

TFEB has DNA-binding and oligomerization properties of a unique helix–loop–helix/leucine-zipper family

David E. Fisher,^{1,2} Cynthia S. Carr,¹ Lana A. Parent,¹ and Phillip A. Sharp¹

¹Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA, ²Departments of Pediatric and Medical Oncology, Dana Farber Cancer Institute & Pediatric Hematology, The Children's Hospital, Harvard Medical School, Boston, Massachusetts USA 02115

The DNA-binding factor TFEB contains adjacent helix–loop–helix (HLH) and leucine zipper (LZ) domains flanked by an upstream basic region. This arrangement of interactive motifs has recently been observed in several other transcription factors and in the Myc family of oncogenes. TFEB was isolated by virtue of its binding to the major late promoter of adenovirus. DNA binding by a soluble protein was achieved by deleting a hydrophobic amino-terminal domain and permitted the structural analysis of the oligomerization and DNA-binding properties of TFEB. TFEB specifically bound DNA as both a homodimer and a heterodimer with another b-HLH–LZ protein TFE3. The LZ domain was essential for homo- or hetero-oligomerization and high-affinity DNA binding. In the absence of DNA a tetramer-sized form of TFEB was observed that dissociates to bind added DNA as a dimer. Binding by TFEB and TFE3 to related, but different, naturally occurring DNA target sequences was observed with distinct binding preferences. Analysis of basic domain residues in this family of proteins revealed a pattern of sequence conservation predictive of an interacting α -helical face. Common oligomerization and DNA-binding features suggest the b-HLH–LZ domain structure to define a distinct family of DNA-binding factors.

[*Key Words*: TFEB; helix–loop–helix; leucine zipper; immunoglobulin enhancer; protein dimer; protein tetramer]

Received May 30, 1991; revised version accepted September 13, 1991.

Discrete families of mammalian DNA-binding transcription factors have emerged from the discoveries of numerous interacting peptide motifs. Two of these, the helix–loop–helix (HLH; Tapscott et al. 1988; Murre et al. 1989a,b) and leucine zipper (LZ; Kouzarides and Ziff 1988; Landschulz et al. 1988; Sassone-Corsi et al. 1988; Gentz et al. 1989; O'Shea et al. 1989; Turner and Tjian 1989), have been the subjects of considerable structural and functional analysis. HLH and LZ are frequently found adjacent to an amino-terminal flanking basic region domain (b) containing an abundance of basic amino acids. Mutational analyses (Davis et al. 1990; Voronova and Baltimore 1990) have demonstrated DNA-binding specificity in the basic region. Factors naturally lacking basic domains were found to be incapable of and inhibitory to DNA binding (Benezra et al. 1990; Ellis et al. 1990). The HLH and LZ domains have been shown to facilitate protein dimerization between different members of the same HLH or LZ family (Kouzarides and Ziff 1988; Landschulz et al. 1988; Sassone-Corsi et al. 1988; Gentz et al. 1989; Murre et al. 1989b; O'Shea et al. 1989). Models of the protein–DNA interactions have suggested that the LZ domain forms an amphipathic α -helical seg-

ment that positions the adjacent basic domain for binding in the major groove of DNA (Landschulz et al. 1988; O'Neil et al. 1990; Talanian et al. 1990).

The original structural studies of HLH and LZ domains were done on proteins containing one or the other of these motifs, but some proteins contain both HLH and LZ domains immediately adjacent to a single basic region. This combined structure has been noted recently in the enhancer binding factors AP-4 (Hu et al. 1990), USF (Gregor et al. 1990), TFE3 (Beckmann et al. 1990), TFEB (Carr and Sharp 1990), c-, L-, and N-Myc (DePinho et al. 1987), and Myn/Max (Blackwood and Eisenman 1991; Prendergast et al. 1991). Similarities in target DNA sequences recognized by these factors, as well as the common structural domains, suggest that this group could represent a distinct family of DNA-binding proteins. Although initial reports have suggested that these proteins bind DNA as dimers (Blackwell et al. 1990; Gregor et al. 1990; Hu et al. 1990; Prendergast and Ziff 1991), structural analyses of c-Myc by cross-linking and gel filtration have demonstrated a predominantly tetrameric form of the protein (Dang et al. 1989). The existence of dimeric and tetrameric forms of μ -enhancer-

binding proteins has also been observed (Peterson and Calame 1989).

The factor TFEB has been shown to bind sequences in the adenovirus major late promoter (MLP; Carr and Sharp 1990). The core sequence CACGTG in MLP is closely related to sequences recognized by members of the HLH family and is similar to common enhancer motifs in the regulation of immunoglobulin transcription (Ephrussi et al. 1985). TFE3, a different transcription factor, shares extensive sequence homology with TFEB in the basic, HLH, and LZ regions. TFE3 was isolated originally by virtue of its binding to the μ E3 immunoglobulin enhancer site and can activate transcription *in vivo* by binding to this site (Beckmann et al. 1990).

Analysis of the binding properties of TFEB to DNA and its ability to interact with TFE3 are demonstrated and provide information on structural features of the b-HLH-LZ family of factors. Heterodimer formation has been demonstrated between TFEB and TFE3 but not among factors lacking the adjacent HLH and LZ domains. A 2 : 1 stoichiometry of protein-DNA binding has been observed together with an apparent propensity to form tetramer-sized species in the absence of DNA. The ability of both TFEB and TFE3 to bind two related, but different, naturally occurring target DNA sequences has been shown along with clear binding preferences. Comparison of basic domain sequences from factors capable of binding to a CACGTG DNA target was made and reveals evidence for a conserved α -helical face that may interact with base pairs in the DNA major groove. The capacity for multiple competing interactions of potential regulatory importance appears to distinguish the b-HLH-LZ proteins as a unique family.

Results

MLP binding by TFEB is enhanced by intact LZ

Previous experiments have shown that proteins encoded by the TFEB cDNA were insoluble. Proteins representing both amino- and carboxy-terminal deletions of TFEB (Fig. 1A) were generated by translation *in vitro* and tested for DNA binding. Deletion of an amino-terminal domain rich in glycine and alanine was necessary to produce a protein capable of binding the MLP probe with electrophoretic mobility-shift assay (EMSA; Fig. 1B). Deletion of either the first 120 (A Δ 120) or 265 (A Δ 265) amino acids produced a protein that formed complexes efficiently with mobilities consistent with their relative molecular masses.

Determination of carboxy-terminal sequences necessary for DNA binding was studied with constructs containing the A Δ 265 amino-terminal deletion (Fig. 1C). Deletion of sequences carboxyl terminus to or including the last leucine of the LZ domain retained strong DNA binding (C Δ 421 and C Δ 415, respectively). Deletion of the entire LZ domain nearly abrogated DNA binding (C Δ 394 and C Δ 390). Nonetheless, a weak band was observed with these smaller deletion mutants, a reproducible finding with both C Δ 394 and C Δ 390 with long exposure of gels (data not shown). Comparable translation efficiency was verified by SDS-PAGE (data not shown). The faint band generated by A Δ 265/C Δ 390 migrates disproportionately rapidly relative to that generated by A Δ 265/C Δ 415, considering that the size difference between the two proteins is only 25 amino acids. Taken together, these data suggest that the leucine zipper is important for high-affinity DNA binding, although low-affinity

