(1989); G. C. Stevens, *Am. Nat.* **133**, 240 (1989); E. R. Pianka, *Trends Ecol. Evol.* **4**, 223 (1989); H. R. Pulliam, *Am. Nat.* **132**, 652 (1988).
19. S. L. Pimm, H. L. Jones, J. M. Diamond, *Am. Nat.* **132**, 757 (1988).
20. K. Ralls, J. D. Ballou, A. Templeton, *Conserv. Biol.* **2**, 185 (1988).
21. D. M. Raup, in *Biodiversity*, E. O. Wilson, Ed. (National Academy Press, Washington, DC, 1988), pp. 51–57.
22. M. L. McKinney, *Nature* **325**, 143 (1987).
23. S. M. Stanley, *Paleontology* **20**, 869 (1977); *Bull. Am. Acad. Arts. Sci.* **40**, 29 (1987); P. D. Taylor, *Hist. Biol.* **1**, 45 (1988).
24. N. Myers, *Deforestation Rates in Tropical Forests and Their Climatic Implications* (Friends of the Earth, London, 1989).
25. E. O. Wilson, *Sci. Am.* **261**, 108 (September 1989). Other, higher estimates are reviewed by W. V. Reid and K. R. Miller [*Keeping Options Alive: The Scientific Basis for Conserving Diversity* (World Resources Institute, Washington, DC, 1989)].
26. J. M. Diamond, *Discover* **11**, 55 (1990).
27. A. H. Knoll, in *Extinctions*, M. H. Nitecki, Ed. (Univ. of Chicago Press, Chicago, 1984), pp. 21–68.
28. P. M. Vitousek, P. R. Ehrlich, A. H. Ehrlich, P. A. Matson, *BioScience* **36**, 368 (1986).
29. World Commission on Environment and Development, *Our Common Future* (Oxford Univ. Press, New York, 1987).
30. N. Myers, *A Wealth of Wild Species* (Westview Press, Boulder, CO, 1983).
31. T. Eisner, *Issues Sci. Technol.* **6** (Winter), 31 (1989).
32. See summary in S. Roan, *Ozone Crisis: The 15 Year Evolution of a Sudden Global Emergency* (Wiley, New York, 1989).
33. R. A. Kerr, *Science* **247**, 1297 (1989).
34. M. Blumthaler and W. Ambach, *ibid.* **248**, 206 (1990).
35. S. H. Schneider, *Global Warming* (Sierra Club Books, San Francisco, 1989); D. Abrahamson, Ed., *The Challenge of Global Warming* (Island Press, Washington, DC, 1989); P. R. Ehrlich and A. H. Ehrlich, *Healing the Planet* (Addison-Wesley, Reading, MA, in press).
36. J. Shukla, C. Nobre, P. Sellers, *Science* **247**, 1322 (1990).
37. J. Charney, P. H. Stone, W. J. Quirk, *ibid.* **187**, 434 (1975).
38. For more details on soils and technical citations on what follows, see P. R. Ehrlich, A. H. Ehrlich, J. P. Holdren, *Ecoscience: Population, Resources, Environment* (Freeman, San Francisco, 1977).
39. P. R. Ehrlich and H. A. Mooney, *BioScience* **33**, 248 (1983).
40. This article is dedicated to some of the key individuals who have influenced our thinking on biodiversity: W. L. Brown, E. Mayr, C. D. Michener, and R. R. Sokal; and the late J. H. Camin, R. W. Holm, and R. H. MacArthur. For helpful comments on the manuscript we are indebted to G. C. Daily, J. M. Diamond, A. H. Ehrlich, T. Erwin, J. Harte, M. E. Harte, C. E. Holdren, J. P. Holdren, T. Lovejoy, J. P. Meyers, S. Conway Morris, D. M. Raup, P. H. Raven, and M. E. Soulé.

Research Article

Convergence of Ets- and Notch-Related Structural Motifs in a Heteromeric DNA Binding Complex

CATHERINE C. THOMPSON, THOMAS A. BROWN, STEVEN L. MCKNIGHT

Analysis of the heteromeric DNA binding protein GABP has revealed the interaction of two distinct peptide sequence motifs normally associated with proteins located in different cellular compartments. The α subunit of GABP contains an 85–amino acid segment related to the Ets family of DNA binding proteins. The ETS domain of GABPα facilitates weak binding to DNA and, together with an adjacent segment of 37 amino acids, mediates stable interaction with GABPβ. The β subunit of GABP contains four imperfect repeats of a sequence present in several transmembrane proteins including the product of the *Notch* gene of *Drosophila melanogaster*. These amino-terminal repeats of GABPβ mediate stable interaction with GABPα and, when complexed with GABPα, directly contact DNA. These observations provide evidence for a distinct biochemical role for the 33–amino acid repeats, and suggest that they may serve as a module for the generation of specific dimerization interfaces.

SINCE THE INITIAL RECOGNITION OF A COMMON PROTEIN sequence motif in the *SWI6* gene product of *Saccharomyces cerevisiae* and the *Notch* gene product of *Drosophila melanogaster* (*1*), similar sequences have been identified in different biologically interesting proteins. The motif, variously termed the cdc10/SWI6 or ankyrin repeat, consists of a 33–amino acid sequence often present in tandem arrays. This motif has been observed in the products of the *Notch*, *lin-12*, and *glp-1* genes, putative transmembrane proteins of *Drosophila melanogaster* and *Caenorhabditis elegans* that transmit signals critical for specification of cell fate (*2*); the product of *fem-1*, a *Caenorhabditis elegans* gene that regulates sex determination (*3*); cdc10, SWI4, SWI6, yeast proteins involved in cell cycle control (*1, 4*); ankyrin, a multifunctional protein of the red blood cell cytoskeleton (*5*); the product of *bcl-3*, a human gene located near a translocation breakpoint associated with some leukemias (*6*); the 105-kD precursor to the active 50-kD subunit of NFκB/KBF1 (*7*); and IκB, a regulatory subunit of NFκB that inhibits DNA binding and has been implicated in cytoplasmic sequestration (*8*). Despite the widespread occurrence of the 33–amino acid motif, its functional role has heretofore remained obscure.

Our interest in the 33–amino acid repeat arose from studies of GA binding protein (GABP), a multisubunit DNA binding protein purified from rat liver nuclei (*9*). GABP was originally identified as a factor that binds to a cis-regulatory element required for VP16-mediated activation of herpes simplex virus (HSV) immediate early genes (*10*). Biochemical and molecular biological experiments have shown that GABP is composed of two distinct polypeptides, both of which are required for avid interaction with DNA (*9, 11*). The amino acid sequence of the GABPα subunit exhibits similarity to the Ets family of nuclear proteins, whereas that of GABPβ contains a tandem series of 33–amino acid, cdc10/SWI6 repeats (*11*). We now demonstrate that it is these two distinct protein sequence motifs that form the heteromeric interface between GABPα and GABPβ.

We view the 33–amino acid repeat as a versatile module for the

The authors are in the Howard Hughes Research Laboratories, Carnegie Institution of Washington, Department of Embryology, Baltimore, MD 21210.

generation of highly specific dimerization interfaces that, through evolution, have been tailored to interact with a variety of different polypeptide substrates. This concept contrasts with the nature of dimerization domains typified by the leucine zipper (*12*) and helix-loop-helix (*13*), which facilitate oligomerization through the interaction of two structurally related polypeptide segments.

