Our results suggest that TFIID may have a broader function in gene expression than previously appreciated (*21, 22*). The possibility that a transcription factor is shared by pol B and pol C correlates well with previous indications that upstream activating factors stimulate transcription by the B and C enzymes (*2, 23*). These upstream activator proteins may act via TFIID (*22, 24*). Participation of the TATA factor in the transcription of pol C genes may not be restricted to the U6 (and 7SK) genes, as upstream TATA-like elements have been found in a number of other genes transcribed by pol C (*1*).

In pol B genes, the interaction of TFIID with the TATA element is thought to promote preinitiation complex formation by favoring subsequent binding of TFIIB (or a preformed pol B–TFIIB complex) (*22, 25*). In the case of the U6 gene, TFIID may favor the assembly of TFIIIB that, by itself, does not bind DNA (*16, 18*). A functional and evolutionary relationship may therefore exist between TFIIIB and TFIIB or some other general pol B factor. The relatedness of class C and class B transcription factors would thus parallel the close evolutionary relationship of pol B and pol C, as these two enzymes were found more closely related to each other than to enzyme A (*26*).

REFERENCES AND NOTES

1. E. P. Geiduschek and G. Tocchini-Valentini, *Annu. Rev. Biochem.* **57**, 873 (1989).
2. B. Sollner-Webb, *Cell* **52**, 153 (1988); S. Murphy, B. Moorefield, T. Pieler, *Trends Genet.* **5**, 122 (1989); J. M. Palmer and W. R. Folk, *Trends Biochem. Sci.* **15**, 300 (1990).
3. S. Murphy, C. Di Liegro, M. Melli, *Cell* **51**, 81 (1987).
4. P. Carbon *et al.*, *ibid.*, p. 71; C. Bark, P. Weller, J. Zabielski, L. Janson, U. Petterson, *Nature* **328**, 356 (1987); G. Das, D. Henning, D. Wright, R. Reddy, *EMBO J.* **7**, 503 (1988); H. Kleinert and B. J. Benecke, *Nucleic Acids Res.* **16**, 1319 (1988).
5. I. W. Mattaj, N. A. Dathan, H. D. Parry, P. Carbon, A. Krol, *Cell* **55**, 435 (1988).
6. S. M. Lobo and N. Hernandez, *ibid.* **58**, 55 (1989); G. R. Kunkel and T. Pederson, *Nucleic Acids Res.* **17**, 7371 (1989).
7. G. Waibel and W. Filipowicz, *Nucleic Acids Res.* **18**, 3451 (1990); *Nature* **346**, 199 (1990).
8. A. Moenne *et al.*, *EMBO J.* **9**, 271 (1990).
9. D. A. Brow and C. Guthrie, *Nature* **334**, 213 (1988).
10. TFIIIB activity was purified successively by phosphocellulose, heparin-ultrogel, and Cibacron blue–agarose chromatography as described (*8*), then subjected to fast protein liquid chromatography (FPLC) on a Superose 12 column (Pharmacia) equilibrated in 20 mM Hepes-KOH, pH 7.9, 0.5 mM EDTA, 10 mM 2-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride, 10% glycerol, and 370 mM KCl. Protein fractions (250 μl) eluted with the same buffer were assayed for transcription factor activity with the $tRNA_3^{glu}$ or the U6 genes. Molecular sizes of proteins were estimated with size markers (Bio-Rad protein standards). Transcription mixtures (*8*) (40 μl) received 7-μl aliquots of different column fractions as indicated, and contained 120 mM KCl (final concentration), purified pol C (50 ng), and pTaq6 plasmid DNA (150 ng) (*9*). For tRNA synthesis, the mixtures were supplemented with factor Tau (75 ng). After a 50-min incubation at 25°C, transcripts were analyzed by polyacrylamide-urea gel electrophoresis and autoradiography (*8*). For the purification of yeast pol C and factor Tau, see O. Gabrielsen, N. Marzouki, A. Ruet, A. Sentenac, P. Fromageot, *J. Biol. Chem.* **264**, 7505 (1989).
11. This nuclease activity cleaves phenol-extracted U6 RNA (114 nt) in transcription buffer. The activity requires Mg^{2+} ions and is inhibited at high salt concentrations (F. Margottin and A. F. Burnol, unpublished data). A similar activity that selectively trims mouse U6 RNA has been described [D. I. Lee, H. Hirai, S. Natori, K. Sekimizu, *J. Biochem.* **105**, 526 (1989)].
12. F. Margottin and G. Dujardin, unpublished observations.
13. B. Cavallini *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9803 (1989).
14. B. Cavalllini *et al.*, *Nature* **334**, 77 (1988).
15. For S1 mapping, the RNA from 200-μl transcription reactions was annealed to a 191-nt Taq I–Eco NI fragment ^{32}P-labeled at the 5′ end on the coding strand (*8*), and samples were processed according to a standard S1 mapping procedure [T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1983)]. Partially purified U6 RNA consisted of a preparation of yeast RNA enriched for small RNA species by gel electrophoresis and contained predominantly 5S RNA.
16. G. A. Kassavetis, B. R. Braun, L. U. Nguyen, E. P. Geiduscheck, *Cell* **60**, 235 (1990).
17. M. S. Klekamp and P. A. Weil, *J. Biol. Chem* **261**, 2819 (1986).
18. G. A. Kassavetis, D. L. Riggs, R. Negri, L. H. Nguyen, E. P. Geiduschek, *Mol. Cell. Biol.* **171**, 2551 (1989).
19. D. Brow and C. Guthrie, *Genes Dev.* **4**, 1345 (1990).
20. In preliminary experiments with a U6 gene that harbors the B-block, we found that peak 2 fractions (TFIID) were still required for U6 RNA synthesis, even in the presence of Tau (F. Margottin and A. F. Burnol, unpublished observations).
21. B. L. Davison, J. M. Egly, E. R. Mulvihill, P. Chambon, *Nature* **301**, 680 (1983); S. Hahn, S. Buratowski, P. A. Sharp, L. Guarente, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5718 (1989).
22. M. Sawadogo and A. Sentenac, *Annu. Rev. Biochem.* **59**, 711 (1990).
23. S. Murphy, A. Pierani, C. Scheidereit, M. Melli, R. G. Roeder, *Cell* **59**, 1071 (1989).
24. M. Ptashne and A. A. F. Gann, *Nature* **346**, 329 (1990).
25. S. Buratowski, S. Hahn, L. Guarente, P. A. Sharp, *Cell* **56**, 549 (1990).
26. S. Mémet, W. Saurin, A. Sentenac, *J. Biol. Chem.* **263**, 10048 (1988).