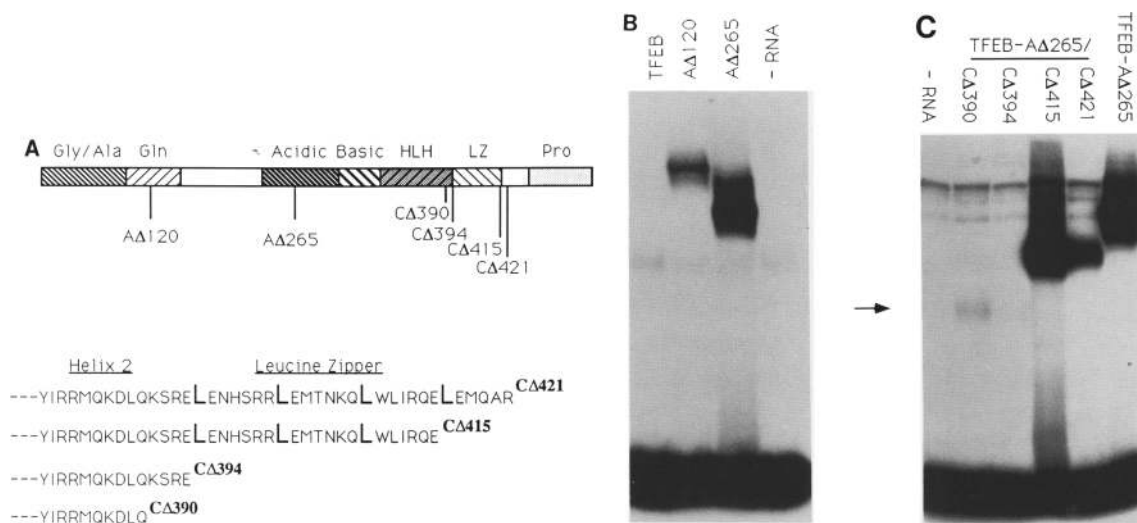


Figure 1. MLP binding by soluble TFEB. (A) Map of TFEB protein indicating glycine- and alanine-rich (Gly/Ala), glutamine-rich (Gln), acidic, basic, helix-loop-helix (HLH), leucine zipper (LZ), and proline-rich (Pro) domains. Amino-terminal (A Δ) and carboxy-terminal (C Δ) deletions are indicated with the respective amino acid positions. Carboxy-terminal deletion mutant amino acid sequences are shown below, with the periodic leucines of the zipper in boldface type. (B) EMSA of binding to MLP DNA probe by TFEB and amino-terminal deletion mutants. Reticulocyte lysate to which no exogenous RNA was added before incubation is indicated (- RNA). (C) EMSA of TFEB carboxy-terminal deletions using the A Δ 265 amino-terminal deletion mutant and MLP probe. (- RNA) No RNA added to reticulocyte lysate. The arrow indicates the position of the weaker gel-shift band in mutants with deleted LZ domains.

Fisher et al.

DNA binding may occur in its absence. Construct A Δ 265/C Δ 390 is also missing part of the putative helix 2 from its HLH domain, perhaps implying its dispensability in this weak binding. The most probable explanation is that the LZ minus mutants bind DNA weakly as monomers, whereas the high-affinity binding requires formation of dimers (see below). High-affinity DNA binding has not been observed with mutants containing in-frame deletion within the HLH motif but leaving the leucine zipper intact (data not shown).

Sequence specificity of binding

Specificity of DNA binding by a soluble TFEB deletion mutant was demonstrated by competition (Fig. 2A) with either unlabeled wild-type MLP probe (core sequence CACGTG) or double point mutant (core sequence GAGGTG). The wild-type probe competed effectively, whereas the point mutant did not compete at identical molar excess (Fig. 2A). It is significant that an intact half-site (GTG), which remained intact in the mutant probe, was unable to compete with wild type, even in vast excess. Specificity of binding was demonstrated further by methylation interference analysis (Fig. 2B). The pattern of guanine residues whose methylation interfered with binding is centered over the core CACGTG sequence and extends to either side in a roughly symmetrical fashion to a pair of flanking guanine residues. This pattern is nearly identical to that determined previously for the binding of MLTF or USF to the MLP site (Carthew et al. 1985; Sawadogo and Roeder 1985; Chodosh et al. 1987) and to the methylation interference pattern seen by Southwestern blotting with a β -gal-

TFEB fusion (Carr and Sharp 1990). Because that fusion protein contained the first 120 amino acids whose deletion produces a soluble DNA-binding protein, those 120 residues are unlikely to make direct DNA contacts.

TFEB homodimerization

TFEB homo-oligomerization and stoichiometry of protein-DNA complexes were examined by resolving the mobilities of complexes formed by a mixture of two different TFEB-derived proteins containing intact b-HLH-LZ domains (Fig. 3A). Titration was carried out to optimize the chances of observing any potentially less favored intermediate forms. Mixing varying proportions of large peptide (TFEB-A Δ 120; Fig. 3A) and small peptide (TFEB-A Δ 265/C Δ 421; Fig. 3B) generated a single new intermediate mobility band (Fig. 3A,B). Weaker binding of the probe by an endogenous reticulocyte component could also be seen (asterisk; Fig. 3A). Formation of these putative heterodimers was equally efficient either by mixing previously translated proteins or cotranslation of the two mRNAs (data not shown), although the former allows for more accurate titration of products. These results suggest that DNA binding by the intact b-HLH-LZ domain occurs as a protein-protein dimer, although a higher-order structure, such as tetramer, cannot be excluded from these results alone if unanticipated constraints prevent the formation of certain oligomer forms.

TFEB and TFE3 can heterodimerize efficiently

The possibility of interactions between TFEB and TFE3 was tested by mixing the two proteins in the presence of

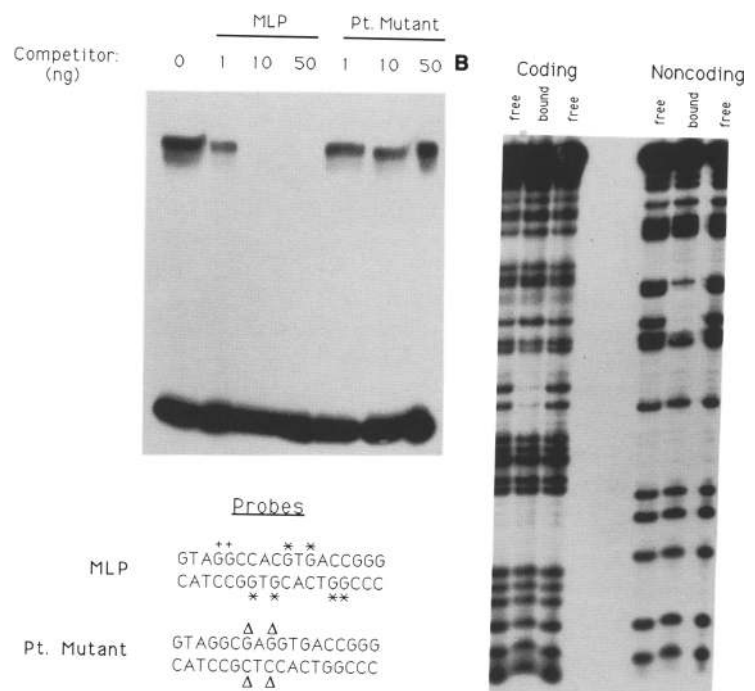


Figure 2. Specificity of binding to MLP. (A) EMSA competition was performed by adding unlabeled DNAs (MLP or double point mutant) of the same molecular masses as the labeled MLP probe. (B) Guanine methylation interference analysis of binding by TFEB A Δ 265/C Δ 421 to single end-labeled MLP probes. Coding and noncoding strands were determined with separately end-labeled probes. DNA sequence that was bound is in each case shown flanked by two lanes containing the unbound probe (free). Bands corresponding to sequences strongly affected (*) and weakly affected (+) by methylation interference are indicated. DNA probe sequences are shown at the bottom of A, with the interference pattern indicated for the MLP probe and the two sites of mutation indicated (gD) on the double point mutant competitor sequence.

