27. I. Rips, J. Klafter, J. Jortner, *ibid.* **88**, 3246 (1988); *ibid.* **89**, 4288 (1988).
28. A. L. Nichols III and D. F. Calef, *ibid.*, p. 3783.
29. See, for example, C. G. Gray and K. E. Gubbins, *Theory of Molecular Fluids* (Clarendon, Oxford, 1984), chap. 5.
30. M. Maroncelli and G. R. Fleming, *J. Chem. Phys.* **89**, 875 (1988).
31. L. Onsager, *Can. J. Chem.* **55**, 1819 (1977).
32. M. Maroncelli and G. R. Fleming, *J. Chem. Phys.* **89**, 5044 (1988).
33. O. A. Karim, A. D. J. Haymet, M. J. Banet, J. D. Simon, *J. Phys. Chem.* **92**, 3391 (1988).
34. A. Chandra and B. Bagchi, *Chem. Phys. Lett.* **151**, 47 (1988).
35. L. D. Zusman, *Chem. Phys.* **49**, 295 (1980).
36. D. F. Calef and P. G. Wolynes, *J. Phys. Chem.* **87**, 3387 (1983); *J. Chem. Phys.* **78**, 470 (1983).
37. J. T. Hynes, *J. Phys. Chem.* **90**, 3701 (1986).
38. H. Sumi and R. A. Marcus, *J. Chem. Phys.* **84**, 4894 (1986); W. Nadler and R. A. Marcus, *ibid.* **86**, 3906 (1987).
39. I. Rips and J. Jortner, *ibid.* **87**, 2090, 6513 (1987); *ibid.* **88**, 818 (1988).
40. M. Sparpaglioni and S. Mukamel, *J. Phys. Chem.* **91**, 3938 (1987); *J. Chem. Phys.* **88**, 1465 (1988); *ibid.*, p. 3263.
41. L. D. Zusman, *Chem. Phys.* **119**, 51 (1988).
42. For an excellent discussion of reaction adiabaticity, see H. Frauenfelder and P. G. Wolynes, *Science* **229**, 337 (1985).
43. The reviews by M. D. Newton and N. Sutin [*Annu. Rev. Phys. Chem.* **35**, 437 (1984)] and R. A. Marcus and N. Sutin [*Biochim. Biophys. Acta* **811**, 265 (1985)] provide good introduction to these theories.
44. M. Opallo and A. Kapturkiewicz, *Electrochim. Acta* **30**, 1301 (1985); M. Opallo, *J. Chem. Soc. Faraday Trans. 1* **82**, 339 (1986).
46. I. Rips and J. Jortner, *Chem. Phys. Lett.* **133**, 411 (1987).
47. E. M. Kosower and D. Huppert, *Annu. Rev. Phys. Chem.* **37**, 127 (1986), and references therein.
48. J. D. Simon and S.-G. Su, *J. Phys. Chem.* **92**, 2395 (1988).
49. M. A. Kahlow, T. J. Kang, P. F. Barbara, *ibid.* **91**, 6452 (1987); T. J. Kang *et al.*, *J. Chem. Phys.* **92**, 6800 (1988).
50. H. Heitele, M. E. Michel-Beyerle, P. Finckh, *Chem. Phys. Lett.* **138**, 237 (1987).
51. S.-G. Su and J. D. Simon, *J. Chem. Phys.* **89**, 908 (1988).
52. R. M. Neison, G. E. McManis, M. N. Golovin, M. J. Weaver, *J. Phys. Chem.* **92**, 3441 (1988); G. E. McManis, M. N. Golovin, M. J. Weaver, *ibid.* **90**, 6563 (1986), and references therein.
53. G. E. McManis and M. J. Weaver, *Chem. Phys. Lett.* **145**, 55 (1988).
54. ———, *J. Chem. Phys.*, in press.
55. B. J. Gertner, K. R. Wilson, J. T. Hynes, *J. Chem. Phys.*, in press; J. P. Bergsma, B. J. Gertner, K. R. Wilson, J. T. Hynes, *ibid.* **86**, 1356 (1987).
56. K. Mansigh and A. Mansigh, *ibid.* **41**, 827 (1964); G. P. Srivastava, P. C. Mathur, K. N. Tripathi, *Indian. J. Pure Appl. Phys.* **9**, 364 (1971).
57. M. Claessens, P. Fiasse, O. Fabre, D. Zimmermann, J. Reisse, *Nouv. J. Chim.* **8**, 357 (1984).
58. U. Kaatze and V. Uhlendorf, *Z. Phys. Chem.* **126**, 151 (1981).
59. F. A. L. Dullien, *AIChE J.* **18**, 62 (1972).

Research Articles

The DNA Binding Domain of the Rat Liver Nuclear Protein C/EBP Is Bipartite

WILLIAM H. LANDSCHULZ, PETER F. JOHNSON, STEVEN L. MCKNIGHT

C/EBP is a rat liver nuclear protein capable of sequence-specific interaction with DNA. The DNA sequences to which C/EBP binds in vitro have been implicated in the control of messenger RNA synthesis. It has therefore been predicted that C/EBP will play a role in regulating gene expression in mammalian cells. The region of the C/EBP polypeptide required for direct interaction with DNA has been identified and shown to bear amino acid sequence relatedness with the product of the *myc*, *fos*, and *jun* proto-oncogenes. The arrangement of these related amino acid sequences led to the prediction of a new structural motif, termed the "leucine zipper," that plays a role in facilitating sequence-specific interaction between protein and DNA. Experimental tests now provide support for the leucine zipper hypothesis.

A FUNDAMENTAL QUESTION IN THE STUDY OF GENE REGULATION has been how regulatory proteins bind DNA selectively. Given the relatively uniform structure of double helical DNA, how do regulatory proteins achieve the binding specificity necessary to execute precise decisions? The simple answer is that regulatory proteins detect differences in the nucleotide sequence of DNA. Whether by establishing a series of direct atomic contacts with a sequence of base pairs, or by detecting subtle, nucleotide sequence–induced deformity in DNA, a regulatory protein is somehow able to lock onto its cognate sites on DNA with unusual avidity.

Recent x-ray crystallographic studies on a related class of repressor proteins from bacteria have begun to provide an account of the molecular forces that mediate sequence-specific interaction between protein and DNA (*1*). These related bacterial proteins adopt, within their respective DNA binding domains, similar three-dimensional structures. The common structural motif consists of two α helices bridged by a sharp β turn (helix-turn-helix). One of the two α helices is oriented in a manner that allows its close apposition to the major groove of DNA. The distinctive binding specificity inherent to different repressors is established by small differences in the shape or projection of the helix-turn-helix motif, by variations in the amino acid side chains that project from it, and by a limited number of additional interactions donated by amino acid residues outside of the helix-turn-helix motif.

The basic principles emerging from studies on bacterial proteins have illuminated the problem of DNA binding selectivity in eukaryotic cells. A structure similar to the helix-turn-helix motif has been hypothesized to form in a class of eukaryotic regulatory proteins that share a highly conserved amino acid sequence termed the homeobox (*2*). Homeobox proteins have now been shown to be capable of sequence-specific interaction with DNA (*3*), and to be dependent on the integrity of the homeobox for this interaction (*4*).

Owing to the widespread occurrence of the helix-turn-helix motif, one might have anticipated that it would constitute the sole three-dimensional structure used to interface protein with DNA—perhaps comparable to the ubiquitous use of immunoglobulins for antigen

The authors are with the Howard Hughes Medical Institute, Carnegie Institution of Washington, Department of Embryology, 115 West University Parkway, Baltimore, MD 21210.

recognition by the immune system. Recent observations are incompatible with this hypothesis. During the past several years the amino acid sequences of a large number of newly discovered, eukaryotic DNA binding proteins have been resolved, and fewer than half exhibit relatedness to the helix-turn-helix motif (*5*). Although high-resolution structural studies have yet to be completed on any of these newly characterized gene regulatory proteins, studies of primary amino acid sequence have led to predictions of several new groups of structural motifs. One of these motifs, termed the "zinc finger," is characterized by an ordered arrangement of histidines or cysteines (or both) that form a tetrahedral coordination complex with a zinc ion (*6*). Another, termed the "helix-loop-helix" motif, predicts the sequential occurrence of two α helices separated by a loop of variable length (7). Finally, a hypothetical motif termed the "leucine zipper" predicts the formation of dimeric molecules held together by hydrophobic interactions between long, solvent-stable α helices of two polypeptide chains (*8*).