27. We thank A. F. Burnol for her help at the final stage of this work, C. Carle, O. Lefebvre, M. Riva, and especially A. Ruet for constant discussions and various materials, and I. Faus for yIID expression vector. G.D. was on secondment from the Centre National de la Recherche Scientifique. Supported by a grant from the Ministère de la Recherche et de la Technologie (F.M.).

28 August 1990; accepted 9 November 1990

A Genetic Model for Interaction of the Homeodomain Recognition Helix with DNA

STEVEN D. HANES AND ROGER BRENT

The Bicoid homeodomain protein controls anterior development in the *Drosophila* embryo by binding to DNA and regulating gene expression. With the use of genetic assays in yeast, the interaction between the Bicoid homeodomain and a series of mutated DNA sites was studied. These experiments defined important features of homeodomain binding sites, identified specific amino acid–base pair contacts, and suggested a model for interaction of the recognition α-helices of Bicoid and Antennapedia-class homeodomain proteins with DNA. The model is in general agreement with results of crystallographic and magnetic resonance studies, but differs in important details. It is likely that genetic studies of protein-DNA interaction will continue to complement conventional structural approaches.

THE GENE *bicoid* ENCODES A PROTEIN morphogen (Bicoid) that is required for anterior development (*1*). Like many regulatory proteins important for development, Bicoid contains a 60–amino acid sequence known as the homeodomain (*2, 3*). Bicoid exerts its effects, in part, by activating expression of zygotic genes such as *hunchback* and *orthodenticle* (*4*). Bicoid binds to the sequence TCTAATCCC and close variants repeated in the 5′ regulatory region of *hunchback* (*5*). The Antennapedia (Antp) class of homeodomain proteins, such as those encoded by *Antennapedia*, *fushi tarazu*, and *Ultrabithorax*, bind the sequence TCAATTAAAT, which was first identified upstream of *engrailed* (*6–8*), a gene involved in segmentation.

The homeodomain contains a structure similar to the helix-turn-helix motif of prokaryotic transcriptional repressors (*9, 10*). Recognition of specific DNA sites by homeodomain proteins depends on the second α-helix (recognition helix) of this motif (*11, 12*). Bicoid and Antp-class proteins use Lys or Gln, respectively, at position 9 of the recognition helix (Fig. 1A) to distinguish between related binding sites (*11, 12*). When Lys^9 in the Bicoid recognition helix is replaced by Gln, the mutant protein (Bicoid-Q_9) no longer recognizes Bicoid sites, but instead recognizes Antp-class sites (*11*).

Bicoid and Antp-class proteins expressed

Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, and Department of Genetics, Harvard Medical School, Boston, MA 02115.

in yeast stimulate transcription of target genes that contain functional binding sites (*11, 13*). In this report, we measured Bicoid-dependent transcription from *GAL1-lacZ* constructs that contained multiple copies of wild-type or mutant binding sites. We changed nucleotides in the Antp-class binding site to those found at equivalent positions in the Bicoid binding site (and vice versa) to generate a series of sites with switched specificities for protein binding. This strategy allowed us to delineate the minimal site requirements for homeodomain recognition and to identify base pairs in each site that determine which class of protein is bound.

Experiments (Table 1) designed on the basis of an alignment of binding sites as they occur upstream of regulated genes (Fig. 1, B and C) indicated the importance of base pairs 7 and 8, but did not produce sites with switched specificities. This prompted us to devise an alternative alignment with the use of the complementary strand of the Antp site (Fig. 1D).

On the basis of the new alignment, we replaced CCC at positions 7, 8, and 9 in the Bicoid site by TGA, which is found at these positions in the Antp site (mutant Bicoid site 9, Table 2A). The mutant site had switched specificity; it was recognized by Bicoid-Q_9, but not by wild-type Bicoid. We then tested a series of sites to determine the individual contributions of base pairs 7, 8, and 9 (sites 10 to 17). Our results showed that base pair 7 was critical for site recognition by Bicoid-Q_9 (see especially sites 9 and 15 to 17). In addition, base pair 8 (sites 12 to 14 and 24) (*14*) and base pair 9 (sites 9 to 12) influenced recognition by Bicoid-Q_9; purines were preferred over pyrimidines on the upper strand.

We performed reciprocal experiments to switch the specificity of an Antp site to that of a Bicoid site (Table 2B). An Antp site that contained CC in positions 7 and 8 was recognized by wild-type Bicoid but not Bicoid-Q_9 (site 18). Results with sites 18 to 25 showed that base pair 7 was also critical for recognition by wild-type Bicoid (see especially sites 18 versus 22) and that pyrimidines (preferably cytosine) were favored over purines on the upper strand at base pair 8 (sites 5 and 18 to 21).

