# Molecular and genetic characterization of GABPβ

Fabienne Charles de la Brousse, Edward H. Birkenmeier,[1] David S. King,[2] Lucy B. Rowe,[1] and Steven L. McKnight[3]

Tularik, Inc., South San Francisco, California 94080 USA

This report outlines three observations relating to GABPβ, a polypeptide constituent of the heterotetrameric transcription factor GABP. Evidence is presented showing that the mouse genome encodes two highly related GABPβ polypeptides, designated GABPβ1-1 and GABPβ2-1. Genomic and cDNA copies of the newly defined *Gabpb2* gene were cloned and characterized, providing the conceptually translated amino acid sequence of GABPβ2-1. The genes encoding these two proteins, as well as GABPα, were mapped to three unlinked chromosomal loci. Although physically unlinked, the patterns of expression of the three genes were strikingly concordant. Finally, the molecular basis of GABPβ dimerization was resolved. Carboxy-terminal regions of the two GABPβ polypeptides, which mediate dimerization, bear highly related primary amino acid sequences. Both sequences are free of α-helix destabilizing residues and, when displayed on idealized α-helical projections, reveal marked amphipathy. Two observations indicate that these regions adopt an α-helical conformation and intertwine as coiled–coils. First, the dimer-forming region of GABPβ2-1 can functionally replace the leucine zipper of a bZIP transcription factor. Second, a synthetic peptide corresponding to this region shows distinctive helical properties when examined by circular dichroism spectroscopy. Finally, evidence is presented showing that GABPβ1-1 and GABPβ2-1 can heterodimerize through this carboxy-terminal domain, but neither protein can heterodimerize via the dimer-forming region of the bZIP protein C/EBPβ.

[*Key Words*: GABPβ; transcription complexes; gene expression; dimerization; α-helical; coiled–coil]

An emerging theme from biochemical studies of eukaryotic gene expression is that transcription factors function as multiprotein complexes. Considerable effort has focused on the molecular interfaces involved in the formation and stabilization of transcriptional regulatory complexes. Surprisingly, small polypeptide segments of transcription factors have proven to retain functional specificity when excised from their intact translation products. Detailed studies of these domains have been particularly informative in defining how transcriptional regulatory complexes assemble.

The importance of protein–protein interactions in gene regulation has been emphasized in studies of the VP16-responsive enhancers of herpes simplex virus type 1 (HSV-1) immediate early (IE) genes. Two types of regulatory motifs are associated with IE enhancers, one characterized by the sequence 5'-TAATGARAT-3' (for review, see Goding and O'Hare 1989) and the other by tandem repeats of the purine-rich hexanucleotide 5'-CGGAAR-3' (Triezenberg et al. 1988). The TAATGARAT

motif is recognized by a protein complex consisting of the virally encoded activator protein VP16 and two host cell factors, Oct-1 and HCF (Kristie and Sharp 1990; Stern and Herr 1991; Wilson et al. 1993; for review, see Thompson and McKnight 1992). The tandem copies of the CGGAAR motif are bound by a protein complex called GA-binding protein or GABP (LaMarco and McKnight 1989; LaMarco et al. 1991). Molecular characterization of GABP revealed the involvement of two distinct polypeptides; GABPβ1, a 42-kD polypeptide containing four ankyrin repeats, and GABPα, an ETS-related DNA-binding protein. GABPα and GABPβ1 associate to form an α2β2 tetrameric complex (Thompson et al. 1991). Biochemical studies of this complex have highlighted a number of molecular interfaces critical for transcriptional activation of HSV-1 IE genes. These include (1) the surfaces involved in the heteromeric interactions between GABPα and GABPβ1 that include the ETS domain of GABPα and the ankyrin repeats of GABPβ1, (2) the molecular interfaces formed upon association of the complex with DNA, and (3) the carboxy-terminal region of GABPβ1 that specifies its homotypic dimerization (Thompson et al. 1991).

Studies focused on the mechanisms of dimerization have become the subject of extensive work since the

Present addresses: [1]Jackson Laboratory, Bar Harbor, Maine 04609 USA; [2]Howard Hughes Medical Institute, Department of Molecular and Cellular Biology, University of California, Berkeley, Berkeley, California 94720 USA.
[3]Corresponding author.

realization that many transcription factors function as obligate dimers. For example, the CCAAT/enhancer binding protein (C/EBP) family of transcription factors consists of at least four members that have the inherent ability to form productive homo- and heterodimeric complexes (Cao et al. 1991; Williams et al. 1991). The C/EBP proteins belong to the basic leucine zipper (bZIP) family of proteins. The bZIP domain represents a bipartite DNA-binding motif consisting of an α-helical dimerization surface (Landschulz et al. 1988; Gentz et al. 1989; Kouzarides and Ziff 1989; Turner and Tjian 1989; Vinson et al. 1989) and a basic region that directly contacts DNA (Agre et al. 1989; Talanian et al. 1990). The leucine zipper of bZIP transcription factors assumes a coiled–coil conformation (O'Shea et al. 1989, 1991; Ellenberger et al. 1992).

Coiled–coil dimers consist of amphipathic α-helices that assemble in a parallel orientation. The amphipathic character is specified by a primary amino acid sequence showing a heptapeptide repeat (residue positions a–g) wherein hydrophobic residues are located at positions a and d and polar residues at the other positions (Crick 1953; for review, see Cohen and Parry 1990). When folded into an α-helical conformation, the hydrophobic and polar residues align on opposite sides of the helix. The helices dimerize in a parallel orientation through a combination of attractive hydrophobic and ionic interactions. Hydrophobic interactions increase on interhelical packing of residues at the a and d positions (O'Shea et al. 1991; Ellenberger et al. 1992; Harbury et al. 1993). Interhelical ion pairs contribute to stabilization and dimerization specificity (O'Shea et al. 1992; Vinson et al. 1993).

For bZIP proteins, there is a characteristic repeat of leucine residues at the d position of the coiled–coil dimerization interface (Landschulz et al. 1989; for review, see Johnson and McKnight 1989). Alignment of various bZIP proteins reveals an invariant spacing of exactly seven residues, or two helical turns, between the basic region and leucine zipper (Vinson et al. 1989). This observation suggests that DNA binding might be dependent on the topological disposition of the DNA contact surface relative to the dimer interface. Consistent with this prediction, analyses of C/EBP and GCN4 proteins that bear either insertions or deletions of residues between the basic region and leucine zipper fail to bind to DNA (Agre et al. 1989; Pu and Struhl 1991).

By defining the molecular interfaces of transcription factors, it has been possible to rationalize how related proteins assemble into multiprotein complexes. Here, we describe the cloning and characterization of a new member of the GABPβ family of proteins, termed GABPβ2-1. The amino acid sequence deduced from the mouse GABPβ2-1 cDNA reveals that its encoded protein is the product of a distinct gene that shares extensive homology with GABPβ1-1 (originally called GABPβ1; LaMarco et al. 1991). The tissue distribution of GABPβ2-1 is highly concordant to that of GABPβ1-1 and GABPα. Moreover, GABPβ2-1, like GABPβ1-1, interacts avidly and selectively with GABPα. Given that GABPβ1-1and GABPβ2-1 share highly related carboxy-terminal dimerization domains, we have investigated the molec-

ular basis of dimerization. Our results indicate that dimerization occurs via the interaction of amphipathic α-helical domains in a manner similar to the "zippering" of bZIP proteins. We show further that recombinant GABPβ1-1 and GABPβ2-1 proteins readily form heterodimers in solution.

## Results

### Isolation of Gabpb1 and Gabpb2 genes

Southern blot analyses to map the chromosomal location of the encoding gene for GABPβ1-1, termed Gabpb1, revealed the presence of related genes (see Chromosomal location of the genes encoding GABPα, GABPβ1-1, and GABPβ2-1 section of Results). Previous work had shown that the mouse Gabpb1 gene is alternatively spliced to give rise to the GABPβ1 and GABPβ2 proteins (LaMarco et al. 1991). To distinguish this gene from the related Gabpb2 gene described here, we refer to the original GABPβ1 and GABPβ2 isoforms as GABPβ1-1 and GABPβ1-2, respectively. This new nomenclature takes into account the gene number and protein isoform number. To clone Gabpb-related genes, a library was constructed from mouse strain 129svJ genomic DNA and screened with two hybridization probes. One probe was prepared from sequences encoding the ankyrin repeats of GABPβ1-1, and the other corresponded to sequences encoding its carboxy-terminal dimerization domain (LaMarco et al. 1991). Two different genomic clones, termed Gabpb1 and Gabpb2, were identified in this screen. Restriction mapping and partial sequencing analyses confirmed the presence of two Gabpb genes in the mouse: Gabpb1 encodes the original GABPβ1-1 and GABPβ1-2 isoforms (LaMarco et al. 1991) while Gabpb2 represents a novel gene.

cDNA clones corresponding to Gabpb2 were cloned from a mouse lung cDNA library (Materials and methods). Four different recombinant phage were identified, isolated, and analyzed by DNA sequencing. These studies revealed that one clone contained a long uninterrupted open reading frame beginning with an AUG near its 5′ terminus and terminating at a UAA stop codon. This GABPβ-related cDNA clone, encoding GABPβ2-1, comprised 1363 nucleotides capable of encoding a polypeptide of 414 residues (Fig. 1). Assignment of the translation initiation codon for GABPβ2-1 was based on sequence similarity with the GABPβ1-1 cDNA, as well as the presence of multiple stop codons in all three reading frames preceding the assigned AUG. As depicted in Figure 1, the conceptually translated GABPβ2-1 polypeptide is highly similar to GABPβ1-1. The proteins show 87% identity within their ankyrin repeats (residues 1–130) and 70% identity within their putative carboxy-terminal dimerization domains (residues 334–382 for GABPβ1-1 and residues 317–366 for GABPβ2-1).

