Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif

Julie A. Nye, Jeannine M. Petersen, Cathy V. Gunther, Matthew D. Jonsen, and Barbara J. Graves

Department of Cellular, Viral, and Molecular Biology, University of Utah School of Medicine, Salt Lake City, Utah 84143 USA

The proto-oncogene ets-1 is the founding member of a new family of eukaryotic transcriptional regulators. Using deletion mutants of murine ets-1 cDNA expressed in Escherichia coli, we show that the DNA-binding domain corresponds closely to the ETS domain, an 85-amino-acid region that is conserved among ets family members. To investigate the specificity of DNA binding of the ETS domain, we mapped the DNA contacts of a monomeric Ets-1 fragment by chemical protection and interference assays. DNA backbone interactions span a 20-nucleotide region and are localized on one face of the helix. Close phosphate and base contacts are restricted to 10 central nucleotides. Contacts map to the major groove in the center of the site. Flanking minor groove interactions also are predicted. To determine the sequence preference in the close contact zone, we selected a pool of high-affinity binding sites using a purified Ets-1 carboxy-terminal fragment. Our Ets-1-selected consensus, 5'-A/GCCGGAA/TGT/C-3', differs from the binding consensus for the Drosophila ETS domain protein E74A, suggesting that specificity of action of ets family members is mediated by the ETS domain. Compared to other well-characterized classes of DNA-binding proteins, Ets-1 produces a unique pattern of DNA contacts. These studies demonstrate that the ETS domain proteins bind DNA in a novel manner.

[Key Words: DNA-binding protein; ETS domain; ets-1 transcription factor]

Received March 5, 1992; revised version accepted April 13, 1992.

Transcription factors are organized into families based on conserved structural motifs that mediate DNA binding [Steitz 1990; Harrison 1991]. Members of a particular family interact with the phosphodiester backbone and the major and minor grooves of the DNA helix in a characteristic manner. Furthermore, in many cases, members of a family bind a set of similar DNA sequences [Johnson and McKnight 1989]. It is expected that different family members will activate or repress a specific set of promoter and enhancer elements. To discriminate between potential targets, the sequence-specific DNA-binding sites of family members must extend beyond the common recognition sequences. There is no code for predicting from the primary sequence which amino acids might mediate specificity of binding of different family members. Thus, every sequence-specific DNA-binding protein represents a unique problem in molecular recognition.

The ets gene family is one of the newest classes of eukaryotic DNA-binding proteins. The ets-1 gene, which contributes sequences to the transforming oncogene of the E26 avian retrovirus [Radke et al. 1982; Leprince et al. 1983], is the founding member of the family. In addition to the original group that was discovered by sequence similarity to v-ets [ets-1, ets-2, erg, and elk-1], the family includes E74 [Burtis et al. 1990], PU.1 [Klemsz et al. 1990], GABPa [LaMarco et al. 1991], Fli-1 [Ben-David et al. 1991], Elf-1 [Thompson et al. 1992], PEA3 [Xin et al. 1992], and SAP-1 [Dalton and Treisman 1992]. Some members of the family have been shown to function as transcriptional activators [Bosselut et al. 1990; Klemisz et al. 1990, Wasylyk et al. 1991; Xin et al. 1992]. The family is based on a conserved domain of ~85 amino acids. The discovery of the DNA-binding competence of Ets-1, E74A, and PU.1 led us with others to propose that the conserved amino acids, designated the ETS domain, function in DNA binding [Karim et al. 1990]. Conservation of the domain among family members ranges between 97 and 38% identity to Ets-1. Isolated positions of complete conservation are distributed over the entire 85 amino acids, and several patches of high conservation are striking. The features of interest include a highly basic region in the carboxy-terminal half of the domain. Eight
basic residues are conserved and four additional basic positions have substitutions of only glutamine and asparagine. The amino-terminal half of the domain contains a highly conserved, 10-amino-acid region that is rich in leucine. Aside from the preponderance of basic residues, the primary sequence of the ETS domain shows little similarity to any well-characterized DNA-binding motifs, suggesting that the family represents a new protein structure.

All ETS domain proteins whose DNA binding has been characterized contact a 5'-GGA-3' sequence motif [Karim et al. 1990; Hipskind et al. 1991; Thompson et al. 1991, 1992, Dalton and Treisman 1992; Xin et al. 1992]. Protein interactions with only 3 bp cannot mediate high-affinity DNA binding. Furthermore, if family members regulate different target genes by distinct binding properties, additional DNA contacts must be involved. This report describes our investigation of the sequence specificity of the ets family. Deletion mutagenesis was used to map the amino acids within Ets-1 that mediate high-affinity, sequence-specific interactions. Our results indicate that the DNA-binding domain of ets-1 maps closely to the ETS domain. Alkylation interference and hydroxyl radical protection assays delineate the zone of contact of the ets-1 ETS domain on the DNA helix. Both major and minor groove interactions are implicated. Finally, a selected-and-amplified-binding (SAAB) analysis specifies the preferred DNA sequences recognized by the ets-1 protein. These studies will help identify transcriptional control elements that are potential targets for the transcriptional activation function of Ets-1. Furthermore, we propose that the contacts of Ets-1 on the DNA helix illustrate the hallmarks of ETS domain interactions with DNA. We demonstrate that the pattern is unique among known binding motifs, establishing the novelty of the ETS domain family.

Results
Mapping the DNA-binding domain within the ets-1 protein

Full-length and amino-terminal deletion mutants of the murine ets-1 protein were expressed in bacteria with a T7 polymerase-dependent expression system [Studier et al. 1990] [Fig. 1A]. To produce full-length protein, we obtained an ets-1 cDNA clone containing a 1300-bp open reading frame (ORF). We placed the ORF into a pET3 vector with the first ATG positioned to be the start codon. The vector expressed a protein of an apparent molecular mass of 54 kD (Fig. 1B, lane 4). Amino-terminal sequencing confirmed the use of the expected start codon in bacteria. Antibodies specific for ets-1 also identified a 54-kD polypeptide in mouse tissue and T-lymphocyte cell lines [data not shown]. These data suggested that the first ATG of the ORF was used in both prokaryotic and eukaryotic cells. We will refer to this 54-kD protein as Ets-1. To produce amino-terminally truncated proteins, we cloned deleted versions of the ets-1 cDNA into pET3 vectors that expressed the ets-1 sequences in-frame with the sequences of T7 gene 10. All truncated proteins were stably expressed in bacteria [Fig. 1B].

The DNA-binding properties of partially purified ets-1 proteins were tested in gel mobility-shift assays. Ets-1 and three truncated proteins, ΔN170, ΔN322, and ΔN336, formed stable complexes with DNA [Fig. 1C]. The smallest deletion mutant, ΔN345, did not bind DNA [data not shown]. Competition experiments tested the specificity of binding. Labeled nucleoprotein complexes were competed by DNA bearing a strong Ets-1-binding site [WT] but were resistant to competition by DNA bearing a mutant site [MT] [Fig. 1C]. We conclude that Ets-1 and the truncated proteins, ΔN170, ΔN322, and ΔN336 form sequence-specific complexes. Furthermore, the DNA-binding domain of Ets-1 lies within the carboxy-terminal 105 amino acids and includes 80 of the 85 amino acids of the ETS domain [Fig. 1A].

In the DNA-binding assays presented in Figure 1, the amount of added extract was adjusted to present comparable binding activities from the various ets-1 polypeptides (>90% shift of probe DNA). This strategy required larger amounts of Ets-1 and ΔN170-containing extracts than of ΔN322- and ΔN336-containing extracts. This differential binding activity could not be explained by a significant difference in the purity of soluble ets-1 polypeptides [Fig. 1B, lanes 5,7,9,11]. Possibly, the extremely truncated ets-1 polypeptides bind DNA more avidly than the larger ets-1 polypeptides. Alternatively, the specific activity of the proteins could differ. We will describe attempts to distinguish these possibilities in the discussion. More importantly, for our continued analyses, these observations led us to assume that the truncated polypeptides, which displayed relatively high binding activity, were appropriate reagents for further binding studies.