We draw a number of conclusions from the above experiments. In all sites with switched specificity, the identity of base pair 7 is decisive; wild-type Bicoid requires C:G, while Bicoid-Q_9 requries T:A, indicating that Bicoid and Antp-class homeodomain

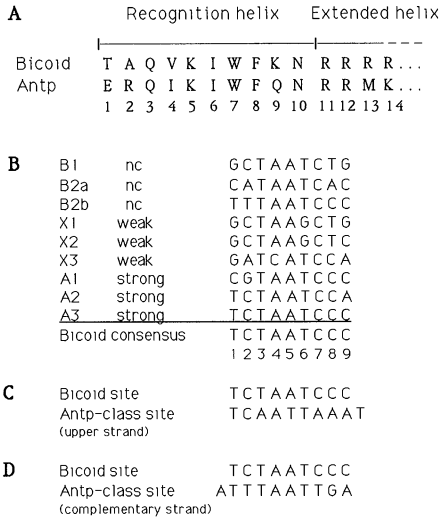

Fig. 1. Sequences of the Bicoid and Antennapedia recognition helices and their cognate binding sites. (**A**) Recognition helices of Bicoid and Antp-class protein also showing the extended helix (helix IV) (*10*). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**B**) Alignment of naturally occurring Bicoid sites found in the regulatory region of *hunchback*. Strong sites A1 to A3 and B1 to B2b and weak sites X1 to X3 have been described (*5*). Strength of sites B1 to B2 was not characterized. (**C**) Consensus Bicoid and Antp-class binding sites aligned 5′ to 3′ as they occur upstream of regulated genes. (**D**) Consensus Bicoid site aligned with the complementary strand of the Antp-class binding site.

Table 1. Recognition of mutant binding sites by wild-type Bicoid and altered specificity Bicoid-Q_9 proteins. Yeast cells were cotransformed with plasmids that encoded Bicoid activator proteins and plasmids that contained a target gene that carried the binding sites to be tested (*27*). Bicoid proteins were expressed as fusions to the *Escherichia coli* LexA protein (*11, 27*). DNA binding–dependent stimulation of *GAL1*-lacZ target genes was measured as described (*11*). Designations are: (+++) >100 units of β-galactosidase activity; (++) 20 to 100 units; (+) 2 to 20 units; (+/−) 0.1 to 2 units; and (−) <0.1 unit. The number of binding sites for constructs 1 to 7 was between 6 and 14.

			Bicoid		Bicoid-Q_9	
		123456789				
Wild-type Bicoid site		TCTAATCCC	+++	(106)	−	(<0.1)
Mutant Bicoid site	(1)	TCTAATAAC	−	(<0.1)	−	(<0.1)
Mutant Bicoid site	(2)	TCTAATCAC	+	(1.2)	−	(<0.1)
Mutant Bicoid site	(3)	TCTAATACC	−	(<0.1)	−	(<0.1)
Mutant Bicoid site	(4)	TCTAATTTC	−	(<0.1)	−	(<0.1)
Mutant Bicoid site	(5)	TCTAATCTC	++	(75)*	−	(<0.1)
Mutant Bicoid site	(6)	TCTAATTCC	−	(<0.1)	−	(<0.1)
Mutant Bicoid site	(7)	TCTAAACCC	−	(<0.1)	−	(<0.1)
Wild-type Antp site		ATTTAATTGA	−	(<0.1)	++	(97)
Mutant Antp site	(8)	ATTTAAATGA	−	(<0.1)	−	(<0.1)

*The target plasmid that bore mutant site 5 contained a large number of binding sites, 14, which may explain the high levels of activation compared to sites 18 and 21 (Table 2).

Table 2. Switching the specificity of Bicoid and Antp binding sites. Recognition of mutant sites by Bicoid and altered specificity Bicoid-Q_9 was determined as described in Table 1. Here, each target gene was constructed (*28*) so that it contained exactly six copies of each binding site oriented 5′ to 3′ as shown in the table. Designations for units of β-galactosidase activity (+, −) are as in Table 1.

			Bicoid		Bicoid-Q_9	
(**A**)		123456789				
Wild-type Bicoid site		TCTAATCCC	+++	(106)	−	(<0.1)
Mutant Bicoid site	(9)	TCTAATTGA	−	(<0.1)	++	(39)
Mutant Bicoid site	(10)	TCTAATTGC	−	(<0.1)	+	(17)
Mutant Bicoid site	(11)	TCTAATTGT	−	(<0.1)	+	(8)
Mutant Bicoid site	(12)	TCTAATTGG	−	(<0.1)	++	(44)
Mutant Bicoid site	(13)	TCTAATTAG	−	(<0.1)	++	(79)
Mutant Bicoid site	(14)	TCTAATTTG	−	(<0.1)	+	(4)
Mutant Bicoid site	(15)	TCTAATCGA	−	(<0.1)	−	(<0.1)
Mutant Bicoid site	(16)	TCTAATAGA	−	(<0.1)	−	(<0.1)
Mutant Bicoid site	(17)	TCTAATGGA	−	(<0.1)	−	(<0.1)
(**B**)		0123456789				
Wild-type Antp site		ATTTAATTGA	−	(<0.1)	++	(97)
Mutant Antp site	(18)	ATTTAATCCA	++	(70)	−	(<0.1)
Mutant Antp site	(19)	ATTTAATCGA	−	(<0.1)	−	(<0.1)
Mutant Antp site	(20)	ATTTAATCAA	−	(<0.1)	+	(7)*
Mutant Antp site	(21)	ATTTAATCTA	+	(10)	−	(<0.1)
Mutant Antp site	(22)	ATTTAATTCA	−	(<0.1)	−	(<0.1)
Mutant Antp site	(23)	ATTTAATTCT	−	(<0.1)	−	(<0.1)
Mutant Antp site	(24)	ATTTAATTCG	−	(<0.1)	+/−	(0.8)
Mutant Antp site	(25)	ATTTAATTCC	+/−	(1.8)	−	(<0.1)