DNA binding properties of GABP. Recombinant cDNA copies of the mRNAs that encode the two subunits of GABP were isolated and characterized (*11*). In order to study the functional properties of the subunits, GABPα and GABPβ1 were expressed in *Escherichia coli* and purified by using conventional chromatographic techniques (*14*). The DNA binding properties of the two individual polypeptides and mixtures thereof were first studied by gel retardation with a DNA substrate derived from the enhancer of an immediate early gene of herpes simplex virus. As in earlier studies (*9, 11*), binding was not observed when DNA was incubated with either of the isolated subunits. When GABPα and GABPβ1 were incubated with DNA at the same time, a DNA-protein complex that retarded mobility relative to that of free DNA was observed (Fig. 1, left).

The multisubunit dependence of DNA binding by GABP was relieved in gel retardation assays conducted at lower ionic strength (Fig. 1, right). GABPα formed two retarded complexes that migrated at positions between free DNA and the complex formed with both subunits. Under these conditions, the mixture of GABPα and GABPβ1 again led to the formation of a slowly migrating DNA-protein complex. The retarded mobility of the latter complex, relative to those formed by GABPα alone, reflected the presence of the GABPβ1 subunit. Antisera specific to each subunit were used to confirm the validity of this interpretation. Antiserum specific to GABPα further retarded the migration of complexes formed with GABPα alone or the mixture of GABPα and GABPβ1. Antiserum specific to GABPβ1 did not affect the mobility of complexes formed between GABPα and DNA, but did retard the complex formed in the presence of both subunits (*15*).

The HSV1-derived DNA fragment used in binding assays of GABP contains three imperfect repeats of the hexanucleotide sequence 5′-CGGAAR-3′, which had been shown earlier to be protected from deoxyribonuclease I (DNase I) digestion when bound by GABP (*9, 10*). DNase I footprinting assays were performed with bacterially synthesized proteins under conditions that allowed interaction of GABPα alone. Addition of 0.15 nM GABPα could protect the repeated hexanucleotide motifs from DNase I digestion. When the GABPβ1 subunit was added, protection was observed with only 0.015 nM GABPα. Furthermore, the pattern of nuclease protection was extended slightly beyond the adenine residues of the third repeat.