Ets-1 binds DNA as a monomer

An important question concerns the oligomerization state of Ets-1 in solution and bound to DNA. We obtained highly purified preparations of Ets-1 and ΔN322 to investigate oligomerization. Gel filtration on Superose 12 with ΔN322 and with Ets-1 showed elution profiles expected for monomeric species [data not shown]. To investigate the oligomerization state of bound Ets-1, the truncated and full-length versions of Ets-1 were mixed and added to DNA. An oligomeric protein–DNA complex formed by Ets-1 and ΔN322 would have intermediate mobility in a gel-shift assay compared with homotypic complexes composed solely of Ets-1 and ΔN322. Complexes of intermediate mobility were not observed [Fig. 2]. We repeated this experiment under conditions in which the mixed proteins were denatured in 6 M urea and renatured before addition of DNA. Again, no apparent heterotypic complexes were observed [data not shown]. Although formally we have not excluded the possibility that an oligomerization domain is missing in ΔN322, we tentatively concluded that Ets-1 binds DNA in specific complexes as a monomer. Additional data in support of this proposal will be presented below.
DNA protection by Ets-1

Ets-1-binding sites are present in the long terminal repeat (LTR) of the Moloney murine sarcoma virus (MSV) [Gunther et al. 1990], the polyoma enhancer [Wasylyk et al. 1991], the enhancer of the T-cell receptor-α subunit gene [Ho et al. 1990], and the promoter of the stromelysin gene [Wasylyk et al. 1990] [Fig. 3]. The only invariant sequence is a conserved triplet, 5'−GGA−3'. Other ETS domain proteins also have a GGA motif in their binding sites [Fig. 3]. To explore the hypothesis that determinants of Ets-1 binding include additional sequences, we mapped the contacts made by Ets-1 on a variety of DNA-binding sites. To facilitate comparison of data from multiple sites, we numbered 20 positions within each Ets-1-binding site. The GGA triplet [designated by positions 10, 11, and 12 on the top strand in each binding site] served as the point of reference.

In a DNase I protection analysis, AN322 protected 16 nucleotides on the GGA strand and 18 nucleotides on the TCC strand [Fig. 4A,B]. In the center of the protected region, a DNase I cleavage site was observed near the
conserved triplet on the TCC strand (Fig. 4B). This analysis was performed on the SC2-binding site described in Figure 8, below. This same pattern of cleavage and protection is observed in AN170-mediated footprints on the MSV LTR promoter site (Gunther et al. 1990). We detected a similar pattern in AN322-mediated footprints on the enhancer of the MSV LTR and in Ets-1-mediated footprints on both the promoter and enhancer sites of the MSV LTR (data not shown). The consistency of the DNase I footprint suggests that Ets-1, AN170, and AN322 contact DNA in an identical manner.

We then used hydroxyl radical protection to study interactions of AN322 with the DNA backbone. Direct protein contact or protein-induced changes in the DNA conformation can affect the susceptibility of deoxyribose within the backbone to hydroxyl radical attack (Tullius et al. 1987). In the absence of protein, hydroxyl radicals cleaved the probe DNA uniformly (Fig. 4A), indicating that the binding-site DNA had no unusual structure. AN322 protected two patches of nucleotides on each strand of the binding site. A quantitative analysis of these data demonstrated that significant protection (>20%) on the GGA strand spanned 4- and 6-nucleotide patches, whereas the two patches on the TCC strand each included 6 nucleotides (Fig. 4C). Two slightly enhanced cleavage sites at positions 12 and 13 on the GGA strand were evident by the apparent negative protection in the histogram.

The display of the data on a helix model illustrates the juxtaposition of protected regions between positions 15 and 20 on the GGA strand across the minor groove from the backbone protection at positions 12-17 on the TCC strand (Fig. 4D). Likewise the region of protection between 6 and 9 was positioned across the minor groove from the protection between positions 3-6. The regions of protection flanked the conserved GGA triplet and mapped to one face of the DNA helix. A protein structure interacting with DNA over 20 nucleotides predominantly on one face of the helix could account for this pattern of hydroxyl radical footprinting. An altered DNA conformation that might accompany protein binding also could explain some of the protection data.

Interference of Ets-1 binding by phosphate ethylation

We performed ethylation interference experiments to assay more directly for protein contacts on the DNA phosphodiester backbone. Ethylation of DNA at phosphates by N-ethyl N-nitrosourea (ENU) (Siebenlist and Gilbert 1980) can interfere sterically with protein-DNA contacts or interfere by disrupting potential electrostatic interactions. In these experiments, probe DNA was ethylated and challenged with Ets-1 polypeptides. The ethylation of two phosphates on the GGA strand and four phosphates on the TCC strand interfered significantly [greater than fivefold effect] with Ets-1 binding (Fig. 5A, B). The positions of the most severe interference flanked the GGA triplet. The helix model (Fig. 5C) illustrates that the sites of the strongest interference lie near the major groove in the center of the binding site. Figure 5B compares ethylation interference data from Ets-1 and AN322 experiments. The positions and ranking of interference were almost identical in the two analyses. These results support our assumption that the carboxy-terminal fragment of Ets-1 and full-length Ets-1 make similar DNA contacts.

Interference of Ets-1 binding by purine methylation

We used methylation interference to probe Ets-1 interactions within the helix grooves. Dimethyl sulfate methylates guanine residues at the N7 position in the major groove and adenine residues predominantly at the N3 position in the minor groove. We studied two different Ets-1-binding sites because this analysis probed only G and A residues. We classified sites of interference as either strongly interfering [5- to 10-fold effects] or mildly interfering [1.5- to 3-fold effects]. In each binding site, methylation of the two G residues within the GGA triplet strongly interfered with binding (Fig. 6A, B). The only other interference in the severe category was at position 8 on the GGA strand of the promoter-binding site (Fig. 6B). The strong guanine methylation interference suggests close DNA contacts within the major groove between positions 8 and 11. Sites of less severe interference mapped to every G residue within positions 8 and 16 on
Selected and amplified binding sites: an Ets-1 consensus site

The protection and interference data demonstrated that backbone interactions spanned a 20-nucleotide region, whereas close phosphate and base contacts mapped to only 10 contiguous nucleotide pairs. We predicted that the most important sequence determinants for Ets-1 DNA binding would lie within the more restricted close contact region. To elucidate the sequence preferences at these positions, we employed a binding-site selection technique (Blackwell and Weintraub 1990). In brief, we generated a pool of DNA molecules (52 bp in length) with a central GGA flanked by seven randomly specified positions, 5' ... NNNGGANNNN ... 3'. High-affinity binding sites for AN322 were selected from the pool by gel mobility-shift assays. Selected DNA molecules were amplified DNA was radiolabeled and used in the next selection assay. The selection and amplification procedure was repeated five times. An additional round of selection was performed under more stringent binding conditions, yet no decrease in the complexity of the consensus was observed.

The random composition of the positions flanking the invariant GGA (bp 7–9; bp 13–15) is shown in the sequence of the starting pool (start, Fig. 71. A set of non-random sequences, a "selected consensus", was obtained after five rounds of selection (Figs. 7 and 8). Sequencing of both strands showed similar preferred sequences. Parallel selection on a separate, identical pool of DNA molecules also derived the same set of preferred sequences (data not shown). The selection experiments detected preferences at only six of the seven positions assayed. In addition, the study indicated that five of these positions exhibited dual preferences. To determine whether the dual preferences represented as many as 32 different sites or only a subset of this maximal number, we cloned and sequenced 15 randomly selected molecules of the round 5 pool. The sequences of the selected and cloned sites [SC sites] are presented in Figure 8. Sequences in lower-case do not match the most frequently selected nucleotide in the consensus. Boxed positions in each site fail to match the consensus. The frequency of nucleotides in only these 15 clones (Fig. 8) was sufficient to suggest a consensus that is similar to the selected consensus. In addition, these individual clones confirmed that the selected pool of sites contains considerable diversity. Only two pairs of sites showed the identical sequence at all six selected positions.