*Activation of mutant site 20 by Bicoid-Q_9 may have resulted fortuitously from the presence of a functional Antp binding site on the complementary strand: TTAAATGATTGA, where the close match to the ideal site (TAATTGA) is underlined. Alternatively, activation may reflect an interaction of Gln^9 with the thymine of base pair 8 in this site as seen in the crystal structure (*21*).

proteins discriminate between their sites by an interaction between recognition helix residue 9 and base pair 7.

Base pairs 8 and 9 also influence recognition, perhaps by providing weak contacts that cannot be strictly base-specific, as several combinations of base pairs are permissible. For example, at position 8, Bicoid-Q_9 tolerates A, G, and, to a lesser extent, T (sites 9, 13, and 14), while wild-type Bicoid tolerates C and, to a lesser extent, T (sites 18 and 21).

Base-specific contacts to positions 1 and 2 are not critical for binding, because Bicoid and Bicoid-Q_9 tolerate either T or C at position 2 (for example, site 18 versus a wild-type Bicoid site). In addition, a site in which positions 1 and 2 are CG instead of TC is still bound by Bicoid (Fig. 1B, site A1) (*5*).

The TAAT motif, which is conserved in nearly all homeodomain binding sites (*3*), is not used to discriminate between Bicoid and Antp binding sites, because complete switches in specificity are obtained with sites in which the TAAT is unchanged (for example, sites 9 and 18). Its importance, however, is inferred from alkylation interference experiments (*15*), in vitro binding studies (*7*), and from our own experiments, showing that T at position 6 is crucial for recognition by both Bicoid and Bicoid-Q_9 (sites 7 and 8). It is likely that the TAAT is necessary to distinguish these sites from nonspecific DNA. Thus, the homeodomain binding site consists of two subsites, a common TAAT core element and specificity-determining bases that lie 3′ to it. In the sites examined here, the closest base pair to the 3′ end of the TAAT element (TAATN) contributed the most to binding specificity. The fact that switched specificity sites are obtained with nonsymmetrically disposed base substitutions in one-half of the site suggests that each site is bound by a single protein monomer. This is consistent with recent biochemical studies (*15*) and with the general lack of symmetry in homeodomain binding sites (*7*).

Our results strongly suggest that interactions between the recognition helix and DNA occur in the major groove. This conclusion follows from the fact that Bicoid-Q_9 distinguishes between T:A and A:T at position 7 (sites 9 versus 16). These base pairs display distinct functional groups in the major groove, but almost identical groups in the minor groove (*16*). Similarly, wild-type Bicoid distinguishes between base pairs C:G and G:C at positions 7 and 8 (sites 17 versus 18), which are nearly indistinguishable in the minor groove (*16*). Finally, contacts made to base pair 6 in each site are also likely to be in the major groove, because Bicoid

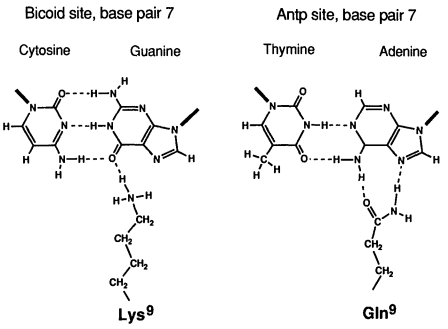

Fig. 2. Proposed amino acid 9–base pair 7 interactions. Lys^9 of the Bicoid recognition helix donates a hydrogen bond to O6 or N7 of guanine in the Bicoid binding site. It is possible that Lys^9 might also donate a hydrogen bond to the guanine of adjacent base pair 8 as shown for λ cro protein (*17*). In Bicoid-Q_9, the Gln at position 9 of the recognition helix donates a hydrogen bond to N7 and accepts a hydrogen bond from N6 of the adenine of base pair 7 in the Antp site. Similar contacts may be possible for Gln at position 9 of all Antp-class homeodomain proteins.

proteins distinguish T:A from A:T base pairs at this position (sites 7 and 8 versus wild type). On the basis of known contacts made by Lys^{32} in λ cro (*17*), Lys^4 and Gln^{44} in λ repressor (*18*), and Gln^{28} in 434 repressor (*19*), we propose the interactions depicted in Fig. 2. Here, Lys^9 contacts guanine of the C:G base pair and Gln^9 contacts adenine of the T:A base pair.

The fact that amino acid 9 in both proteins interacts with the same base pair in both Bicoid and Antp sites implies that their recognition helices contact DNA in the same way. The Antp-class protein, fushi-tarazu (Ftz), which has Gln at position 9, makes an equivalent contact to base pair 7 (*20*), suggesting that Bicoid and Antp-class proteins recognize DNA in a fundamentally similar way.