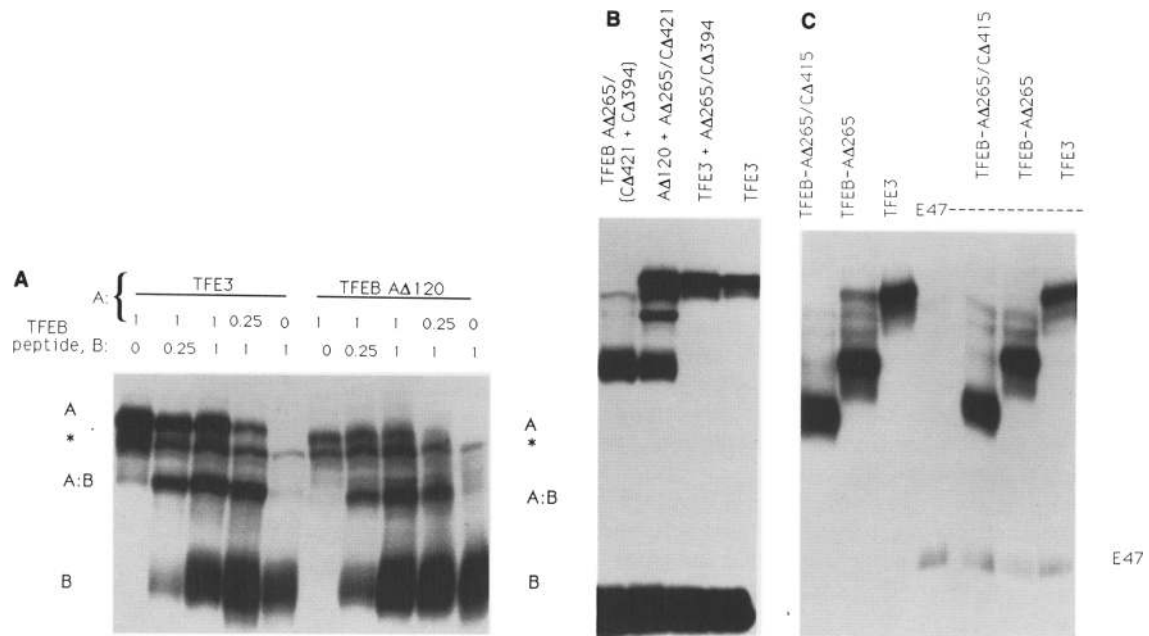


Figure 3. Protein–protein interactions of TFEB. (A) Large peptides, labeled A [TFE3 intact (Beckmann et al. 1990) or TFEB AΔ120] were produced by *in vitro* translation and mixed in varying amounts with separately translated small peptide, labeled B (TFEB AΔ265/CA421). Quantities of reticulocyte lysate-generated protein are indicated (0, 0.25, or 1 μ l). Homodimer gel-shift bands are indicated (A and B) as are the heterodimer bands (A : B) for both the TFE3(A) : TFEB(B) and TFEB(A) : TFEB(B) mixes. An endogenous reticulocyte lysate band is indicated by an asterisk (*). Free probe was allowed to run off the bottom of the gel. (B) Inability to heterodimerize in the absence of intact LZ. TFEB deletion mutants containing (TFEB AΔ265/CA421) or lacking (TFEB AΔ265/CA394) intact LZs were mixed with each other and intact TFE3, all generated by *in vitro* translation. Mobility of the LZ minus mutant (AΔ265/CA394) band, seen on long exposure, is faster than the LZ-containing constructs (see Fig. 1C). (C) Inability to heterodimerize with the HLH protein E47. TFEB deletion mutants and TFE3 were mixed individually with *in vitro*-translated E47S protein. Free probe was allowed to run off the bottom of the gel.

MLP probe (Fig. 3A). Again, a single new band of intermediate mobility was seen. Analogous to the TFEB–TFEB complexes described above, those complexes formed between TFEB and TFE3 are likely to represent protein dimers. The heterodimer formation was equally efficient by cotranslation or mixing of separately translated proteins (data not shown). The interactions of TFEB with TFE3 reveal a similarity of two b-HLH–LZ family members both in forming protein–protein oligomers as well as in binding the same DNA sequence.

Reversibility of the dimerization process was studied by mixing one protein with DNA probe before the addition of the second protein or, alternatively, by mixing the two proteins before the addition of the DNA probe. The order of addition had little effect on the formation of the heterodimer (data not shown). Specifically, preincubation of TFE3 alone with DNA probe did not affect its efficient formation of intermediate mobility heterodimers upon addition of the smaller TFEB deletion mutant, and vice versa. These results suggest that protein oligomerization and DNA binding are both largely reversible over a short time.

The importance of the LZ domain in dimerization was demonstrated by mixing proteins containing or lacking this domain in the presence of DNA probe. As shown in Figure 3B, addition of the smaller LZ minus mutant

AΔ265/CA394 did not produce a more rapidly migrating intermediate band when mixed either with TFEB–AΔ265/CA421 (containing intact LZ) or with intact TFE3. Furthermore, the presence of the LZ minus protein (as verified by SDS-PAGE; data not shown) did not inhibit DNA binding by proteins with an intact LZ, as the intensity of the TFE3 homodimer band was not diminished. Under identical conditions the TFEB protein AΔ120 produced an intermediate mobility heterodimer when mixed with the same smaller TFEB mutant AΔ265/CA421. These results suggest that the LZ domain is important for the formation of DNA-binding oligomers.

The HLH protein E47S was tested for possible binding to the MLP sequence and found to give a reproducible gel-shift complex (Fig. 3C). Although E47 was isolated by virtue of its binding to the κ E2/ μ E5 sequence (Murre et al. 1989a), the MLP sequence falls into the same CANNTG family. This binding by E47S to MLP also provided the unique opportunity to test for possible heterodimerization by members of two different families, both capable of binding the same DNA sequence. As seen in Figure 3C, although E47S as well as TFEB and TFE3 were all capable of binding the MLP probe in the same reaction mixtures, no intermediate mobility complexes were observed to suggest heterodimer formation

or interference with homodimers. The same results were obtained by cotranslation of mRNAs. Failure to form DNA-binding heterodimers or inhibit TFEB DNA binding has also been observed with the LZ proteins Fos and Jun and the b-HLH-LZ protein AP-4 (data not shown).

Relative binding affinities for MLP and μ E3

The relative affinities of TFEB and TFE3 for binding the MLP (core sequence CACGTG) and μ E3 (core sequence CATGTG) sites were also tested. TFEB and TFE3 were found to bind the μ E3 probe separately as homodimers, although with lower efficiency than the MLP probe (data not shown). In addition, when mixed together, TFE3 and TFEB formed a heterodimer band with the μ E3 probe (data not shown). Binding to μ E3 probe was observed by using one-tenth the amount of poly[d(I-C)] nonspecific competitor as compared with assays for binding the MLP probe. To assay more quantitatively for relative affinity, unlabeled μ E3, MLP, and double point mutant DNAs were tested for competition with formation of protein-DNA complexes containing the b-HLH-LZ intact peptide TFEB A Δ 265/C Δ 421 and labeled MLP probe (Fig. 4A). A clear preference was observed for binding to MLP DNA. Although addition of 10 ng of unlabeled MLP essentially eliminated binding to the labeled probe, 100 ng of unlabeled μ E3 DNA was required for the same extent of competition. The double point mutant DNA (core sequence GAGGTG) competed more weakly than either MLP or μ E3 DNA. The same preferential avidity for MLP as compared with μ E3 templates was observed with the intact TFE3 protein (Fig. 4B).

Dimer- and tetramer-sized protein oligomers

Given the apparent tetrameric state of c-Myc, as ana-

lyzed previously by cross-linking and gel filtration (Dang et al. 1989), it was of interest to determine the state of TFEB in the absence of DNA probe. This was accomplished by fractionating *in vitro*-translated TFEB (A Δ 265/C Δ 421) by gel filtration and assaying all fractions by electrophoretic mobility-shift assay (EMSA). Three distinct peaks of DNA-binding activity were identified (Fig. 5A). The most prominent peak has mobility somewhat slower than aldolase (158 kD). A second peak has mobility just slower than BSA (67 kD). And a third smaller peak has mobility near that of chymotrypsinogen (25 kD). The TFEB construct used for these studies migrates on SDS-PAGE as a 25-kD protein (data not shown). The gel filtration peaks seen here would appear to be candidates for tetramer, dimer, and monomer. By using the gel filtration elution data, Stokes radii of the three peaks were determined to be 48.2, 31.5, and 22.4 Å (Fig. 5C, right).