We now report on the mode of interaction between C/EBP and DNA in the context of the leucine zipper hypothesis. The DNA binding domain of C/EBP exhibits the two hallmarks common to leucine zipper proteins: a helix-permissive region containing a heptad repeat of leucines, and a highly basic region located immediately adjacent to the leucine repeat helix. We have proposed earlier that the helical leucine repeat region constitutes a dimerization interface that serves to bring into close juxtaposition the basic regions of two polypeptide chains. The two appropriately assembled basic regions are suggested to form the determinant of the protein required for direct interaction with DNA (*8*). In order to test this hypothesis, we have used methods of in vitro mutagenesis to introduce specific amino acid substitutions into the C/EBP polypeptide. Mutated versions of C/EBP have been synthesized in bacterial cells and tested in biochemical assays that measure the capacity of the protein either to form dimers or to bind specifically to DNA.

A rapid assay for C/EBP dimers. As originally conceived, the leucine zipper model proposed that two monomeric chains of the C/EBP polypeptide would be linked via hydrophobic interactions between two amphipathic α helices (*8*). The prediction that C/EBP monomers establish dimeric interactions has been substantiated by glutaraldehyde-mediated cross-linking experiments. Purified C/EBP protein was exposed exposed to low concentrations of glutaraldehyde for varying lengths of time, and the reaction products were separated on a denaturing polyacrylamide electrophoresis gel (*9*). These procedures were initially carried out on a fragment of C/EBP that consisted of the carboxyl-terminal 14 kD of the protein. Earlier experiments had shown that this fragment of C/EBP retains full DNA binding activity (*10*). Glutaraldehyde treatment of the 14-kD fragment led, in a time-dependent manner, to the formation of a molecule exhibiting an apparent molecular mass roughly double that of the starting material (Fig. 1A).

We interpret the 28-kD molecule to represent a covalently cross-linked dimer. Aside from the facts that cross-linking occurred at low concentrations of protein (5 μ*M*) and glutaraldehyde (0.01 percent), there were two other reasons to believe that it reflected a stable association between C/EBP monomers. (i) The products of the cross-linking reaction were exclusively monomeric or dimeric. Even after 6 minutes of exposure to glutaraldehyde, almost no multimeric forms, other than presumed dimers, were generated. If the cross-linking reaction reflected no more than a random collision of polypeptide chains, we might have anticipated the generation of higher order multimers (trimers and tetramers). (ii) Cross-linking was restricted to C/EBP polypeptide chains even when the reaction was carried out in the presence of an excess of contaminating bacterial proteins (see below). Since glutaraldehyde forms adducts via the free amino groups of various amino acid side chains (*11*), its selective action under dilute conditions probably reflects an intimate interaction between individual C/EBP polypeptide chains.

In order to investigate the aggregate nature of C/EBP more closely, we purified a form of the protein that could be distinguished from the 14-kD species. This second C/EBP polypeptide contained an additional 40 amino acids on the NH_2-terminal side of the 14-kD species, had an apparent molecular mass of 18 kD in its monomeric form, and, when subjected to glutaraldehyde, shifted to an apparent mass of 36 kD (Fig. 1B). The 14- and 18-kD species were mixed, incubated for 30 minutes at 37°C, and subjected to cross-linking. The reaction products of the mixture of proteins generated a new multimeric species that migrated on the electrophoresis gel at a position between the dimeric products generated from either single protein species (Fig. 1C). We interpret this 32-kD species to reflect a heterodimeric complex consisting of one 14-kD subunit and one 18-kD subunit. The fact that the presumed heterodimeric complex (32 kD) predominated relative to the 28-kD and 36-kD forms is consistent with a binomial subunit distribution (one 14-14 kD homodimer: two 14-18 kD heteromers: one 18-18 kD homodimer).

The availability of two distinguishable forms of C/EBP made it possible to do a mixing experiment to examine the kinetics of subunit interchange. Purified 14- and 18-kD forms of C/EBP were mixed at a concentration of 2.5 μ*M* each; the mixture was then incubated for varying intervals at 24°C and subjected to glutaraldehyde-mediated cross-linking. The presumed heterodimeric species (32 kD) was detected as early as 1 minute after 14-kD and 18-kD preparations were mixed. Within 3 minutes the 32-kD species reached a plateau relative to the 28-kD and 36-kD species. In summary, these experiments indicate that at 5 μ*M*, C/EBP monomers exist in close association, that the carboxyl-terminal 14 kD of the protein houses the determinant of C/EBP required for this association, and that associated subunits interchange within several minutes of incubation at 24°C.

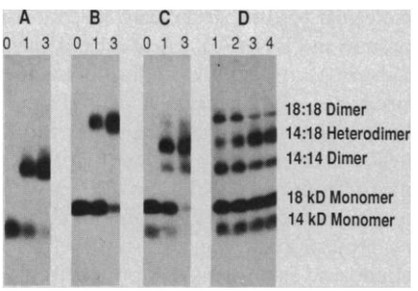

Fig. 1. Evidence supporting the existence of C/EBP dimers. Two fragments of the C/EBP polypeptide were purified after overexpression in bacterial cells (*32*). Both fragments extended to the natural COOH-terminus of C/EBP and were capable of sequence-specific interaction with DNA. One fragment consisted of the COOH-terminal 14 kD of C/EBP, and the other consisted of the COOH-terminal 18 kD of the protein (Fig. 2). Each fragment was incubated with 0.01 percent glutaraldehyde at a protein concentration of 5 μ*M*. Incubations were carried out at room temperature for either 1 or 3 minutes. Cross-linking was stopped by the addition of 0.1*M* tris-HCl (added in the form of SDS gel sample buffer), and samples were subjected to electrophoresis on an 8 to 18 percent polyacrylamide gradient gel in the presence of SDS (*19, 20*). Both the 14 kD fragment (**A**) and the 18 kD fragment (**B**) generated, in a time-dependent manner, slower migrating species interpreted to represent cross-linked dimers (numbers above gel lanes refer to the time of exposure, in minutes, to glutaraldehyde). When the 14-kD and 18-kD samples were mixed, incubated at 37°C for 30 minutes, and tested in the cross-linking assay, a new species migrated at a position between the two presumed homodimers (**C**). The electrophoretic behavior of this new species was consistent with that expected of a heterodimer. The kinetics of subunit interchange were examined by mixing the 14-kD and 18-kD fragments at 24°C, then subjecting the mixed proteins to 0.01 percent glutaraldehyde for 1 minute. Evidence of the presumed heterodimeric species was observed even when glutaraldehyde was added during the first minute after mixing (left lane, **D**). By 3 minutes after mixing, the proportion of heterodimeric species reached its maximal level.

Mutations in the leucine repeat helix interfere with dimer formation and DNA binding. All determinants of C/EBP required for its primary interaction with DNA reside within the carboxyl-terminal 14 kD of the protein (*10*). The leucine zipper model predicted the existence of two functional components within the C/EBP DNA binding domain: (i) a highly basic region necessary for establishing close contact with DNA, and (ii) an α helix-permissive region, characterized by a heptad repeat of leucines, associated with establishing a dimerization interface. The basic region of C/EBP can be divided into three clusters wherein the net positive charge over windows of five amino acids is at least +3 (*10*). These are designated basic regions 1, 2, and 3 (BR1 to 3). BR1, BR2, and BR3 all occur within a 36–amino acid region on the NH_2-terminal side of the leucine repeat (Fig. 2).