To understand the significance of the degenerate consensus, we determined the relative affinity of the SC sites. By quantitative gel mobility-shift assays, the equilibrium dissociation constant of AN322 for the SC2 site was calculated to be $4.9 \times 10^{-11}$ (for details of this determination, see Materials and methods). The apparent affinities of the other 14 clones were determined by competition titrations (Lui-Johnson et al. 1986; Li et al.}

---

**Table:**

| Protein             | Consensus Site          | References                                 |
|---------------------|-------------------------|--------------------------------------------|
| Ets-1               | ACAGGATATC (this report)|                                            |
| MSV LTR enhancer    | ACAGGATATCG             | Gunther et al. (1990)                      |
| MSV LTR promoter    | AGAGGATATCG             |                                            |
| Polyoma enhancer    | AGAGGACCTT              | Klemes et al. (1990)                       |
| T-Cell Receptor α   | AGAGGATATGG             |                                            |
| Stromelysin         | GCAGGAAGCA              |                                            |
|                     | CCAGGAAATG              | Wasylyk et al. (1991)                      |
| PU.1                | CCCGGGAAGTA             | Urness and Thummel (1990)                  |
| MHCII               | AGAGGATATG              |                                            |
| GABPα               | AGCGGAGGCG              |                                            |
| HSV ICP4            | AGCGGAAACC              |                                            |
| SAP-1               | AGAGGATGTC              | Dalton and Treisman (1992)                 |
| Elk-1               | AGAGGATGTC              | Hipskind et al. (1991)                     |
| c-Fos               | AGAGGATGTC              |                                            |

**Figure 3.** DNA-binding sites for ETS domain proteins. Sites have been aligned by the invariant GGA triplet [underline]. In all cases, mutational analyses or methylation interference have shown that the GGA nucleotides are involved in DNA binding.
The clones (SC1–SC15) are ranked according to their relative affinities and organized into four groups in Figure 8. The affinities of the binding sites varied over two orders of magnitude. These quantitative studies demonstrate that the relative match to the selected consensus correlates with the relative affinity of a particular site. Thus, the selected consensus will be useful in predicting the strength of other Ets-1-binding sites. This collection of sites also should predict the location of possible Ets-1 binding sites in viral and cellular enhancers. Interestingly, previously reported Ets-1 binding sites (Fig. 3) do not match the SC clones with the highest affinities, suggesting that biological regulatory mechanisms utilize sites with less than optimal binding affinities.

The selected consensus and the strongest binding sites included a minimal twofold symmetry, 5'-CCGG-3'. Sites SC9 and SC12 contained the larger palindromic sequence 5'-TCCGGA-3', but these clones did not show the highest affinities. Binding sites for proteins that bind as homodimers often have relative affinities that are proportional to the degree of symmetry of the binding-site sequence. Thus, the selection experiment supports the proposal that a single Ets-1 polypeptide recognizes a single GGA motif.

Two groups have recently reported a selected consensus for Ets-1: A/GG/CC/aGGAAGT/c (Fisher et al. 1991); ACC/AGGAA/TA/G (Woods et al. 1992). These complementary studies strengthen our conclusions. Although no quantitative analysis of individual clones was presented in these studies, the selected binding sites display sequences that would position them high in our affinity hierarchy (Fig. 8). Both studies used full-length Ets-1 (either expressed in SF9 cells or prepared from T-cell nuclear extracts), thus supporting our assumption that the DNA binding of AN322 expressed in bacteria accurately represents Ets-1 activity. In each study, the number of randomly specified positions in the starting
Figure 5. Ethylation interference analysis of Ets-1 binding. (A) Autoradiograph of analysis performed with the 54-kD 
ets-1 protein on the SC2 site. DNA fragments were labeled and ethylated, to a limited degree, with ENU and incubated with protein. Bound and unbound DNA fractions separated on a nondenaturing gel were cleaved with alkali. Positions of cleavage were mapped by electrophoresis on a denaturing gel (shown). The alkali cleavage of phosphotriester bonds generates fragments terminating in either a hydroxyl group or an ethylated phosphate. These fragments display slightly different mobility [visible as doublets in the lower portion of the gel]; the fragment with the ethylated phosphate migrates more slowly than the other cleavage product. Lanes G contain the products of chemical sequencing reactions for G residues that serve as markers. The markers migrate slightly faster than both of the alkali cleavage products. C designates lanes loaded with DNA that was modified and cleaved in the absence of protein. Lanes U show the cleavage products of unbound DNA; lanes B represent the assay of bound DNA. (B) Quantitative analysis of the degree of 
sensitivity to ethylation, Ets-1 (open bars) and AN322 (solid bars). The relative autoradiographic density of lanes loaded with unbound and bound DNA is plotted as a function of sequence position for both strands (note difference in scale). Bars are positioned to designate 5'-ethylated phosphate of a numbered base pair. Phosphates at which ethylation had no effect on binding have a unity value. (C) B-form 
DNA helix, with triangles marking the sites of ethylation interference that show greater than a fivefold effect.

pool was greater than the seven positions assayed in our 
study. Nevertheless, preferred sequences were restricted 
to the same positions observed in our study (positions 
7–15). Thus, binding-site analyses indicate that Ets-1 
recognizes specific DNA sequences over only a 9-bp re-

Discussion

Protein determinants for Ets-1 DNA binding

The deletion mutagenesis demonstrated that the DNA-
binding domain of Ets-1 includes the conserved 85-
amino-acid ETS domain. The smallest mutant that 
bound DNA, ΔN336, retained 80 amino acids of the ETS 
domain plus 25 carboxyl-terminal residues. Parallel stud-
ies performed by Lim and colleagues (1992) on chicken 
Ets-1 are consistent with these conclusions. An Ets-1 
polypeptide containing only the 85 amino acids of the 
ETS domain also binds DNA with high affinity (J.M. Pe-
tersen, unpubl.). Domain mapping studies of other ets 
family members have led to similar findings. A polypep-
tide containing the ETS domain of GABPα plus 3 amino 
acids on the amino terminus and 7 amino acids on the 
carboxyl terminus mediates DNA binding (Thompson et al. 1991). Likewise, truncated PEA3 that retains the ETS 
domain plus 18 amino acids on the carboxyl terminus 
binds DNA (Xin et al. 1992).

Two independent approaches indicated that the ETS 
domain of a single ets-1 polypeptide recognizes a single 
GGA motif. In mixing experiments, no heterotypic com-
plexes of truncated and full-length ets-1 polypeptides 
bind DNA. In the consensus selection, the relative af-
inities of binding sites did not correlate with the degree 
of symmetry of the DNA sequence. Association of one 
ETS domain with a single GGA motif is also proposed for 
DNA binding by GABP (Thompson et al. 1991). In this 
case, high-affinity DNA binding is mediated by two 
identical α-subunits that each contain an ETS domain.
Each α-subunit is modeled to contact a single GGA motif.

These protein studies suggested that truncated ets-1 polypeptides were suitable reagents for mapping the DNA contacts of Ets-1. Consistent with this proposal, Ets-1 and ΔN322 showed identical DNA contacts in both DNase I protection and ethylation interference experiments. However, we observed that the protein preparations of the truncated ets-1 polypeptides display greater DNA-binding activity than protein preparations containing equivalent amounts of full-length Ets-1. For example, a 100-fold molar excess of Ets-1 over ΔN322 was necessary to shift approximately equal amounts of probe in binding assays of highly purified protein preparations (Fig. 2). Ets-1 could have a weaker affinity for DNA than the ΔN322 species. Alternatively, a significant fraction of Ets-1 could be inactive in DNA binding.