Glycerol gradient sedimentation of TFEB A Δ 265/C Δ 421 in the absence of DNA was also carried out and assayed by EMSA (data not shown). This also revealed prominent dimer- and tetramer-sized peaks (Fig. 5C, left) with corresponding S values of 4.56 and 6.64, respectively. By using the combined S value and Stokes radius determinations, native molecular mass of these two peaks was calculated to be 58 and 129 kD (see Materials and methods). These sizes are consistent with dimer and tetramer formation by TFEB in the absence of DNA.

Figure 5A also reveals that the DNA-protein complexes formed by the different oligomerized forms of TFEB all migrate identically by using native gel electrophoresis. This observation suggests that each complex contains the same protein-DNA stoichiometry. Because the fraction of TFEB containing a putative tetramer generates a DNA-protein complex that migrates as an ap-

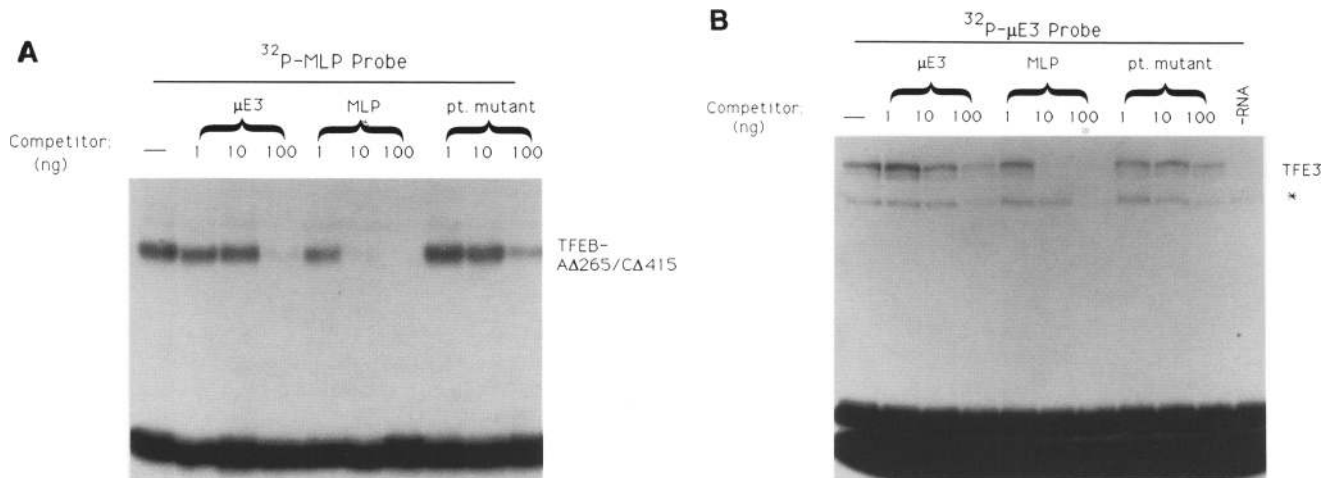


Figure 4. Preferential binding by TFEB and TFE3 for the MLP site over the μ E3 site. (A) *In vitro*-translated TFEB A Δ 265/C Δ 415 was incubated with 32 P-labeled MLP probe in the presence of increasing quantities of unlabeled competitor DNAs with 1 μ g of poly[d(I-C)]. The unlabeled μ E3 competitor was derived by *Bam*HI cleavage of the plasmid pBS-SK- μ E3. (B) *In vitro*-translated TFE3 was incubated with 32 P-labeled μ E3 probe in the presence of unlabeled MLP, μ E3, and double point mutant competitor DNAs and 0.1 μ g of poly[d(I-C)]. The μ E3 labeled and unlabeled sequences were derived by *Xba*I and *Eco*RV cleavage of pBS-SK- μ E3. Reticulocyte lysate to which no exogenous RNA was added before *in vitro* translation is indicated (-RNA). An endogenous reticulocyte lysate band is indicated by an asterisk (*).