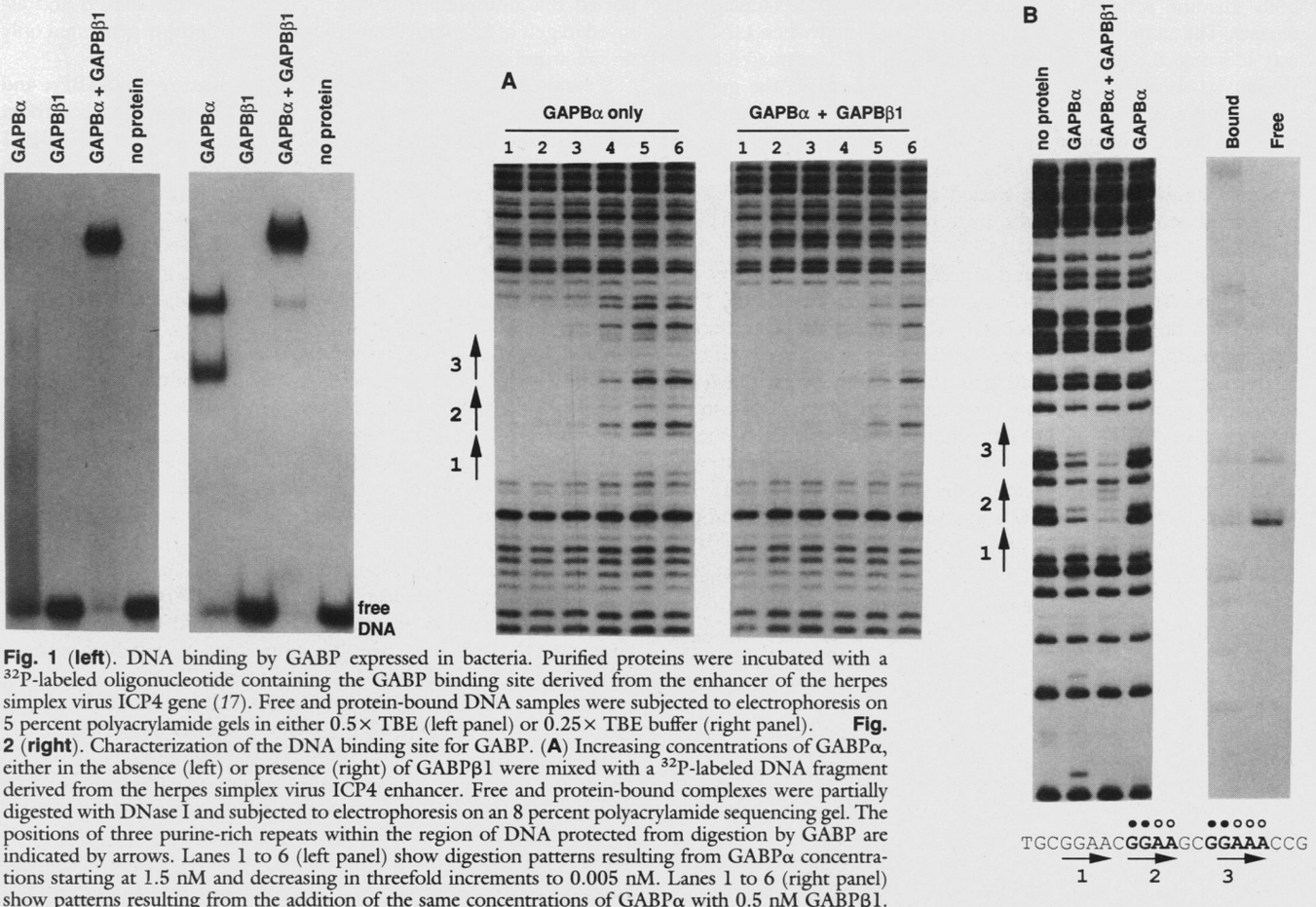


Fig. 1 (**left**). DNA binding by GABP expressed in bacteria. Purified proteins were incubated with a ^{32}P-labeled oligonucleotide containing the GABP binding site derived from the enhancer of the herpes simplex virus ICP4 gene (*17*). Free and protein-bound DNA samples were subjected to electrophoresis on 5 percent polyacrylamide gels in either 0.5× TBE (left panel) or 0.25× TBE buffer (right panel). **Fig. 2** (**right**). Characterization of the DNA binding site for GABP. (**A**) Increasing concentrations of GABPα, either in the absence (left) or presence (right) of GABPβ1 were mixed with a ^{32}P-labeled DNA fragment derived from the herpes simplex virus ICP4 enhancer. Free and protein-bound complexes were partially digested with DNase I and subjected to electrophoresis on an 8 percent polyacrylamide sequencing gel. The positions of three purine-rich repeats within the region of DNA protected from digestion by GABP are indicated by arrows. Lanes 1 to 6 (left panel) show digestion patterns resulting from GABPα concentrations starting at 1.5 nM and decreasing in threefold increments to 0.005 nM. Lanes 1 to 6 (right panel) show patterns resulting from the addition of the same concentrations of GABPα with 0.5 nM GABPβ1. (**B**) Methylation protection (left panel) and interference (right panel) of DNA binding by GABP. The same DNA fragment used in (A) was incubated with GABPα, GABPβ1, or an equimolar mixture of the two subunits, and exposed to dimethyl sulfate (DMS). Partially methylated DNA was recovered, cleaved with piperidine, and analyzed on an 8 percent polyacrylamide sequencing gel. For methylation interference assays, DNA was partially methylated, incubated with an equimolar mixture of GABPα and GABPβ1, and subjected to electrophoresis on a nondenaturing polyacrylamide gel as described in Fig. 1. Free and protein-bound DNA species were recovered, cleaved with piperidine, and subjected to electrophoresis on an 8 percent polyacrylamide sequencing gel. Nucleotide residues closely contacted by GABP are shown in the lower part of (B). Filled circles identify guanine residues protected from DMS by GABP and when methylated inhibited DNA binding by GABP. Open circles identify adenine residues where methylation is enhanced in the presence of both GABPα and GABPβ1.

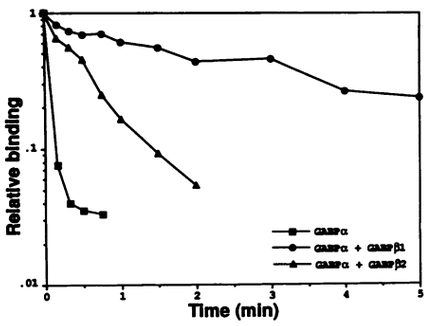


Fig. 3. Measurements of DNA binding stability of complexes formed by GABP subunits. A ^{32}P-labeled oligonucleotide containing a GABP binding site (Fig. 1) was incubated with GABPα alone, or together with equimolar amounts of the two β subunits. After a 10-minute incubation at 24°C, protein-DNA complexes were treated with a 500-fold excess of unlabeled oligonucleotide. Protein-bound and free DNA were separated by nondenaturing gel electrophoresis as described in Fig. 1A. Protein-bound and free DNA's were located by autoradiography, excised, and quantified by scintillation spectrometry. Results are presented as the fraction of probe bound, normalized to 1.0 at time = 0.

Table 1. Determination of molecular masses of GABP subunits. Purified GABPα, GABPβ1, and GABPβ2 produced in *E. coli* were analyzed by gel filtration and sedimentation velocity (*18, 19*). K_{av} values were calculated from the elution volume of a Superose-6 FPLC column. Apparent molecular sizes were determined from a plot of K_{av} as a function of the log of the molecular size of the protein standards. Stokes radii were determined from a plot of $(-\log K_{av})^{1/2}$ as a function of the Stokes radii of the standards. Sedimentation coefficients together with measured Stokes radii were used to calculate native molecular mass. *S*, sedimentation coefficient.

Protein	K_{av}	Apparent molecular mass (kD)	Stokes radius	*S*	Corrected molecular mass (kD)
GABPα	0.449	158	49.3	3.1	66
GABP	0.377	349	63.0	3.1	82
GABPβ2	0.486	106	44.0	2.5	46
GABPα + GABPβ1	0.255	1300	87.7	4.5	170
GABPα + GABPβ2	0.367	390	64.8	3.8	104

The segment of DNA protected from DNase I digestion by GABP encompassed all three hexanucleotide repeats, yet was not centered over the repeats. Methylation protection and interference assays were undertaken in order to further define the sites of contact established between GABP and DNA. Methylation protection assays with GABPα showed a pattern of protection that included both guanine residues of the second and third hexanucleotide repeats. The same sets of guanines were protected when GABPβ1 was added to the binding reaction. In addition, accentuated methylation was observed at adenine residues adjacent to the guanine dinucleotides of the second and third repeats. Sites of methylation interference were mapped by separating protein-bound DNA molecules from those inactivated by partial methylation. Methylation of guanine dinucleotides in the second and third hexanucleotide repeats inhibited binding by the mixture of GABPα and GABPβ1 (Fig. 2B).

Methylation protection and interference assays indicate that GABP binds to sites on DNA corresponding to two of the three purine-rich hexanucleotide repeats. The pattern of protection of guanine residues by GABPα was similar to that observed with the mixture of both subunits, indicating that GABPα, when added at a sufficiently high concentration, can bind specifically to DNA in the absence of GABPβ1. Such observations offer an explanation for the two retarded bands observed when DNA was incubated with GABPα under conditions of low ionic strength (Fig. 1, right). The less retarded of the two bands is interpreted to represent a complex wherein GABPα is associated with only one of the two hexanucleotide repeats, while the more retarded complex contains GABPα subunits associated with two hexanucleotide repeats. Binding assays with DNA probes containing a single hexanucleotide repeat supported this interpretation. When incubated with GABPα and assayed in gels of low ionic strength, such DNA probes generated only one retarded complex (*16*).

Several observations indicated that the mixture of GABPα and GABPβ1 forms a complex that binds DNA more stably than the α subunit alone (*9, 11*). To further investigate the effect of GABPβ1 on DNA binding by GABPα, we measured the rate at which variously mixed proteins dissociate from DNA (Fig. 3). The dissociation rate of GABPα alone was too rapid to be measured accurately. Less than 10 percent of the DNA remained bound to GABPα after a 10-second challenge with excess, unlabeled competitor DNA. In contrast, when both GABPα and GABPβ1 were present, the dissociation rate was much slower ($T_{1/2}$, 1.5 minutes). Similar assays were performed with a mixture of GABPα and GABPβ2, which in earlier experiments did not form a stable complex with DNA (*11*). When used at nanomolar concentrations, GABPβ2 was capable of forming a DNA binding complex with GABPα (as shown below). The β2 isoform of GABP also stabilized DNA binding by GABPα, yet yielded a complex that dissociated more rapidly ($T_{1/2}$, 30 seconds) than that formed with GABPβ1. Because the β1 and β2 isoforms differ only at their COOH-termini (*11*), this part of the protein may participate in stabilizing DNA binding.

The foregoing data indicate that the β1 and β2 isoforms of GABP do not bind to DNA alone, but associate with the α subunit to augment DNA binding. The question arises as to whether the β subunits cause a conformational change in α, leading to its more avid interaction with DNA, or whether the β subunits might, in association with GABPα, establish direct contact with DNA.