Our preliminary investigation of this phenomenon has...
led us to consider another explanation. The DNA-binding activity of preparations of highly purified Ets-1 is increased significantly by partial proteolytic digestion. Amino-terminal sequencing of proteolytic fragments indicated that the higher DNA-binding activity is mediated by 12- to 14-kD polypeptides that contain the ETS domain [M. Jonsen and B. Graves, unpubl.]. The DNA binding of chicken Ets-1 can also be activated by either deletion mutagenesis or partial proteolysis [Lim et al. 1992]. In these studies, chicken Ets-1 was synthesized in reticulocyte lysates; thus, the relatively low binding activity of Ets-1 is not restricted to protein produced in bacteria. These additional studies indicate that Ets-1 contains a well-folded DNA-binding domain and suggest that the remainder of the protein is also properly folded. To explain the difference in activity of the preparations of full-length and truncated proteins, we speculate, as proposed by Lim and colleagues [1992], that the amino-terminal portion of Ets-1 negatively regulates the DNA-binding activity of the carboxy-terminal ETS domain. Interactions with other proteins might alleviate this inhibition in vivo, increasing the DNA-binding activity of full-length Ets-1 to the level observed for ΔN322. In this model, consistent with our findings, active Ets-1 would make the same DNA contacts as truncated ets-1 polypeptides. The DNA binding of four other ETS domain proteins has been shown to be enhanced by the association with other unrelated proteins. The most active binding species of GABP contains two β-subunits in addition to the ETS domain bearing α-subunits. PU.1 recruits an additional factor into a binding complex on the immunoglobulin κ enhancer [Pongubala et al. 1992]. Finally, both Elk-1 and SAP-1 have been shown to bind DNA in association with the serum response factor [SRF] [Hippskind et al. 1991; Dalton and Treisman 1992]. We are searching for functionally analogous components that could modulate the putative inhibitory domain of Ets-1 in eukaryotic cells.

DNA determinants for Ets-1 binding

Protection and interference data are summarized schematically on a planar representation of a DNA helix in Figure 9. Data from all analyses have been aligned by the conserved GGA triplet. The hydroxyl radical protection analysis (data shown in heavy lines) showed that 20 nucleotides were affected by Ets-1 binding. Four clusters of sugar residues were protected from attack. The ethylation interference experiments (sites of interference marked by triangles) confirmed backbone contacts within a central 8-bp region. All possible backbone interactions mapped to one face of the helix. [We designated this surface of the helix the front face.] Methyl-ation interference experiments (sites of interference marked by circles and squares) identified putative contacts within 10 contiguous bp (positions 7–16). The sites of strongest interference caused by base methylation also mapped to the front face of the helix. As discussed below, the configuration of chemical contacts, as well as the selected sequence preferences, suggest that Ets-1 contacts nucleotides in the major groove. Furthermore, on each side of the major groove zone, there are interactions near the minor groove.

Major groove contact zone

Three types of data predicted that positions 8–12 are contacted in the major groove. [1] The strong sites of ethylation interference and hydroxyl radical protection flanked the major groove of the DNA helix that centers over position 10; [2] putative base contacts in the major groove were suggested by the strong guanine methylation interference at positions 8, 10, and 11; and [3] the SAAB experiment demonstrated sequence determinants at positions 8 and 9 by selecting an invariant C for position 8 and showing a strong preference for a C at position 9. Furthermore, six SC sites showing the strongest Ets-1 binding had a CpG dinucleotide at positions 8 and 9. A C preference at positions 8 and 9 is also indicated by mutagenesis of the Ets-1 binding site in the stromelysin promoter [Wasylyk et al. 1991]. The absolute conservation of the GGA at positions 10, 11, and 12, as well as the sensitivity of these
Figure 9. Summary of protection and interference data. The DNA helix is diagrammed in a planar projection (Siebenlist and Gilbert 1980) [B-form helix with radius: 8.9 Å, helical pitch: 10.5 bp/turn]. The DNA phosphodiester backbone is represented by diagonal lines rising 3.26 Å/bp. The base pairs are represented as horizontal lines across the minor groove (base-pairing in the major groove is not pictured). The front and back designations were chosen with respect to the Ets-1 mode of binding. The coordinate system, which assigned the GGA triplet to helix. This protection could accompany protein interactions with base pairs within the minor groove or protein contacts that are positionally only on the backbone. To attempt to distinguish these alternatives, we must consider the nature of the sequence determinants in this region. In selection experiments, dual preferences were selected between positions 13–15, whereas no preference was detected at position 16. Position 7 was also identified as a degenerate sequence determinant. The importance of sequence composition in these regions is also suggested by the mutational analysis of the stromelysin promoter [Wasylyk et al. 1991], although the consensus degeneracy in these regions was not anticipated by the mutational analysis. On one hand, DNA-binding specificity could be mediated by base pair contacts in the minor groove in which there is minimal distinction between base pairs [Seeman et al. 1976]. For example, either a C : G or a T : A base pair at position 15 could participate in a hydrogen bond with Ets-1 via the oxygen of the pyrimidine ring. Methylation interference analysis also suggested base contacts in the minor groove at position 7 and positions 13–15. Alternatively, the nonstringent sequence requirements may reflect a selection for a sequence that favors a DNA conformational change. In this context, we note the conservation of a purine–pyrimidine step at positions 14–15 on the GGA strand in the high-affinity SC sites, a feature noted for its effect on DNA conformation [Calladine 1982]. In conclusion, base contacts with limited sequence specificity and sequence-dependent DNA conformation changes may both contribute to the interactions in the minor groove contact zone.

Several observations indicated that DNA conformational change is a feature of Ets-1-binding specificity. The bound DNA, for example, was sensitive to cleavage by hydroxyl radical and DNase I on the back face of the helix. Accessibility or reactivity to cutting reagents may be enhanced by an altered conformation of the minor groove between positions 8 and 13. An interesting feature of the chemical mapping data was the protection from hydroxyl radical that extended 5 bp in each direction beyond the region containing close base and phosphate contacts. This hydroxyl radical protection could be the result of protein binding in the central major groove zone and indirectly narrowing the flanking minor grooves. We have not detected a dramatic protein-induced bend in the Ets-1-binding site by gel mobility-shift.
assays with permuted binding site probes [J. Petersen and C. Hull, unpubl.]. However, slight differences in the mobility of protein–DNA complexes in these assays are consistent with a subtle conformational change.

Unique mode of binding for ETS domain family

Chemical mapping studies have been performed on proteins that represent many of the structural motifs for sequence-specific DNA binding. In several cases, crystallographic studies have confirmed the features of DNA–protein interactions defined by chemical approaches [λ repressor [Jordan and Pabo 1988]; EcoRI [McClarin et al. 1986]; MATo2 [Wolberger et al. 1991]]. Members of a family of DNA-binding proteins that share a structural motif generally display a similar pattern of contacts. For a new family of DNA-binding proteins, such as the ETS domain group, these types of data can establish the uniqueness of the family and provide a means for identifying new members. The distinctive features of Ets-1 binding to DNA include interactions with four segments of the phosphodiester backbone, a central major groove contact zone of 5 bp, and two putative minor groove contact zones. Although detailed mapping studies have not been performed on other ETS domain proteins, the DNase I hypersensitive site [between positions 8 and 9 on the TCC strand] that we noted within a number of Ets-1 binding sites also is observed in footprint patterns of other members of the ets family [Urness and Thummel 1990]. We speculate that all ETS domain proteins bind DNA in a similar manner.

