dev.* **27,** 683–698

17. Vageli, D., Ioannou, M. G., and Koukoulis, G. K. (2009) Transcriptional activation of hTERT in breast carcinomas by the Her2-ER81-related pathway. *Oncol. Res.* **17,** 413–423

18. Shin, S., Oh, S., An, S., and Janknecht, R. (2013) ETS variant 1 regulates matrix metalloproteinase-7 transcription in LNCaP prostate cancer cells. *Oncol. Rep.* **29,** 306–314

19. Chi, P., Chen, Y., Zhang, L., Guo, X., Wongvipat, J., Shamu, T., Fletcher, J. A., Dewell, S., Maki, R. G., Zheng, D., Antonescu, C. R., Allis, C. D., and Sawyers, C. L. (2010) Etv1 is a lineage survival factor that cooperates with KIT in gastrointestinal stromal tumours. *Nature* **467,** 849–853

20. Janknecht, R. (2005) EWS-ETS oncoproteins: the linchpins of Ewing tumors. *Gene* **363,** 1–14

21. Peter, M., Couturier, J., Pacquement, H., Michon, J., Thomas, G., Magdelenat, H., and Delattre, O. (1997) A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene* **14,** 1159–1164

22. Hendricks, T., Francis, N., Fyodorov, D., and Deneris, E. S. (1999) The ETS domain factor Pet-1 is an early and precise marker of central serotonin neurons and interacts with a conserved element in serotonergic genes. *J. Neurosci.* **19,** 10348–10356

23. Hendricks, T. J., Fyodorov, D. V., Wegman, L. J., Lelutiu, N. B., Pehek, E. A., Yamamoto, B., Silver, J., Weeber, E. J., Sweatt, J. D., and Deneris, E. S. (2003) Pet-1 ETS gene plays a critical role in 5-HT neuron development and is required for normal anxiety-like and aggressive behavior. *Neuron* **37,** 233–247

24. Wang, Y. C., Zuraek, M. B., Kosaka, Y., Ota, Y., German, M. S., Deneris, E. S., Bergsland, E. K., Donner, D. B., Warren, R. S., and Nakakura, E. K. (2010) The ETS oncogene family transcription factor Fev identifies serotonin-producing cells in normal and neoplastic small intestine. *Endocr. Relat. Cancer* **17,** 283–291

25. Yan, C., and Higgins, P. J. (2013) Drugging the undruggable: transcription therapy for cancer. *Biochim. Biophys. Acta* **1835,** 76–85

26. Erkizan, H. V., Kong, Y., Merchant, M., Schlottmann, S., Barber-Rotenberg, J. S., Yuan, L., Abaan, O. D., Chou, T. H., Dakshanamurthy, S., Brown, M. L., Uren, A., and Toretsky, J. A. (2009) A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. *Nat. Med.* **15,** 750–756

27. Rahim, S., Beauchamp, E. M., Kong, Y., Brown, M. L., Toretsky, J. A., and Üren, A. (2011) YK-4-279 inhibits Erg and Etv1 mediated prostate cancer cell invasion. *PLoS One* **6,** e19343

28. Donaldson, L. W., Petersen, J. M., Graves, B. J., and McIntosh, L. P. (1996) Solution structure of the ETS domain from murine Ets-1: a winged helix-turn-helix DNA binding motif. *EMBO J.* **15,** 125–134

29. Gajiwala, K. S., and Burley, S. K. (2000) Winged helix proteins. *Curr. Opin. Struct. Biol.* **10,** 110–116

30. 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

31. Liang, H., Mao, X., Olejniczak, E. T., Nettesheim, D. G., Yu, L., Meadows, R. P., Thompson, C. B., and Fesik, S. W. (1994) Solution structure of the ets domain of Fli-1 when bound to DNA. *Nat. Struct. Biol.* **1,** 871–875

32. Wei, G. H., Badis, G., Berger, M. F., Kivioja, T., Palin, K., Enge, M., Bonke, M., Jolma, A., Varjosalo, M., Gehrke, A. R., Yan, J., Talukder, S., Turunen, M., Taipale, M., Stunnenberg, H. G., *et al.* (2010) Genome-wide analysis of ETS-family DNA-binding *in vitro* and *in vivo*. *EMBO J.* **29,** 2147–2160