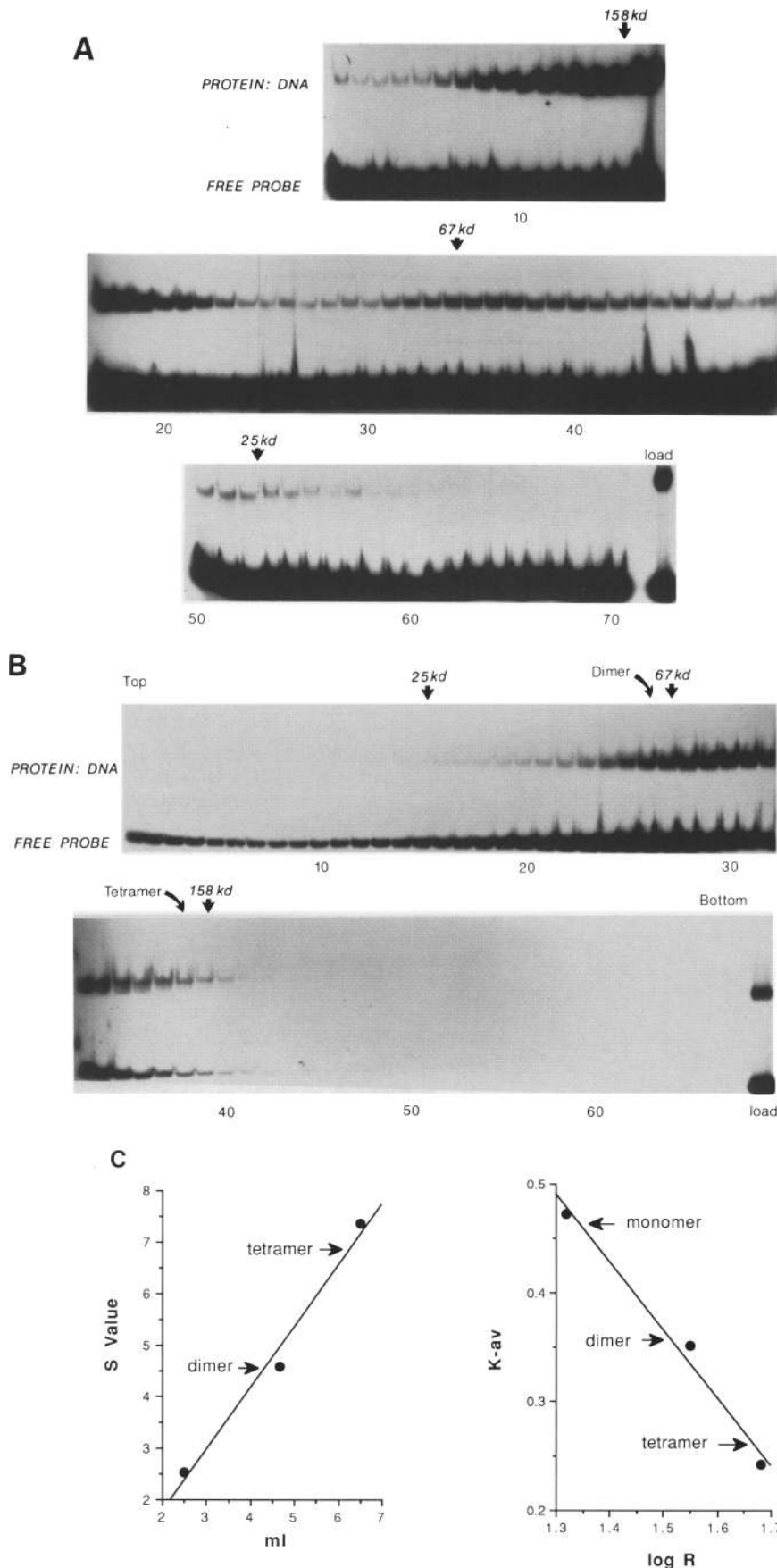


Figure 5. Gel filtration and glycerol gradient centrifugation analysis of protein in the absence or presence of DNA probe. (A) In vitro-translated TFEB A Δ 265/C Δ 421 was fractionated by gel filtration with a Superose 12 FPLC column. Labeled MLP probe was added to each fraction and complexes were visualized by EMSA. A lane showing the column load material is indicated. Protein standards are chymotrypsinogen (25 kD), BSA (67 kD), and aldolase (158 kD). Approximately tetramer-sized (lane 16), dimer-sized (lane 36), and monomer-sized (lane 51) peaks of DNA-binding protein were resolved, though all three migrate identically on the native gel when bound to DNA probe. Molecular mass calculations were made as described in Materials and methods. (B) In vitro-translated TFEB A Δ 265/C Δ 421 was mixed with 32 P-labeled MLP probe and fractionated by glycerol gradient centrifugation. Complexes were visualized after native gel electrophoresis of each fraction. Load material is indicated in a separate lane. The positions of dimer- and tetramer-sized peaks, as resolved in the absence of DNA during parallel gradient centrifugation (data not shown), are indicated. The same protein standards were employed as in A. (C) Standard curves are shown for gel filtration (right) and glycerol gradient sedimentation (left) of TFEB A Δ 265/C Δ 421 in the absence of added DNA. (●) Migration of standards listed above. Positions of the various oligomer-sized peaks are indicated by arrows. Combined S value and Stokes radius [R] determinations allowed for calculation of native molecular mass corrected for shape (see text).

Fisher et al.

parent dimer (see Fig. 3A), it appears that the tetrameric form of TFEB protein dissociates to bind as dimer with the addition of the DNA probe. An implication of this result is that the tetrameric form probably cannot bind MLP DNA directly, at least under the standard assay conditions. The overall amount of DNA–protein complex detected across the gradient (relative to the total protein loaded) indicates nearly complete recovery of the protein in a DNA-binding form. As such, it is likely that nearly all of the protein in the tetramer-sized fractions bound DNA. This suggests that dimer–DNA interactions are much more stable than interactions that hold the tetramer together (e.g., dimer–dimer) for this protein and MLP probe.

As an independent test of whether TFEB dimers and tetramers are able to bind DNA, glycerol gradient sedimentation was carried out with TFEB in the presence of radioactive DNA probe, followed by native gel electrophoresis of all fractions (Fig. 5B). A single peak of protein–DNA complexes was identified and has an intermediate mobility relative to the dimer- and tetramer-sized complexes resolved in a parallel sedimentation experiment (Fig. 5C). This conversion of two peaks to one by the addition of DNA is consistent with the gel filtration studies showing the same-sized gel-shift complex when bound to DNA from either monomer-, dimer-, or tetramer-sized fractions. These data suggest that only the dimeric size form is capable of binding DNA directly.

Potential conserved α -helical face of MLP-binding basic domains

Because multiple different members of this family of DNA-binding proteins appear capable of binding to CACGTG DNA targets, sequence comparison of corresponding basic regions was carried out on the basis of the method used by O'Neil et al. (1990). Amino acid sequences of eight proteins that were likely able to bind the MLP core sequence were aligned in their basic domains. A consensus emerged and is illustrated in Figure 6, where the number of different residues is plotted for each position in the ~20-amino-acid region. A clear cyclical pattern emerged, with greatest conservation of sequence occurring at an average distance of 3–4 amino acids. This is consistent with the highly conserved amino acids being oriented on the same face of an α -helix so that these residues may make contact with the major groove of the DNA template. Although the most conserved positions all contain charged amino acids, several are acidic and others are basic. The AP-4 protein (a different b-HLH–LZ protein) differs at several consensus residues and matches at others. In even greater contrast, the consensus pattern for the Fos/Jun/GCN4 family is shown and appears strikingly different for these LZ proteins that bind different DNA target sequences.

Discussion

The presence of multiple potentially interactive domains in TFEB and similar proteins has prompted the

analysis of the specificity of its protein–protein and protein–DNA complexes. Deletion of the amino-terminal domain of TFEB that is rich in glycine and alanine was necessary to observe DNA binding. This is probably due to the relative insolubility of the more intact form of TFEB. The biological importance of the domain rich in glycine and alanine is unclear, but a similar alanine-rich domain has been shown to function as a transcriptional repressor in the *Drosophila* Krüppel protein (Licht et al. 1990).

The TFEB protein and a number of related factors contain a basic region immediately adjacent to a HLH domain and a LZ domain. Mutational analysis shows that the LZ is essential for high-affinity DNA binding. A low-affinity complex was observed if the LZ region was deleted, which probably represents the binding of a monomeric protein. Similar observations have been made for USF (Gregor et al. 1990) and AP-4 (Hu et al. 1990), both of which form homodimers similarly dependent on the integrity of the LZ domain. Both proteins also bind, perhaps as monomers, when the LZ is deleted. The heterodimerization and DNA binding of c-Myc, Max, and Myn (the mouse homolog of Max) also appear to require intact LZ domains (Blackwood and Eisenman 1991; Prendergast et al. 1991).

Methylation interference analysis of the solubilized TFEB–DNA complex revealed that the critical guanine contacts were distributed symmetrically over the CACGTG core sequence and two flanking guanine residues. The residues affected by methylation interference analysis are virtually identical to those affected by the intact insoluble TFEB protein by Southwestern technique (Carr and Sharp 1990), suggesting that the amino-terminal glycine/alanine-rich domain does not contribute directly to DNA-binding specificity. The pattern is also very similar to that of USF–MLTF factor binding to the MLP site (Carthew et al. 1985; Sawadogo and Roeder 1985; Chodosh et al. 1987) and analogous to the binding of TFE3 to the μ E3 site (Beckmann et al. 1990). Thus, proteins with significantly different (USF) or similar (TFEB and TFE3) basic region amino acid sequences are able to bind the same target DNA with virtually identical interference patterns. Such observations suggest that a search for conserved residues may yield more information on possible contact residues for this family of proteins (see below). The interference patterns are also consistent with dimeric binding, which places two basic domains in opposite orientations over the center of the core hexanucleotide sequence. An implication of this structure is that an altered binding specificity might arise by heterodimer formation with a different member of the family containing a compatible LZ domain. In addition, with multiple proteins in the b-HLH–LZ family capable of binding the CACGTG core, dimeric binding allows for a diversity of other functional domains (e.g., *trans*-activating motifs) to be bound to the DNA target by heterodimerization.

Among proteins in the newly emerging b-HLH–LZ family a strikingly restricted ability to heterodimerize has been seen, even among proteins capable of binding

the same target DNA (Beckmann and Kadesch 1991). Only Myc and Max have been seen previously to form heterodimers (Blackwood and Eisenman 1991; Prendergast et al. 1991), although unlike TFEB and TFE3, the Myc/Max proteins have not been demonstrated to homodimerize efficiently. The TFEB–TFE3 complex bound DNA with an affinity comparable to either homodimer, and its formation was dependant on the integrity of the LZ. TFEB and TFE3 may bind DNA efficiently as separate or mixed species, thereby participating potentially in diverse regulatory pathways as each carries unique additional domains to any bound DNA target sequence. An acidic *trans*-activating domain has been demonstrated for TFE3 (Beckmann et al. 1990), whereas TFEB (Carr and Sharp 1990) contains three potential *trans*-activating domains and a glycine/alanine-rich region bearing some homology to transcriptional repressor sequences identified in the Krüppel protein (Licht et al. 1990). Apparently, TFEB did not interact with E47 by using EMSA, in which both proteins were capable of binding the same target DNA sequence. Although this observation does not address whether TFEB and E47 can dimerize in the absence of DNA, it does suggest that either their HLH domains are inherently incompatible or that the formation of heterodimers positions the adjacent basic domains inappropriately for DNA binding.