To demonstrate the uniqueness of the ETS domain mode of binding, we briefiy compare the contact pattern of Ets-1 to the patterns generated by four other structural motifs that bind DNA in the major groove. Contacts made by homeo domain proteins have been mapped by ethylation interference and hydroxyl radical protection (Sauer et al. 1988; Afolter et al. 1990). Contacts are made on three segments of the phosphodiester backbone, outlining the major groove and the contiguous minor groove on one face of the helix. The X-ray crystal structure of the engrailed and MATo2 homeo domain bound to DNA shows the major groove filled by the helix–turn–helix motif and interactions of a flexible arm within the minor groove [Kissinger et al. 1990; Wolberger et al. 1991]. The ethylation interference data for the helix–turn–helix motif of λ repressor illustrate the contact pattern of dimeric helix–turn–helix proteins. The phosphate contacts map to four segments of the phosphodiester backbone on one face of the helix [Johnson 1980], highlighting a central minor groove zone and two regions of major groove contact. Crystallographic studies demonstrate the presence of the two helix–turn–helix motifs within the major groove [Jordan and Pabo 1988]. A different pattern of protection is caused by basic leucine zipper [bZIP] proteins. The major groove contact zone spans 12 contiguous nucleotides as evidenced by ethylation interference over this region on both strands [Nye and Graves 1990]. There is only minor protection of the DNA helix from hydroxyl radical cleavage, suggesting no minor groove interactions [Vinson et al. 1989]. The C6 zinc finger motif found in GAL4 and other related DNA-binding proteins has also been studied by chemical mapping techniques (Carey et al. 1989; Halvorsen et al. 1991). These proteins display a unique pattern of hydroxyl radical protection and ethylation interference along an extensive region in the center of the binding site. The pattern outlines a central minor groove contact region. Putative base contacts map to the major groove on both sides of this central zone.

These comparisons of ethylation interference and hydroxyl radical protection data demonstrate that the ETS domain of Ets-1 contacts DNA in a novel manner. The uniqueness of the contact pattern observed with Ets-1 is consistent with the distinctive character of the primary sequence of the ETS domain. None of the landmark features of the structural motifs described above have been noted in the ETS domain sequence [Karim et al. 1990]. It is now clear from the binding site analysis that the ETS domain identifies a new family of DNA-binding proteins.

The ETS domain mode of DNA binding accommodates associated proteins

As described above, several ETS domain proteins bind DNA in association with other proteins. The ETS domain protein and the associated protein could bind the major groove on the same face of the DNA helix. We have detected this arrangement of binding sites in the enhancer of the murine leukemia virus [MLV] LTR. The major groove contacts of Ets-1 and the core-binding-protein CBF [Wang and Speck 1992], are positioned 10 bp apart on the DNA helix, as determined by ethylation and methylation interference experiments [N.A. Speck and B.J. Graves, unpubl.].

A different configuration of neighboring binding sites appears to be represented by the binding of GABP. The two α-subunits of GABP each contain an ETS domain. Thompson et al. [1991] proposed that GABP recognizes two GGA motifs that are separated by one-half turn of the DNA helix, ... GGANNNGGA ... . This spacing places the major groove interactions of the two ETS domains on opposite faces of the DNA helix (front and back face in a planar helix, Fig. 9). This orientation could accommodate a 9-bp contact zone for each ETS domain, as only 5 bp are proposed to be contacted in the major groove. The 3 bp that lie between the two GGA motifs would be recognized in the minor groove by the first ETS domain and in the major groove by the second ETS domain. Phosphodiester backbone interactions spanning 20 bp were observed in the binding of Ets-1 to the DNA helix. Even this large stretch of putative contacts could be accommodated by positioning GABP on two faces of the helix.

Sequence specificity of DNA binding of the ETS domain proteins

All ets family members bind sites containing a central
GGA triplet. The flanking sequences may mediate the specificity of binding of different ETS domain proteins. For example, a set of sequences could be uniquely preferred by Ets-1 and not tolerated in sites bound by other family members. Sites recognized by other family members are listed in Figure 3. The selected consensus for Ets-1 shows a preference for RCC to the left of the specificity of binding of different ETS domain proteins.

The selected consensus for Ets-1 shows a preference for RCC to the left of the specificity of binding of different ETS domain proteins. Ets-1 data from Figs. 7 and 8. E74A data from Figs. 4 and 6. Expression and partial purification of Ets-1 polypeptides

For expression, pET plasmids were transformed into E. coli strain BL21 (DE3) containing the pLysS plasmid (Studier et al. 1990). Expression of foreign proteins was induced with 1 mM IPTG for 1–2 hr at 37°C (OD600 = 0.7–0.9). Expression proteins were partially purified by the method of Landschulz et al. [1989]. In brief, bacterial pellets from 50-ml cultures were resuspended in 2 ml of lysis buffer [5 mM urea, 10 mM Tris–Cl (pH 7.9), 0.15 mM NaCl, and 0.1% Triton X-100]. Cells were sonicated discontinuously for 1 min at 30 W. Samples were centrifuged at 10,000g for 5 min at 4°C. Soluble proteins were applied to a DEAE-cellulose column equilibrated in urea–lysis buffer without the Triton X-100. Unbound proteins were dialyzed against TGMK100 [25 mM Tris–Cl (pH 7.9), 10% glycerol, 5 mM MgCl2, 1 mM EDTA, 100 mM KCl, 0.1 mM PMSF, 1 mM dithiothreitol (DTT)]. After dialysis, samples were centrifuged at 10,000g for 30 min, and the supernatant was frozen in liquid nitrogen and stored at −80°C.

For large-scale purification of ΔN322, a bacterial pellet from a 125-ml culture was resuspended in 10 ml of lysis buffer [50 mM Tris–Cl (pH 8.0), 0.2 mM NaCl, 0.1 mM PMSF, 1 mM DTT, 1 μg/ml of leupeptin]. Cells were sonicated discontinuously for 3 min on ice. The sample was then centrifuged for 20 min at 27,000g. Soluble proteins were precipitated by adding 0.5 grams of [NH4]2SO4/ml of extract and gentle stirring for 1 hr. The precipitate was collected by centrifugation for 15 min at 27,000g. The pellet was resuspended in chromatography buffer, TGEK,00 [10 mM Tris–Cl (pH 7.9), 10% glycerol, 0.1 mM EDTA, and 100 mM KCl] and dialyzed against the same buffer. The sample was cleared by centrifugation at 48,000g for 30 min and applied to a 5-ml DEAE–cellulose column (equilibrated with TGEK,00). Unbound protein was then applied to a 1-ml Mono-S column (Pharmacia). Bound protein was eluted in a 20-ml linear gradient from 0 to 350 mM KCl. Fractions containing DNA-binding activity eluted between 200 and 250 mM KCl. Active fractions were frozen in liquid nitrogen and stored at −80°C. In peak fractions, ΔN322 was 80–90% pure as estimated by Coomassie blue staining of an SDS-PAGE gel. The extinction coefficient of the ΔN322 was estimated by the method of Gill and von Hippel [1989] to be 26,152 M−1 cm−1 at 278 nm. We used this value to interpret optical density measurements and estimate the

**Materials and methods**

**Construction of bacterial expression plasmids**

The ets-1 cDNA was constructed from two mouse thymus cDNA clones (Gunther et al. 1990). The cDNA sequences of one clone, including 509 bp of the untranslated leader (5' UTR) and 690 bp of the amino-terminal coding sequences, were ligated at a native HpaI site to 631 bp of carboxy-terminal coding sequences and 13 bp of 3' UTR of a second clone. Initially, the entire 1849 bp were cloned into the BamHI site of pET3 vector (which contains 11 codons of T7 gene 10). To mediate expression of only ets-1 codons, the gene 10 codons, the 5' UTR and Ets-1 codons 1–3 were removed by restriction with Ndel and EagI and replaced with an Ndel–EagI adapter (5'-TATCGAAGCC-3', 5'-GGCGCGCCTTCA-3'). This synthetic DNA contained a methionine codon in the context of an Ndel site and the first three ets-1 codons.