To determine whether GABPβ1 contacted DNA when complexed with GABPα, we exposed DNA-protein complexes to ultraviolet (UV) light under conditions expected to permit covalent cross-linking between DNA and intimately bound proteins (*17*). The GABP subunits were incubated with ^{32}P-labeled DNA that contained the purine-rich hexanucleotide repeats, exposed to UV light, and subjected to electrophoresis on a denaturing polyacrylamide gel. When DNA was incubated with GABPα and exposed to UV light, a cross-linked product bearing an electrophoretic mobility close to that of GABPα was observed (Fig. 4). The appearance of this product was dependent on the presence of GABPα, and on exposure to UV light increased with time. Cross-linking was elimi-

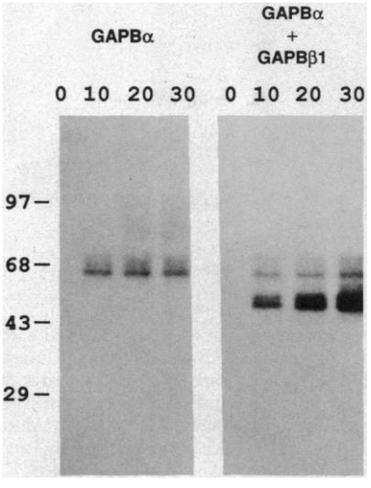


Fig. 4. UV-mediated cross-linking of GABP subunits to DNA. Isolated or mixed GABP subunits were incubated with a ^{32}P-labeled oligonucleotide containing a GABP binding site (*16*) and then exposed to UV light as indicated. Samples were denatured by boiling in SDS sample buffer and subjected to electrophoresis on a denaturing 12.5 percent polyacrylamide gel. The gel was then dried and exposed to x-ray film. Time of exposure to UV light (minutes) is indicated above each gel lane. Migration of molecular size markers (kilodaltons) is shown on the left.

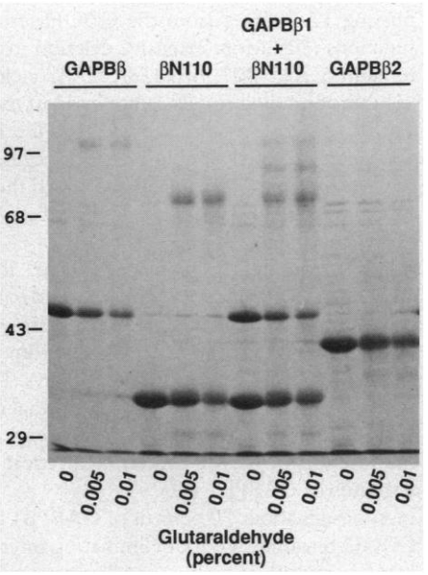


Fig. 5. Glutaraldehyde cross-linking of GABPβ1 and GABPβ2. Bacterially synthesized proteins were incubated in phosphate-buffered saline with varying concentrations of glutaraldehyde as indicated for 5 minutes at room temperature. Samples were denatured by boiling in SDS sample buffer and subjected to electrophoresis on a denaturing 10 percent polyacrylamide gel. The gel was then stained with Coomassie brilliant blue. Proteins present in cross-linking reactions are indicated above each lane. βN110 is a truncated version of GABPβ1 missing 110 NH_2-terminal residues (see Fig. 7B).

nated by the inclusion of excess, unlabeled DNA that contained the purine-rich hexanucleotide repeats, but not by excess nonspecific DNA (*16*).

No evidence of protein-DNA cross-linking was observed when ^{32}P-labeled DNA was mixed with GABPβ1 and exposed to UV light (*16*). However, when the mixture of GABPα and GABPβ1 was complexed with DNA and irradiated, new cross-linked products were observed. In addition to GABPα, two closely migrating polypeptide bands, slightly larger than the native size of GABPβ1, became covalently attached to the ^{32}P-labeled DNA substrate (Fig. 4). Although GABPβ1 did not bind DNA on its own, when present in a ternary complex it appeared to be even more susceptible to UV-mediated cross-linking than GABPα. These data provide evidence that the β1 subunit of GABP associates closely with DNA when complexed with GABPα.

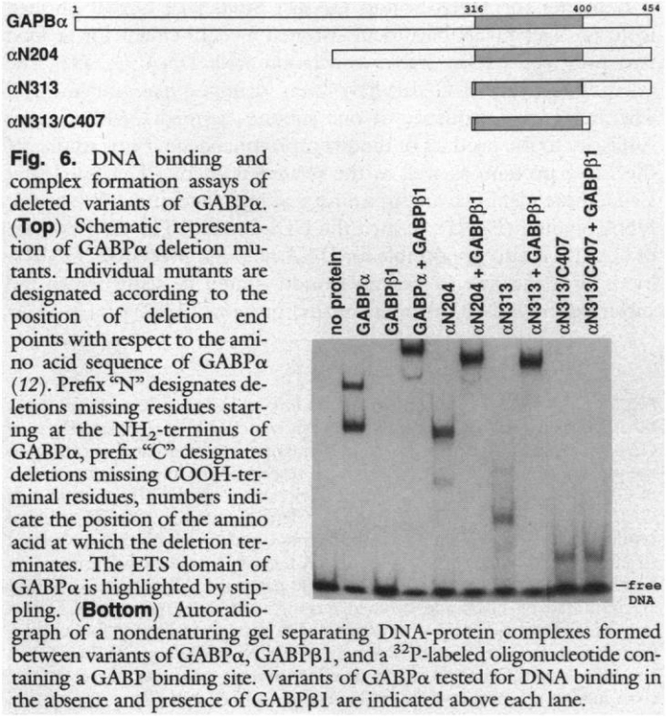


Fig. 6. DNA binding and complex formation assays of deleted variants of GABPα. (**Top**) Schematic representation of GABPα deletion mutants. Individual mutants are designated according to the position of deletion end points with respect to the amino acid sequence of GABPα (*12*). Prefix "N" designates deletions missing residues starting at the NH_2-terminus of GABPα, prefix "C" designates deletions missing COOH-terminal residues, numbers indicate the position of the amino acid at which the deletion terminates. The ETS domain of GABPα is highlighted by stippling. (**Bottom**) Autoradiograph of a nondenaturing gel separating DNA-protein complexes formed between variants of GABPα, GABPβ1, and a ^{32}P-labeled oligonucleotide containing a GABP binding site. Variants of GABPα tested for DNA binding in the absence and presence of GABPβ1 are indicated above each lane.

Formation of a stable complex between GABPα and GABPβ in the absence of DNA. Having found that the α and β subunits of GABP formed a heteromeric complex when exposed to their specific DNA substrate, we used gel filtration chromatography to determine whether these subunits might associate in the absence of DNA (*18*). When chromatographed separately, GABPα and GABPβ1 eluted as single peaks at K_{av} (partition coefficient) values of 0.45 and 0.38, respectively. However, co-chromatography of the subunits resulted in a single peak at K_{av} 0.26 (Table 1). Analysis of column fractions by polyacrylamide gel electrophoresis confirmed that the peak at K_{av} 0.26 contained both subunits.

Evidence confirming the stable association of GABP subunits in the absence of DNA was also obtained from measurements of sedimentation velocity (Table 1). The gel filtration and sedimentation properties of a protein or protein complex are affected both by size and molecular shape. However, use of the analytical methods of Siegel and Monty (*19*) made it possible to calculate native molecular masses of the various protein species. The calculated sizes of GABPα and GABPβ2 corresponded closely to their predicted sizes (51.3 and 37 kD, respectively). In contrast, GABPβ1 exhibited a native molecular size (82 kD) roughly twice that expected (41.3 kD).