The observations that TFEB, TFE3, and USF bind to both the μ E3 and MLP sites (Fig. 4A; Beckmann et al. 1990; Gregor et al. 1990) led to an analysis of affinity preferences for the two DNA sequences. The μ E3 core hexanucleotide (CATGTG) differs at only one position from the MLP core (CACGTG), although they differ more extensively in sequences outside the core region that are important in DNA binding. Both TFEB and TFE3 bind with approximately an order of magnitude preference for the MLP site as compared with the μ E3 site. The near-identical DNA-binding preferences for TFE3 and TFEB are consistent with there being DNA contacts with basic regions of these peptides, as the two proteins share 19 of 20 amino acids in the basic region and are less homologous elsewhere. The varied affinities for particular sites may also reflect important physiologic differences in comparing viral (MLP) and cellular (μ E3) promoter/enhancer sequences. It is interesting that TFE3 apparently fails to *trans*-activate through the MLP site under conditions where it *trans*-activates readily through the lower affinity μ E3 site (Beckmann et al. 1990; Gregor et al. 1990). This suggests that the particular sequences to which a factor is bound are important for activation of transcription, possibly due to the character of the binding complexes. The promiscuity of DNA recognition was also demonstrated for E47 (Fig. 3C), which was shown to recognize the CACGTG core sequence whose two central positions differ from the κ E2/ μ E5 sequence to which it avidly binds (Murre et al. 1989a). Such multiplicity of binding to related naturally occurring DNA sequences may be a general property of basic region containing transcription factors implying an additional regulatory mechanism for transcriptional control.

Protein–DNA stoichiometry of 2 : 1 has been seen for numerous transcription factors that bind to palindromic DNA sequences (e.g., HLH or LZ families). Evidence that b-HLH–LZ proteins bind as dimers has been presented for AP-4 and TFE3 with mixed heterodimers and cross-linking (Hu et al. 1990; Beckmann and Kadesch 1991), as well as for USF with mixed heterodimers (Gregor et al. 1990). Mixed heterodimer formation and glycerol gradient centrifugation were used here and revealed a single band of intermediate mobility consistent with 2 : 1 protein–DNA stoichiometry of binding.

With two potential dimerization interfaces (HLH and LZ) one might predict proteins of this family to be capable of forming tetramers (or even higher-order complexes). The report that bacterially produced c-Myc (also a b-HLH–LZ protein) behaves as tetramer during gel filtration and chemical cross-linking (Dang et al. 1989) supports this prediction. On the other hand, concerns that the observed c-Myc tetramers could have arisen as a consequence of bacterial overproduction or purification have been raised, as dimeric DNA binding has been demonstrated for many of the proteins in this family (Fig. 3A; Blackwell et al. 1990; Gregor et al. 1990; Hu et al. 1990; Blackwood and Eisenman 1991; Prendergast et al. 1991). The existence of dimer and tetramer forms of a μ E3-binding protein in a plasmacytoma cell extract has also been demonstrated previously by gel filtration and EMSA (Peterson and Calame 1989). To study the state of TFEB in native, dilute conditions, gel filtration was performed with *in vitro*-translated protein in the absence of DNA. Three discrete peaks corresponding in size to predicted monomers, dimers, and tetramers were observed, with the tetramer peak containing the majority of the protein. The tetramer-sized mobility of this large peak is consistent with the bacterially produced tetramer form of c-Myc (Dang et al. 1989) that was seen in the absence of DNA. The tetrameric and dimeric states of TFEB were also observed by using glycerol gradient centrifugation. From the combined *S* value and Stokes radius, native molecular masses for these complexes were calculated. Molecular masses of 58 and 129 kD were obtained and strongly suggest dimeric and tetrameric forms for these complexes. Because the current study was not performed with purified protein, it is impossible to be certain that the tetramer-sized activity was not caused by complexes to other proteins in the translation mix, although the experiments with purified c-Myc (Dang et al. 1989) are complementary in this regard. In addition, selective disappearance of the tetramer-sized complexes with the addition of DNA (Fig. 5B) argues against nonspecific interactions in reticulocyte lysate.

It is noteworthy that the EMSA mobility of the protein–DNA complexes generated by the dimer- and tetramer-sized peaks was identical. Both fractions apparently generate homodimer protein–DNA complexes. The tetramer form of TFEB probably cannot bind DNA directly but dissociates into DNA-binding dimers. This suggests a potential regulatory competition between dimer–DNA and dimer–dimer (tetramer)-forming pathways. Modifications that shift the stability of the alter-


```

TFEB.....LAKERQKKDNHNLIERRRRFN
TFE3.....LLKERQKKDNHNLIERRRRFN
MAX.....VESDADKRAHNALEKRRRDH
L-Myc.....DTEDEVTKRKNHNFLEKRRND
c-Myc.....NTEENVKRRTHNVLEQRNRNE
N-Myc.....DSEDSERRRNHNILERQRRND
USF.....TTRDEKRRRAQHNEVERRRRDK
E47.....DLRDRERRMANNARERVRVRD

```

```

          E  KK  H      R
consensus.....**--**--**--N**ER*R--**
          D  RR  N      V

```

```
AP-4.....DQERRIRREIANSNERRRMQS
```

```

c-Fos.....RERNKMAAAKCRNRRRELTDT
Jun-D.....RLRNRIAAKCRKRKLERISR

```

```

          R      C  R  E
consensus.....*RN**AA**R*R-*R***
          L      S  K  Q

```

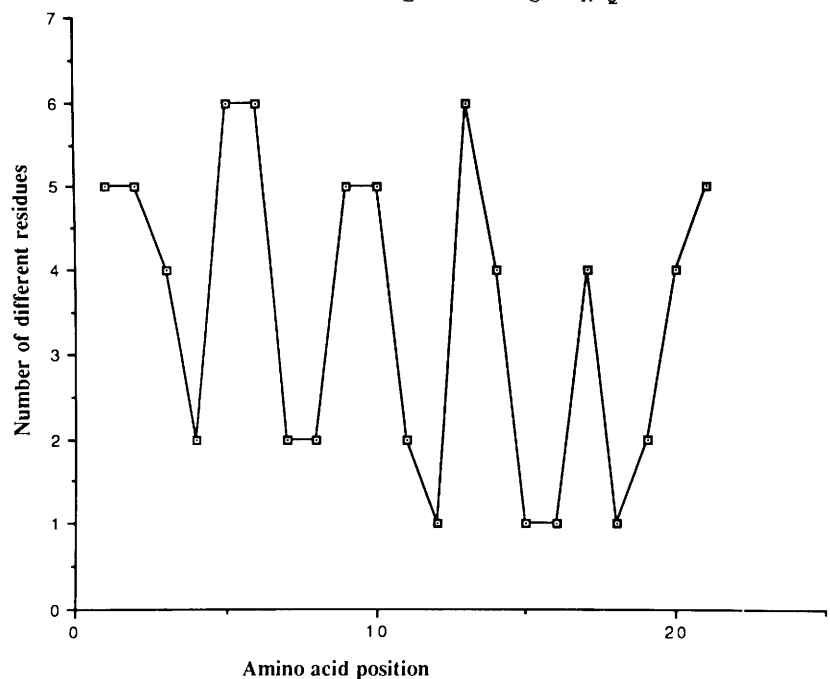


Figure 6. Basic domain alignments for MLP-binding proteins. Amino acid residues from the basic domains of the MLP-binding (or likely binding) proteins TFEB (Carr and Sharp 1990), TFE3 (Beckmann et al. 1990), Max (Blackwood and Eisenman 1991), L-, c-, N-Myc (DePinho et al. 1987), USF (Gregor et al. 1990), and E47 (Murre et al. 1989a) were aligned, and a consensus pattern was determined. Relatively nonconserved residues are shown by asterisks (*). The sequences of AP-4 basic domain (Hu et al. 1990), as well as the c-Fos, Jun-D, and GCN4 family consensus sequence (derived from O'Neil et al. 1990), are shown in contrast. The number of different residues at each position for the MLP-binding proteins is plotted. Nadirs correspond to conserved residues in the consensus sequence.

native arms of this competition would regulate the activity of TFEB. Although interactions with the MLP site appear strong enough to dissociate all tetramer, the apparent weaker affinity for the μ E3 site could be the result of close K_d values for the protein dimer- μ E3 versus dimer-dimer competitive interactions. The tetramer may therefore serve as a repository for non-DNA-binding protein. Thus, although interactions among different b-HLH-LZ proteins appear to be highly restricted in terms of DNA binding, the existence of interactions among non-DNA-binding oligomers is an important possibility that remains to be tested.