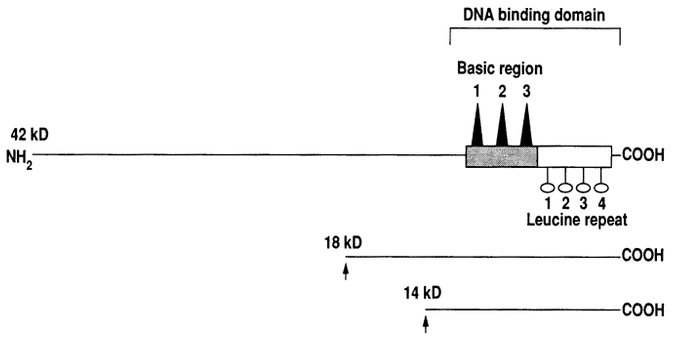

Fig. 2. Anatomy of the C/EBP polypeptide chain. The intact, 42-kD form of C/EBP consists of 359 amino acids and is displayed horizontally with the left and right ends corresponding, respectively, to the NH_2- and COOH-termini of the protein. The 18-kD and 14-kD fragments used in the cross-linking assays of Fig. 1 are shown below the intact protein. The DNA binding domain is located within the COOH-terminal 14 kD of the protein, and is composed of a highly basic region and a leucine repeat region. The basic region can be further divided into three subcomponents wherein windows of five amino acids exhibit a net positive charge of at least +3. The three basic clusters (designated 1, 2, and 3) occur in a region immediately NH_2-terminal to the leucine repeat. Of these, cluster 2 displays a high degree of sequence similarity to a basic region located immediately, on the NH_2-terminal side, to the leucine repeat region of the Fos transforming protein. The leucine residues numbered 1, 2, 3, and 4 correspond, respectively, to amino acid residues 317, 324, 331, and 338 of the C/EBP polypeptide chain (*10*).

The segment of C/EBP that contains its heptad leucine repeats exhibits three properties compatible with α-helical structure. First, it is free of amino acid residues (prolines and glycines) that are seldom found in α helices. Second, when displayed on an idealized α helix, its amino acid sequence segregates hydrophobic amino acid residues to one side of the putative helix and hydrophilic residues to the opposite side. Finally, this 28–amino acid segment of C/EBP exhibits an unusually high proportion of positively and negatively charged amino acid residues juxtaposed in the $i + 3/4$ manner required for the formation of helix-stabilizing salt bridges (*8*). Some of these same properties also occur in the sequences of several other nuclear proteins, including Fos, Myc, Jun, and GCN4 (*8, 12*). Perhaps most surprising, however, was the observation that each of the five proteins has a continuous, heptad array of leucine residues within their respective helix-permissive segment. On the basis of these observations, it was proposed that the leucine repeat helices would serve the same function in each of these proteins, and that function would be to create a dimerization interface termed the leucine zipper (*8*).

A fundamental premise of the leucine zipper model is that a dimerization interface is formed between polypeptide chains via hydrophobic interactions between two α helices. This premise is based, in principle, on the "coiled-coil" interactions that facilitate the interweaving of structural proteins such as lamins and intermediate filaments. However, two aspects of the "zipper helices" differ from the amino acid sequence arrangement of prototypical coiled-coil proteins. First, with only one exception (a methionine residue located at an integral repeat in one of the Myc proteins), leucine residues occur invariantly at every seventh amino acid. Prototypical coiled-coil proteins do not exhibit a strict heptad array of leucines. Second, prototypical coiled-coil proteins exhibit hydrophobic residues at two positions within the heptad repeat in an $i + 3$ or $i + 4$

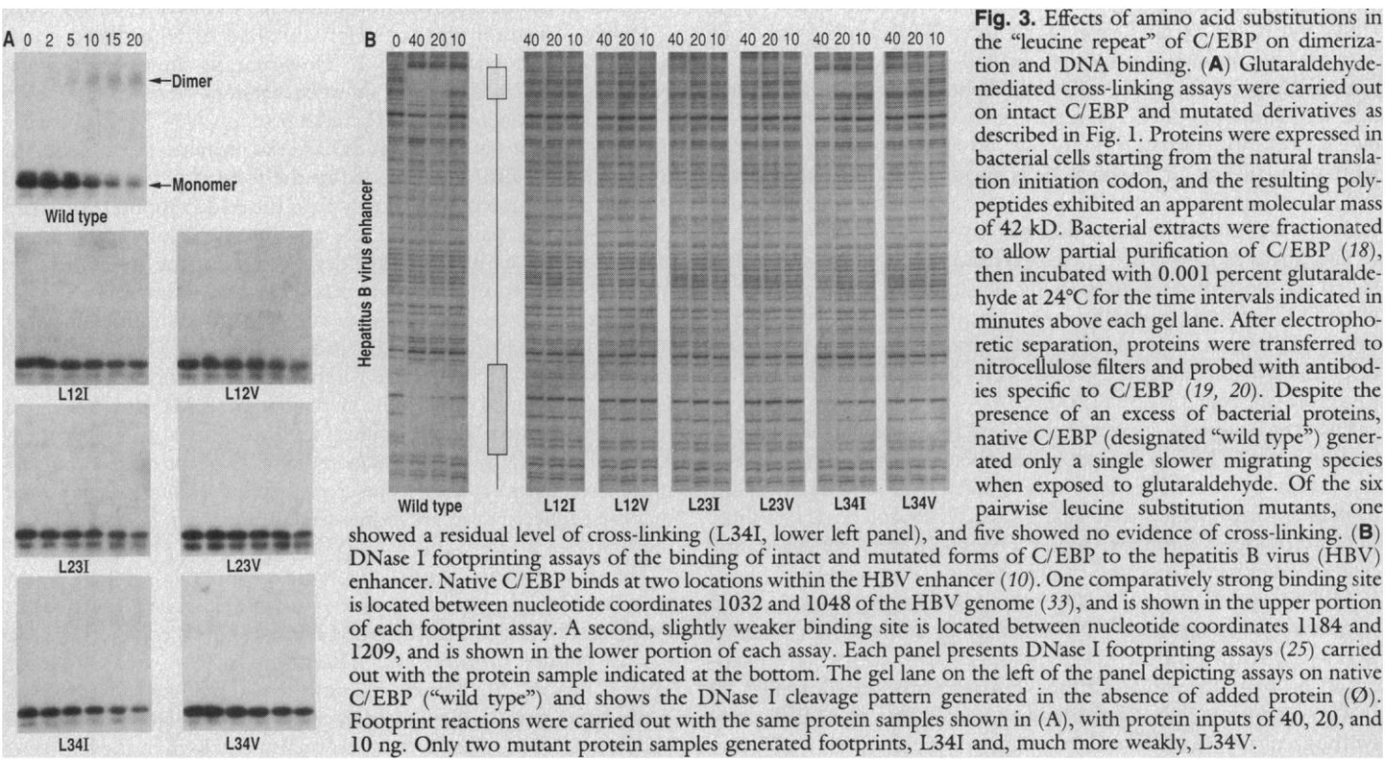

Fig. 3. Effects of amino acid substitutions in the "leucine repeat" of C/EBP on dimerization and DNA binding. (**A**) Glutaraldehyde-mediated cross-linking assays were carried out on intact C/EBP and mutated derivatives as described in Fig. 1. Proteins were expressed in bacterial cells starting from the natural translation initiation codon, and the resulting polypeptides exhibited an apparent molecular mass of 42 kD. Bacterial extracts were fractionated to allow partial purification of C/EBP (*18*), then incubated with 0.001 percent glutaraldehyde at 24°C for the time intervals indicated in minutes above each gel lane. After electrophoretic separation, proteins were transferred to nitrocellulose filters and probed with antibodies specific to C/EBP (*19, 20*). Despite the presence of an excess of bacterial proteins, native C/EBP (designated "wild type") generated only a single slower migrating species when exposed to glutaraldehyde. Of the six pairwise leucine substitution mutants, one showed a residual level of cross-linking (L34I, lower left panel), and five showed no evidence of cross-linking. (**B**) DNase I footprinting assays of the binding of intact and mutated forms of C/EBP to the hepatitis B virus (HBV) enhancer. Native C/EBP binds at two locations within the HBV enhancer (*10*). One comparatively strong binding site is located between nucleotide coordinates 1032 and 1048 of the HBV genome (*33*), and is shown in the upper portion of each footprint assay. A second, slightly weaker binding site is located between nucleotide coordinates 1184 and 1209, and is shown in the lower portion of each assay. Each panel presents DNase I footprinting assays (*25*) carried out with the protein sample indicated at the bottom. The gel lane on the left of the panel depicting assays on native C/EBP ("wild type") and shows the DNase I cleavage pattern generated in the absence of added protein (Ø). Footprint reactions were carried out with the same protein samples shown in (A), with protein inputs of 40, 20, and 10 ng. Only two mutant protein samples generated footprints, L34I and, much more weakly, L34V.