### Tissue distribution of the GABPβ proteins and evidence of their affinity for GABPα

The ankyrin repeats of GABPβ1-1 are sufficient for avid

```
GABPβ2-1  MSLVDLGKKL LEAARKGQDD EVRILMANGA PFTTDWLGTS   40
GABPβ1-1  MSLVDLGKKL LEAARAGQDD EVRILMANGA PFTTDWLGTS   40

          PLHLAAQYGH MSTHEVLLRA GVSRDARTKV DRTPLHMAAA   80
          PLHLAAQYGH HSTHEVLLRA GVSRDARTKV DRTPLHMAAS   80

          LGHVHIVELL VRSGADVNAK DMLQMTALHW ATEHHHRIVV  120
          EGHANIVEML LKHGADVNAK DMLKMTALHW ATEHNHQEVV  120

          ELLIKYGADV YAHSKFIKSA FDIAMEKNNT EILVMLQHAM  160
          ELLIKYGADV HTQSKFIKIA FDISIDNGNE DLAEILQHAM  160

          QNQMNTNHER ANEVANPVTV TAPHIFTSGE VINLASFVSS  200
          QNQINTNFES PDTVTIHAAT PQFIIGPQGV VNLTDETGVS  200

          ANTKATSAHL ---------- ---------- ----EEMEE  215
          AVQFGNSSTS VLATLAALAE ASAPLSNSSE TPVATEEVVT  240

          GNSLDSSTQQ VVDSGGQHVI TIVTDGVHLG NIQTSLPACG  255
          AESVDGAIQQ VVDSGGQQVI TIVTDGIQLG NLH-SHPTSG  279

          IGQPFIVTMQ DGQQVLTVPA GQVAEETHIE IEEEEEEKLP  295
          MGQPHIVTMP DGQQVLTVPA TDIAEETVIS EE------P  312

          LVKRPRMAEM TNRVEEMKEG SEREILQLQL HEANRHAQHY  335
          PAKRDQMEII ESRVHETAET EEREALQKQL HEANRHAQKY  351

          RHQLLKKEQE AELYRLRLEA MAQQQTNGVE VDVTVVEEVA  375
          RQQLLKKEQE AEHYRDKLEA MTRIQTNKEA V--------  382

          EVDAVVVTEG DEVERATQVM KSGRTTEPHT NVSIETISS  414
          ---------- ---------- ---------- ---------  382
```

**Figure 1.** Amino acid sequence comparison of GABPβ2-1 and GABPβ1-1. The predicted amino acid sequences are displayed with gaps allowing optimized alignment. Identical residues are boxed. The amino-terminal ankyrin repeats (residues 1–130) and the dimerization domains of GABPβ1-1 (residues 334–382) and GABPβ2-1 (residues 317–366) show 87% and 70% identity, respectively. Located on the amino-terminal side of the dimerization domains are small regions rich in acidic residues as well as regions showing a disproportionately high frequency of glutamine residues. These latter regions have been implicated in the transcriptional activation function of the GABP complex (C. Thompson, unpubl.).

and specific interaction with GABPα (Thompson et al. 1991). Because GABPβ1-1 and GABPβ2-1 share 87% identity in the respective ankyrin repeats, we imagined that GABPβ2-1 might also bind GABPα. Likewise, the high level of amino acid sequence similarity between carboxy-terminal segments of GABPβ1-1 and GABPβ2-1 led to the prediction that GABPβ2-1 might be capable of forming homodimeric, and together with GABPβ1-1, heterodimeric complexes. To begin to address the biological relevance of closely related GABPβ proteins, the tissue distributions of the GABPβ1-1 and GABPβ2-1 proteins were compared by Western blotting. Whole cell extracts from mouse tissues were electrophoresed on SDS–polyacrylamide gels and transferred to Immobilon membranes (Fig. 2A). Figure 2B shows immunostaining patterns using antisera specific to GABPβ1-1 (top) and GABPβ2-1 (bottom). In both cases, the major immunoreactive protein corresponded to the size of the primary translation product. Aberrantly sized, immunoreactive

bands were observed in both Western blots, some of which likely reflect the products of alternatively spliced GABPβ mRNAs (LaMarco et al. 1991; Virbasius et al. 1993; Watanabe et al. 1993; F.C. de la Brousse and S.L. McKnight, unpubl.). These analyses revealed that the tissue distributions of GABPβ1-1 and GABPβ2-1 were largely concordant, with the highest levels detected in thymus, spleen, kidney, and intestine. It is important to note that the high levels of GABPβ1-1 and GABPβ2-1 in thymus and spleen may simply reflect the relatively high ratio of nucleus to cytoplasm typical of these tissues. Taking this variable into account, the data provided in Figure 2 may be taken to indicate a relative constancy of GABPβ1-1 and GABPβ2-2 abundance in most tissues that were examined.

Previous studies have shown that GABPβ1-1 must obligately associate with GABPα to bind to DNA (for review, see Thompson and McKnight 1992; also see Watanabe et al. 1990; Virbasius et al. 1993). To determine whether GABPβ2-1 also interacts selectively with GABPα, radiolabeled GABPβ1-1 and GABPβ2-1 were used to probe Western blots of whole tissue extracts. As shown in Figure 2C, predominantly one protein band of ~58 kD was detected with both probes. Both the size and tissue distribution of this protein matched those characteristics of endogenous GABPα (Brown and McKnight 1992). In addition, our analyses also revealed that *Escherichia coli*-expressed GABPα reacted with similar avidity and specificity with the two GABPβ probes. Finally, and as shown previously for GABPβ1-1 and GABPα (Thompson et al. 1991), mobility retardation assays using recombinant GABPα and GABPβ2-1 proteins indicated that a region including the ETS domain of GABPα and a region including the ankyrin repeats (residues 1–157) of GABPβ2-1 are necessary and sufficient for strong binding to the CGGAAR motif (data not shown).

*Chromosomal location of the genes encoding GABPα, GABPβ1-1, and GABPβ2-1*

*GABPα* A PCR fragment that encompassed the entire insert of 2.1 kb from the GABPα cDNA was labeled by random priming and used as a probe in Southern blots (Taylor and Rowe 1989). This probe hybridized to four to six bands in each restriction enzyme digest of mouse genomic DNA, and there were few differences between standard inbred mouse strains. *Sst*I digests gave three dark bands (4.5, 4.0, and 3.5 kb in C57BL/6J and AKR/J; 4.5, 4.1, and 3.5 kb in DBA/2J) and three lighter bands (4.3, 2.4, and 2.2 kb, nonpolymorphic) when hybridized to this probe. We typed the 4.0/4.1-kb *Sst*I polymorphism in the BXD and AKXD RI strain sets, and the data are shown in Table 1. This polymorphism maps to chromosome 16 with 1 recombinant (BXD-2) of 49 RI strains typed for *App* and *Gabpa*. *Msp*I digests gave six or seven bands among the progenitors of the RI strains, many of which were polymorphic. C57BL/6J and SWR/J share an identical *Msp*I band pattern with bands at 7.0, 6.5, 6.2, 3.5, 2.0, and 0.7 kb. In contrast, BALB/cByJ and C57/L
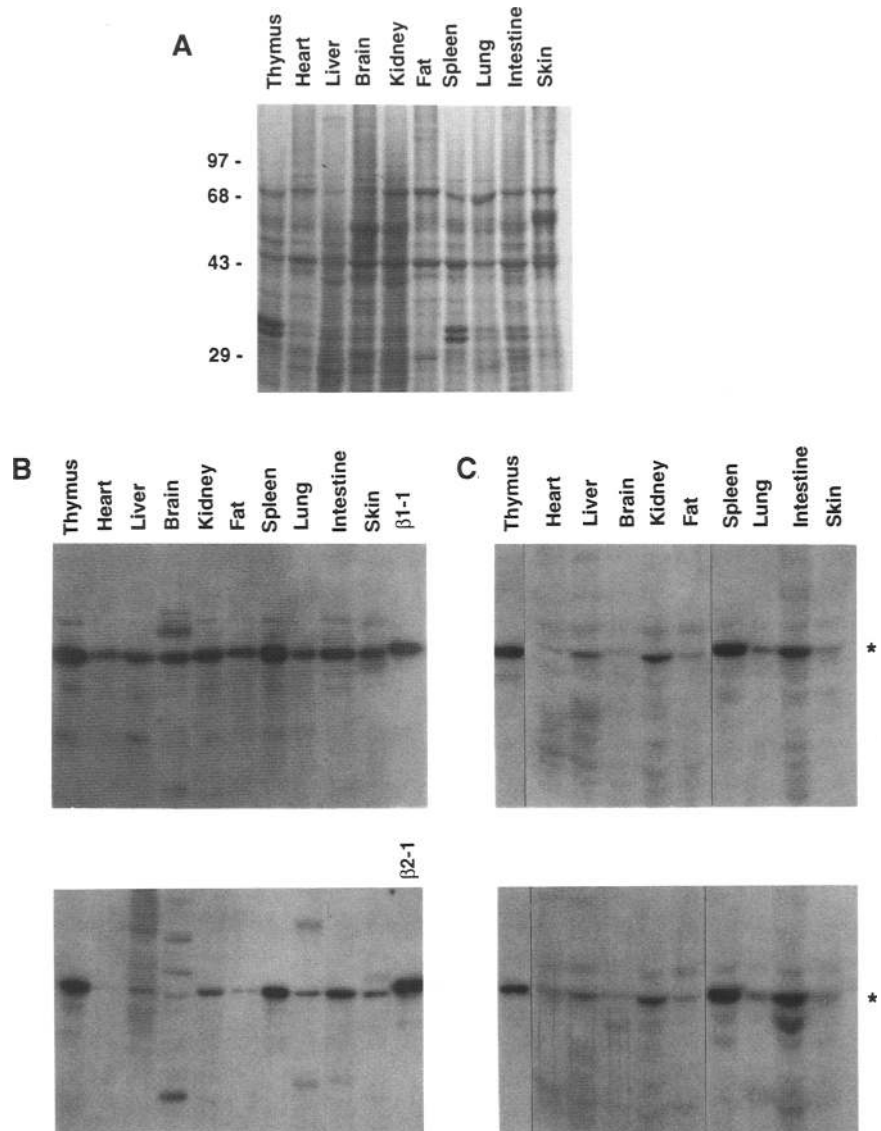
**Figure 2.** Tissue distribution of the GABPβ proteins and evidence of their affinity for GABPα. (A) Coomassie-stained SDS–polyacrylamide gel of proteins extracted from various mouse tissues. (B) Western blots probed with rabbit antisera raised against GABPβ1-1 and GABPβ2-1 recombinant proteins. Equal amounts of proteins were loaded as in A, transferred to Immobilon membranes, and subjected to immunochemical staining with polyclonal antiserum to GABPβ1-1 (top) and GABPβ2-1 (bottom). Two nanograms of recombinant GABPβ1-1 and GABPβ2-1 proteins were used as controls (far right lanes). In C, proteins transferred to Immobilon were probed with radiolabeled GABPβ1-1 (top) and GABPβ2-1 (bottom). Migration of purified recombinant GABPα is denoted by an asterisk (*).