What other amino acids in the recognition helix contact DNA? We reasoned that contact amino acids should be near residue 9, on the same side of the α-helix and able to hydrogen-bond with base pairs in DNA. If such amino acids were conserved, they might contact conserved features of each site. For Bicoid, candidate residues that are conserved include Lys^5, Asn^{10}, and Arg^{12} (Lys^{46}, Asn^{51}, and Arg^{53} in Antp) (*10*) (Fig. 1A.). A contact by Lys^5 seemed unlikely, because replacement by Ala did not affect binding (*11*).

We tested whether Asn^{10} and Arg^{12} made essential DNA contacts by introducing $Asn^{10} \rightarrow$ Ala, $Asn^{10} \rightarrow$ Gln, or $Arg^{12} \rightarrow$ Ala substitutions into wild-type Bicoid. As in previous experiments (*11*) (see also legend to Table 1), Bicoid proteins were expressed as fusions to the bacterial LexA protein. All three mutants were deficient for recognition of both the Bicoid site and a switched-specificity Antp-class site (Table 3). The mutant proteins were apparently specifically defective in homeodomain-DNA binding; they were stably expressed, as demonstrated by immunoblot analysis, and they activated LexA operator-containing target genes, indicating that their nuclear transport, oligomerization, and gene activation functions were maintained (*20*). While such loss of function experiments cannot establish precise contacts, the results suggest that Asn^{10} and Arg^{12} contact conserved features of the two sites. If these contacts are sequence-specific, they would be made to the conserved TAAT element, which would indicate that the recognition helix is aligned with its NH_2-terminus toward the 3′ end of the binding site (Fig. 3).

Our model suggests that like prokaryotic helix-turn-helix proteins, the homeodomain recognition helix makes specific DNA contacts in the major groove. However, the geometry of this interaction is quite different; unlike prokaryotic proteins, residues toward the COOH-terminus make these contacts. Our experiments show that Lys^9 of the Bicoid recognition helix contacts base pair 7 of the Bicoid site (TCTAATCCC), while Gln substituted at this position (as in Antp-class proteins) contacts the equivalent

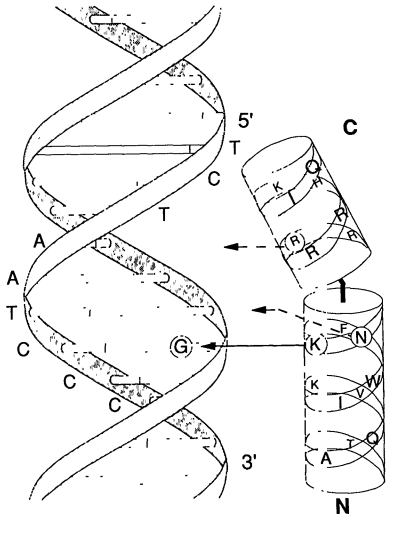

Fig. 3. Mapping contact amino acids onto the Bicoid binding site. Contact between Lys^9 in the Bicoid recognition helix and base pair 7 is shown by a solid line. In Antp-class proteins position 9 is Gln and would contact an adenine at position 7 of the Antp site (not shown). Hypothetical contacts between recognition helix Asn^{10}, Arg^{12}, and the TAAT motif are indicated by dashed arrows. The Bicoid binding site (5′-TCTAATCCC-3′) is shown as B-form DNA. Amino acids are drawn on a recognition helix based on the proposed NMR model of the Antp homedomain. In this model, the α-helix extending beyond residue 10 is kinked at a 30° angle from the rest of the recognition helix (*10*), although a recent cocrystal structure shows that it might be straight (*21*).

Table 3. Elimination of DNA binding by mutations in Asn^{10} and Arg^{12}. Mutant proteins (*29*) were tested for recognition of Bicoid binding sites (TCTAATCCC) and switched-specificity (sw) Antp-class sites (site 18, ATTTAATCCA) as described in Table 1. Results given in units of β-galactosidase activity have been normalized to activity obtained when assayed on *LexA* operator-containing target genes.

Producer plasmid	Target plasmid	
	Bicoid site	Antp site (sw)
Bicoid wild-type	119	116
$Asn^{10} \rightarrow$ Ala	6	1
$Asn^{10} \rightarrow$ Gln	2	0.5
$Arg^{12} \rightarrow$ Ala	3	1

base pair in the Antp-class site (TTTAATTGA). They show that conserved residues Asn^{10} and Arg^{12} contact DNA and suggest that the recognition helix is aligned COOH-terminus to NH_2-terminus (C to N), 5′ to 3′ with the binding site. In this model, NH_2-terminal residues 1 and 2 of the recognition helix do not contact DNA, and we suggest that their partial sequence conservation (*3*) may reflect interactions with a limited set of cellular proteins that modulate homeodomain DNA binding or gene regulation.

This picture is in good agreement with a recently reported crystal structure of an Engrailed-DNA complex (*21*) and a model of an Antennapedia-DNA complex proposed from nuclear magnetic resonance (NMR) experiments (*22*). In all three models (*23*), the COOH-terminus of the recognition helix inserts in the major groove, and the helix is tilted in a C to N, 5′ to 3′ configuration. In all three models, Asn^{10} and Arg^{12} (Asn^{51} and Arg^{53} in Engrailed and Antp) are in position to contact DNA. In the cocrystal, Asn^{10} contacts a base in the core of the site (TAAT), while Arg^{12} makes a backbone contact outside of the TAAT. In the NMR model, contacts by Asn^{10} and Arg^{12} were not observed, perhaps because not all of the protein-DNA NOEs (nuclear Overhauser effects) were identified.