Alignment of basic region domains has been used previously to demonstrate contacting residues in the Fos/Jun/GCN4 LZ (b-LZ) family (O'Neil et al. 1990). A similar alignment of basic regions for the family of proteins recognizing the core CACGTG sequence shows a strik-

ing similarity of pattern, although with different conserved residues (Fig. 6). Whereas the spacing of consensus residues is consistent with alignment on one face of an α -helix, it has not been proven that this face directly contacts DNA. Intermediate amino acids may also play important roles in determining properties such as binding preferences and α -helix stability/inducibility. It is interesting that E47, which binds to MLP, but more weakly than TFEB or TFE3 (see Fig. 3C), is the only protein to vary slightly at two of the consensus sites. If the conserved face does contact DNA, these varied residues may be responsible for affinity differences between E47 and other MLP-binding proteins. The pattern of conserved residues in the HLH-LZ family extends approximately three turns of an α -helix—more than adequate to interact with a prototype half-site. The "induced helical fork" model of b-LZ proteins (O'Neil et al. 1990) suggests

an important role for DNA in establishing the α -helical structure of the basic domain (O'Neil et al. 1990; Patel et al. 1990; Talanian et al. 1990; Weiss et al. 1990). Such induced α -helix could interfere with structures necessary for tetramer formation as well, either through interactions with the adjacent HLH domain or through steric limitations. Identification of structural features common to the family of proteins containing b-HLH-LZ domains will be one step toward an understanding of their complex regulatory networks.

Materials and methods

Plasmids

The EcoRI cDNA fragment from λ TFEB (Carr and Sharp 1990) was made blunt-ended and subcloned into the *Sma*I site of the vector pBS-ATG (Baldwin et al. 1990) and used as template in reactions. Constructs A Δ 120 and A Δ 265 were subcloned into pBS-ATG by using *Sph*I and *Tth*III sites of TFEB. Carboxy-terminal ends were defined by cleaving with EcoRI, *Apa*LI (CA421), *Pfl*MI (CA415), *Bsp*MI (CA394), or *Bst*XI (CA390). TFE3 was transcribed from the plasmid β G- λ 3 (Beckmann et al. 1990), which was cleaved with *Bam*HI. E47S transcripts were made from the plasmid pE47S (Murre et al. 1989a), which was cleaved with EcoRI.

Probes

MLP-binding site (core sequence, CACGTG) and double point mutant (core sequence, GAGGTG) were isolated by *Hind*III and EcoRI digestion of the respective pUC plasmids (Carr and Sharp 1990) and Klenow-filled and end-labeled with [α -³²P]dATP. μ E3 probe was isolated after cleavage with either *Bam*HI or *Xba*I and EcoRV from a construct in pBS-SK (Stratagene) containing the oligonucleotide GATCCAAGCAGGTCATGTGGCAAGGCTC cloned into its *Bam*HI site (pBS-SK- μ E3) and end-labeled as above. Unlabeled probes were identical to their labeled counterparts and were gel purified and quantitated by OD₂₆₀.

In vitro transcription/translation, EMSA, and methylation interference

In vitro transcriptions and translations were done by using T3 RNA polymerase (TFEB and E47S constructs) or T7 RNA polymerase (TFE3) and reticulocyte lysate according to the recommendations of the manufacturer (Promega). EMSA was carried out in 20 μ l with 1 μ l of reticulocyte translation product in binding buffer [5% glycerol, 100 mM KCl, 10 mM MgCl₂, 10 mM Tris (pH 7.4), 1 mM DTT] plus 2 \times 10⁴ to 5 \times 10⁴ cpm of ³²P-labeled probe and 1 μ g (for MLP) or 0.1 μ g (for μ E3) of poly[d(I-C)]. Binding was carried out at room temperature for 20 min. The native polyacrylamide gels (6%) were prepared in the TGE buffer system (25 mM Tris base, 190 mM glycine, 1 mM EDTA). Dimethyl sulfate-methylation interference analysis was performed as described (Baldwin 1991) by using single end-labeled probes.

Glycerol gradient centrifugation and gel filtration

Translated reticulocyte lysate (20 μ l) programmed by TFEB A Δ 265/CA421 was diluted 10-fold with binding buffer in the presence or absence of 10 μ g of poly[d(I-C)] and 10⁶ cpm of ³²P-labeled MLP probe. Samples were loaded onto prechilled 10–30% glycerol gradients [in binding buffer with 1 μ g of d(I-C)

per 20 μ l] and spun in a SW41 rotor at 38,000 rpm for 45 hr at 4°C. Gradients were fractionated and resolved either by addition of 5 \times 10⁴ cpm of ³²P-labeled MLP followed by EMSA or directly by native gel electrophoresis. Mobility standards were chymotrypsinogen (M_r = 25 kD, S = 2.54, R = 20.9 Å), BSA (M_r = 67 kD, S = 4.58, R = 35.5 Å), and aldolase (M_r = 158 kD, S = 7.35, R = 48.1 Å), and peaks were quantitated by OD₂₈₀. S values for the dimer and tetramer peaks were determined from the standard curves shown in Figure 5C. Gel filtration was carried out by using identically prepared reticulocyte lysate diluted into binding buffer and fractionated on an automated Superose-12 FPLC column (Pharmacia). Fractions were assayed by addition of ³²P-labeled MLP (5 \times 10⁴ cpm) and poly[d(I-C)] (1 μ g) and resolved by native gel electrophoresis. Blue dextran was included, together with those standards used for glycerol gradients. K_{av} was determined for each peak by using the formula $K_{av} = (V_e - V_o)/(V_t - V_o)$, and V_e is the peak fraction exclusion volume, V_o is the exclusion volume for blue dextran, and V_t is the column volume. The Stokes radius (R) of each oligomer form was determined from the protein standard curve of K_{av} versus log(R) (Cantor and Schimmel 1980) and then used to determine the frictional coefficient ($f = 6\pi\zeta R$), where ζ is solvent viscosity. Partial specific volume, ω , was calculated to be 0.719 based on the amino acid composition (Creighton 1984). Native molecular mass was then calculated by the equation $S = [M(1-\omega\rho)]/f$, where M is molecular mass/Avogadro's number, and ρ is solvent density.

Acknowledgments

We thank Drs. H. Beckmann and T. Kadesch for the gift of the TFE3-containing plasmid and Drs. C. Murre and D. Baltimore for the E47S plasmid. We also thank Drs. K. LeClair, A. Gil, R. Meyers, R. Clerc, M. Brown, D. Potter, T. Kristie, and J. Parvin for useful discussions and advice and Dr. Preston Hensley (SmithKline Beecham Labs) for the calculation of ω . D.F. acknowledges postdoctoral support from the Howard Hughes Medical Institute (HHMI-70401-501304). This work was supported by U.S. Public Health Service grant PO1-CA42063 and partially by Cancer Center Support (core) grant P30-CA14051 from the National Institutes of Health and by a Cooperative Agreement CDR-8803014 from the National Science Foundation and Ajinomoto Co., Inc., of Japan, to P.A.S.