have 7.0-, 6.5-, 6.2-, 3.5-, 3.3-, and 0.7-kb bands. In the SWXL and CXB RI strain sets, the two polymorphic bands (3.3 and 2.0 kb) appeared allelic and showed linkage to chromosome 16 markers. Among the SWXL RI strains, there was one recombinant (SWXL-14) of the seven strains typed (1/7) for Pmv35, and no recombinants (0/7) with Pmv14. CXB-I was the sole recombinant (1/7) for PMV35, and CXB-D was the sole recombinant (1/7) for Pmv16 among the seven CXB RI strains tested for these markers (Frankel et al. 1989).

MspI digests were also analyzed on the testcross (MEV×CAST/Ei)×BXD-32, where MEV and BXD-32 share the C57BL/6J band pattern, and CAST/Ei had bands of 6.5, 6.2, 4.3, 3.5, 3.3, 3.0, and 0.7 kb. The three CAST/Ei bands (4.3, 3.3, and 3.0 kb) cosegregated. The two MEV bands (7.0 and 2.0 kb) also appeared to cosegregate on the basis of band intensities suggesting that they were present in only one copy when the CAST/Ei

bands were present. All of these polymorphic bands showed linkage to Igl (17/89 recombinants), ApoD (16/89), and Pitl (12/89) on chromosome 16 (Warden et al. 1992). From the testcross data the gene order was ambiguous because the order (centromere, Igl, ApoD, Pitl, Gapba) resulted in six apparent double crossovers, and the order (centromere, Igl, ApoD, Gapba, Pitl) gave a different six apparent doubles. Nevertheless, the location of this gene to chromosome 16 was well established by these data, and it appears that most of the bands hybridizing to this probe originate from a single chromosomal locus that maps between App and D16Mit6. We conclude that Gabpa maps to a single locus within 3.2 cM (95% upper confidence limit) of App on mouse chromosome 16.

GABPβ An insert from a cDNA clone encoding GABPβ1-1 was amplified by PCR with flanking primers,

**Table 1.** *Strain distribution patterns among the RI strains for loci linked to the GABP gene family*

| | | RI strain | | |
|---|---|---|---|---|
| Chr | Locus | AKXD | BXD | BXH |
| | | 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2<br>1 2 3 6 7 8 9 0 1 2 3 4 5 6 7 8 0 1 2 3 4 5 6 7 8 | 0 0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 3 3 3<br>1 2 5 6 8 9 1 2 3 4 5 6 8 9 0 1 2 3 4 5 7 8 9 0 1 2 | 0 0 0 0 0 0 1 1 1 1 1<br>2 3 4 6 7 8 9 0 1 2 4 9 |
| 2 | 82m^a | | B B B B D B D D B B D D D D D D B D D B D D D D B D | B B B B B H B H B H H B |
| 2 | Hdc^b | | B B B B D D B D B B D D D D D D B D D B D D D D B B | B B H B B H B H B H H B |
| 2 | Gabpb1 | D . D A D A A D D A D A D A D A D D A D A D A D A A A | B B B B D D B D B B D D D D D D B D D B D D D D B B | B B H B B H B H B H H B |
| 2 | II1a^c | D D D A D A A D D A D A D A D A D D A D A D A D A A A | B B B B D D B D B B D D D D D . . B D D B D D D D B | B B H H B H B H B H H B |
| 4 | Pmv19^d | | D B D B B D B D B D D B D D B D D D B D D D D D B D | H B H B B B H H H B B H |
| 4 | Glut1^e | D D D D D A D A D A D . D D . D D A A A D D D D D | D B D B B D B D B D B D D D D D B D D D B D D D D B D | H B H H B B H B H B B H |
| 4 | Gabpb1-ps | D . D D D D A D A D A D D D D D D D A A D A D D D D | D B D B B D B D B D D D D D D D B D D D B D D D D B D | H B H H B B H B H B B H |
| 4 | Ckb-rs1^f | D D D D A D D A D A D A D D D D D D D A A D A D D D D | D B D B D D B D D D D D D D D D B D B D B D D D D B D | H B H H B B H B H B B H |
| 16 | Pmv14^d | D D D A D D A D A A D D A A D A A D A A D A D D A | B D B B D B D D D B D D D D B D D B B D D D D B D B B D | |
| 16 | App^g | D D A . D D A D A A D A A A D A A D A A D A D D A | B D D B B B D B B D D D B D D D B D D D B D D D B D B B D | |
| 16 | Gabpa | D D A A D D . D A A D A A A D A A D A A D A D D A | B B D B B B D B D B D D D B D D D B D D D B D B B D | |
| 16 | D16Mit6^h | | . B B B B D D B B D D B B D D B B B B D D D B D . . . | |

^aData from D'Eustachio et al. (1987).
^bData from Frankel et al. (1989).
^cData from Elliot et al. (1992).
^dData from Cho et al. (1991).
^eData from Cheng et al. (1988).
^fData from Chorney et al. (1982).
^gData from Martin et al. (1984).
^hData from Dietrich et al. (1992).

and the resulting fragment was used for random hexamer labeling and probing Southern blots of mouse genomic DNAs. A survey of restriction enzymes showed one to four bands per lane, with moderate polymorphism among standard inbred strains. The enzyme *Bst*EII was chosen for segregation analysis because four of the five hybridizing bands were polymorphic. C57BL/6J and AKR/J had bands of 13, 9.4, 6.9, and 2.8 kb, while C3H/HeJ and DBA/2J had 17-, 10-, 6.7-, 3.8-, and 2.8-kb bands. The data for the AKXD, BXD RI strain sets tested with this probe are given in Table 1. The 6.7-kb (C57BL/6J and AKR/J) to 6.9-kb (C3H/HeJ and DBA/2J) bands mapped to chromosome 4, and because of their relatively lighter intensity on the autoradiographs, we assume that they represent a pseudogene, *Gabpb1-ps*. All of the remaining polymorphic bands cosegregated in the 62 RI strains typed. This locus, *Gabpb1*, maps to chromosome 2 with 2/59 recombinants with *Il1a* and 0/38 (95% confidence limits 0–2.69 cM) with *Hdc*.

As described earlier in this report, a second cDNA clone was isolated that encodes a protein highly similar to GABPβ1-1. We amplified a 300-bp fragment of the GABPβ2-1 cDNA for use as a probe in Southern blots. We surveyed C57BL/6J and SPRET/Ei DNAs digested with various restriction enzymes, and detected one prominent band in each digest, with low frequency of polymorphism. We screened our (C57BL/6J×SPRET/Ei)F1×SPRET/Ei backcross panel DNAs (Rowe et al. 1994) digested with *Hin*dIII (C57/BL6J band 6.7 kb; SPRET/Ei band 6.8 kb) and mapped this polymorphism to the mid-region of chromosome 3 (Fig. 3). This map position is not the same as the previously identified *Gabpb1* or *Gabpb1-ps*, so this locus was named *Gabpb2*. It thus appears that there are at least three unlinked loci of related sequence to GABPβ.

## The carboxy-terminal dimerization domains of GABPβ1-1 and GABPβ2-1 function as coiled–coil α-helices

Previous gel filtration, sedimentation, and cross-linking studies have shown that purified GABPβ1-1 associates as a homodimer. Thompson et al. (1991) identified a region close to the carboxyl terminus of GABPβ1-1 that facilitates homotypic dimerization. Dimerization of GABPβ1-1 was shown to be independent of its association with GABPα or the presence of DNA. This carboxy-terminal domain, which is encoded by an alternatively spliced exon, is absent in the GABPβ1-2 protein (Thompson et al. 1991). The GABPβ2-1 mRNA, like GABPβ1-1, contains a similarly spliced exon.