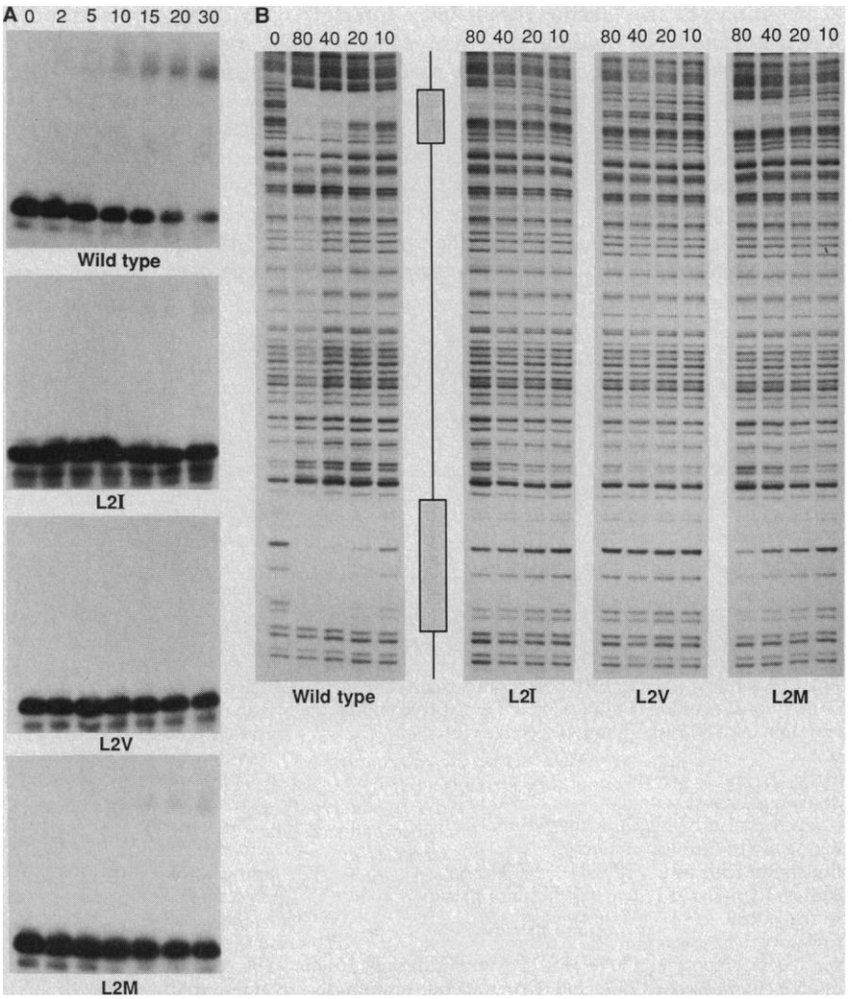

Fig. 4. Effects of amino acid substitutions at leucine position #2 on dimerization and DNA binding. (**A**) Glutaraldehyde-mediated cross-linking assays were carried out in intact C/EBP and variant proteins carrying single amino acid substitutions at residue 324 (leucine residue 2 as shown in Fig. 2). Proteins were expressed in bacteria and subjected to cross-linking assays as described in Fig. 3A. Of the three proteins bearing single amino acid substitutions, one failed to generate cross-linked product (L2V), one generated a small amount (L2I), and one generated a moderate amount (L2M). (**B**) DNase I footprinting assays were carried out on the HBV enhancer as described in Fig. 3B. The same protein samples shown (A) were used in footprinting reactions at input levels of 80, 40, 20, or 10 ng. Relative to intact C/EBP, one mutant failed to show evidence of sequence-specific DNA binding (L2V), one showed modest binding (L2I), and one showed moderate binding (L2M).

arrangement (*13*). This is not an invariant property of leucine zipper proteins. Four regulatory proteins postulated to contain the zipper (three members of the Myc family and Fos) display only one hydrophobic residue (leucine) per heptad repeat.

Leucine differs from two other hydrophobic residues that participate frequently in coiled-coil interactions (valine and isoleucine) in that it does not contain a methyl group attached to the β carbon atom of its side chain. Methionine is also free of a methyl group at this position. Since both leucine and methionine contain bulky hydrophobic groups appended to the tips of their respective side chains, it was proposed that these two residues would be particularly suited to the hydrophobic packing achieved by this category of proteins. Therefore, isoleucine would be unsuitable because of the single methyl group appended to its β carbon, and that valine would be particularly incompatible because of its shorter length and its two β-carbon–appended methyl groups.

The above hypotheses were tested by substituting specific amino acids into the C/EBP polypeptide, and then examining the effects of such mutations on both dimerization and DNA binding. Initially, pairs of leucine residues were changed either to isoleucine or valine. Leucine residues numbered 1 and 2 (Fig. 2) were concomitantly changed to either isoleucine or valine, leading to mutated variants of C/EBP termed L12I and L12V (*14*). Likewise, residues 2 and 3 were changed to produce L23I and L23V, and residues 3 and 4 were changed to produce L34I and L34V. Mutations were introduced by oligonucleotide-directed mutagenesis and confirmed by DNA sequencing (*15*). Native and mutant forms of C/EBP were synthesized in bacterial cells starting at the natural translation initiation codon (*16, 17*), resulting in the production of protein that exhibited the same electrophoretic mobility as intact 42-kD C/EBP.

Dimer formation was tested by glutaraldehyde-mediated cross-linking as described in Fig. 1. However, in these experiments, parental 42-kD C/EBP and mutated variants were only partially purified from contaminating bacterial proteins (*18–20*). Despite the presence of a substantial excess of bacterial proteins, cross-linking of the parental C/EBP molecule led to the formation of only a single species that migrated more slowly than the 42-kD monomeric chain (Fig. 3A, upper panel). Since the apparent molecular mass of this additional species was roughly double that of the monomer, we interpret it to represent a covalently cross-linked dimer.

All six of the mutated variants failed to cross-link to the extent observed with unaltered C/EBP protein (Fig. 3A). No evidence of cross-linking could be detected for five of the mutants (L12I, L12V, L23I, L23V, and L34V). The sixth mutant, L34I, did generate a small amount of a slower migrating species in a time-dependent manner. We provisionally conclude that each of the double mutants is impaired in its capacity to be cross-linked by glutaraldehyde, and that five of the mutants are more severely affected in this regard than the sixth.

The same protein samples that were used in cross-linking assays were tested in deoxyribonuclease I (DNase 1) footprinting assays for specific binding to the hepatitis B virus (HBV) enhancer. Native C/EBP purified from rat liver nuclei binds at two sites within the HBV enhancer (*10*). The stronger of the two binding sites is located in the upper part of each footprint track, and the slightly weaker site is located in the lower part of each track. As shown in the left panel

of Fig. 3B, C/EBP produced in bacterial cells binds to these same two sites within the HBV enhancer. Four of the mutated proteins (L12I, L12V, L23I, and L23V) failed to bind to either site, one (L34V) bound very weakly to the upper site, and one (L34I) bound to both sites with slightly less avidity than parental, bacterially produced C/EBP.