33. Szymczyna, B. R., and Arrowsmith, C. H. (2000) DNA binding specificity studies of four ETS proteins support an indirect read-out mechanism of protein-DNA recognition. *J. Biol. Chem.* **275,** 28363–28370

34. Hollenhorst, P. C., Chandler, K. J., Poulsen, R. L., Johnson, W. E., Speck, N. A., and Graves, B. J. (2009) DNA specificity determinants associate with distinct transcription factor functions. *PLoS Genet.* **5,** e1000778

35. Baillat, D., Laitem, C., Leprivier, G., Margerin, C., and Aumercier, M. (2009) Ets-1 binds cooperatively to the palindromic Ets-binding sites in the p53 promoter. *Biochem. Biophys. Res. Commun.* **378,** 213–217

36. Coyne, H. J., 3rd, De, S., Okon, M., Green, S. M., Bhachech, N., Graves, B. J., and McIntosh, L. P. (2012) Autoinhibition of Etv6 (TEL) DNA binding: appended helices sterically block the ETS domain. *J. Mol. Biol.* **421,** 67–84

37. 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

38. Bojović, B. B., and Hassell, J. A. (2001) The PEA3 Ets transcription factor comprises multiple domains that regulate transactivation and DNA binding. *J. Biol. Chem.* **276,** 4509–4521

39. Charlot, C., Dubois-Pot, H., Serchov, T., Tourrette, Y., and Wasylyk, B. (2010) A review of post-translational modifications and subcellular localization of Ets transcription factors: possible connection with cancer and involvement in the hypoxic response. *Methods Mol. Biol.* **647,** 3–30

40. Wu, J., and Janknecht, R. (2002) Regulation of the ETS transcription factor ER81 by the 90-kDa ribosomal S6 kinase 1 and protein kinase A. *J. Biol. Chem.* **277,** 42669–42679

41. Baert, J. L., Beaudoin, C., Coutte, L., and de Launoit, Y. (2002) ERM transactivation is up-regulated by the repression of DNA binding after the PKA phosphorylation of a consensus site at the edge of the ETS domain. *J. Biol. Chem.* **277,** 1002–1012

42. Chinenov, Y., Schmidt, T., Yang, X. Y., and Martin, M. E. (1998) Identification of redox-sensitive cysteines in GA-binding protein-α that regulate DNA binding and heterodimerization. *J. Biol. Chem.* **273,** 6203–6209

43. Hogart, A., Lichtenberg, J., Ajay, S. S., Anderson, S., NIH Intramural Sequencing Center, Margulies, E. H., and Bodine, D. M. (2012) Genome-wide DNA methylation profiles in hematopoietic stem and progenitor cells reveal overrepresentation of ETS transcription factor binding sites. *Genome Res.* **22,** 1407–1418

44. Nickel, J., Short, M. L., Schmitz, A., Eggert, M., and Renkawitz, R. (1995) Methylation of the mouse M-lysozyme downstream enhancer inhibits heterotetrameric GABP binding. *Nucleic Acids Res.* **23,** 4785–4792

45. Lucas, M. E., Crider, K. S., Powell, D. R., Kapoor-Vazirani, P., and Vertino, P. M. (2009) Methylation-sensitive regulation of TMS1/ASC by the Ets factor, GA-binding protein-α. *J. Biol. Chem.* **284,** 14698–14709

46. Polansky, J. K., Schreiber, L., Thelemann, C., Ludwig, L., Krüger, M., Baumgrass, R., Cording, S., Floess, S., Hamann, A., and Huehn, J. (2010) Methylation matters: binding of Ets-1 to the demethylated Foxp3 gene contributes to the stabilization of Foxp3 expression in regulatory T cells. *J. Mol. Med.* **88,** 1029–1040

47. Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33,** 151–152

48. Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A., and Gileadi, O. (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J. Struct. Biol.* **172,** 3–13