The dimerization domains of the GABPβ1-1 and GABPβ2-1 proteins have very similar amino acid sequences. Because both sequences are free of α-helix destabilizing residues (proline and glycine), we examined these sequences as idealized α-helices. A striking array of conserved hydrophobic residues was observed in the four/three repeat pattern typical of proteins that associate as coiled–coils. Figure 4 compares the leucine zipper
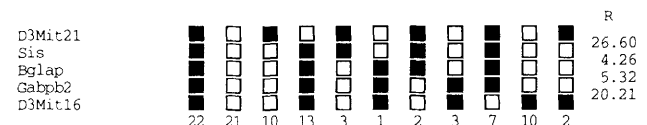


**Figure 3.** Haplotype figure for key markers on chromosome 3 in the cross (C57BL/6J×SPRET/Ei)F1×SPRET/Ei. (■) The CX57BL/6J allele; (□) the SPRET/Ei allele. More proximal markers are listed at the *top*. The number below each column of boxes indicates the number of N2 animals with that haplotype. The percent recombination between adjacent listed loci is given at *right*.
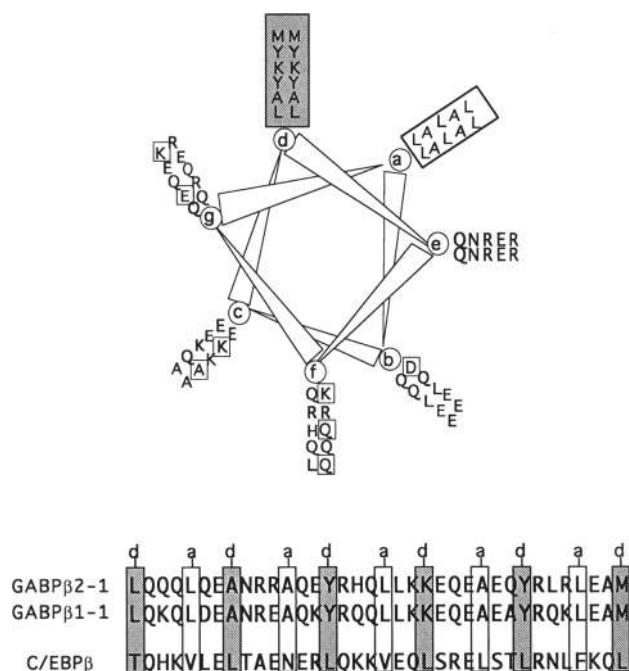
GABPβ2-1  L QQQL QE ANRR AQE YRHQ LLKKEQE AE QYRLR LEAM

GABPβ1-1  L QKQL DE ANRE AQKY RQQL LLKKEQE AE AYRQK LEAM

C/EBPβ  T QHKVL EL TAEN ERLQKKVEQ LSRE LSTL RNL FKQL

**Figure 4.** The dimerization domains of GABPβ1-1 and GABPβ2-1 projected onto idealized α-helices. The view down the helical axis begins with leucine residue 321 of GABPβ2-1 and leucine residue 337 of GABPβ1-1. This projection reveals an amphipathic distribution of hydrophilic and hydrophobic residues. Residues at positions b, c, e, f, and g that diverge between the deduced protein sequences are boxed on the GABPβ1-1 sequence. (*Bottom*) A comparison of the amino acid sequences of the GABPβ1-1 and GABPβ2-1 dimerization domains with the leucine zipper sequence of C/EBPβ (Descombes et al. 1990; Cao et al. 1991; Williams et al. 1991).

of C/EBPβ (Cao et al. 1991) with the hypothetical dimerization regions of GABPβ1-1 and GABPβ2-1. To facilitate visualization of the four/three repeat, residues at positions **a** and **d** are boxed. The view down the α-helical axis begins at the **d** position and reveals an arrangement wherein hydrophilic residues occur on one helical face and hydrophobic residues on the other face. This polarized distribution of hydrophilic and hydrophobic residues suggests that the dimerization domains of GABPβ1-1 and GABPβ2-1 function by forming coiled-coils.

As a first means of testing this hypothesis, we examined whether C/EBPβ could function in DNA-binding assays with its leucine zipper exchanged for the dimerization domain of GABPβ2-1. A truncated C/EBPβ protein, termed C1i0C1, was expressed in *E. coli* from a cDNA that was engineered to contain an *XhoI* restriction site in the spacer region between the basic region and leucine zipper domains (Fig. 5). This type of *XhoI* recombinant, which was used by Agre et al. (1989) to analyze the specificity of DNA binding by bZIP proteins, provided the means to create a hybrid protein, termed C1i0β2, that contained the basic region of C/EBPβ fused to the dimerization domain of GABPβ2-1. The C1i0C1

and C1i0b2 proteins were expressed in bacterial cells, partially purified by chromatography, and tested for sequence-specific DNA binding in DNase I footprinting assays. The DNA probe used in these studies was a 170-bp fragment from the HSV thymidine kinase (tk) promoter modified to contain a high affinity C/EBP DNA-binding site (Vinson et al. 1989). As seen in Figure 5B, both C1i0C1 and C1i0b2 protected the C/EBP recognition sequence from digestion by DNase I. bZIP proteins, such as C/EBPβ, require a dimerization domain capable of orienting two polypeptide chains in parallel (Gentz et
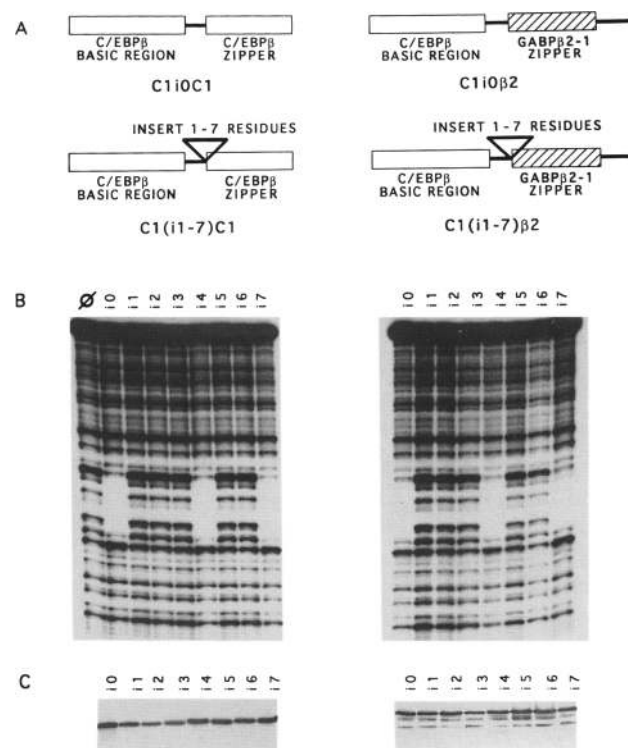


**Figure 5.** The leucine zipper of C/EBPβ can be functionally replaced by the dimerization domain of GABPβ2-1. (*A*) Schematic diagram of C/EBPβ protein and insertion variants (*left*) and chimeric C/EBP–GABP fusions (*right*). For C1i0C1, the segment of C/EBPβ corresponding to the basic region was linked to the leucine zipper via an *XhoI* restriction enzyme recognition site. The chimeric C1i0β2 protein contains the C/EBPβ basic region (C1) fused to the GABPβ2-1 dimerization domain (β2). The spatial register between the basic region and dimerization domains of C1i0C1 and C1i0β2 is identical to native C/EBPβ. Lower diagrams depict C/EBPβ and C/EBPβ–GABPβ2-1 chimeric proteins containing one to seven residues inserted between the basic region and the dimerization domain. (*B*) DNase I footprinting assays of parental and insertion mutants of C1i0-i7C1 and C1i0-i7β2. The DNA probe consists of an HSV *tk* promoter fragment modified to contain a single high affinity C/EBP DNA-binding site. End-labeled probe DNA was incubated with equimolar concentrations of C1i0-i7C1 (10 mM) and C1i0-i7β2 (30 mM) recombinant proteins, digested with DNase I, and sized on sequencing gels. (*C*) Western blots to monitor amount of C1i0-i7C1 (*left*) and chimeric C1i0-i7β2 proteins (*right*) used in the DNase I footprinting assays.

al. 1989; Landschulz et al. 1989; O'Shea et al. 1989, 1991; Turner and Tjian 1989; Oas et al. 1990; Rasmussen et al. 1991); the data shown in Figure 5B are consistent with the interpretation that the GABPβ dimerization domains assemble in parallel as well.

Sequence analyses of bZIP protein domains have revealed an invariant spacing of seven residues between the basic region and the first leucine of the zipper (Vinson et al. 1989; O'Neil et al. 1990; for review, see McKnight 1991). Mutagenesis studies of C/EBPα and GCN4 have also shown the importance of correct spacing between the basic region and the leucine zipper (Agre et al. 1989; Pu and Struhl 1991). Moreover, X-ray crystallographic studies of the GCN4 bZIP region (Ellenberger et al. 1992) have demonstrated that α-helical structure persists between the zipper and the basic region, thus offering a logical explanation for the highly conserved spacing between the two domains.

If the dimerization domain of GABPβ2-1 functions as a coiled–coil like the leucine zipper of bZIP proteins, it should function only when properly aligned with the basic region. We investigated this spacing dependency by systematically expanding, from one to seven residues, the spacer regions of C1i0C1 and C1i0β2 (Fig. 5A) and comparing their sequence-specific DNA-binding abilities. To ensure that approximately equal amounts of protein were used in each DNA-binding assay, the various proteins were run on SDS–polyacrylamide gels and quantitated by Coomassie staining and Western blotting (Fig. 5C). As shown in Figure 5B, only C/EBPβ variants C1i4C1 and C1i7C1 were capable of binding to the C/EBP recognition site. Variants containing inserts of one, two, three, five, or six residues between the basic and leucine zipper failed to bind to the target DNA. An identical profile of DNA binding was found for the C1β2 insertion mutants (Fig. 5B, right) further revealing the functional similarities between the C/EBPβ leucine zipper and the GABPβ dimerization domain. When tested in glutaraldehyde cross-linking assays, all of the insertion mutants behaved as competent dimers (data not shown). As such, we conclude that the inactive variants fail to bind DNA because of an incorrect alignment of the basic region relative to the dimer interface.