The complex formed between GABPα and GABPβ1 eluted from the gel filtration column prior to the largest molecular mass standard. Thus, the calculated molecular mass of this complex (170 kD) represents a provisional assignment. Since GABPβ1 existed as a stable dimer on its own, the very large complex formed between it and GABPα was tentatively identified as a tetramer composed of two molecules of each subunit. Polyacrylamide gel analysis of the constituents of the GABPα-GABPβ1 complex were consistent with this interpretation, showing that the subunits existed in equal stoichiometries (*20*). This interpretation was also consistent with the properties of the complex formed between GABPα and GABPβ2. The native molecular mass of the latter complex (104 kD) corresponded to the expected size of a heterodimer. Thus, the differences in the sizes of the heteromeric complexes formed between GABPα and the two different β subunits may reflect the dimer-forming property of GABPβ1.

Gel filtration and gradient sedimentation assays indicated that GABPβ1 might exist as a dimer. This interpretation was tested using glutaraldehyde cross-linking. Bacterially expressed GABPβ1 and GABPβ2 were exposed to glutaraldehyde and subjected to electrophoresis on a denaturing polyacrylamide gel. Incubation of GABPβ1 with glutaraldehyde led to the formation of a second polypeptide band with an apparent size roughly double that of the monomeric form of the protein (Fig. 5). Similar experiments with GABPβ2 did not yield an analogous product.

Additional evidence suggesting that GABPβ1 exists as a dimer resulted from cross-linking experiments with the intact polypeptide and a truncated form lacking 110 residues at the NH_2-terminus (βN110). Cross-linking of the truncated protein yielded an additional species roughly twice the size of the monomeric form. When the truncated protein was mixed with intact GABPβ1 and exposed to glutaraldehyde, three cross-linked protein species were observed. Two species corresponded to cross-linked, homodimeric complexes that had been observed after glutaraldehyde treatment of native GABPβ1 and the NH_2-terminal truncated derivative. The third species migrated between the presumed homodimeric forms and probably represented a heteromeric complex consisting of one GABPβ1 polypeptide and one truncated polypeptide. Therefore, both molecular size measurements and cross-linking assays showed that GABPβ1, but not GABPβ2, exists as a stable homodimer.

Mapping of functional domains of GABPα and GABPβ.

Experimental results described thus far indicate that GABPα should contain at least two functional components, one that facilitates DNA binding and another that allows complex formation with GABPβ. The GABPβ1 polypeptide should contain at least three components, facilitating self-dimerization, heterodimerization with GABPα, and direct contact with the DNA substrate. Recombinant copies of the genes that encoded each subunit were systematically deleted in an effort to localize these components (*21*).

Deletion variants of GABPα that were missing as many as 313 residues from the NH_2-terminus retained the capacities to bind DNA and form a complex with GABPβ1 (Fig. 6). A GABPα variant missing 17 residues from the COOH-terminus also retained both functions (*20*). More extensive deletion from the COOH-terminus, to amino acid 407 (αN313/C407), yielded a protein that was capable of binding to DNA, but had lost the ability to complex with GABPβ1. These results showed that the ETS domain of GABPα was sufficient for DNA binding. The region of GABPα required to form a complex with GABPβ1 included the ETS domain, as well as 37 amino acids located on the immediate COOH-terminal side of the ETS domain.

In order to define regions of GABPβ1 that interact with GABPα and contact DNA, we produced deleted variants and tested them in gel retardation and UV cross-linking assays (Fig. 7). Variants that lacked up to 228 residues (βC154) from the COOH-terminus of GABPβ1 were functional in both assays. Deletion of an additional 33 residues (βC121) yielded a protein that did not function in either assay. The boundary defined by these experiments corresponded to the location of the 33–amino acid repeat nearest to the COOH-terminus of GABPβ1 (*11*).

Although about 70 percent of GABPβ1 could be deleted from its COOH-terminus without eliminating interaction with GABPα and DNA, removal of only a small segment from the NH_2-terminus resulted in detrimental effects. A variant of GABPβ1 that lacked 19 NH_2-terminal residues (βN19) was slightly less effective in converting GABPα-derived complexes into the very slowly migrating heteromeric complex. When tested in the UV cross-linking assay, βN19 yielded a reduced amount of cross-linked product relative to the intact β1 isoform. Variants that lacked 47 and 67 residues (βN47 and βN67) were progressively more defective in the complex formation assay and failed to be cross-linked to the radioactive DNA probe as efficiently as the intact protein. Finally, variants missing 80 or more residues from the NH_2-terminus were completely defective in both assays. The progressive loss of function observed in deleted forms of GABPβ1 corresponded to the progressive loss of the 33–amino acid repeats (Fig. 7). βN19 was truncated within the first of the repeats, βN47 within the second, βN67 after the second, and βN80 within the third repeat. The functional properties of the GABPβ1 deletion mutants indicate that the 33–amino acid repeats are important both for complex formation with GABPα and DNA contact.

A model for DNA-bound GABP. Studies of GABP purified from rat liver nuclei have demonstrated a requirement for at least two polypeptides for stable association with DNA (*9, 11*). The relevant subunits of GABP have been identified here and in (*11*). The amino acid sequence of one subunit, termed GABPα, bears similarity to the product of the Ets proto-oncogene. Prior studies of the Ets-1 protein, as well as the related proteins PU.1/Spi-1 and E74A, have identified an 85–amino acid region required for specific DNA binding (*22–25*), termed the ETS domain. The ETS domain of GABPα is also responsible for DNA binding. Moreover, GABPα binds to a purine-rich hexanucleotide related in sequence to the substrates bound by other Ets-related proteins (*22–25*).