There are important differences between the models. In each model, the side chain of recognition helix residue 9 (Gln^{50} in Engrailed and Antp) makes a different major groove contact. In the cocrystal, this Gln is in position to hydrogen bond with the adenine of base pair 7, as we have deduced, but instead makes van der Waals contact with thymine of base pair 8 (TAATNN). We think it unlikely that any such contact determines specificity in vivo, because our functional assays show sites that do not have thymine at position 8 are recognized by Bicoid-Q_9 (sites 9 to 12 and 14) and, conversely, sites that do have thymine at position 8 but have the wrong base at position 7 are not recognized by Bicoid-Q_9 (sites 1 and 2). Moreover, we note that the proposed bidentate hydrogen bond contact between Gln and adenine (Fig. 2) would contribute more free energy to binding. As the authors have suggested (*21*), the contact observed in the cocrystal might have been caused by distortions in the DNA caused by adventitious binding of a second protein monomer to the end of the fragment; alternatively, it may result from the solvent used to induce crystal formation, or it may be a nonspecific contact observed only at the high protein concentration in the cocrystal.

In the NMR model, residue 9 contacts base pair 7, but the site contains G:C instead of T:A at this position. This interaction may be specific, but, because cytosine can only make a single hydrogen bond contact with Gln, it is unlikely to be as energetically favorable as an interaction with adenine. In fact, Ftz, which contains Gln at residue 9, binds in vitro to sites with G:C at position 7 (TAATG) about one-tenth as strongly as it binds to sites with T:A at this position (TAATT) (*24*), and at the protein concentration found in vivo, neither Bicoid-Q_9 (site 17) nor Ftz (*20*) recognizes the G:C-containing site. On the basis of these results, we suggest that, in *Drosophila*, T:A will be the preferred base pair.

Finally, both the cocrystal structure and NMR model show several DNA contacts not addressed in our experiments [for example, a contact between Ile^6 in the recognition helix (Ile^{47} in Engrailed and Antp) and the TAAT, and contacts made by amino acids outside of the recognition helix]. These findings illustrate an important shortcoming of our approach; that is, it cannot detect interactions that we do not explicitly investigate. However, when used to test specific structural ideas, we believe that simple genetic and biochemical assays will continue to provide independent insight into structural issues (*25*) and occasional correction (*26*) of conclusions derived from conventional structural methods.

REFERENCES AND NOTES

1. H.-G. Frohnhöfer and C. Nüsslein-Volhard, *Nature* **324**, 120 (1986).
2. W. J. Gehring, *Science* **236**, 1245 (1987); S. Hayashi and M. P. Scott, *Cell* **63**, 883 (1990).
3. M. P. Scott, J. W. Tamkun, G. W. Hartzell III, *Biochim. Biophys. Acta* **989**, 25 (1989).
4. W. Driever and C. Nüsslein-Volhard, *Cell* **54**, 95 (1988); R. Finkelstein and N. Perrimon, *Nature* **346**, 485 (1990).
5. W. Driever and C. Nüsslein-Volhard, *Nature* **337**, 138 (1989); W. Driever, G. Thoma, C. Nüsslein-Volhard, *ibid.* **340**, 363 (1989).
6. C. Desplan, J. Theis, P. H. O'Farrell, *ibid.* **318**, 630 (1985); T. Hoey and M. Levine, *ibid.* **332**, 858 (1988).
7. C. Desplan, J. Theis, P. H. O'Farrell, *Cell* **54**, 1081 (1988).
8. M. Müller *et al.*, *EMBO J.* **7**, 4299 (1988).
9. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984); A. Laughon and M. P. Scott, *Nature* **310**, 25 (1984).
10. Y. Q. Qian *et al.*, *Cell* **59**, 573 (1989).
11. S. D. Hanes and R. Brent, *ibid.* **57**, 1275 (1989).
12. J. Treisman *et al.*, *ibid.* **59**, 553 (1989).
13. V. D. Fitzpatrick and C. J. Ingles, *Nature* **337**, 666 (1989); M.-L. Samson, L. Jackson-Grusby, R. Brent, *Cell* **57**, 1045 (1989); G. Struhl, K. Struhl, P. M. Macdonald, *ibid.*, p. 1259.
14. Although we tested the interaction of Bicoid-Q_9 with base pair 7 and 9 in the context of the TGA triplet, we did not do so for base pair 8. This was because one of the three possible position 8 changes in this context would have fortuitously created a binding site of a different sequence on the complementary strand, complicating our analysis. In the TGA context, the G → A transition TCTAATTGA (site 9) to TCTAATTAA would create the sequence TAATTAG on the opposite strand, which is itself a strong site (see site 13). Instead, we examined base pair 8 in the TNG context, in which the G → A transition, TCTAATTGG (site 12) to TCTAATTAG (site 13) created the same site on both strands (TAATTAG).
15. M. Affolter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4093 (1990).
16. N. C. Seeman, J. M. Rosenberg, A. Rich, *ibid.* **73**, 804 (1976).
17. D. H. Ohlendorf *et al.*, *Nature* **298**, 718 (1982).
18. S. R. Jordan and C. O. Pabo, *Science* **242**, 893 (1988).
19. A. K. Aggarwal, D. W. Rodgers, M. Drotter, M. Ptashne, S. C. Harrison, *ibid.*, p. 899.
20. S. D. Hanes and R. Brent, unpublished data.
21. C. R. Kissinger *et al.*, *Cell* **63**, 579 (1990).
22. G. Otting *et al.*, *EMBO J.* **9**, 3085 (1990).
23. The experiments presented herein were conducted independently and without prior knowledge of the results of crystallographic (*21*) and NMR (*22*) studies. Therefore, we first describe a model consistent with our genetic data and then compare it with models derived from the direct structural studies.
24. L. Pick, A. Schier, M. Affolter, T. Schmidt-Glenewinkel, W. J. Gehring, *Genes Dev.* **4**, 1224 (1990).
25. R. P. Wharton and M. Ptastne, *Nature* **316**, 601 (1985); A. Hochschild and M. Ptashne, *Cell* **44**, 925 (1986); A. Hochschild, J. Douhan II, M. Ptashne, *ibid.* **47**, 807 (1986); R. P. Wharton and M. Ptashne, *Nature* **326**, 888 (1987); N. Lehming *et al.*, *EMBO J.* **6**, 3145 (1988); N. Lehming, J. Sartorius, B. von Wilcken-Bergmann, B. Kisters-Woike, B. Müller-Hill, *ibid.* **9**, 615 (1990).
26. D. Staacke, B. Walter, B. Kisters-Woike, B. von Wilcken-Bergmann, B. Müller-Hill, *EMBO J.* **9**, 1963 (1990).
27. Plasmids that direct the synthesis of Bicoid and Bicoid-Q_9 proteins in yeast have been described (*11*). Bicoid proteins were expressed as fusions to the NH_2-terminus of the bacterial LexA protein [R. Brent and M. Ptashne, *Cell* **40**, 729 (1985)]; this moiety does not contribute to DNA-binding specificity for either Bicoid or Antp sites (*11*). Target genes were constructed by insertion of Xho I-compatible oligonucleotides into pLR1Δ1 [R. W. West, R. R. Yocum, M. Ptashne, *Mol. Cell. Biol.* **4**, 2467 (1986)] at the unique Xho I site 167 base pairs upstream of the transcription start site of *GAL1-lacZ*. Mutant site target genes (sites 1 to 7) were constructed by insertion of 26–base pair complementary oligonucleotides of the following general form: upper strand, 5′-TCGAA(TCTAATNNC)TA-(TCTAATNNC)T-3′; lower strand, 5′-TCGAA(G-/hdNNATTAGA)TA(GNNATTAGA)T-3′. The binding sites are shown in parentheses separated by a 2–base pair spacer sequence. These constructs contained between three and seven oligonucleotide inserts (6 to 14 binding sites) in various orientations and thus are not directly comparable. Mutant 8 contains a single insert of the complementary 54-base oligonucleotides, 5′-TCGAG(TCATTTAAATGA)$_4$C-3′ and 5′-TCGAG(TCATTTAAATGA)$_4$C-3′.
28. Target gene constructs used in Table 2 were constructed with oligonucleotide pairs that carried six binding sites. Orientations of sites with respect to