The identical behavior of the two sets of insertion mutants is best explained by a structural relatedness between the C/EBPβ and GABPβ dimerization domains. To investigate this possibility further, a peptide corresponding to the GABPβ2-1 dimerization domain (residues 317–360) was synthesized, purified, and assessed for helical content by measuring its circular dichroism (CD) spectrum under various conditions. The CD spectrum (Fig. 6A) showed the strong negative bands at 208 and 222 nm and a positive band at 195 nm characteristic of α-helices (Chen et al. 1974; Woody 1985). The comparative intensity at λ222 nm in PBS relative to 1:1 PBS/ 2,2,2-trifluoroethanol (TFE) indicated α-helical contents of ~ 84% (5 μM at 0°C) to 88% (20 μM at 0°C). Thermal melting experiments (Fig. 6B) showed that the α-helical state adopted by the GABPβ2-1 peptide was moderately stable, with the midpoint of temperature dependence
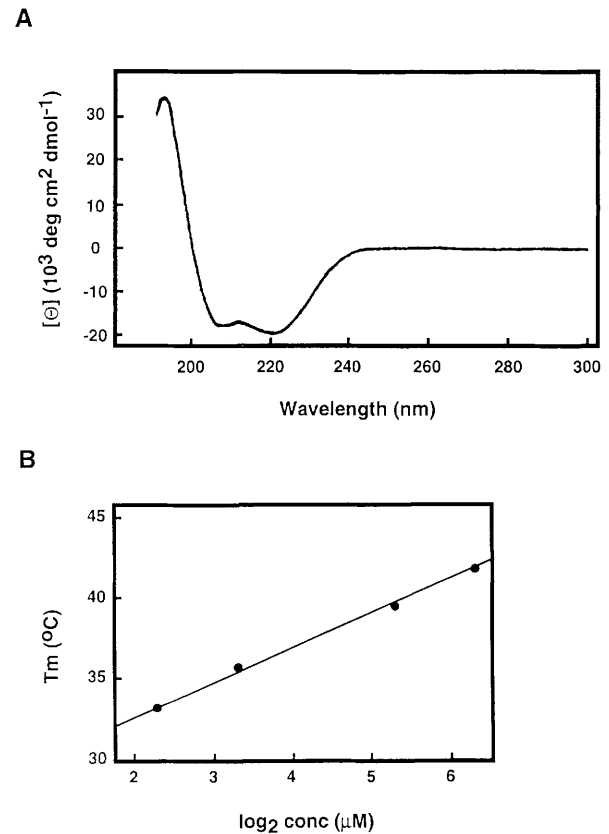
Figure 6. Circular dichroism spectra of a synthetic peptide corresponding to the dimerization domain of GABPβ2-1. (A) CD spectrum at 0°C of the peptide at a concentration of 10 mM in PBS prepared with potassium phosphate and 150 mM KF in a 1 cm path length cell. The amino acid sequence of the synthetic peptide corresponds to residues 317–360 of GABPβ2-1. (B) Thermal unfolding measurements of the peptide as a function of peptide concentration. The Tm increases linearly with $\log_2$ peptide concentration between 5 and 40 μM.

(Tm) at λ222 nm ranging from 33°C to 43°C for concentrations of 5–40 μM (reversible transition). Both the molar ellipticity and the concentration-dependent Tm increase could be interpreted in accordance with a monomer–dimer equilibrium (O'Neil and deGrado 1990).

## Specificity of the GABPβ1-1 and GABPβ2-1 dimerization domains

Sequence similarities within the dimerization domains of GABPβ1-1 and GABPβ2-1, coupled with the similar expression patterns of these proteins (Fig. 2B), led us to ask whether they might cross-dimerize. Partially purified recombinant proteins composed of the C/EBPβ basic region, fused to either the leucine zipper of C/EBPβ (C1C1) or the dimerization domains of GABPβ1-1 (C1β1) or GABPβ2-1 (C1β2), were tested for oligomerization in a chemical cross-linking assay. As seen in Figure 7A, incubation of the individual proteins in the presence of glutaraldehyde resulted in the appearance of an addi-
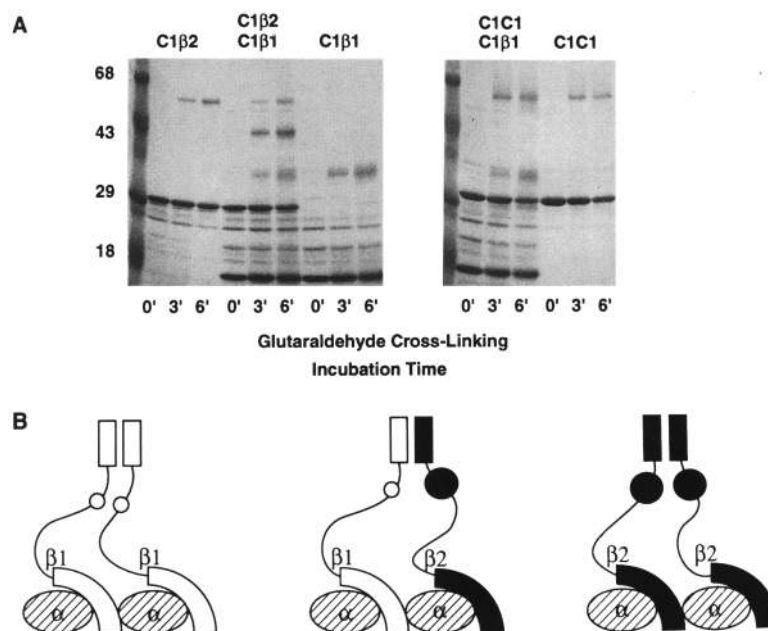
**Figure 7.** (A) Glutaraldehyde cross-linking of the C/EBPβ (C1C1) and C/EBPβ–GABPβ chimeric proteins (C1β1 and C1β2). Bacterially expressed proteins were incubated with 0.005% glutaraldehyde at room temperature for the time indicated (min). Reactions were stopped by adding an equal volume of 2× SDS sample buffer, boiled for 3 min and subjected to SDS-PAGE. The gel was stained with Coomassie brilliant blue. (B) Model illustrating how GABP complexes containing either a GABPβ1-1 (β1) homodimer, a GABPβ2-1 (β2) homodimer, or a GABPβ1-1/GABPβ2-1 (β1/β2) heterodimer might bind different GABP response elements. Parallel rectangles represent coiled–coil dimerization domains of GABPβ subunits. Open and solid circles represent GABPβ activation domains. Curved regions of GABPβ subunits represent ankyrin repeats of GABPβ that mediate contact with GABPα (hatched ellipses).

tional protein band roughly twice the size of the monomeric species. Heterodimerization between the various polypeptides was tested by this same assay. When the chimeric C1β1 and C1β2 proteins were mixed, three glutaraldehyde-induced species were observed. The upper and lower of these cross-linked products are C1β2 and C1β1 homodimers, respectively, whereas the intermediate band was interpreted as a heterodimer (Fig. 7A, middle lanes). When the C/EBPβ protein fragment (C1C1) was incubated with C1β1 and exposed to glutaraldehyde, no heterodimeric product was observed (Fig. 7A, right). Thus, whereas GABPβ1-1 and GABPβ2-1 proteins can form homodimers, they do not appear to be capable of dimerizing with the leucine zipper of C/EBPβ.

## Discussion

The studies reported herein identify a novel gene that encodes a protein termed GABPβ2-1. The deduced amino acid sequence of GABPβ2-1 is highly related to that of GABPβ1-1 (GABPβ1; LaMarco et al. 1991), with most striking similarity localized to the amino-terminal ankyrin repeats and the carboxy-terminal dimerization domain (Thompson et al. 1991). Previous studies addressing the molecular and biochemical properties of GABP have identified these same two domains as being necessary for the formation of a heterotetrameric GABPα2/β2 complex. One essential region includes the amino-terminal ankyrin repeats of GABPβ1-1; this region is necessary for interaction with GABPα and subsequent contact with DNA. The other domain involved in complex formation is the carboxy-terminal region that facilitates homotypic dimerization (for review, see Thompson and McKnight 1992).

Using antisera raised against purified recombinant proteins, we observed concordant tissue distribution for GABPβ2-1 and GABPβ1-1. The GABPβ tissue distributions were quite similar to that of GABPα. The genes encoding these three proteins map to separate chromosomes: *Gabpa* maps to chromosome 16; GABPβ1-1 is encoded by the *Gabpb1* gene located on chromosome 2; and GABPβ2-1 is encoded by the *Gabpb2* gene located on chromosome 3. Despite being physically unlinked, expression of the three genes appears coordinately regulated. One possible means of coordinate regulation would entail positive autoregulatory feedback, wherein the heteromeric GABP complex transcriptionally activates each of the encoded genes. It should be possible to test this hypothesis by studying the regulatory DNA sequences that control the expression of the *Gabpa*, *Gabpb1*, and *Gabpb2* genes.

Two observations indicate that the carboxy-terminal regions of GABPβ1-1 and GABPβ2-1 assemble dimers as coiled–coils. First, the GABPβ2-1 dimerization domain can functionally replace the leucine zipper of C/EBPβ. The relevance of this observation is supported by the fact that the topological relationship between the GABPβ2-1 dimerization domain and the C/EBPβ basic region in the chimeric protein is subject to the same restrictions as that of native C/EBPβ. Second, CD spectroscopic studies of the GABPβ2-1 dimerization domain reveal that it is predominantly α-helical and that the $T_m$ and ellipticity at $\lambda222$ nm are concentration-dependent. The dimer-forming region of the GABPβ family of proteins likely adopts a parallel, coiled–coil conformation, such as those described in detailed structural studies of bZIP proteins (O'Shea et al. 1991; Ellenberger et al. 1992).