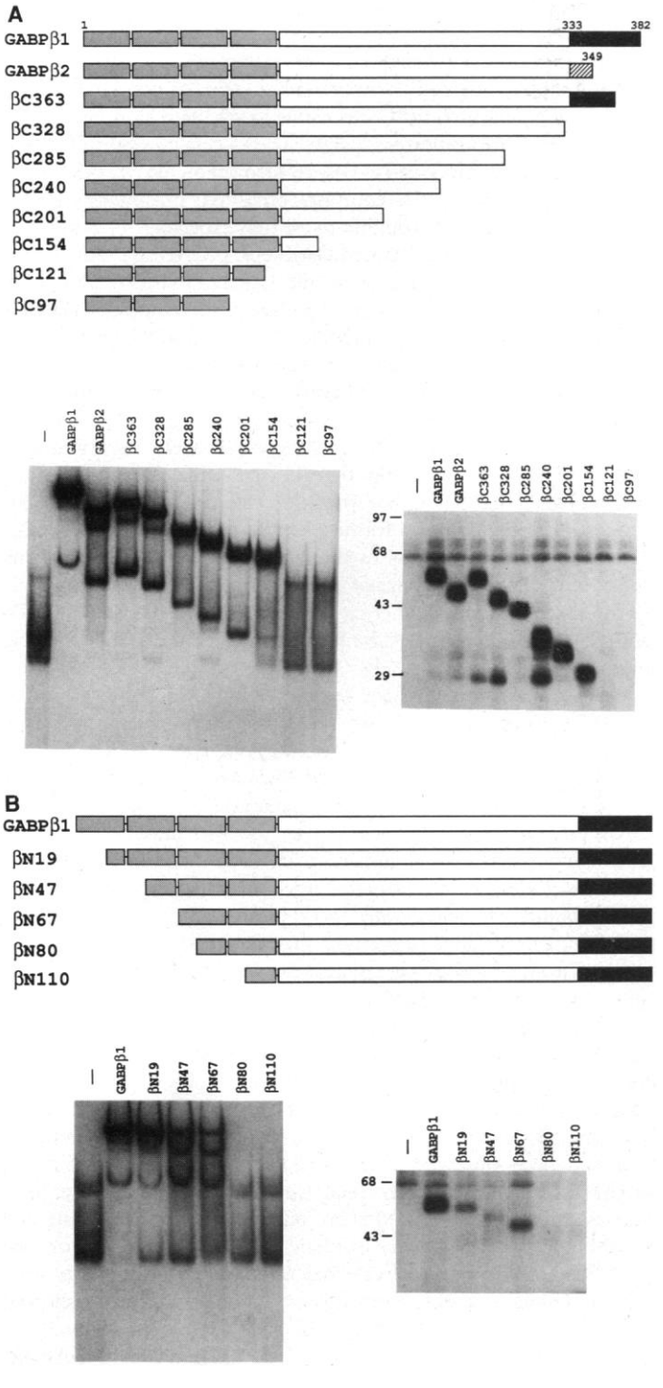


Fig. 7. Complex formation and UV cross-linking assays of deleted variants of GABPβ1. Top panels of (**A**) and (**B**) show schematic representations of GABPβ1 deletion mutants. Individual mutants are designated according to the positions of deletion end points with respect to the amino acid sequence of GABPβ1 (7). Prefix "N" designates deletions from the NH_2-terminus of GABPβ1 (B), prefix "C" designates deletions missing COOH-terminal residues. Repeated sequences 33 or 32 amino acids in length that are related to similarly sized repeats in the Notch protein of *Drosophila melanogaster* are highlighted by grey stippling. The unique parts of GABPβ1 and GABPβ2 are indicated by black and hatched rectangles at their respective COOH-termini. Deleted variants were synthesized in bacteria and tested for complex formation with GABPα as shown in the lower left of (A) and (B). Free DNA has been run off the gel. Each deletion mutant was also tested in UV cross-linking assays shown in the lower right of (A) and (B).

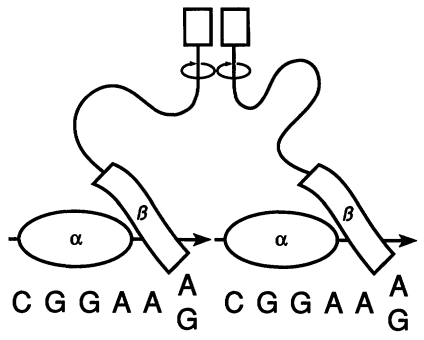


Fig. 8. Model of complex formed between GABP and DNA. The sequence of the GABP binding site is shown as two hexanucleotide repeats of the sequence 5′-CGGAAR-3′. Oval spheres directly above guanine residues correspond to GABPα subunits, elongated rectangles correspond to 33–amino acid repeats of GABPβ subunits. Smaller rectangles shown at top correspond to the region of GABPβ1 required for formation of homodimers. Circular arrows designate flexible regions inferred to occur between the dimer forming region of GABPβ1 and the 33–amino acid repeats located at its NH_2-terminus.

Both forms of the second subunit of GABP, β1 and β2, were capable of forming a stable complex with GABPα in the absence of DNA. The region of GABPα required for association with GABPβ1 included the ETS domain and a 37–residue region located on the immediate COOH-terminal side of the ETS domain. Both β subunits, when associated with GABPα, were capable of being cross-linked to the specific DNA substrate. Association with GABPα and direct DNA contact were specified by a tandem series of 33–amino acid repeats located within identical NH_2-terminal regions of GABPβ1 and GABPβ2. Finally, evidence that GABPβ1, unlike GABPβ2, exists as a homodimer has been presented.

Our observations have been incorporated into a provisional model of the complex formed when GABPα and GABPβ1 associate with a directly repeated set of purine-rich hexanucleotides (Fig. 8). Each hexanucleotide repeat is hypothesized to be contacted by both GABPα and GABPβ1. The linear order of contact of the two subunits was deduced from three observations. First, GABPα was capable of protecting both guanines from methylation by dimethyl sulfate. Second, the DNase I footprint generated by the mixed subunits, relative to that resulting from GABPα alone, was extended in a direction toward the adenine residues of the hexanucleotide repeat. Third, addition of GABPβ1 caused an enhanced pattern of methylation of adenine residues relative to the pattern generated in binding reactions that contained GABPα alone.

The most stable complexes of GABP and DNA formed with the mixture of GABPα and GABPβ1. The β1 subunit, unlike β2, exists as a stable homodimer. Moreover, when mixed with GABPα, the β1 subunit generated a large complex probably consisting of two polypeptides of each subunit. We hypothesize that this heteromeric tetramer binds in a concerted manner to two purine-rich hexanucleotide repeats. If this idea is correct, a flexible region should exist between the dimerization domain of GABPβ1 and the surfaces located near its NH_2-terminus that facilitate interaction with GABPα and DNA. Thus, linked polypeptides would be capable of binding simultaneously to a DNA substrate that is not rotationally symmetric.

The interfaces identified in the complex formed between GABP and DNA can be traced to two protein sequence motifs; the ETS domain of GABPα and the 33–amino acid repeats of GABPβ. These observations, in revealing the convergence of two otherwise disparate sequence motifs, raise the possibility that other Ets-related proteins might participate in multiprotein complexes. They further provide a function for the 33–amino acid repeat motif that has been identified in numerous regulatory proteins. The four repeats located close to the NH_2-terminus of GABPβ are necessary for stable association with GABPα. Moreover, they constitute the contact surface of GABPβ that becomes cross-linked to DNA by UV light.

These observations provoke consideration of the potential functions of the 33–amino acid repeats associated with plasma membrane proteins such as the products of the *Notch* gene of *Drosophila melanogaster* and the *glp-1* and *lin-12* genes of *Caenorhabditis elegans* (*2*). Genetic studies have implicated each of these proteins in cell signaling events that regulate key developmental processes. Studies of ankyrin, a constituent of red blood cell membranes that has 22 of these repeats, have suggested a role for the repeats in protein-protein interactions (*5*). On the basis of observations outlined herein we propose that the 33–amino acid repeats of Notch, Glp-1, and Lin-12 will constitute interaction surfaces that bind other proteins. The substrates sequestered by specific interaction with transmembrane proteins may be transcription factors or rate-limiting components of a signal transduction pathway that are inactive until released.

This type of mechanism may be involved in the regulation of NKκB. NFκB is a transcription factor that is retained in the cytoplasm by binding to another protein (IκB) (*26*), but can be released and translocated to the nucleus in response to extracellular signals (*27*). The inactive precursor of NFκB, as well as IκB, contains tandem 33–amino acid repeats similar to those in GABP (*7, 8*), suggesting the possibility that the repeats may be responsible for confining NFκB in the cytoplasm. Given our results with GABP, we propose that the repeats in IκB may mediate interaction with NFκB rather than with a cytoplasmic "anchor," and may accomplish cytoplasmic sequestration by masking a nuclear localization signal.