49. Barik, S. (1996) Site-directed mutagenesis in vitro by megaprimer PCR. *Methods Mol. Biol.* **57,** 203–215

50. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 293–302

51. Read, R. J., and Sussman, J. (2007) in *Evolving Methods for Macromolecular Crystallography: The Structural Path to the Understanding of the Mechanism of Action of CBRN Agents* (Read, R. J., and Sussman, J., eds) pp. 41–51, Springer, Dordrecht, The Netherlands

52. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132

53. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Sto-

roni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674

54. Sheldrick, G. M. (2008) A short history of SHELX. *Acta Crystallogr. A* **64,** 112–122

55. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.* **59,** 2023–2030

56. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501

57. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255

58. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221

59. Greenall, A., Willingham, N., Cheung, E., Boam, D. S., and Sharrocks, A. D. (2001) DNA binding by the ETS-domain transcription factor PEA3 is regulated by intramolecular and intermolecular protein-protein interactions. *J. Biol. Chem.* **276,** 16207–16215

60. Strauss, J. K., and Maher, L. J., 3rd (1994) DNA bending by asymmetric phosphate neutralization. *Science* **266,** 1829–1834

61. Rooman, M., Liévin, J., Buisine, E., and Wintjens, R. (2002) Cation-π/H-bond stair motifs at protein-DNA interfaces. *J. Mol. Biol.* **319,** 67–76

62. Wintjens, R., Liévin, J., Rooman, M., and Buisine, E. (2000) Contribution of cation-π interactions to the stability of protein-DNA complexes. *J. Mol. Biol.* **302,** 395–410

63. Batchelor, A. H., Piper, D. E., de la Brousse, F. C., McKnight, S. L., and Wolberger, C. (1998) The structure of GABPα/β: an ETS domain-ankyrin repeat heterodimer bound to DNA. *Science* **279,** 1037–1041

64. 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

65. 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

66. Pio, F., Kodandapani, R., Ni, C. Z., Shepard, W., Klemsz, M., McKercher, S. R., Maki, R. A., and Ely, K. R. (1996) New insights on DNA recognition by ets proteins from the crystal structure of the PU.1 ETS domain-DNA complex. *J. Biol. Chem.* **271,** 23329–23337

67. Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* **305,** 386–389

68. Pufall, M. A., Lee, G. M., Nelson, M. L., Kang, H. S., Velyvis, A., Kay, L. E., McIntosh, L. P., and Graves, B. J. (2005) Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science* **309,** 142–145

69. Carrère, S., Verger, A., Flourens, A., Stehelin, D., and Duterque-Coquillaud, M. (1998) Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. *Oncogene* **16,** 3261–3268

70. Babayeva, N. D., Baranovskaya, O. I., and Tahirov, T. H. (2012) Structural basis of Ets1 cooperative binding to widely separated sites on promoter DNA. *PLoS One* **7,** e33698

71. Evans, E. L., Saxton, J., Shelton, S. J., Begitt, A., Holliday, N. D., Hipskind, R. A., and Shaw, P. E. (2011) Dimer formation and conformational flexibility ensure cytoplasmic stability and nuclear accumulation of Elk-1. *Nucleic Acids Res.* **39,** 6390–6402

72. Fisher, R. J., Koizumi, S., Kondoh, A., Mariano, J. M., Mavrothalassitis, G., Bhat, N. K., and Papas, T. S. (1992) Human ETS1 oncoprotein. Purification, isoforms, −SH modification, and DNA sequence-specific binding. *J. Biol. Chem.* **267,** 17957–17965

73. Prinz, W. A., Aslund, F., Holmgren, A., and Beckwith, J. (1997) The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem.* **272,** 15661–15667

74. Hollenhorst, P. C., Shah, A. A., Hopkins, C., and Graves, B. J. (2007) Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev.* **21,** 1882–1894

75. Goel, A., and Janknecht, R. (2003) Acetylation-mediated transcriptional activation of the ETS protein ER81 by p300, P/CAF, and HER2/Neu. *Mol. Cell. Biol.* **23,** 6243–6254