Identical DNA-binding behavior was observed when the C/EBPβ and C/EBP/GABPβ chimeras were modified to contain insertions between the basic region and dimerization domain. In both cases, insertions of one, two, three, five, or six residues yielded proteins incapa-

ble of sequence-specific DNA binding. Insertion of four or seven residues, however, yielded active proteins. Previous studies by Pu and Struhl (1991) showed that spacing changes of seven residues between the basic region and leucine zipper of GCN4 yield active proteins. Surprisingly, the +4 variants studied herein also yield functional proteins. Because the α-helical repeat is 3.5 residues per turn, one would predict that derivatives bearing insertions of either three or four residues would function suboptimally. One possible explanation for the discrepant behavior of the +3 and +4 insertions comes from the observations of Ellenberger et al. (1992). On examining the crystal structure of GCN4 bound to its recognition site on DNA, they noted that the coiled–coil is slightly unwound in a right-handed fashion, thus disposing the basic regions for optimal contact with DNA. The +4 mutation, relative to the canonical 3.5-residue repeat, should dispose the DNA contact surface in an "unwound" direction by one-half a helical turn. Likewise, the +3 mutation should dispose the contact surface in an "overwound" direction of one-half turn. Thus, the strong binding by the +4 variant may reflect its ability to properly align the basic region for contact with DNA.

The GABPβ1-1 and GABPβ2-1 proteins share very similar biochemical properties. They both bind GABPα through a similar set of ankyrin repeats. Likewise, they form both homo- and heterodimers via a carboxy-terminal domain that appears to function as a coiled–coil. Whereas the proteins might carry out independent functions, it is also possible that they are functionally redundant. One reason for speculating that the two GABPβ proteins might differ functionally relates to transcription activation. Earlier studies by Watanabe et al. (1990) showed that GABPβ specifies the transcriptional activation function of the GABP complex. The amino acid sequences of GABPβ1-1 and GABPβ2-1 proteins differ substantially in their interior regions where the activation function of GABPβ1-1 has been mapped (C. Thompson, unpubl.). It is therefore possible that GABPβ1-1 and GABPβ2-1 differ in their regulatory activities. If this is the case, it is also possible that the GABPβ1-1/GABPβ2-1 heterodimer is functionally distinct from either homodimer (see Fig. 7B). Additional studies, including targeted gene disruption, will be required to address this point.

Coiled–coil oligomerization domains have now been observed in four types of transcription factors: (1) bZIP (O'Shea et al. 1989; Vinson et al. 1989); (2) basic helix–loop–helix (bHLH)/ZIP (Hu et al. 1990; Beckmann and Kadesch 1991; for review, see Baxevanis and Vinson 1993); (3) heat shock factor (HSF; Rabindran et al. 1991, 1993); and (4) GABPβ. In the former two cases, the dimerization domain precisely dictates alignment of the DNA contact surfaces. The close apposition and defined orientation of the dimerization and DNA-binding domains of bZIP and bHLH proteins imposes stringent requirements for binding to dyad symmetric recognition sequences. In the latter two cases, HSF and GABPβ, the spacer regions between the oligomerization and DNA contact surfaces differ markedly and therefore provide

flexibility between the two domains. HSF, for example, binds stably to either dyad symmetric or tandemly repeated recognition sites (Perisic et al. 1989). Likewise, the orientation of GABP-binding sites (CGGAAR) can be organized either head to tail or head to head (Bolwig et al. 1992). Moreover, these primary recognition sites can be separated by up to 76 bp as in the promoter of the mouse cytochrome oxidase gene (Virbasius et al. 1993). We believe that the binding by the GABP complex to distantly spaced sites is made possible because of flexibility of the spacer region between the carboxy-terminal dimer-forming region and the amino-terminal GABPα/GABPβ interface. For GABPβ1-1, the distance between the dimerization domain and the ankyrin repeats is 201 amino acids. Assuming that for each GABPβ subunit in the complex, this region could extend fully unstructured in opposing directions, the complex could span a distance of 144 nm. For canonical B-form DNA, such a distance would allow DNA recognition sites to be separated by as many as 426 bp. We conclude by emphasizing that an understanding of the molecular architecture of transcription factors, as exemplified in the case of GABP, can allow prediction of concrete parameters regarding their mechanisms of action.

## Materials and methods

### Isolation and sequencing of GABPβ2-1 cDNA

Genomic DNA clones corresponding to GABPβ1-1 (GABPβ1; LaMarco et al. 1991) and GABPβ2-1 were isolated from a phage λGEM-12 library. The library was prepared according to the manufacturer's recommendations (Promega) from mouse strain 129svJ genomic DNA (Jackson Laboratories) partially digested with Sau3AI. Approximately $1.5 \times 10^6$ plaques (8× mouse genome) were screened on duplicate nitrocellulose filters (Schleicher & Schuell) with $^{32}$P-random primed (Feinberg and Vogelstein 1984) GABPβ1-1 DNA probes. Restriction mapping and partial sequencing of the genomic clones identified a 300-bp PstI fragment spanning ankyrin repeats 2 and 3 (see Fig. 1; LaMarco et al. 1991) in Gabpb2. This DNA fragment was then used as a probe to screen $1 \times 10^6$ plaques from a mouse lung cDNA library (Stratagene). Filters were prehybridized, then hybridized at 55°C or 65°C and washed at the same temperatures for the genomic and cDNA screens, respectively (Peden et al. 1982). After subcloning of genomic DNA fragments in pBluescript vectors, or excision from cDNA recombinant bacteriophages (Stratagene), dideoxynucleotide chain termination sequencing (Sanger et al. 1977) was performed with a commercial sequencing kit (U.S. Biochemical).

### Recombinant C/EBPβ (C1i0-i7C1) and C/EBPβ–GABPβ (C1i0-i7β2 and C1β1) hybrid proteins

Expression vectors producing high levels of recombinant C/EBPβ (C1i0-i7C1) or C/EBPβ chimeric proteins (C1β1 and C1i0-i7β2) were constructed essentially as described for C/EBPα and C/EBPα–GCN4 chimeric proteins (Agre et al. 1989). The parental recombinant C/EBPβ expression plasmid was prepared by polymerase chain reaction (PCR) methods that allowed precise joining of 5' sequences encoding the basic region with 3' sequences encoding the leucine zipper via an XhoI restriction enzyme recognition site.

C/EBPβ 5' sequences were first amplified from a C/EBPβ mammalian expression plasmid (Cao et al. 1991) with PCR primers FCT57 and FCT44. The 5' PCR primer FCT57 spans nucleotides 278–297 and includes a BamHI restriction enzyme recognition site. The 3' PCR primer FCT44 is complementary to nucleotides 841–854 and carries an XhoI site that replaces nucleotides 855–860. PCR amplification of the C/EBPβ cDNA with FCT57 and FCT44 primers generated a 594-bp fragment that is flanked by BamHI and XhoI sites and codes for the C/EBPβ basic region. For most constructs, this 5' fragment was modified furhter by religating the parental PCR product after deletion of an internal StyI restriction fragment in order to generate the amino-terminal CliO product (see Fig. 5). The CliO fragment of C/EBPβ encodes amino acids 58–116 joined to 195–251. The 3' sequences encoding either the C/EBPβ leucine zipper or GABPβ1-1 or GABPβ2-1 dimerization domains were also generated by PCR amplification using oligonucleotides that permitted insertion of terminal restriction enzyme recognition sites. C/EBPβ leucine zipper sequences were amplified using FCT52, which places the XhoI site immediately 5' of nucleotides 861–881, and FCT53, which is complementary to nucleotides at the carboxyl terminus of the C/EBPβ open reading frame and carries a terminal EcoRI recognition site.

Sequences that code for the GABPβ1-1 and GABPβ2-1 dimerization domains were also generated by PCR amplification of cDNA clones using cognate primers (FCT46 and FCT59 for GABPβ1-1, and FCT48 and M13 rev for GABPβ2-1). Restriction enzyme-digested PCR products were joined in a three-fragment ligation reaction (5' BamHI–XhoI-, 3' XhoI–EcoRI- and BamHI–EcoRI-cut vector) for cloning into the T7 promoter-based E. coli expression plasmid TP-7 (Wilson et al. 1993). Insertion mutant genes (i1–i7) were derived from C/EBPβ-containing hybrid genes prepared by PCR using FCT57 and any of the seven FCTi primers to replace the parental CliO fragment. Linker sequences used in the insertion mutants were identical for Cli0-i7C1 and Cli0-i7β2 and included GABPβ2-1 residues that are just amino-terminal to the dimerization domain (i1 = L, i2 = EL, i3 = REL, i4 = EREL, i5 = SEREL, i6 = RSEREL, i7 = LRSEREL).

PCR conditions for generating recombinant sequences were as follows: One cycle of denaturation at 94°C for 5 min; annealing at 52°C for 2 min; and extension at 72°C for 10 min; followed by denaturation at 94°C for 1 min; annealing at 52°C for 2 min; and extension at 72°C for 3 min repeated for 34 cycles. Standard reaction conditions were used (Sambrook et al. 1989) except that the amplification reactions also included 10% dimethylsulfoxide (vol/vol). All PCR-generated constructs were sequenced to ensure correct ligation of the fragments, presence of the linker insertion mutations (i1–i7), and integrity of the intended coding sequence. Primers used to generate and sequence Cli0-i7C1, C1β1, and Cli0-i7β2 constructs are shown: FCT44, 5'-GG CTC GAG CTC CAG GTT GCG CAT CT-3'; FCT44i1, 5'-CC CTC GAG TAG CTC CAG GTT GCG CAT CT-3'; FCT44i2, 5'-CC CTC GAG TAG CTC CTC CAG GTT GCG CAT CT-3'; FCT44i3, 5'-CC CTC GAG TAG CTC TCT CTC CAG GTT GCG CAT CT-3'; FCT44i4, 5'-CC CTC GAG-TAG CTC TCT TTC CTC CAG GTT GCG CAT CT-3'; FCT44i5, 5'-CC CTC GAG TAG CTC TCT TTC ACT CTC CAG GTT GCG CAT CT-3'; FCT44i6, 5'-CC CTC GAG TAG CTC TCT TTC ACT TCT CTC CAG GTT GCG CAT CT-3'; FCT44i7, 5'-CC CTC GAG TAG CTC TCT TTC ACT TCT TAG CTC CAG GTT GCG CAT CT-3'; FCT46, 5'-GG CTC GAG AAA CAG CTG GAT GAA GC-3'; FCT48, 5'-GG CTC GAG CAA CAG CTC CAG GAG GC-3'; FCT52, 5'-GG CTC GAG CAC AAG GTG CTG GAG CT-3'; FCT53, 5'-GG GAA TTC CTA GCA GTG GCC CGC CG-3'; FCT57, 5'-CC GGA TCC CAC GAG CGC GCC ATC GAC TTC A-3'; FCT59, 5'-

CC GAA TTC AGT TTT AAA TAC ATT TC-3'; M13 rev, 5'-AAC AGC TAT GAC CAT G-3'.