The pET3 vector system (Studier et al. 1990) also was used to express the mutants ets-1 polypeptides in *Escherichia coli*. In this case, the vectors retained 11 codons of the S10 gene (to provide a start codon), as well as the ets-1 stop codon and 59 bp of 3' UTR. The polypeptide ΔN170 was expressed from pET-Ets-1 (Gunther et al. 1990). For production of ΔN322, ΔN326, and ΔN345, the cDNA expressing ΔN170 was deleted by exonuclease III–S1 nuclease treatment. A synthetic BamHI linker was ligated to the deleted cDNAs for cloning into pET3. Deletion end points were determined by didexoxy sequencing. Designating the start codon as the origin, the exact deletion end points for expression of ΔN170, ΔN322, ΔN336, ΔN345 were nucleotides 508, 963, 1006, and 1033, respectively.

**Figure 10.** Comparison of the selected consensuses for two ETS domain proteins. Ets-1 data from Figs. 7 and 8. E74A data from a selection procedure that began with 25 bp of random sequence (Urness and Thummel 1990). Two cycles of selection were performed by E74A affinity chromatography. DNA molecules that bound tightly were selected for cloning and sequencing. The consensus was derived from a statistical analysis of the sequences of 34 clones. The positions of differences between the two consensuses. Vertical lines and minor groove and major groove labels designate proposed mode of recognition of DNA by ETS domain proteins.
\[\Delta N322\] concentration in stock solutions. Estimated yield from the 125-ml culture was 0.85 mg.

For further purification of Ets-1, a bacterial pellet from 100-ml culture was subjected to the urea-lysis protocol described above except for the following modifications. The scale of the lysis step and DEAE-chromatography was increased twofold and the dialysis was against TGEK\textsubscript{eqq}. Cleared sample was applied to a 1.0-ml Mono-S column [Pharmacia] equilibrated in TGEK\textsubscript{eqq}. Bound protein was eluted with a 20-ml linear gradient from 0.1 to 1 M KCl. Ets-1 eluted in fractions containing 150–200 mM KCl. Ets-1 was estimated to be 95% pure based on Coomassie blue staining of a SDS-PAGE gel. Protein concentration was estimated as described above for \[\Delta N322\], based on an extinction coefficient of 87,470 M cm\(^{-1}\).

DNA-binding site probes
Probes were labeled on 5'-extended ends with \([\gamma-\text{32P}]\text{ATP}\) and T4 polynucleotide kinase [New England Biolabs] after treatment with calf intestinal phosphatase [Sigma]. The 3'-recessed ends were labeled by fill in with an appropriate [\(\alpha-\text{32P}\)]dNTP and E. coli DNA polymerase, Klenow fragment [New England Biolabs]. Probes were resolved on 4–6% acrylamide/0.25% agarose gels and electroeluted onto NA-45 paper (Schleicher & Schuell) in Western blot buffer (12.5 mM Tris-Cl, 96 mM glycine). The paper was then incubated in 1 M NaCl/0.05 M arginine at 65°C for 2 hr. The DNA solution was diluted 1 : 1 with TE [10 mM Tris-Cl (pH 8.0), 1 mM EDTA], and DNA was ethanol precipitated.

The MSV LTR enhancer site probe was prepared from a plasmid bearing the LTR sequences between viral genome positions 201 and 221 [Van Beveren et al. 1981]. This plasmid was constructed by synthesizing the complementary oligonucleotides 5'-GATCCCAAACAGGATATCTGTGGTAAGCA-3' and 5'-GATCCCCAACAGATATCTGGTAAGCA-3'. The oligonucleotides were annealed and cloned into the BamHI site of pKS (Stratagene). Probe DNA for the enhancer site was a 95-bp fragment. The MSV LTR promoter site probe was prepared from peMSV(-74/-31) bearing the MSV LTR sequences described previously (Tullius et al. 1987), with the following modifications. The scale of the lysis step and DEAE-chromatography was increased twofold and the dialysis was against TGEK\textsubscript{eqq}. Cleared sample was applied to a 1.0-ml Mono-S column [Pharmacia] equilibrated in TGEK\textsubscript{eqq}. Bound protein was eluted with a 20-ml linear gradient from 0.1 to 1 M KCl. Ets-1 eluted in fractions containing 150–200 mM KCl. Ets-1 was estimated to be 95% pure based on Coomassie blue staining of a SDS-PAGE gel. Protein concentration was estimated as described above for \[\Delta N322\], based on an extinction coefficient of 87,470 M cm\(^{-1}\).

**DNA-binding reactions and gel mobility-shift assays**
DNA-binding reactions for gel mobility-shift assays contained radiolabeled probes (6–10 fmols, unless otherwise noted) and protein preparations in a final volume of 20 \(\mu\)l of binding buffer [65 mM KCl, 25 mM Tris-Cl (pH 7.9), 6 mM MgCl\(_2\), 0.25–0.5 mM EDTA, 10% glycerol]. Bovine serum albumin was included ([0.15–0.2 mg/ml] in all reactions. Poly[dil-Cl] was included as a nonspecific competitor as needed. After incubation on ice for 20 min, DNA–protein complexes were resolved on 5% polyacrylamide gels [acrylamide/bis, 30 : 0.8, 45 mM Tris-borate (pH 8.3), 1 mM EDTA] at 4°C. Gels were prerun for 1 hr at constant voltage of 17 V/cm. The samples were loaded with voltage set at 7.7 V/cm and resolved by electrophoresis for 2–3 hr at 17 V/cm.

**DNase I and hydroxyl radical protection assays**
The hydroxyl radical protection assays were performed as described previously [Tullius et al. 1987], with the following modifications. Binding reactions with a final volume of 35 \(\mu\)l were prepared as described above except that the glycerol level was reduced to <0.5%. The cutting reaction was initiated by the addition of 15 \(\mu\)l of premixed reagents. The final concentrations of \(H_2O_2\), sodium ascorbate, iron (II), and EDTA in the binding reaction were 0.01%, 3 mM, 0.3 mM, and 0.6 mM, respectively. Cutting was stopped after 5 min at 20°C by the addition of 21 \(\mu\)l of stop solution [0.1 M thiourea, 0.2 M EDTA]. DNase I protection assays were performed as described previously [Graves et al. 1986]. Cleavage products from both hydroxyl radical and DNase I protection assays were resolved on 9% sequencing gels. Autoradiographs were obtained by exposure to X-ray film with the aid of an intensifying screen. Autoradiographs were scanned with a Bio-Rad densitometer, and film densities were plotted with 1-D Analyst Macintosh (Bio-Rad) software. Quantitative analyses were performed using peak heights of absorbance as an estimate of band density.

**Alklylation interference assays**
Alklylation interference assays were performed as described previously [Nycz and Graves 1990]. End-labeled DNA (0.3–0.5 pmole of radiolabeled probe and 2.5 \(\mu\)g of poly[dil-Cl]) was methylated with dimethyl sulfate [Aldrich] [Maxam and Gilbert 1980] or ethylated with ENU [Sigma] [Siebenlist and Gilbert 1980]. To ethylate probe, DNA was incubated in 40 \(\mu\)l of 0.3 M sodium cacodylate [pH 7.3] and 40 \(\mu\)l of ethanol saturated with ENU (–0.3 grams of ENU/ml of 95% ethanol) for 20 min at 50°C. The reaction was stopped by ethanol precipitation with ammonium acetate and incubation at –70°C. Binding reactions performed in buffer conditions described above contained 35–105 fmols of modified DNA and sufficient protein to achieve a 10–30% shift of the probe. Bound and unbound DNA fractions were recovered by electroelution as described above, with the addition of 5 \(\mu\)g of tRNA as carrier in the ethanol precipitation. Phosphotriester bonds were cleaved by resuspension of DNA in 10 mM sodium phosphate [pH 7.0], 1 mM EDTA, followed by the addition of an equal volume of 0.3 M NaOH and a 30-min incubation at 90°C. DNA was cleaved at methylated guanine residues by incubation in 1 mM piperidine for 30 min at 90°C [Maxam and Gilbert 1980]. Cleavage at methylated adeines occurred during the NA45 paper elution procedure. Modified control DNA was not electrophoresed prior to strand cleavage. Equal amounts of radioactivity from bound, unbound, and control samples were resolved on sequencing gels. Gels were exposed to film, and autoradiographic densities were quantified as described above. Value ranges for each interfering site were observed in at least two independent experiments.