the transcription start site are 5′ to 3′ as shown in Table 2. β-Galactosidase activity for mutant sites 9 to 25, wild-type Bicoid, and wild-type Antp sites are directly comparable because the number and orientation of binding sites were identical in each construct as confirmed by DNA sequencing. The general form for each pair of oligonucleotides is as follows:

Bicoid sites: upper strand, 5′-TCGAA(TCTAATNNNTA)$_5$(TCTAATNNN)T-3′; lower strand, 5′-TCGAA(NNNATTAGA)(TANNNATTAGA)$_5$T-3′.

Antp sites: upper strand, 5′-TCGAA(TCATTTAATNNN)$_6$T-3′; lower strand, 5′-TCGAA(NNNATTAAATGA)$_6$T-3′.

29. Asn^{10} → Ala; Arg^{12} → Ala; and Asn^{10} → Gln substitutions in the Bicoid recognition helix were created with an overlapping polymerase chain reaction (PCR) strategy as described [R. M. Horton *et al.*, *BioTechniques* **8**, 538 (1990)]. The *bicoid*-containing Eco RI fragment of pSH11-1 (*11*) was subcloned into pUC119 [J. Vieira and J. Messing, *Methods Enzymol.* **153**, 3 (1987)]. Two inner oligonucleotide primers were used to introduce the mutations in two separate PCR reactions extending in opposite directions from the site of the base changes. The purified products of the first sets of reactions were annealed and used for a second round of PCR with only outer primers to amplify an intact Sal I–Sal I fragment, which was then recloned into the *bicoid* backbone in pUC119. Vent polymerase was used for all PCR reactions (New England Biolabs). Inner primer oligonucleotides were as follows:

Asn^{10} → Ala: 5′-GGTGAAGATCTGGTTTAAGGCCCGTCGGCG-3′ and 5′-CGCCGACGGGCCTTAAACCAGATCTTCACC-3′.

Arg^{12} → Ala: 5′-GTGAAGATCTGGTTTAAGAACCGTGCGCGTCGTC-3′ and 5′-GACGACGCGCACGGTTCTTAAACCAGATCTTCACC-3′.

Asn^{10} → Gln: 5′-GGTGAAGATCTGGTTTAAGCAACGTCGGCGTC-3′ and 5′-CGACGCCGACGTTGCTTAAACCAGATCTTCACC-3′.

Outer primer oligonucleotides were: 5′-AGCGGATAACAATTTCACACAGGA-3′ (reverse primer for pUC series) and 5′-GGCCGCCATTGACATTGGTCGACCCAG-3′ [base pairs 746 to 772 of the *bicoid* cDNA, as in T. Berleth *et al.*, *EMBO J.* 7, 1749 (1988)]. Base changes that created the desired amino acid substitution and a silent mutation that created a diagnostic Bgl II site are underlined. After reconstruction of the mutant genes in pUC119, the Eco RI *bicoid* fragments were reinserted into yeast expression vector, pSH2-1(*11*). Base substitutions were confirmed by DNA sequencing. Independently derived clones for each mutant protein behaved identically in a yeast transcription assay.