Recombinant C/EBPβ-derived proteins were produced as heart muscle kinase (HMK) tagged polypeptides in HMS174 (DE3) or BL21 strains of E. coli. Expression was induced with 2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 2 hr at 37°C. At the end of the IPTG induction, cells were harvested by centrifugation and resuspended in one-tenth the culture volume of phosphate-buffered saline (PBS, pH 7.9) supplemented with 5 M urea, 1 mM EDTA, and 0.2% Triton X-100 (Agre et al. 1989). Cells were subjected to two to three cycles of freezing in liquid nitrogen, thawing at 37°C, and sonicating for 1 min. After centrifugation at 16,000g for 30 min, lysates were loaded on fast-flow Q–Sepharose columns (Pharmacia). The flowthrough fraction was applied to an 8-ml Mono S FPLC column (Pharmacia) and the recombinant proteins were eluted from the column with a 100–600 mM KCl gradient. Recombinant proteins partially purified by this method comprised 30–90% of the soluble material. Coomassie blue staining of SDS–polyacrylamide gels, immunoblotting, and Bradford protein quantitation methods were used to estimate the amount of recombinant protein present in each Mono S fractions.

Glutaraldehyde cross-linking experiments were performed as described in Landschulz et al. (1989). Mono S-purified recombinant proteins diluted to concentrations of 1–2 mM in PBS containing 2 mM DTT were incubated without or with 0.005% glutaraldehyde for either 3 or 6 min at room temperature. Cross-linking reactions were stopped by the addition of an equal volume of 2× SDS protein sample buffer (Sambrook et al. 1989). Protein samples were subjected to electrophoresis on SDS–polyacrylamide gels (Laemmli 1970). Monomeric and dimeric protein species were visualized by Coomassie blue staining.

## DNase I footprinting assays

The DNA probe used in the DNase I footprinting assays was a 170-bp EcoRI–XbaI fragment that included the 136-bp EcoRI–BglII DNA fragment of the HSV tk promoter modified to contain a high affinity C/EBP DNA-binding site (Agre et al. 1989; Vinson et al. 1989). The DNA probe used for DNase I footprinting was prepared by digesting at the unique EcoRI cleavage site located 64 bp upstream of the C/EBP-binding site, filling in with [α-$^{32}$P]dATP and [α-$^{32}$P]PTTP (3000Ci/mmole, Amersham), then digesting at the 3' XbaI restriction site. Binding reactions included the recombinant proteins at concentrations of 10 mM for Cli0-i7C1 and 30 mM for Cli0-i7β2 and the probe DNA (1.5×10$^4$ cpm). Reactions were incubated at room temperature for 20 min in a 30-μl reaction mixture consisting of 50 mM KCl, 25 mM Tris-HCl, (pH 7.9), 3 mM MgCl$_2$, 1 mM CaCl$_2$, 0.5 μg of poly[d(I-C)] (Pharmacia), 1 mg/ml of bovine serum albumin (Boehringer Mannheim), and 2 mM DTT. After incubation, binding reactions were exposed to RNase-free DNase I (Boehringer Mannheim) for 1 min at room temperature and stopped by precipitation with ethanol containing 2 μg of yeast tRNA and 1 M ammonium acetate. Precipitated reaction samples were recovered by centrifugation and resuspended in formamide gel loading buffer (Sambrook et al. 1989), boiled for 3 min, and loaded onto a 6% polyacrylamide sequencing gel.

## Preparation of antibodies to GABPβ2-1

A GABPβ2-1 cDNA clone generated by PCR to introduce a BamHI restriction site at its 5' terminus and an Asp718 site at its 3' terminus was subcloned into the multiple cloning site of TP-7 (Studier and Moffat 1986; Wilson et al. 1993). The recom-

binant GABPβ2-1 protein was expressed in *E. coli* strain HMS174 (DE3) by induction with IPTG for 3 hr at 37°C. Purification of GABPβ2-1 was essentially as described for GABPβ1-1 (GABPβ1; Thompson et al. 1991). Briefly, recombinant GABPβ2-1 was solubilized from the *E. coli* particulate fraction in buffer A (10 mM Tris-HCl at pH 7.9, 100 mM KCl, 0.5 mM EDTA, 1 mM MgCl₂, 1 mM DTT, 1 mM CaCl₂, 0.2 mM phenylmethanesulfonyl fluoride) supplemented with 7 M urea. The urea-solubilized fraction was dialyzed overnight against buffer B [25 mM Tris-HCl at pH 8.0, 75 mM NaCl, 1 mM EDTA, 10% (vol/vol) glycerol, 1 mM DTT]. The dialyzed sample was centrifuged at 16,000g for 40 min and the supernatant was loaded on a fast-flow Q–Sepharose column (Pharmacia). A gradient of 75–500 mM NaCl was applied to elute the protein fractions. The final protein preparation used for rabbit immunization was estimated to be at least 90% pure by Commassie staining of the column fractions sized on SDS–polyacrylamide gels. Polyclonal antisera raised against GABPβ2-1 were produced by the Berkeley Antibody Company following standard procedures. Rabbit polyclonal antisera raised previously against GABPβ1-1 (Thompson et al. 1991) were observed to exhibit some cross-reactivity with the GABPβ2-1 protein. This cross-reactivity of the antisera, however, was minimal and only observed when high titer dilutions of the antisera >1:100 were used to probe very high amounts (>1 μg) of bacterially expressed polypeptides.

### Western blotting and probing with GABPβ1-1 and GABPβ2-1 antisera

Mouse tissues were homogenized in boiling 4× SDS sample buffer without bromophenol blue (Sambrook et al. 1989). The insoluble material was separated from the soluble fraction by centrifugation at 12,000g for 5 min. Western blotting and antigen–antibody complex detection were performed as described (Sambrook et al. 1989) using alkaline phosphatase-conjugated anti-rabbit antibodies (Promega) for detection of the immunoreactive recombinant proteins or ¹²⁵I-coupled to protein A (Amersham), followed by autoradiography for visualization of immunoreactive material in tissue extracts. Rabbit polyclonal antibodies directed against GABPβ1-1, GABPβ2-1, and C/EBPβ were used for these analyses. Antiserum specific to the carboxy-terminal sequences of C/EBPβ has been described previously (Cao et al. 1991). For antibody probing of Western blots, approximately equal amounts of protein were electrophoretically separated on 10% polyacrylamide–SDS gels and transferred onto Immobilon membranes (Millipore). Soluble recombinant HMK epitope-tagged GABPβ1-1 and GABPβ2-1 were labeled with [γ-³²P]ATP (Amersham) using purified HMK (Sigma). Prior to exposure to the labeled recombinant protein, Western blots were subjected to a cycle of denaturation and renaturation starting with 5 M urea (Vinson et al. 1988). Blocking, binding, and washing conditions have been described previously (Brown and McKnight 1992).

### Peptide synthesis and circular dichroism measurements

The synthetic peptide corresponding to the GABPβ2-1 dimerization domain (ERELLQQQLQEANRRAQEYRHQLLKKEQE-AEQYRLRLEAMAQQQ) was synthesized on an Applied Biosystems 431A peptide synthesizer using HBTU-activated FMOC amino acids and self-devised cycles. After amino-terminal acetylation, the peptide was cleaved with reagent K (King et al. 1990) and purified by reversed-phase HPLC, eluting with a linear gradient of acetonitrile in water (both containing 0.1% trifluoroacetic acid). Peptide identity and purity (96%) was as-

sessed by electrospray–ionization mass spectrometry (Hewlett-Packard 5989A; calculated mass 5553.17; observed mass 5553.08).

Circular dichroism spectra were measured on an Aviv 62DS spectropolarimeter equipped with a thermoelectric temperature control unit, using 0.5- and 1.0-cm path-length cells. Spectra were taken in PBS (150 mM NaCl, 10 mM sodium phosphate at pH 7) at 0°C for 15 sec per data point, unless otherwise noted. Thermal melting was measured by monitoring the signal at λ222 nm over a temperature range of 0°–80°C, and the Tm was obtained from the second derivative of the melting curve.

## References

Agre, P., P. Johnson, and S.L. McKnight. 1989. Cognate DNA binding specificity retained after leucine zipper exchange between GCN4 and C/EBP. *Science* **246:** 922–925.

Baxevanis, A.D. and C.R. Vinson. 1993. Interactions of coiled coils in transcription factors: Where is the specificity? *Curr. Opin. Genet. Dev.* **3:** 278–285.

Beckmann, H. and T. Kadesch. 1991. The leucine zipper of TFE3 dictates helix-loop-helix dimerization specificity. *Genes & Dev.* **5:** 1057–1066.

Bolwig, G.M., J.T. Bruder, and P. Hearing. 1992. Different binding site requirements for binding and activation for the bipartite enhancer factor EF-1A. *Nucleic Acid Res.* **20:** 6555–6564.

Brown, T.A. and S.L. McKnight. 1992. Specificities of protein-protein and protein-DNA interaction of GABPα and two newly defined ets-related proteins. *Genes & Dev.* **6:** 2502–2512.