**Selected and amplified binding assays**
The selection and amplification of Ets-1-binding sites were performed with the following oligonucleotides: template N1, 5'-CTCGAGATGCATGACANNNNTTCTGGCCGTC-GACTAGATCT-3', primer A, 5'-GCCAGATCTAGTC-GACCGG-3', primer B, 5'-GGCTCGAGATGCATGAC-3'. The central 20 nucleotides of N1 DNA represented MSV LTR enhancer sequences on the noncoding viral strand between viral genome positions 200 and 221 [Van Beveren et al. 1981] except for the 7 positions with random sequence. The flanking regions of N1 contained restriction enzyme recognition sequences XhoI plus SphI within the left 13 nucleotides and Scal plus BgIII within the right 13 nucleotides. Primer A was complementary to the 3' end of the N1 plus a 3-nucleotide 5' extension. Primer B was identical to the 5' end of N1 with a 3-nucleotide 5' extension. Each oligonucleotide was synthesized, and full-length products were resolved from shorter products by gel purifica-
Nye et al.

Sizing the complementary strand with E. coli DNA polymerase (Klenow fragment) and isolating the product on a 6% agarose gel. DNA was electrophoresed, radiolabeled with [γ-32P]ATP and T4 polynucleotide kinase, and applied to a Bio-Rad spin 6 column to separate from nucleotides.

DNA-binding reactions and mobility-shift assays were performed as described above. Each contained 1.0 ng of probe [1.45 x 10^-9 M] and 0.36 ng of highly-purified ΔN322 [1.3 x 10^-9 M]. The wet gel was exposed to X-ray film for 8-20 hr at 4°C. DNA in complex with a monomeric protein species was gel purified by electrophoresis with 5 μg of tRNA in the final ethanol precipitation step. In the first round, the radioactivity of the DNA-protein complex was not detectable on X-ray film. To locate the position of nucleoprotein complexes, a binding reaction was electrophoresed in parallel that contained a 1000-fold higher concentration of ΔN322. A ladder of complexes was observed in this marker lane. The fastest complex provided a marker for the mobility of a monomeric complex. In subsequent rounds, a radioactive signal was visible at this position.

One-fifth of the gel-purified DNA was used in a PCR reaction at a final buffer concentration of 10 mM Tris-Cl (pH 8.4), 50 mM KCl, 1.5 mM MgCl2, 100 μg/ml of gelatin, 200 μM each dNTP, 2.5 μM each primer (A and B), and 2.5 units of Taq polymerase (Perkin Elmer). Reactions were incubated at 95°C (20 sec), 55°C (30 sec), 72°C (20 sec) for 35 cycles, and 72°C (3 min). The PCR product was resolved on a 6% polyacrylamide/0.25% agarose gel and recovered by electroelution. Labeling of the DNA and mobility-shift assays was performed in an identical manner for each of four subsequent selection rounds. The sixth round of selection was performed under more stringent binding reaction conditions (1.45 x 10^-9 M AN322). After each round of selection, approximately one-tenth of the PCR product was obtained (Kd = 2.8 x 10^-11 M) was close to the value obtained by the DNA titration experiment (Kd = 4.9 x 10^-11 M). The DNA titration design and Scatchard analysis allowed us to measure the values for total protein concentration, P0, and Kd for SC2 on the same mobility-shift gel using the same protein sample and radiolabeled probe that was used in the following competition experiments. Therefore, we used the Kd = 4.9 x 10^-11 value in calculation of relative affinities.

Competition was performed with a constant amount of radiolabeled SC2 DNA [1.24 x 10^-10 M] and ΔN322 [1.37 x 10^-9 M]. Competitor DNA ranged between 2 and 6 ng. To obtain competitor DNA, 1-2 μg of 61-bp XhoI-ApaI restriction fragments for each SC site were gel purified. Aliquots of the stock solutions were electrophoresed in parallel with DNA standards whose concentrations were determined spectrophotometrically. The DNA concentration of a stock competitor solution was determined by video densitometry performed on photographic negatives of ethidium-stained 6% polyacrylamide/0.25% agarose gels.

The strategy for our quantitative analysis assumed that a coupled equilibrium exists between added ΔN322 and the two binding sites (Liu-Johnson et al. 1986; Li et al. 1989). To calculate the dissociation equilibrium constant for a SC-binding site, X, the DNA titration experiment assumed that a coupled equilibrium exists: X = [P][D]/[PD], reactant and product concentrations were derived from the following measurable parameters: P0, the total added protein; D0, total added competitor site; Kd, dissociation constant for SC2-binding site; [PD] = P - [P]; D322, SC2 DNA in complex; [D322], unbound SC2 DNA. The latter two values were obtained from phosphorimaging analysis of the mobility-shift gels to determine the fraction of the total added SC2 site DNA in bound and unbound fractions. With these values, we calculated as follows: [P] = [PD322][Kd322]/[D322], [PD] = P - [P]; [D] = D0 - [PD] and thus Kd.

The GenBank accession number for murine ets-I cDNA is X53953.

Acknowledgments

We thank Diane Richardson for assistance in protein preparation. We thank Tom Alber and Brenda Bass for critical comments on the manuscript. This work was supported by National Institutes of Health [NIH] grant GM38663 and the March of Dimes Birth Defects Foundation Basic Research grant 1-0138 awarded to B.J.G.C. C.G. was funded as a predoctoral trainee on the NIH Cancer Training Grant CA09602.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Affolter, M., A. Percival-Smith, M. Müller, W. Leupin, and W.J. Gehring. 1990. DNA binding properties of the purified Antennapedia homeodomain. Proc. Natl. Acad. Sci. 87: 4093-4097.

Ben-David, Y., E.B. Giddens, K. Letwin, and A. Bernstein. 1991. Erythroleukemia induction by Friend murine leukemia virus: Insertional activation of a new member of the ets gene family, Fli-I, closely linked to c-ets-1. Genes & Dev. 5: 908-918.

Blackwell, T.K. and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A pro-

988 GENES & DEVELOPMENT
tein complexes revealed by binding site selection. Science 250: 1104–1110.

Bosselut, R., J.F. Duvall, A. Gegenonne, M. Bailly, A. Hemar, J. Brady, and J. Ghysdael. 1990. The product of the c-ets-1 proto-oncogene and the related Ets2 protein act as transcriptional activators of the long terminal repeat of human T cell leukemia virus HTLV-I. EMBO J. 9: 3137–3142.

Brenowitz, M., D.F. Senear, M.A. Shea, and G.K. Ackers. 1986. Quantitative DNase footprint titration: A method for studying protein-DNA interactions. Methods Enzymol. 130: 132–181.

Burtis, K.C., C.S. Thummel, C.W. Jones, F.D. Karim, and D.S. Fisher. 1986. c-ets-1 DNA binding domain and bending angle of E. coli CAP protein. Science 234: 1526–1541.

Calladine, C.R. 1982. Mechanics of sequence-dependent stacking of bases in B-DNA. J. Mol. Biol. 161: 343–352.

Carey, M., H. Kakidani, J. Leatherwood, F. Mostashari, and M. Ptashne. 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. J. Mol. Biol. 209: 423–432.

Dalton, S. and R. Treisman. 1992. Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell 68: 597–612.

Fisher, R.J., G. Mavrothalassitis, A. Kondoh, and T.S. Papas. 1991. High-affinity DNA-protein interactions of the cellular ETS protein: The determination of the ETS binding motif. Oncogene 6: 2249–2254.