Four observations emerged from the aforementioned experiments. (i) Relatively conservative changes in the leucine repeat region of C/EBP (L to I and L to V) exert substantial effects on the ability of the protein to be cross-linked by glutaraldehyde, as well as on its ability to bind DNA. (ii) Mutations that severely affect cross-linking eliminate sequence-specific DNA binding. The single mutant that exhibited a residual propensity to be cross-linked, L34I, was the only mutant capable of significant interaction with both binding sites on the HBV enhancer. (iii) Equivalent pairwise changes to either isoleucine or valine exerted a more severe effect when introduced at positions 1 + 2 and 2 + 3 than at positions 3 + 4. (iv) The pairwise change of residues 3 + 4 to valine exerted a more severe effect on cross-linking and DNA binding than that exerted by the change to isoleucine.

The severity of effects characteristic of L12I, L12V, L23I, and L23V indicate that leucine residue 2 might play a particularly important role in the formation of C/EBP dimers. Thus, a second, related series of mutagenesis experiments was conducted wherein leucine residue 2 was alone changed either to isoleucine, valine, or methionine (L2I, L2V, and L2M). All three of these amino acid substitutions interfered with the ability of C/EBP to be cross-linked by glutaraldehyde (Fig. 4A). However, the three substitutions did not affect cross-linking to the same degree. Cross-linking was not evident with L2V, some cross-linking was observed with L2I, and a moderate degree was observed with L2M. Precisely the same hierarchy of effects was observed when these three mutant proteins were tested for sequence-specific binding to DNA (Fig. 4B). L2V failed to bind to either the upper or lower site on the HBV enhancer, L2I bound weakly to the upper site, and L2M bound to an intermediate degree to both sites.

Two results from these single amino acid substitutions are consistent with observations on the series of doubly substituted variants. (i) The capacity of mutated protein to be cross-linked correlates with its capacity to bind DNA. (ii) A valine substitution at leucine residue 2 exerts a more detrimental effect than an isoleucine substitution. The single amino acid substitution experiments (Fig. 4) also show that a methionine substitution at leucine residue 2, whether assayed for dimer formation or DNA binding, is less detrimental than substitutions by either valine or isoleucine.

Mutations in the basic region interfere with DNA binding. A polypeptide segment of C/EBP rich in basic amino acids occurs on the amino terminal side of the leucine repeat region (*10*). Fourteen basic amino acids are distributed in three clusters over a 36–amino acid segment. The original leucine zipper hypothesis predicted that this region of the protein would be responsible for intimate contact with DNA, and that the role of the zipper was to bring two such basic regions into close apposition (*8*). In order to study the role of the basic region of C/EBP, mutations were introduced within the three clusters of basic amino acids shown in Fig. 2. The mutations involved the substitution of either two or four basic amino acids by uncharged residues (*21*). Mutants were again generated by oligonucleotide-directed mutagenesis, expressed in bacterial cells, and tested in assays for both dimerization and DNA binding.

All three basic region (BR) mutant proteins generated slower migrating species when tested in the glutaraldehyde-mediated cross-linking assay (Fig. 5A). Relative to unaltered C/EBP, two mutants, BR1 and BR2, showed a slight reduction in the proportion of protein that became cross-linked. The third mutant, BR3, was indistinguishable from native C/EBP in its behavior in the cross-linking assay. Unlike the leucine substitution mutants characterized in the preceding section, the BR mutants of C/EBP failed to show concordance between the cross-linking and DNA binding assays. BR3, the mutant that tested fully positive in the cross-linking assay, failed to bind either site within the HBV enhancer (Fig 5B). The BR1 and BR2 mutants, which performed equally in the cross-linking assay, behaved differently in the DNA binding assay. BR1 retained modest DNA binding activity, whereas BR2 was completely inactive in this regard.

Observations from the study of the three BR mutants allow tentative identification of a region of the protein that is required for interaction with DNA, yet dispensable for dimerization. The most reliable localization of such a region is afforded by the BR3 mutant. The protein product of this mutant was cross-linked after exposure to glutaraldehyde to the same extent as that produced by the unaltered C/EBP gene. The fact that the product of BR3 could perform one of the functions inherent to C/EBP (dimerization) argues against the possibility that the detrimental effect of the mutation on DNA binding reflects gross destabilization or deformation of the protein. Moreover, it, as well as the other BR mutants, was capable of being expressed in bacterial cells as intact polypeptides.

An assumption inherent to the leucine zipper hypothesis is that the role of the leucine repeat helix is to bring into close juxtaposition the basic regions of two polypeptide chains (*8*). Data obtained thus

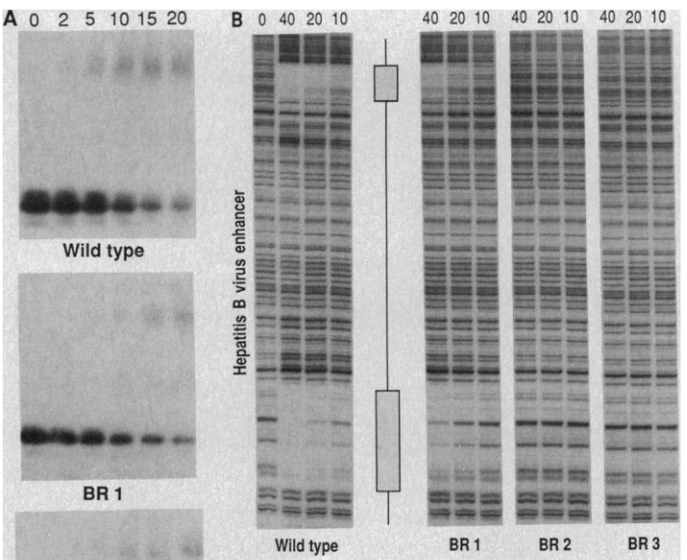

Fig. 5. Effects of amino acid substitutions in the "basic region" of C/EBP on dimerization and DNA binding. (**A**) Glutaraldehyde-mediated cross-linking assays were carried out on intact C/EBP and variant proteins carrying small clusters of amino acid substitutions (*21*) in the three "basic regions" (BR) of C/EBP (see Fig. 2). Proteins were expressed in bacteria and subjected to cross-linking assays as described in Fig. 3A. Of the three BR mutants, one performed normally in the cross-linking assay (BR3), and two exhibited a slightly reduced propensity to be cross-linked (BR1 and BR2). (**B**) DNase I footprinting assays were carried out on the HBV enhancer as described in Fig. 3B. The same protein samples shown in (A) were used in footprinting reactions at input levels of 40, 20, and 10 ng. Relative to intact C/EBP, two BR mutants failed to show evidence of sequence-specific DNA binding with the HBV enhancer (BR2 and BR3), and one showed a modest reduction in binding (BR1).

far are supportive of this hypothesis. If the region of the protein required for dimerization is mutated, DNA binding activity is lost despite the fact that the basic region is left intact (Figs. 3 and 4). Likewise, if the basic region is mutated, DNA binding activity is lost despite maintenance of a functional dimerization domain (Fig. 5). The availability of a mutant that retains the capacity to dimerize, yet lacks a functional basic region, permitted additional testing of the leucine zipper hypothesis. Protein produced from the BR3 mutant was mixed, under conditions that facilitate subunit interchange (see Fig. 1), with that produced by the intact C/EBP gene. If BR3 protein is capable of forming mixed dimers with intact C/EBP, the heterodimeric product should contain one intact basic region and one mutated basic region. The zipper model predicts that such a molecule would be incapable of sequence-specific interaction with DNA.

In order to conduct the experiment outlined above, we mixed various amounts of BR3 protein with a standard amount of unaltered C/EBP. The mixture was incubated for an interval sufficient for subunit interchange (30 minutes) and then tested in DNA binding assays. Protein extract prepared from naive bacterial cells was used to compensate for mixtures containing differing amounts of added BR3 protein, and, as a control, the same experiments were done with a mutated variant of C/EBP incapable of dimerization (L12V). The BR3 protein began to exhibit deleterious effects on the binding capacity of unaltered C/EBP when added at a 2:1 molar ratio. The mixture that contained a 6:1 ratio of BR3 to unaltered C/EBP showed no evidence of sequence-specific interaction with either binding site on the HBV enhancer. When L12V protein was tested in such an assay, no effect was observed on sequence-specific DNA binding by C/EBP even in the presence of an 18:1 molar ratio.