In conclusion, we return to the central theme of this study—the reliance of competent DNA binding complexes on multiple subunits. Why have transcriptional regulatory proteins of eukaryotic cells evolved in this manner? By separating functional components onto different polypeptide chains, critical subunits might be differentially expressed or sequestered, generating useful strategies for regulation. For example, differential expression of the mRNA's encoding the β1 and β2 subunits of GABP would be expected to have a substantial impact on the function of the resulting complex. It is also possible that subunits of one complex might cross-mix with those of a second complex as is the case for leucine zipper and helix-loop-helix proteins. It will be of interest, in this regard, to determine whether GABPβ is capable of associating with Ets-related proteins other than GABPα.

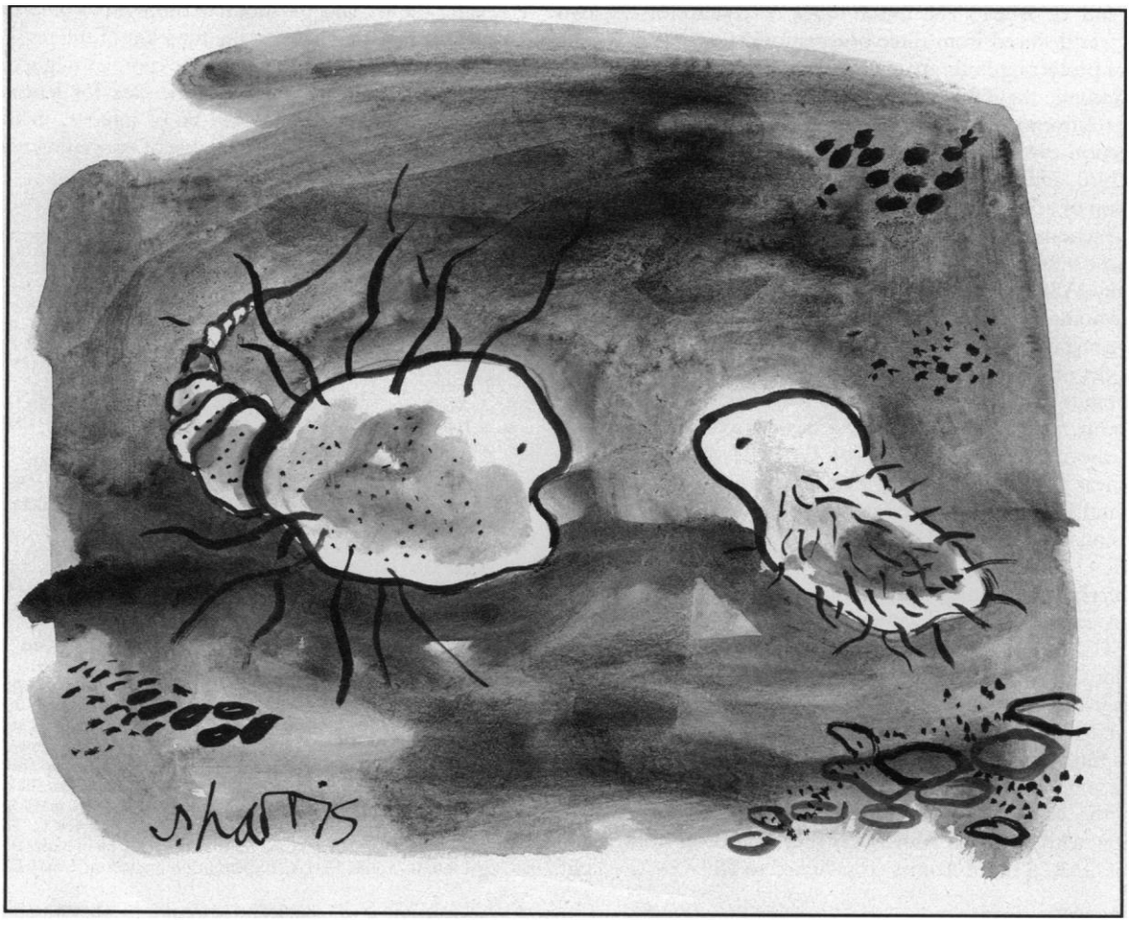
REFERENCES AND NOTES

1. L. Breeden and K. Nasmyth, *Nature* **329**, 651 (1987).
2. K. A. Wharton, K. M. Johansen, T. Xu, S. Artavanis-Tsakonas, *Cell* **43**, 567 (1985); S. Kidd, M. R. Kelley, M. W. Young, *Mol. Cell Biol.* **6**, 3094 (1986); I. Greenwald, *Cell* **43**, 583 (1985); J. Yochem, K. Weston, I. Greenwald, *Nature* **335**, 547 (1988); J. Yochem and I. Greenwald, *Cell* **58**, 553 (1989); J. Austin and J. Kimble, *ibid.*, p. 565.
3. A. M. Spence, A. Coulson, J. Hodgkin, *Cell* **60**, 981 (1990).
4. S. J. Aves, B. Durkacz, A. Carr, P. Nurse, *EMBO J.* **4**, 457 (1985); B. J. Andrews and I. Herskowitz, *Nature* **342**, 830 (1989).
5. S. E. Lux, K. M. John, V. Bennett, V., *ibid.* **344**, 36 (1990).
6. H. Ohno, G. Takimoto, T. W. McKeithan, *Cell* **60**, 991 (1990).
7. M. Kieran *et al.*, *ibid.* **62**, 1007 (1990); V. Bours, J. Villalobos, P. R. Burd, K. Kelly, U. Siebernlist, *Nature* **348**, 76 (1990).
8. S. Haskill *et al.*, *Cell* **65**, (1991).
9. K. L. LaMarco and S. L. McKnight, *Genes Dev.* **3**, 1372 (1989).
10. S. L. Triezenberg, K. L. LaMarco, S. L. McKnight, *ibid.*, p. 1372.
11. K. L. LaMarco, C. C. Thompson, B. P. Byers, E. M. Walton, S. L. McKnight, *Science* **253**, 789 (1991).
12. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *ibid.* **240**, 1759 (1988).
13. C. Murre, P. S. McCaw, D. Baltimore, *Cell* **56**, 777 (1989).
14. F. W. Studier and B. A. Moffatt, *J. Mol. Biol.* **189**, 113 (1986). Polymerase chain reaction was used to introduce a Bam HI site at the 5′ end of the open reading frames encoding GABPα or GABPβ1. The cDNA's were inserted into a modified pT5 vector. Proteins were expressed as described [J. D. Shuman, C. R. Vinson, S. L. McKnight, *Science* **249**, 771 (1990)]. GABPα was precipitated from the soluble fraction by the addition of one volume of 2M ammonium sulfate in buffer A (10 mM tris-HCl, pH 7.9, 100 mM KCl, 0.5 mM EDTA, 1 mM $MgCl_2$, 2 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride) with 2 mM $CaCl_2$. The precipitate was collected by centrifugation and resuspended in 25 ml of buffer B [25 mM tris-HCl, pH 8.0, 0.75 mM EDTA, 10 percent (v/v) glycerol, 1 mM DTT] with 75 mM

NaCl and dialyzed against the same buffer. The dialyzate was chromatographed on a Q-Sepharose Fast Flow (Pharmacia) column; GABPα was eluted with a 75 to 500 mM NaCl gradient in buffer B. Peak fractions were pooled, dialyzed against buffer B, and chromatographed on a salmon sperm DNA–Sepharose column. GABPα was eluted with a 0 to 400 mM NaCl gradient. GABPα accounted for more than 90 percent of the total protein by Coomassie blue staining. GABPβ1 was solubilized from the particulate fraction of bacterial extracts by sonification in buffer A with 7 M urea. The urea-solubilized fraction was dialyzed against buffer B with 75 mM NaCl and centrifuged at 16,300*g* for 1 hour. The supernatant was applied to a Q-Sepharose column and eluted with a gradient of 75 to 500 mM NaCl. GABPβ1 accounted for more than 90 percent of total protein by Coomassie blue staining.