Gill, S.C. and P.H. von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182: 319–326.

Graves, B.J., P.F. Johnson, and S.L. McKnight. 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell 44: 556–576.

Gunther, C.V., J.A. Nye, R.S. Bryner, and B.J. Graves. 1990. Sequence-specific DNA binding of the proto-oncoprotein Ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. Genes & Dev. 4: 667–679.

Halvorsen, Y.-D.C., K. Nandabalan, and R.C. Dickson. 1991. Identification of base and backbone contacts used for DNA sequence recognition and high-affinity binding by LAC9, a transcriptional activator containing a C6 zinc finger. Mol. Cell. Biol. 11: 1777–1784.

Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. Nature 353: 715–719.

Hipskind, R.A., V.N. Rao, C.G.F. Mueller, E.S.P. Reddy, and A. Nordheim. 1991. Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62TCF. Nature 354: 531–534.

Ho, I.-C., N.K. Bhat, L.R. Gottschalk, T. Lindsten, C.B. Thompson, T.S. Papas, and J.M. Leiden. 1990. Sequence-specific binding of human Ets-1 to the T cell receptor α gene enhancer. Science 250: 814–818.

Johnson, A. 1980. “Mechanism of action of the Cro protein.” Ph.D. thesis, Harvard University, Cambridge, MA.

Johnson, P.F. and S.L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. 58: 799–839.

Jordan, S.R. and O.C. Pabo. 1988. Structure of the lambda complex at 2.5 Å resolution: Details of the repressor-operator interactions. Science 242: 893–899.

Karim, F.D., J.D. Urness, C.S. Thummel, M.J. Klemisz, S.R. McKercher, A. Celdea, C. Van Beveren, R.A. Maki, C.V. Gunther, J.A. Nye, and B.J. Graves. 1990. The ETS-domain: A new DNA binding motif that recognizes a purine-rich core DNA sequence. Genes & Dev. 4: 1451–1453.

Kissinger, C.R., B. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. Pabo. 1990. Crystal structure of an engrafted homodomain-DNA complex at 2.8 Å resolution: A framework for understanding homodomain-DNA interactions. Cell 63: 579–590.

Klemsz, M.J., S.R. McKercher, A. Celada, C. Van Beveren, and R.A. Maki. 1990. The macrophage and B cell specific transcription factor PU.1 is related to the ets oncogene. Cell 61: 113–124.

LaMarco, K., C.C. Thompson, B.P. Byers, E.M. Walton, and S.L. McKnight. 1991. Identification of ets- and notch-related subunits in GA binding protein. Science 253: 789–792.

Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science 243: 1681–1688.

Leprience, D., A. Gegenonne, J. Coll, C. de Taisne, A. Schneecberger, C. Lagrou, and D. Stiehelin. 1983. A putative second cell-derived oncogene of the avian leukemia retrovirus E26. Nature 306: 395–397.

Li, R., J. Knight, G. Bream, A. Stenlund, and M. Botchan. 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. Genes & Dev. 3: 510–526.

Lim, F., N. Kraut, J. Frampton, and T. Graf. 1992. DNA binding by c-Ets-1, but not v-Ets, is repressed by an intramolecular mechanism. EMBO J. 11: 643–652.

Liu-Johnson, H., M. Gartenberg, and D. Crothers. 1986. The DNA binding domain and bending angle of E. coli CAP protein. Cell 47: 995–1005.

Maxam, A. and W. Gilbert. 1980. Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol. 65: 499–560.

McClarin, J.A., C.A. Frederick, B.-C. Wang, P. Greene, H.W. Boyer, J. Grable, and J.M. Rosenberry. 1986. Structure of the DNA-Eco RI endonuclease recognition complex at 3 Å resolution. Science 234: 1526–1541.

Nye, J.A. and B.J. Graves. 1990. Alkylation interference identifies essential DNA contacts for sequence-specific binding of the eukaryotic transcription factor C/EBP. Proc. Natl. Acad. Sci. 87: 3992–3996.

Pongubala, J.M.R., S. Nagulapalli, M.J. Klemisz, S.R. McKercher, R.A. Maki, and M.L. Atchison. 1992. PU.1 recruits a second nuclear factor to a site important for immunoglobulin κ’ enhancer activity. Mol. Cell. Biol. 12: 368–378.

Radke, K., H. Beug, S. Kornfeld, and T. Graf. 1982. Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the myb gene. Cell 31: 643–653.

Sauer, R.T., D.L. Smith, and A.D. Johnson. 1988. Flexibility of the yeast α2 repressor enables it to occupy the ends of its operator, leaving the center free. Genes & Dev. 2: 807–816.

Seeman, N.C., J.M. Rosenberg, and A. Rich. 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. 73: 804–808.

Siebenlist, U. and W. Gilbert. 1980. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. 77: 122–126.

Stetzig, T.A. 1990. Structure studies of protein-nucleic acid interaction: The sources of sequence-specific binding. Quart. Rev. Biophys. 23: 205–280.

Studier, F.W., A.H. Rosenberg, J.I. Dunn, and J.W. Dubendorff. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 183: 60–89.

Thompson, C.C., T.A. Brown, and S.L. McKnight. 1991. Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. Science 253: 762–768.

Thompson, C.B., C. Wang, I.-C. Ho, P.R. Bohijan, B. Petryniak,
Nye et al.

C.H. June, S. Meisfeldt, L. Zhang, G.J. Nabel, B. Karpinski, and J.M. Leiden. 1992. cis-acting sequences required for inducible Interleukin-2 enhancer function bind a novel ets-related protein, Elf-1. *Mol. Cell. Biol.* 12: 1043–1053.

Tullius, T.D., B.A. Dombroski, M.E.A. Churchill, and L. Kam. 1987. Hydroxyl radical footprinting: A high-resolution method for mapping protein-DNA contacts. *Methods Enzymol.* 155: 537–558.

Urness, L.D. and C.S. Thummel. 1990. Molecular interactions within the ecdysone regulatory hierarchy: DNA binding properties of the Drosophila ecdysone-inducible E74A protein. *Cell* 63: 47–61.

Van Beveren, C., F. van Straaten, J.A. Galleshaw, and I.M. Verma. 1981. Nucleotide sequence of the genome of a murine sarcoma virus. *Cell* 27: 97–108.

Vinson, C.R., P.B. Sigler, and S.L. McKnight. 1989. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science* 246: 911–916.

Wang, S. and N.A. Speck. 1992. Purification of core-binding factor, a binding that binds the conserved core site in murine leukemia virus enhancers. *Mol. Cell. Biol.* 12: 89–102.

Wasylyk, B., C. Wasylyk, P. Flores, A. Begue, D. Leprince, and D. Stehelin. 1990. The c-ets proto-oncogenes encode transcription factors that cooperate with c-Fos and c-Jun for transcriptional activation. *Nature* 346: 191–193.

Wasylyk, C., A. Gutman, R. Nicholson, and B. Wasylyk. 1991. The c-Ets oncoprotein activates the stromelysin promoter through the same elements as several non-nuclear oncoproteins. *EMBO J.* 10: 1127–1134.

Wolberger, C., A.K. Vershon, B. Liu, A.D. Johnson, and C.O. Pabo. 1991. Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* 67: 517–528.

Woods, D.B., J. Ghysdael, and M.J. Owen. 1992. Identification of nucleotide preferences in DNA sequence recognized specifically by c-Ets-1 protein. *Nucleic Acids Res.* 20: 699–704.

Xin, J., A. Cowie, P. Lachance, and J.A. Hassell. 1992. Molecular cloning and characterization of PEA3, a new member of the Ets oncogene family that is differentially expressed in mouse embryonic cells. *Genes & Dev.* 6: 481–496.
Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif.

J A Nye, J M Petersen, C V Gunther, et al.

*Genes Dev.* 1992, 6:
Access the most recent version at doi:10.1101/gad.6.6.975