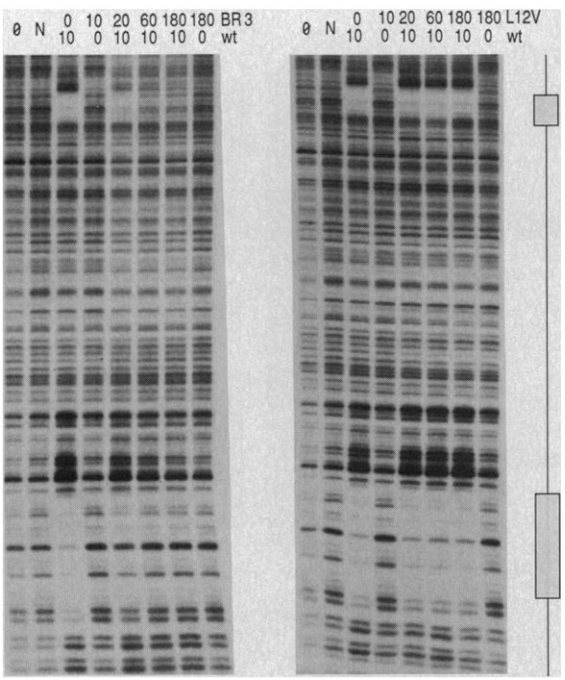

Fig. 6. The protein product of the BR3 mutant poisons sequence-specific binding activity of unaltered C/EBP. Protein synthesized in bacteria from the BR3 mutant was mixed at varying proportions (2:1, 6:1, and 18:1) with that produced from the unaltered C/EBP gene. Extracts prepared from naive (*18*) bacterial cells were added to compensate for the addition of varying amounts of contaminating bacterial proteins After a 30-minute incubation interval, samples were tested for sequence-specific binding to the HBV enhance (left panel). An identical series of experiments was carried out using protein synthesized in bacteria from the L12V mutant (right panel). Each set of trials contain footprint assays in the absence of added protein (Ø), in the presence of naive extract (N), in the presence of 10 ng of unaltered C/EBP alone (0/10), and in the presence of 180 ng of the test protein species (180/0). Lanes depicting footprints generated by the three ratios of mixed protein samples are designated 20/10, 60/10, and 180/10. These assays showed that the BR3 protein exerted a deleterious effect on the capacity of unaltered C/EBP to bind the HBV enhancer, and that the L12V protein did not.

Provisional model for the C/EBP DNA binding domain. Our experiments are consistent with the major tenets of the leucine zipper hypothesis. They provide evidence of a dimeric interaction between C/EBP polypeptide chains and localize regions of the protein that must remain intact to support both dimerization and sequence-specific interaction with DNA. Binding of C/EBP to specific sites on DNA requires the interplay of two components; a dimerization interface (the leucine zipper) and a DNA binding surface (the basic region). Function of the basic region is dependent on an intact dimerization interface, whereas dimerization can readily occur in the absence of an intact basic region. Thus, the most fundamental conclusion to emerge from these studies is that the DNA binding domain of C/EBP is bipartite.

Close examination of the effects of mutations within the leucine repeat region of C/EBP provides additional information pertinent to the zipper hypothesis. When different aliphatic amino acids were substituted at the same position within the leucine repeat region, different effects were observed on both dimer formation and DNA binding. For example, the properties of L34I were different from those of L34V. In both cross-linking and DNA binding assays L34I tested as a weak positive, whereas L34V was negative. Apparently, isoleucine is a more acceptable substitute to leucine at these positions than valine (Fig. 3). Similar results were observed in comparing the effects of single amino acid substitutions at leucine position number 2. Substitution by valine at this position (L2V) completely eliminated both dimerization and DNA binding, substitution by isoleucine (L2I) limited both activities rather severely, and substitution by methionine (L2M) limited both activities by a more modest degree (Fig. 4). These results are consistent with a hierarchy of side chain acceptability at the heptad integral of the leucine repeat wherein leucine > methionine > isoleucine > valine. We believe that these observations underscore the intimate nature of the hydrophobic side chain packing that will occur between zippered helices, and we await the high-resolution structural studies that should resolve the underlying molecular nature of the zipper.

Of additional potential significance is a comparison of the effects of the same amino acid substitution at different positions within the zipper. When pairwise amino acid changes were introduced at leucine residues across the putative dimerization interface, equivalent substitutions did not lead to equivalent effects. For example, L34I maintained a residual capacity to form dimers, and bound to DNA rather respectably. On the contrary, L12I and L23I failed in both assays. Likewise, L34V showed very modest DNA binding activity, whereas L12V and L23V did not. We interpret this asymmetry of effects to favor a parallel helical arrangement for the leucine zipper of C/EBP (Fig. 7). In an antiparallel arrangement, equivalent effects would be expected to result from equivalent substitutions at either terminus of the dimerization interface, since the intact portion of one helix would be expected to interact with the mutated portion of its mate. In a parallel arrangement, it is possible that the two leucines closest to the NH_2-terminus might be capable of facilitating proper apposition of the basic region on their own (Fig. 7). Indeed, Yanofsky and colleagues have identified a fungal regulatory protein that exhibits amino acid sequence similarity to the basic regions of other zipper proteins, yet only contains leucine residues at the two positions closest to the basic region (*22*). Perhaps the most conclusive evidence favoring a parallel helical arrangement has been obtained by Kim and colleagues, who have

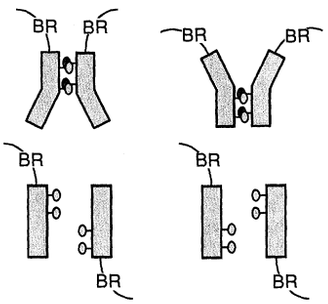

Fig. 7. Hypothetical effects of leucine substitution on dimerization. Potential interactions between mutated forms of C/EBP are shown in the context of a parallel association of leucine repeat helices (top) or an antiparallel association (bottom). Left figure displays hypothetical interaction of two C/EBP polypeptides mutated at leucine residues 3 and 4 (Fig. 2). Right figure displays hypothetical interaction of polypeptides mutated at leucine residues 1 and 2.

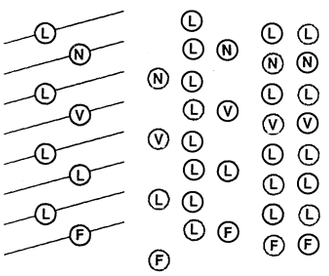

Fig. 8. Parallel arrangement of α helices favors lateral packing of repeating leucine residues. A schematic α helix corresponding to the leucine repeat region of C/EBP is shown on the left. Only the $i + 3$ and $i + 7$ amino acids are shown. The top leucine represents amino acid 317 of the C/EBP polypeptide, and the bottom phenylalanine represents amino acid residue 342 (*11*). The middle diagram displays potential interhelical side chain packing of two leucine repeat helices arranged in a parallel orientation (residues of one helix stippled, the other solid). According to this arrangement, which aligns leucines longitudinally, the $i + 3$ hydrophobic residues would contribute minimally to hydrophobic packing. The right diagram displays potential interhelical packing resulting from the lateral alignment of repeating leucine residues. This arrangement maximizes hydrophobic contributions of $i + 3$ residues.

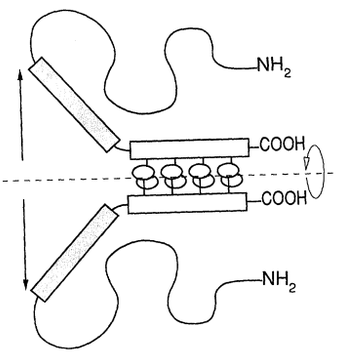

Fig. 9. Schematic model of the DNA binding domain of C/EBP. Two C/EBP polypeptide chains are shown in a parallel dimeric conformation generated by specific interactions between the leucine repeat region of each subunit. Putative leucine repeat helices are indicated as open rectangles, with protruding leucine side chains accounting for the dimerization interface (the leucine zipper). Stippled rectangles adjacent to the paired leucine repeats correspond to the C/EBP basic region, which is predicted to establish direct interaction with DNA. Axis of rotational symmetry is indicated by the dashed line, and arrows at the top correspond to the dyad halves of a hypothetical binding site on DNA.

carried out cross-linking experiments using synthetic peptides corresponding to the leucine repeat helix of GCN4 (*23*).