Cao, Z., R.M. Umek, and S.L. McKnight. 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. *Genes & Dev.* **5:** 1538–1552.

Chen, Y.-H., J.T. Yang, and K.H. Chau. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry* **13:** 3350–3359.

Cheng, S.V., J.H. Nadeau, R.E. Tanzi, P.C. Watkins, J. Jagadesh, B.A. Taylor, L. Hanles, N. Sacchi, and J.F. Gusella. 1988. Comparative mapping of DNA markers from the familial Alzheimer disease and Down Syndrome regions of human chromosome 21 to mouse Chromosomes 16 and 17. *Proc. Natl. Acad. Sci.* **85:** 6032–6036.

Cho, M., V. Villani, and P. D'Eustachio. 1991. A linkage map of distal mouse Chromosome 12. *Mamm. Genome* **1:** 30–36.

Chorney, M., F.-W. Shen, J. Michaelson, and E.A. Boyse. 1982. Monoclonal antibody to an alloantigenic determinant on beta2-microglobulin B2m of the mouse. *Immunogenetics* **16:** 91–93.

Cohen, C. and D. Parry. 1990. α-Helical coiled coils and bundles: How to design an α-helical protein. *Protein* **7:** 1–15.

Crick, F.H.C. 1953. The packing of α-helices: Simple coiled-coils. *Acta Crystallogr.* **6:** 689–697.

Descombes, P., M. Chojkier, S. Lichsteiner, E. Falvey, and U. Schibler. 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. *Genes & Dev.* **4:** 1541–1551.

D'Eustachio, P., P.S. Jedidi, R.C. Fuhlbrigge, R.W. Gray, and D.D. Chaplin. 1987. Interleukin-1 alpha and beta genes: Linkage on chromosome 2 in the mouse. *Immunogenetics* **26:** 339–343.

Dietrich, W., H. Katz, S.E. Lincoln, H.-S. Shin, J. Friedman, N. Dracopoli, and E.S. Lander. 1992. A genetic map of the mouse suitable for typing interspecific crosses. *Genetics* **131:** 423–447.

Ellenberger, T.E., C.J. Brandl, K. Struhl, and S.C. Harrison. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α-helices: Crystal structure of the protein-DNA complex. *Cell* **71:** 1223–1237.

Elliott, R.W., J. Pazik, and C. Hohman. 1992 . High resolution blot analysis of alleles for Glut-1 in RI strains. *Mouse Genome* **90:** 428–429.

Feinberg, A.P. and B. Vogelstein. 1984. A technique for radiolabeling DNA fragments separated by gel electrophoresis. *Anal. Biochem.* **37:** 266–267.

Frankel, W.N., J.P. Stoye, B.A. Taylor, and J.M. Coffin. 1989. Genetic identification of endogenous polytropic proviruses using recombinant inbred mice. *J. Virol.* **63:** 3810–3821.

Gentz, R., F.J. Rauscher III, C. Abate, and T. Curran. 1989. Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. *Science* **243:** 1695–1699.

Goding C. and P. O'Hare. 1989. Herpes simplex virus Vmw65-octamer-binding protein interaction: A paradigm for combinatorial control of transcription. *Virology* **173:** 363–367.

Harbury, P.B., T. Zhang, P.S. Kim, and T. Alber. 1993. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262:** 1401–1407.

Hu, Y.-F., B. Lüscher, 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.

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

King, D.S., C.G. Fields, and G.B. Fields. 1990. A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. *Int. J. Prot. Peptide Res.* **36:** 255–266.

Kouzarides, T. and E. Ziff. 1989. Leucine zippers of Fos, Jun and GCN4 dictate dimerization specificity and thereby control DNA binding. *Nature* **340:** 568–571.

Kristie, T.M. and P.A. Sharp. 1990. Interactions of the Oct-1 POU subdomains with specific DNA sequences and the HSV α-transactivator protein. *Genes & Dev.* **4:** 2383–2396.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head bacteriophage T4. *Nature* **227:** 680–685.

LaMarco, K.L. and S.L. McKnight. 1989. Purification of a set of cellular polypeptides that bind to the purine-rich cis-regulatory element of herpes simplex virus immediate early genes. *Genes & Dev.* **3:** 1372–1383.

LaMarco, K.L., 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., P. 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.

———. 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. *Science* **243:** 1681–1688.

Martin, S.A.M., B.A. Taylor, T. Watanabe, and G. Bulfield. 1984. Histidine decarboxylase phenotypes of inbred mouse strains: A regulatory locus Hdc determines kidney enzyme concentration. *Biochem. Genet.* **22:** 305–322.

McKnight, S.L. 1991. Molecular zippers in gene regulation. *Sci. Am.* **264:** 32–39.

Oas, T.G., L.P. McIntosh, E.K. O'Shea, F.W. Dahlquist, and P.S. Kim. 1990. Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. *Biochemistry* **29:** 2891–2894.

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'Neil, K.T. and W.F. deGrado. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly-occuring amino acids. *Science* **250:** 646–651.

O'Shea, E., R. Rutkowski, and P. Kim. 1989. Evidence that the leucine zipper is a coiled-coil. *Science* **243:** 538–542.

———. 1992. Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* **68:** 699–708.

O'Shea, E., J., Klemm, P. Kim, and T. Alber. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel, coiled-coil. *Science* **254:** 539–544.

Peden, K., P. Mounts, and G.S. Hayward. 1982. Homology between mammalian cell DNA sequences and human herpes virus genomes detected by a hybridization procedure with high complexity probe. *Cell* **31:** 71–80.

Perisic, O., H. Xiao, and J.T. Lis. 1989. Stable binding of Drosophila heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. *Cell* **59:** 797–806.

Pu, W.T. and K. Struhl. 1991. The leucine zipper symmetrically positions the adjacent basic regions for specific DNA binding. *Proc. Natl. Acad. Sci.* **88:** 6901–6905.

Rabindran, S.K., G. Giorgi, J. Clos, and C. Wu. 1991. Molecular cloning and expression of a human heat shock factor, HSF1. *Proc. Natl. Acad. Sci.* **88:** 6906–6910.

Rabindran, S.K., R.I. Haroun, J. Clos, J. Wisniewski, and C. Wu. 1993. Regulation of heat shock factor trimer formation: Role of a conserved leucine zipper. *Science* **259:** 230–234.

Rasmussen, R., D. Benvegnu, E.K. O'Shea, P.S. Kim, and T. Alber. 1991. X-ray scattering indicates that the leucine zipper is a coiled coil. *Proc. Natl. Acad. Sci.* **88:** 561–564.

Rowe, L.B., J.H. Nadeau, R. Turner, W.N. Frankel, V.A. Letts, J.T. Eppig, M.S.H. Ko, S.J. Turston, and E.H. Birkenmeier. 1994. Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. *Mamm. Genome* **5:** 253–274.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.

Stern, S. and W. Herr. 1991. The herpes simplex virus transactivator VP16 recognizes the Oct-1 homeo domain: Evidence for a homeo domain recognition subdomain. *Genes & Dev.* **5:** 2555–2566.

Studier, F.W. and B.A. Moffat. 1986. Use of bacteriophage T7

RNA polymerase to direct high-level expression of cloned genes. *J. Mol. Biol.* **189:** 113–130.

Talanian, R., C. McKnight, and P. Kim. 1990. Sequence-specific DNA binding by a short peptide dimer. *Science* **249:** 769–771.

Taylor, B.A. and L.B. Rowe. 1989. A mouse linkage testing stock possessing multiple copies of the endogenous ecotropic murine leukemia virus genome. *Genomics* **5:** 221–232.

Thompson, C.C. and S.L. McKnight. 1992. Anatomy of an enhancer. *Trends Genet.* **8:** 232–236.

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

Triezenberg, S.J., K.L. LaMarco, and S.L. McKnight. 1988. Evidence of DNA: Protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes & Dev.* **2:** 730–742.

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.

Vinson, C.R., K.L. LaMarco, P.F. Johnson, W.H. Landschulz, and S.L. McKnight. 1988. In situ detection of sequence-specific DNA-binding activity specified by a recombinant bacteriophage. *Genes & Dev.* **2:** 801–806.

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

Vinson, C.R., T. Hai, and S.M. Boyd. 1993. Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: Prediction and rational design. *Genes & Dev.* **7:** 1047–1058.

Virbasius, J.V., C.A. Virbasius, and R.C. Scarpulla. 1993. Identity of GABP with NRF-2, a multisubunit activator of cytochrome oxidase expression, reveals a cellular role for an ETS domain activator of viral promoters. *Genes & Dev.* **7:** 380–392.

Warden, C.H., A. Diep, B.A. Taylor, and A.J. Lusis. 1992. Localization of the gene for apolipoprotein D on mouse Chromosome 16. *Genomics* **12:** 851–852.

Watanabe, H., T. Wada, and H. Handa. 1990. Transcription factor E4TF1 contains two subunits with different functions. *EMBO J.* **9:** 841–847.

Watanabe, H., J.-I. Sawada, K.-I. Yano, K. Yamaguchi, M. Goto, and H. Handa. 1993. cDNA cloning of transcription factor E4TF1 subunits with Ets and Notch motifs. *Mol. Cell. Biol.* **13:** 1385–1391.

Williams, S.C., C.A. Cantwell, and P.F. Johnson. 1991. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro. *Genes & Dev.* **5:** 1553–1567.

Wilson, A.C., K. LaMarco, M.G. Peterson, and W. Herr. 1993. The VP16 accessory protein HCF is a family of polypeptides processed from a large precursor protein. *Cell* **74:** 115–125.

Woody, R.W. 1985. Circular dichroism of peptides. In *The peptides*, Volume 7 (ed. S. Udenfried, J. Meienhofer, and V.J. Hruby), pp. 15–114. Acadamic Press, Orlando, FL.