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

- Baldwin, A.S. 1991. Methylation interference assay for analysis of DNA-protein interactions. In *Current protocols in molecular biology* (ed. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl), vol. 2, pp. 12.3.1–12.3.6. Greene Publishing Associates/Wiley-Interscience, New York.
- Baldwin, A.S., K.P. LeClair, H. Singh, and P.A. Sharp. 1990. A large protein containing zinc finger domains binds to related sequence elements in the enhancers of the class I major histocompatibility complex and kappa immunoglobulin genes. *Mol. Cell. Biol.* **10**: 1406–1414.
- Beckmann, H.L. and T. Kadesch. 1991. The leucine zipper of TFE3 dictates helix-loop-helix dimerization specificity. *Genes & Dev.* **5**: 1057–1066.
- Beckmann, H.L., L.K. Su, and T. Kadesch. 1990. TFE3: A helix-

Fisher et al.

- loop-helix protein that activates transcription through the immunoglobulin enhancer μ E3 motif. *Genes & Dev.* **4**: 167–179.
- Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Weintraub. 1990. The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**: 49–59.
- Blackwell, T.K., L. Kretzner, E.M. Blackwood, R.N. Eisenman, and H. Weintraub. 1990. Sequence-specific DNA binding by the c-Myc protein. *Science* **250**: 1149–1151.
- Blackwood, E.M. and R.N. Eisenman. 1991. Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **251**: 1211–1217.
- Cantor, C.R., and P.R. Schimmel. 1980. *Biophysical chemistry*. Part II: Techniques for the study of biological structure and function. p. 675. W. H. Freeman, New York.
- Carr, C.S. and P.A. Sharp. 1990. A helix-loop-helix protein related to immunoglobulin E box-binding proteins. *Mol. Cell Biol.* **10**: 4384–4388.
- Carthew, R.W., L.A. Chodosh, and P.A. Sharp. 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell* **43**: 439–448.
- Chodosh, L.A., R.W. Carthew, J.G. Morgan, G.R. Crabtree, and P.A. Sharp. 1987. The adenovirus major late transcription factor activates the rat gamma-fibrinogen promoter. *Science* **238**: 684–688.
- Creighton, T.E. 1984. *Proteins: Structures and molecular properties*. W.H. Freeman, New York.
- Dang, C.V., M. McGuire, M. Buckmire, and W.M.F. Lee. 1989. Involvement of the "leucine zipper" region in the oligomerization and transforming activity of human c-Myc protein. *Nature* **337**: 664–666.
- Davis, R.L., P.F. Cheng, A. Lassar, and H. Weintraub. 1990. The MyoD DNA-binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**: 733–746.
- DePinho, R., K.S. Hattton, A. Tesfaye, G.D. Yancopoulos, and F.W. Alt. 1987. The human *myc* gene family: Structure and activity of *L-myc* and an *L-myc* pseudogene. *Genes & Dev.* **1**: 1311–1326.
- Ellis, H.M., D.R. Spann, and J.W. Posokony. 1990. Extramacrochaete, a negative regulator of sensory organ development in *Drosophila*, defines a new class of helix-loop-helix proteins. *Cell* **61**: 27–38.
- Ephrussi, A., G.M. Church, S. Tonegawa, and W. Gilbert. 1985. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* **227**: 134–140.
- Gentz, R., F.J. Rauscher, C. Abate, and T. Curran. 1989. Parallel association of Fos and Jun leucine zippers juxtaposes DNA-binding domains. *Science* **243**: 1695–1699.
- Gregor, P.D., M. Sawadogo, and R.G. Roeder. 1990. The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes & Dev.* **4**: 1730–1740.
- Hu, Y.F., B. Luscher, A. Admon, N. Mermod, and R. Tjian. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. *Genes & Dev.* **4**: 1741–1752.
- Kouzarides, T. and E. Ziff. 1988. The role of the leucine zipper in the Fos-Jun interaction. *Nature* **336**: 646–651.
- Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**: 1759–1764.
- Licht, J.D., M.T. Gossel, J. Figge, and U.M. Hansen. 1990. *Drosophila Kruppel* protein is a transcriptional repressor. *Nature* **346**: 76–79.
- Murre, C., P.S. McCaw, and D. Baltimore. 1989a. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and Myc proteins. *Cell* **56**: 777–783.
- Murre, C.S., P.S. McCaw, H. Vassin, M. Caudy, L.Y. Jan, Y.N. Jan, C.V. Cabrera, J.N. Buskin, S. Hauschka, A.B. Lassar, H. Weintraub, and D. Baltimore. 1989b. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**: 537–544.
- O'Neil, K.T., R.H. Hoess, and W.F. DeGrado. 1990. Design of DNA-binding peptides based on the leucine zipper motif. *Science* **249**: 774–778.
- O'Shea, E.K., R. Rutkowski, and P.S. Kim. 1989. Evidence that the leucine zipper is a coiled coil. *Science* **243**: 538–542.
- Patel, L., C. Abate, and T. Curran. 1990. Altered protein conformation on DNA binding by Fos and Jun. *Nature* **347**: 572–575.
- Peterson, C.L. and K. Calame. 1989. Proteins binding to site C2 (μ E3) in the immunoglobulin heavy-chain enhancer exist in multiple oligomeric forms. *Mol. Cell Biol.* **9**: 776–786.
- Prendergast, G.C. and E.B. Ziff. 1991. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science* **251**: 186–189.
- Prendergast, G.C., D. Lawe, and E.B. Ziff. 1991. Association of Myn, the murine homolog of Max, with c-Myc stimulates methylation sensitive DNA binding and Ras cotransformation. *Cell* **65**: 395–407.
- Sassone-Corsi, P., L.J. Ransone, W.W. Lamph, and I.M. Verma. 1988. Direct interaction between Fos and Jun nuclear oncoprotein: Role of the "leucine zipper" domain. *Nature* **336**: 692–695.
- Sawadogo, M. and R.G. Roeder. 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43**: 165–175.
- Talanian, R.V., C.J. McKnight, and P.S. Kim. 1990. Sequence-specific DNA binding by a short peptide dimer. *Science* **249**: 769–771.
- Tapscott, S.J., R.L. Davis, M.J. Thayer, P.F. Cheng, H. Weintraub, and A.B. Lassar. 1988. MyoD: A nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science* **242**: 405–411.
- Turner, R. and R. Tjian. 1989. Leucine repeats and an adjacent DNA-binding domain mediate the formation of functional cFos-cJun heterodimers. *Science* **243**: 1689–1694.
- Voronova, A. and D. Baltimore. 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci.* **87**: 4722–4726.
- Weiss, M.A., T. Ellenberger, C.R. Wobbe, J.P. Lee, S.C. Harrison, and K. Struhl. 1990. Folding transition in the DNA-binding domain of GCN4 on specific binding on DNA. *Nature* **347**: 575–578.



TFEB has DNA-binding and oligomerization properties of a unique helix-loop-helix/leucine-zipper family.

D E Fisher, C S Carr, L A Parent, et al.

Genes Dev. 1991, **5**:

Access the most recent version at doi:[10.1101/gad.5.12a.2342](https://doi.org/10.1101/gad.5.12a.2342)

References

This article cites 35 articles, 21 of which can be accessed free at:
<http://genesdev.cshlp.org/content/5/12a/2342.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner for Dharmacon Reagents and Horizon. On the left, it says 'Dharmacon Reagents' with the tagline 'Custom synthesis, RNAi, and CRISPR solutions'. In the center, the text 'Infinite Reliability' is displayed in large white font. To the right, the 'horizon' logo is shown in white, with 'a PerkinElmer company' underneath. A 'More' button is visible in the bottom right corner of the banner. The background features a colorful, abstract image of what appears to be a DNA helix or a similar biological structure.