76. Macauley, M. S., Errington, W. J., Schärpf, M., Mackereth, C. D., Blaszczak, A. G., Graves, B. J., and McIntosh, L. P. (2006) Beads-on-a-string, characterization of ETS-1 sumoylated within its flexible N-terminal sequence. *J. Biol. Chem.* **281,** 4164–4172

77. Legrand, A. J., Choul-Li, S., Spriet, C., Idziorek, T., Vicogne, D., Drobecq, H., Dantzer, F., Villeret, V., and Aumercier, M. (2013) The level of Ets-1 protein is regulated by poly(ADP-ribose) polymerase-1 (PARP-1) in cancer cells to prevent DNA damage. *PLoS One* **8,** e55883

78. Wang, H., McIntosh, L. P., and Graves, B. J. (2002) Inhibitory module of Ets-1 allosterically regulates DNA binding through a dipole-facilitated phosphate contact. *J. Biol. Chem.* **277,** 2225–2233

79. Laskowski, R. A., Gerick, F., and Thornton, J. M. (2009) The structural basis of allosteric regulation in proteins. *FEBS Lett.* **583,** 1692–1698

80. Popovych, N., Sun, S., Ebright, R. H., and Kalodimos, C. G. (2006) Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13,** 831–838

81. Regan, M. C., Horanyi, P. S., Pryor, E. E., Jr., Sarver, J. L., Cafiso, D. S., and Bushweller, J. H. (2013) Structural and dynamic studies of the transcription factor Erg reveal DNA binding is allosterically autoinhibited. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 13374–13379

82. Toone, W. M., Morgan, B. A., and Jones, N. (2001) Redox control of AP-1-like factors in yeast and beyond. *Oncogene* **20,** 2336–2346

83. Danciu, T. E., and Whitman, M. (2010) Oxidative stress drives disulfide bond formation between basic helix-loop-helix transcription factors. *J. Cell. Biochem.* **109,** 417–424

84. Nagahara, N. (2011) Intermolecular disulfide bond to modulate protein function as a redox-sensing switch. *Amino Acids* **41,** 59–72

85. Bannister, A. J., Cook, A., and Kouzarides, T. (1991) *In vitro* DNA binding activity of Fos/Jun and BZLF1 but not C/EBP is affected by redox changes. *Oncogene* **6,** 1243–1250

86. Tron, A. E., Bertoncini, C. W., Chan, R. L., and Gonzalez, D. H. (2002) Redox regulation of plant homeodomain transcription factors. *J. Biol. Chem.* **277,** 34800–34807

87. Wu, H., Xiao, Y., Zhang, S., Ji, S., Wei, L., Fan, F., Geng, J., Tian, J., Sun, X., Qin, F., Jin, C., Lin, J., Yin, Z. Y., Zhang, T., Luo, L., Li, Y., Song, S., Lin, S. C., Deng, X., Camargo, F., Avruch, J., Chen, L., and Zhou, D. (2013) The Ets transcription factor GABP is a component of the Hippo pathway essential for growth and antioxidant defense. *Cell Rep.* **3,** 1663–1677

88. Martin, M. E., Chinenov, Y., Yu, M., Schmidt, T. K., and Yang, X. Y. (1996) Redox regulation of GA-binding protein-α DNA binding activity. *J. Biol. Chem.* **271,** 25617–25623

89. Cooper, C. D., Newman, J. A., and Gileadi, O. (2014) Recent advances in the structural molecular biology of Ets transcription factors: interactions, interfaces and inhibition. *Biochem. Soc. Trans.* **42,** 130–138

90. Newman, J. A., Cooper, C. D., Aitkenhead, H., and Gileadi, O. (2015) Structural insights into the auto-regulation and cooperativity of the human transcription factor Ets-2. *J. Biol. Chem.* **290,** 8539–8549

91. Surh, Y. J., Kundu, J. K., Na, H. K., and Lee, J. S. (2005) Redox-sensitive transcription factors as prime targets for chemoprevention with anti-inflammatory and antioxidative phytochemicals. *J. Nutr.* **135,** 2993S–3001S