Realization that the zippered helices of C/EBP may pack in a parallel arrangement prompts reconsideration of the interhelical interactions between leucine residues. We originally hypothesized that the repeating leucines of one helix would interdigitate in a longitudinal array with leucines of the matching helix. However, this arrangement fails to take advantage of the hydrophobic amino acid residues that occur at the $i + 3$ position (Fig. 8B). An alternative mode of helix packing, which arranges invariant leucines in a lateral configuration, maximizes hydrophobic interactions (Fig. 8C). Had helices packed in an antiparallel manner, longitudinal interdigitation of leucines, rather than lateral, would have maximized hydrophobic interactions generated by residues located at the $i + 3$ position. Although these interpretations are consistent with results already described (*23, 24*), we emphasize the fact that they remain hypothetical in the absence of high-resolution structural studies. We also reiterate two of the basic concepts outlined in our original hypothesis; first, that the leucine zipper is founded upon the classical coiled-coil intertwining of α helices, and second, that the side chain structure of leucine is uniquely suited for interhelical packing.

On the basis of the results and interpretations outlined above we propose the rudimentary model of the C/EBP DNA binding domain shown in Fig. 9. Two leucine repeat helices are predicted to dock in an orientation parallel to their respective helical dipoles. This orientation of helix packing results in an axis of rotation parallel to the paired helices. It also leads to close apposition between the basic regions of each polypeptide chain. It may be important that the binding sites for several leucine zipper proteins are dyad symmetrical, with half sites directly abutted (*25, 26*) in a manner analogous to the dyad recognition sites of type II restriction enzymes (*27*). It is possible that DNA binding proteins that utilize the leucine zipper motif dock in such a way as to allow the basic regions to track in opposite directions around dyad half sites. Such a model is consistent with the crucial dependence of DNA binding on dimerization and offers a reasonable explanation for the inhibitory effect of the BR3 polypeptide on DNA binding by native C/EBP (Fig. 6).

The results of mutagenesis studies on the basic region of C/EBP underscore the importance of BR2 and BR3 (Fig. 5). When either of these clusters of basic amino acids is altered by site-directed mutagenesis, DNA binding is eliminated. The critical nature of these segments of C/EBP was, to some extent, anticipated. Earlier studies had shown that BR2 of C/EBP shares extensive amino acid sequence similarity with an analogous region of the Fos transforming protein (*10*). Moreover, sequence relatedness in regions corresponding to BR2 and BR3 had been noted in comparisons of Fos, Jun, and GCN4 (*12*), as well as in the current report by Turner and Tjian (*28*). Recent studies of a protein that regulates sulfur metabolism in *Neurospora crassa* indicate that basic region mutants such as BR2 and BR3 may have precedents generated by classical genetic screening. Marzluf and colleagues have characterized several mutant forms of a fungal regulatory protein, termed CYS3, that exhibit the two hallmarks of the leucine zipper motif (basic region and leucine repeat). Two lesions that inactivate the function of the *cys*3 gene have been identified as amino acid substitutions in the basic region of that protein (*29*).

Since the heptad array of leucines seems to be an invariant feature of a related family of DNA binding proteins, and since the leucine side chain appears to play a pivotal role in dimerization (Figs. 3 and 4), the question arises as to whether heterodimeric species might form by virtue of the interaction of zippers from two different polypeptide chains (*8*). Recent studies on Fos and Jun, transforming proteins that form a stable complex both in vivo and in vitro (*30*), and exhibit the zipper motif (*10*), are consistent with this hypothesis (*24, 28, 31*). The heterotypic mixing of different proteins might create binding specificities or regulatory activities different from parental, homodimeric complexes. This idea, reminiscent of the combinatorial use of light and heavy chain polypeptides in immunoglobulins, offers a means of diversifying the activities of a finite number of genetically encoded regulatory proteins.

REFERENCES AND NOTES

1. A. K. Aggarwal, D. W. Rodgers, M. Drottar, M. Ptashne, S. C. Harrison, *Science* **242**, 899 (1988); S. R. Jordan and C. O. Pabo, *ibid.* **242**, 893 (1988); Z. Otwinowski *et al.*, *Nature* **335**, 321 (1988).
2. B. W. Matthews *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **67**, 427 (1982); A. Laughon and M. P. Scott, *Nature* **310**, 25 (1984); J. C. W. Shepherd, W. McGinnis, A. E. Carrasco, E. M. De Robertis, W. J. Gehring, *ibid.*, p. 70.
3. A. D. Johnson and I. Herskowitz, *Cell* **42**, 237 (1985).
4. C. Desplan, J. Theis, P. H. O'Farrell, *ibid.* **54**, 1081 (1988).
5. P. F. Johnson and S. L. McKnight, *Annu. Rev. Biochem.* **58**, 799 (1989).

6. R. S. Brown, C. Sander, P. Argos, *FEBS Lett.* **186**, 271 (1985); J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **4**, 1609 (1985).
7. C. Murre, P. S. McCaw, D. Baltimore, *Cell*, in press.
8. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
9. A. Joachimiak, R. L. Kelley, R. P. Gunsalus, C. Yanofsky, P. B. Sigler, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 668 (1983).
10. W. H. Landschulz, P. F. Johnson, E. Adashi, B. J. Graves, S. L. McKnight, *Genes Dev.* **2**, 786 (1988).
11. F. A. Quiocho, *Methods Enzymol.* **44**, 546 (1976).
12. P. K. Vogt, T. J. Bos, R. F. Doolittle, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3316 (1987).
13. G. E. Schulz and R. H. Schirmer, *Principles of Protein Structure* (Springer-Verlag New York, 1979).
14. 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.
15. Oligonucleotide-directed mutagenesis was carried out according to the methods of M. J. Zoller and M. Smith, *Methods Enzymol.* **100**, 469 (1983). Mutagenic oligonucleotides (Research Genetics) were annealed with single-stranded M13 mp18 or mp19 vector DNA containing the complementary strand of the C/EBP gene (*10*). The annealed oligonucleotide was then extended with the use of either the Klenow fragment of *Escherichia coli* DNA polymerase (Boeringer Mannheim) or bacteriophage T4 DNA polymerase (Sequenase, U.S. Biochemical). The reaction products were ligated and transformed into *E. coli* strain 71.18. Positive bacteriophage plaques were identified by in situ hybridization with the cognate mutagenic oligonucleotide as a probe. Positive plaques were subjected to two rounds of plaque purification, then grown in liquid culture to allow purification of single-stranded DNA. All putative mutations were confirmed by sequencing through the region of mutation. Mutants are denoted by the single amino acid code for the original residue, followed by its position as numbered in Fig. 2, and by the substituted amino acid. The exact nucleotide changes introduced were (i) the TTG and CTG codons for leucine residues 317 and 324 (*10*) were both changed to ATC in mutant L12I, or GTT in mutant L12V; (ii) the CTG codons for leucine residues 324 and 331 were both changed to ATC in mutant L23I, or GTT in mutant L23V; (iii) the CTG codons for leucine residues 331 and 338 were both changed to ATC in mutant L34I, or GTT in mutant L34V; and (iv) the CTG codon for leucine residue 324 was changed to ATC in mutant L2I, GTT in mutant L2V, or ATG in mutant L2M.
16. Unaltered and mutated forms of C/EBP were synthesized in bacteria with the use of the phage T7 expression system (F. W. Studier, A. H. Rosenberg, J. J. Dunn, *Methods Enzymol.*, in press). The C/EBP gene was engineered for bacterial expression with the plasmid vector pT5 (S. Eisenberg, Synergen), and a synthetic adapter. The vector was modified from the pET vectors (17). pT5 contains the transcriptional and translational start sites for the ϕ10 gene of bacteriophage T7 immediately upstream of a polylinker. The sequence of the synthetic adapter was