30. We are very grateful to E. Golemis and A. Hochschild for helpful discussions and to J. Szostak, A. Ellington, A. Ephrussi, and R. Finley for comments on the manuscript. Supported by the Anna Fuller Fund (S.D.H.), the Pew Scholars Program (R.B.), and Hoechst AG.

13 November 1990; accepted 7 December 1990

T Cell Receptor Peptide Therapy Triggers Autoregulation of Experimental Encephalomyelitis

HALINA OFFNER,* GEORGE A. HASHIM, ARTHUR A. VANDENBARK

Encephalitogenic T cells specific for myelin basic protein share common $V_\beta8$ peptide sequences in their T cell receptor (TCR) that can induce autoregulatory T cells and antibodies that prevent clinical signs of experimental autoimmune encephalomyelitis (EAE). It is not known, however, if TCR peptides can treat established disease. To test its therapeutic value, TCR-$V_\beta8$-39-59 peptide was injected into rats with clinical signs of EAE. This treatment reduced disease severity and speeded recovery, apparently by boosting anti-$V_\beta8$ T cells and antibodies raised naturally in response to encephalitogenic $V_\beta8^+$ T cells. These results demonstrate that synthetic TCR peptides can be used therapeutically, and implicate the TCR-$V_\beta8$-39-59 sequence as a natural idiotope involved in EAE recovery. Similarly, human TCR peptides may be effective in enhancing natural regulation of autoreactive T cells that share common V genes.

THROUGH MECHANISMS THAT ARE not yet fully understood, rat and mouse T cells that arise in response to immunization with guinea pig or rat basic protein (Gp-BP; Rt-BP) preferentially utilize the $V_\beta8.2$ gene and to a lesser extent the $V_\alpha2$ gene in their TCR (*1–6*). The presence of common V region sequences on encephalitogenic T cells allowed us (*7*) and others (*8*) to pre-immunize Lewis rats with synthetic TCR peptides to induce anti-receptor immunity. Our use of the synthetic $V_\beta8$ peptide corresponding to residues 39 to 59 completely protected the animals from the subsequent induction of clinical EAE. Similarly, most signs of EAE could be suppressed if the TCR-$V_\beta8$-39-59 peptide was given during the induction phase but prior to onset of clinical disease (*9*). The protective mechanisms involved both major histocompatibility complex (MHC) class I–restricted T cells (*7*) and antibodies (*9*) specific for the TCR-$V_\beta8$-39-59 peptide that appeared to be directed at a "processed" MHC-associated fragment of the natural TCR V_β chain expressed on the surface of $V_\beta8^+$ T cells. Presumably, interaction of the regulatory T cell or antibody with the $V_\beta8^+$ encephalitogenic T cell perturbed membrane signaling pathways, thus altering effector cell functions.

Although many approaches have been described to prevent encephalomyelitis (EAE), effective treatment of established clinical signs has been much more difficult. To test its therapeutic potential, the TCR-$V_\beta8$-39-59 peptide was injected by several different routes into Lewis rats with moderate signs of EAE. As controls, rats with EAE were injected in parallel with a synthetic peptide corresponding to the TCR-$V_\beta14$-39-59 sequence (not utilized by encephalitogenic T cells), or saline.

Intradermal (id) injection of 50 μg of the TCR-$V_\beta8$-39-59 peptide in saline reduced the clinical severity of EAE from grade 3.2 (paralysis of hind limbs) in the control rats, to grade 1.5 (wobbly gait) within 48 hours, and to grade 0.2 (limp tail in some rats) within 72 hours (Fig. 1A). The TCR-$V_\beta8$-39-59 peptide treatment speeded overall recovery time from 6.3–6.6 days in control rats to 3.1 days (Fig. 1A). A lower dose (10 μg) of the id-injected peptide was only slightly less effective, with a recovery time of 4 days (Fig. 1A). The second route tested, subcutaneous injection of 100 to 500 μg of the TCR-$V_\beta8$-39-59 peptide in saline, produced a nearly identical resolution of clinical EAE, with a recovery time of 3.5 and 4 days, respectively (Fig. 1B). A third regime, injection of the TCR-$V_\beta8$-39-59 peptide in complete Freund's adjuvant (CFA), also arrested disease progression within 24 hours and caused a rapid resolution of the remaining clinical signs from 6.5–6.6 days (controls) to 3.5 days (Fig. 1C).

The rapid clinical resolution of EAE after injection of the TCR-$V_\beta8$-39-59 peptide suggested a recall response similar to that induced in man by tetanus or rabies booster shots. Such a recall response would imply the presence of a preexisting immunity to the TCR-$V_\beta8$-39-59 peptide. Indeed, one might rationalize that the induction of $V_\beta8^+$ encephalitogenic T cells during EAE could stimulate regulatory T cells and antibodies directed at the $V_\beta8$ molecule and more specifically at the TCR-$V_\beta8$-39-59 peptide.

One simple method of assessing preexisting T cell responses in vivo is to measure ear swelling [delayed hypersensitivity (DH)] 24 to 48 hours after an id injection of antigen.

H. Offner and A. A. Vandenbark, Neuroimmunology Research 151D, Veterans Administration Medical Center, and Departments of Neurology, and Microbiology and Immunology (AAV), Oregon Health Sciences University, Portland, OR 97201.
G. A. Hashim, Department of Microbiology and Surgery, St. Luke's Roosevelt Hospital, and Columbia University, New York, NY 10025.

*To whom correspondence should be addressed.