15. Polyclonal antisera were generated by injecting rabbits with purified GABPα or GABPβ1. Antisera were added to gel-shift reactions at a dilution of 1:20. Pre-immune sera did not affect the migration of protein-DNA complexes.
16. C. C. Thompson and S. L. McKnight, unpublished observations.
17. L. A. Chodosh, in *Current Protocols in Molecular Biology*, M. Ausubel *et al.*, Eds. (Wiley, New York, 1988), vol. 2, p. 12.5.1. Cross-linking with UV was performed with an oligonucleotide composed of a GA binding site flanked by 10 bp of nonspecific sequence (5′-AACCAAGCTT*GCGGAACGGAAGCGGAAACCG*-3′). Oligonucleotides were labeled to high specific activity by the fill-in reaction with the Klenow fragment of DNA polymerase I in the presence of all four ^{32}P-labeled dNTP's. DNA binding was performed in a 96-well culture dish; the reactions were then exposed to UV. Samples were boiled in SDS-sample buffer and subjected to electrophoresis on SDS-polyacrylamide gels. Cross-linked species were visualized by autoradiography.
18. Gel filtration chromatography was performed with a Superose-6 column (10 by 300 cm; Pharmacia) in buffer B with 0.4 M NaCl. The column was calibrated with molecular mass standards thyroglobulin, apoferritin, catalase, bovine serum albumin, and ribonuclease. Each protein (50 to 100 μg) was chromatographed at 0.5 ml/min. Elution volume was converted to K_{av} by the equation $K_{av} = (V_e - V_o)/(V_t - V_o)$ where V_o = void volume = 8.1 ml; V_t = total bed volume = 24.0 ml; V_e = eluted volume. The Stokes radius was calculated from a plot of $(-\log K_{av})^{1/2}$ as a function of the Stokes radius [G. K. Ackers, *Adv. Prot. Chem.* **24**, 343 (1970)]. The GABPα eluted as a single peak at 15.2 ml; GABPβ1 at 14.0 ml; GABPβ2 at 15.8 ml. A mixture of equal amounts of GABPα and β1 chromatographed as a single peak at 12.1 ml. The mixture of GABPα and GABPβ2 chromatographed as a single peak at 13.9 ml.
19. R. G. Martin and B. N. Ames, *J. Biol. Chem.* **236**, 1372 (1961). Sedimentation coefficients were determined on 4.5-ml 10 to 30% glycerol gradients in 25 mM tris-HCl, pH 8.0, 75 mM NaCl, 0.75 mM EDTA, 1 mM DTT. Each protein (30 μg) was loaded in 0.1 ml together with catalase, bovine serum albumin (BSA), and cytochrome c as internal standards. Gradients were centrifuged at 4°C for 40 hours at 39,000 rpm (SW50.1 rotor). Fractions (0.25 ml) were collected and analyzed by SDS-PAGE with Coomassie blue staining. The *S* value for each sample was determined by its sedimentation relative to the BSA and cytochrome c standards. Native molecular masses were derived with the use of the Stokes radius and measured sedimentation coefficients as described [L. M. Siegel and K. L. Monty, *Biochim. Biophys. Acta* **112**, 346 (1966)]. Partial specific volume was calculated from the predicted amino acid sequences of each GABP subunit [E. J. Cohn and J. T. Edsall, in *Amino Acids and Peptides as Ions* and *Dipolar Ions*, E. J. Cohn and J. T. Edsall Eds. (Reinhold, New York, 1943), pp. 370–381].
20. T. A. Brown and S. L. McKnight, unpublished observations.
21. Deletion mutants of GABPα were generated by polymerase chain reaction and expressed in pT5 (*14*). Soluble bacterial extracts containing deleted variants of GABPα were used for binding reactions. NH_2-terminal deletions of GABPβ1 were generated by exonuclease III digestion, followed by digestion with S1 nuclease and ligation of Bam HI linkers. All deletions were sequenced and subcloned into the appropriate pET3 vector [A. H. Rosenberg *et al.*, *Gene* **56**, 125 (1987)] to maintain the proper reading frame. The COOH-terminal deletions were generated with the use of 3′ deletions of the cDNA inserted in Bluescript (Stratagene) and subcloning Bst EII–Asp718 or Sac I–Asp718 fragments into the pET-GABPβ1 plasmid that had been digested with the appropriate enzymes. Translation termination codons were provided by vector sequences so that, in some cases, extra amino acids were appended to the open reading frame. All GABPβ derivatives were insoluble and were resolubilized in 8 M urea; the solubilized protein was dialyzed against 10 mM tris, pH 8.0, 75 mM KCl or NaCl, 1 mM DTT, 0.2 mM PMSF, 1 mM benzamidine, 10 percent glycerol before being used in binding reactions. All derivatives were expressed equivalently as determined by Coomassie staining of SDS polyacrylamide gels.
22. C. V. Gunther, J. A. Nye, R. S. Bryner, B. J. Graves, *Genes Dev.* **4**, 667 (1990).
23. M. J. Klemsz, S. R. McKercher, A. Celada, C. Van Beveren, R. A. Maki, *Cell* **61**, 113 (1990); M. G. Goebl, *ibid.*, p. 1165.
24. L. D. Urness and C. S. Thummel, *ibid.* **63**, 47 (1990).
25. F. D. Karim *et al.*, *Genes Dev.* **4**, 1451 (1990).
26. P. Bauerle and D. Baltimore, *Cell* **53**, 211 (1988); *Science* **242**, 540 (1988).
27. R. Sen and D. Baltimore, *Cell* **47**, 921 (1986).
28. We thank Se-Jin Lee for sharing intellectual and material resources, J. Shuman for helpful discussions, P. Lamb for advice on UV cross-linking, W. Kupiec for help with computer graphics, and our colleagues at the Carnegie Embryology Department for advice and encouragement. This work was supported by the Howard Hughes Medical Institute and the Carnegie Institution of Washington. C.T.T. was supported by a postdoctoral fellowship from the Jane Coffin Childs Fund for Medical Research.

3 June 1991; accepted 2 July 1991



"I don't know why--I just* like *to attack the upper respiratory tract."