5′-GATCCGGAGGATTAAGC-3′
3′-GCCTCCTAATTCGGTAC-5′

The adapter consists of (i) a 5′ overlap compatible with the Bam HI site within the pT5 polylinker; (ii) a bacterial ribosome binding site (RBS); (iii) a translation termination codon five nucleotides downstream from the RBS; (iv) a translation initiation codon 11 nucleotides downstream from the RBS; and (v) a 5′ overlap compatible with the Nco I restriction site that occurs at a position overlapping the translation initiation codon of C/EBP (*10*). pT5 plasmids harboring unaltered and mutated forms of the C/EBP gene were transformed into *E. coli* strain BL21 (DE3) pLysS (*17*). This strain contains the T7 gene encoding the bacteriophage RNA polymerase under control of the *E. coli* lac promoter, introduced into the chromosome as a defective bacteriophage lambda prophage. Isopropylthiogalactoside (IPTG) treatment was used to induce the T7 RNA polymerase, which in turn induced transcription from the ϕ10 promoter of the pT5 vector.
17. W. F. Studier and B. A. Moffatt, *J. Mol. Biol.* **189**, 113 (1986).
18. Mutant and wild-type forms of C/EBP were produced in bacteria (*16*). After IPTG induction, cells were harvested by centrifugation, and resuspended in phosphate-buffered saline (PBS) containing 5*M* urea, 1 m*M* benzamidine, 1 m*M* EDTA, and 0.2 percent Triton X-100. After resuspension, the cells were frozen at −80°C, quick-thawed at 37°C, and sonicated. The lysate was passed over a DEAE-cellulose column equilibrated with PBS containing 5*M* urea, 1 m*M* benzamidine, and 1 m*M* EDTA. The flowthrough fraction was collected and dialyzed against PBS. SDS polyacrylamide gel (*19*) analysis showed that the C/EBP polypeptide comprised from 2 to 6 percent of total protein. The extracts were further purified for DNAase I footprint reactions by retaining the soluble fraction after heat treatment at 70°C for 5 minutes. The concentration of C/EBP or mutated variant was determined by immunoblot analysis (*20*). 100 ng of C/EBP or mutant protein was used in each cross-linking reaction, and amounts ranging in twofold dilution from 80 to 10 ng, were used in DNAase I footprint assays. "Naive" extract, used to compensate for addition of either BR3 or L12V extract in Fig. 6, was prepared from a bacterial strain carrying the pT5 vector without a C/EBP insert.
19. U. K. Laemmli, *Nature* **227**, 680 (1970).
20. Western blot analysis was carried out according to H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350 (1979). The antiserum used to detect bacterially synthesized C/EBP was raised as described in (*10*) against a peptide corresponding to amino acids Ala253 to Gly265 of C/EBP. Antibody-antigen complexes were detected by the use of ^{125}I-labeled protein A and visualized by autoradiography.
21. Mutants are denoted by BR as well as by the position of the cluster of basic residues that was mutated (Fig. 2). The exact nucleotide changes introduced were: (i) the two AAG codons for the lysines at positions 275 and 276 (*10*) were changed to ACT and CAG for mutant BR1; (ii) the CGG, CGG, CGG, CGC codons encoding arginines at positions 286, 288, 289, and 291 were changed to CCG, AGC, CAG, and GGC for mutant BR2; and (iii) the CGC, AAG, CGA, and AAA codons encoding arginine 297, lysine 298, arginine 300, and lysine 302 were changed to GGC, ACA, GGC, and AAC for mutant BR3.
22. J. L. Paluh, M. J. Orbach, T. L. Legerton, C. Yanofsky, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3728 (1988).
23. E. K. O'Shea, R. Rutkowski, P. S. Kim, *Science* **243**, 538 (1989).
24. R. Gentz, F. Rauscher III, C. Abate, T. Curran, *ibid.*, p. 1695.
25. B. J. Graves, P. F. Johnson, S. L. McKnight, *Cell* **44**, 565 (1986).
26. D. E. Hill, I. A. Hope, J. P. Macke, K. Struhl, *Science* **234**, 451 (1986).
27. J. A. McClarin *et al.*, *ibid.* p. 1526.
28. R. Turner and R. Tjian, *ibid.* **243**, 1689 (1989).
29. Y.-H. Fu, J. V. Paietta, D. G. Mannix, G. A. Marzluf, *Mol. Cell. Biol.* **9**, 1120 (1989).
30. F. J. Rauscher III *et al.*, *Science* **240**, 1010 (1988); F. J. Rauscher III, P. J. Voulalas, B. R. Franza, T. Curran, *Genes Dev.* **2**, 1687 (1988).
31. Y. Nakabeppu, K. Ryder, D. Nathans, *Cell* **55**, 907 (1988); T. D. Halazonetis, K. Georgopoulos, M. E. Greenberg, P. Leder, *ibid.* **55**, 917 (1988); T. Kouzarides and E. Ziff, *Nature* **336**, 646 (1988); P. Sassone-Corsi, W. W. Lamph, M. Kamps, I. Verma, *ibid.* **336**, 692 (1988).
32. Construction of the 18-kD fragment was as described for full-length C/EBP (*16*) except that the adaptor was blunt-end ligated to the cDNA insert of bacteriophage L20 (*10*). The insert of L20 was liberated by Eco R1 digestion and treated with S1 nuclease. A large-scale preparation of the 18-kD form of C/EBP was grown in a 3-liter fermentor. The culture was grown to an optical density (600 nm) of 10 and induced for 1 hour with 1 m*M* IPTG. Cells were harvested by centrifugation, resuspended in TEK100 (10 m*M* tris, *p*H 7.9, 0.5 m*M*, EDTA, 100 m*M* KCl, 1 m*M* Benzamidine, 1 m*M* DTT, and 0.2 m*M* PMSF), and lysed in a French press. NP40 and ammonium sulfate were added to 1% and to 0.4*M*, respectively, and the lysate was centrifuged to remove cell debris. Ammonium sulfate was then added to 50% (w/v), and precipitated protein was resuspended and dialysed against TEK100. The extract was heated for 5 minutes at 80°C, and precipitated proteins were discarded. Protein was applied to a Cibacron Blue column (Bio-Rad), developed with a linear KCl gradient, and observed to elute at 0.9*M* KCl. The pool of activity was loaded onto a double strand DNA–Sepharose column [P. F. Johnson, W. H. Landschulz, B. S. Graves, S. L. McKnight, *Genes Dev.* **1**, 133 (1987)] and developed with a linear salt gradient. Approximately 20 mg of material eluted from the column at 0.5*M* KCl. According to SDS gel electrophoresis, more than 95 percent of this protein corresponded to the 18-kD form of C/EBP. In order to produce the 14-kD species, an Nco I site was introduced into the C/EBP gene by oligonucleotide-directed mutagenesis. This procedure resulted in the substitution of leucine residue 252 by methionine (*10*), thereby allowing synthesis of the 14-kD form of C/EBP.
33. S. Jameel and A. Siddiqui, *Mol. Cell. Biol.* **6**, 710 (1986); Y. Shaul and R. Ben-Levy, *EMBO J.* **6**, 1913 (1987).
34. We thank P. Sigler for advice on cross-linking assays and help in development of a purification strategy for C/EBP synthesized in bacteria, R. Kingsbury for technical help, J. Shuman for providing the purified proteins used in cross-linking assays, S. Eisenberg for the pT5 expression vector, W. Studier for the host strains developed for T7 expression, C. Norman and C. Jewell for help in preparation of the manuscript, and K. LaMarco, A. Friedman, J. Shuman, and C. Vinson for comments and encouragement. Supported by NIH grants, the Howard Hughes Medical Research Institute, and the Carnegie Institution of Washington (S.L.M.); a stipend from the Life and Health Insurance Medical Research Fund (W.H.L.); stipends from the Damon Runyon–Walter Winchell Cancer Fund and Carnegie Corporation (P.F.J.); and a Faculty Research Award from the American Cancer Society (S.L.M.).

28 January 1989; accepted 3 March 1989