Max: A Helix-Loop-Helix Zipper Protein That Forms a Sequence-Specific DNA-Binding Complex with Myc

ELIZABETH M. BLACKWOOD AND ROBERT N. EISENMAN

The myc protooncogene family has been implicated in cell proliferation, differentiation, and neoplasia, but its mechanism of function at the molecular level is unknown. The carboxyl terminus of Myc family proteins contains a basic region helix-loop-helix leucine zipper motif (bHLH-Zip), which has DNA-binding activity and has been predicted to mediate protein-protein interactions. The bHLH-Zip region of c-Myc was used to screen a complementary DNA (cDNA) expression library, and a bHLH-Zip protein, termed Max, was identified. Max specifically associated with c-Myc, N-Myc, and L-Myc proteins, but not with a number of other bHLH, bZip, or bHLH-Zip proteins. The interaction between Max and c-Myc was dependent on the integrity of the c-Myc HLH-Zip domain, but not on the basic region or other sequences outside the domain. Furthermore, the Myc-Max complex bound to DNA in a sequence-specific manner under conditions where neither Max nor Myc exhibited appreciable binding. The DNA-binding activity of the complex was dependent on both the dimerization domain and the basic region of c-Myc. These results suggest that Myc family proteins undergo a restricted set of interactions in the cell and may belong to the more general class of eukaryotic DNA-binding transcription factors.

THE PRODUCTS OF THE *MYC* FAMILY OF PROTOONCOGENES, including c-Myc, N-Myc, and L-Myc proteins, function in cell proliferation, differentiation, and neoplastic disease (*1*). However, there is as yet no consensus as to the molecular mechanism by which Myc mediates its biological effects. The Myc proteins are nuclear phosphoproteins with short half-lives and nonspecific DNA-binding activities (*2*). Functionally important regions exist at both the amino and carboxyl termini of the c-Myc protein (*3–5*). Indeed, the carboxyl-terminal 85 amino acids of the Myc family proteins share significant sequence similarity with two classes of transcription factors, the basic region helix-loop-helix (bHLH) and basic region leucine zipper (bZip) proteins, both of which have basic regions adjacent to their dimerization domains. The bHLH family includes over 60 proteins in vertebrates, yeast, plants, and insects; many, if not all, exhibit nuclear localization, are sequence-specific DNA-binding proteins, and function as transcriptional regulators (*6*). The region of sequence similarity shared by Myc and other proteins in this class is a critical determinant of function and contains a stretch of basic amino acids followed by two putative amphipathic α helices that flank an Ω-type loop (*7, 8*). Studies of several other bHLH proteins have demonstrated that the HLH region mediates formation of homo- or heterodimers, which in turn permits the basic regions to form a DNA contact surface (*9–11*). Myc family proteins differ from the bHLH family in that adjacent and carboxyl-terminal to their bHLH motif is another α helix that contains a heptad repeat of leucine residues. This structure is characteristic of the dimerization domains of the bZip family of transcriptional regulators (*12*). The array of nonpolar amino acids forms a hydrophobic face along the amphipathic helix, facilitating specific association of bZip proteins through a parallel coiled-coil interaction (*13*). Dimerization is critical for DNA binding (*14, 15*).

For c-Myc there is substantial evidence that the bHLH region and the adjacent leucine zipper motif are functionally important. Deletions within these regions result in loss or alteration of transforming activity (*3, 16*) as well as reduction of the capacity to autoregulate endogenous *myc* expression and to inhibit cell differentiation (*4, 5*). In addition, a bacterially expressed fusion protein that contains the bHLH-Zip domains of c-Myc has sequence-specific DNA-binding activity (*17*) .

The biological importance of and structural similarities in the carboxyl terminus of c-Myc suggest that Myc functions as a component of an oligomeric complex. While Myc self-association has been demonstrated with relatively high concentrations of bacterially expressed Myc protein (*18*), coprecipitation, chemical crosslinking, and dimerization motif chimeras fail to demonstrate homodimerization of Myc under physiological conditions (*1, 19, 20*). Because functionally relevant interactions occur among members of the bHLH and bZip classes (*9, 15, 21, 22*), and c-Myc has not yet been found to associate with members of either group (*10, 15, 16*), we hypothesized that Myc function may depend on heterotypic interaction with an unknown protein. We now describe the cloning of such a Myc binding factor.

Functional cloning of a Myc binding protein. Biologically

E. M. Blackwood is in the Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98104 and at the Department of Pathology, University of Washington, School of Medicine, Seattle, WA 98195. R. N. Eisenman is in the Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98104.

interactive proteins have been identified by functional cloning (*23*). This work encouraged us to use the c-Myc bHLH-Zip region to identify proteins from a λgt11 cDNA expression library that interact with Myc. We prepared a construct that consisted of the carboxyl-terminal 92–amino acid residues of human c-Myc fused to the carboxyl terminus of glutathione *S*-transferase (GST-MycC92) (Fig. 1A). This bacterially expressed fusion protein was soluble, easily purified, and contained 17 tyrosines as potential iodination sites (only one of which lies within the Myc segment). Furthermore, this protein, which was used to identify a specific DNA-binding sequence for c-Myc (*17*), contains the complete bHLH-Zip region and thus should have the minimal structure required for DNA binding and protein interaction.

GST-MycC92 was expressed in *Escherichia coli*, purified by glutathione-agarose affinity chromatography and ^{125}I-labeled to high specific activity (*24*). For cloning we used a random-primed λgt11 expression library derived from a baboon lymphoblastoid cell line (*25*). Phage from this library produce nearly full-length β-galactosidase proteins fused with the open reading frames of the directionally cloned cDNA's. More than 10^6 plaques were screened for their ability to interact with ^{125}I-labeled GST-MycC92 (*26*). Several potential positive plaques were identified, two of which (Max11 and Max14) survived multiple rounds of plaque purification (Fig. 1B). Because the observed binding might have been mediated by the GST sequences in the fusion protein, the plaques were probed with GST ^{125}I-labeled to the same specific activity as GST-MycC92. Only the GST fusion protein that contained c-Myc, and not GST alone, reacted with the Max14 plaques (Fig. 1B). In addition, affinity-purified antibodies to the 12 carboxyl-terminal amino acids of human c-Myc (anti-Myc) (*27*) partially blocked the binding of GST-MycC92 to the plaques in a manner that was prevented by addition of the peptide immunogen (Fig. 1B). To confirm that the association of GST-MycC92 with Max11 and Max14 was attributable to specific protein-protein interaction, Max11 and Max14 lysogen proteins were fractionated by SDS–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to nitrocellulose filters, and subjected to protein blotting with ^{125}I-labeled GST-MycC92. While GST-MycC92 failed to bind to β-galactosidase alone, it did bind to β-galactosidase fusion proteins in both Max11 and Max14 lysates (*16*). These results indicate that the Myc-containing segment of GST-MycC92 specifically interacts with the protein products encoded by Max11 and Max14 cDNA's.

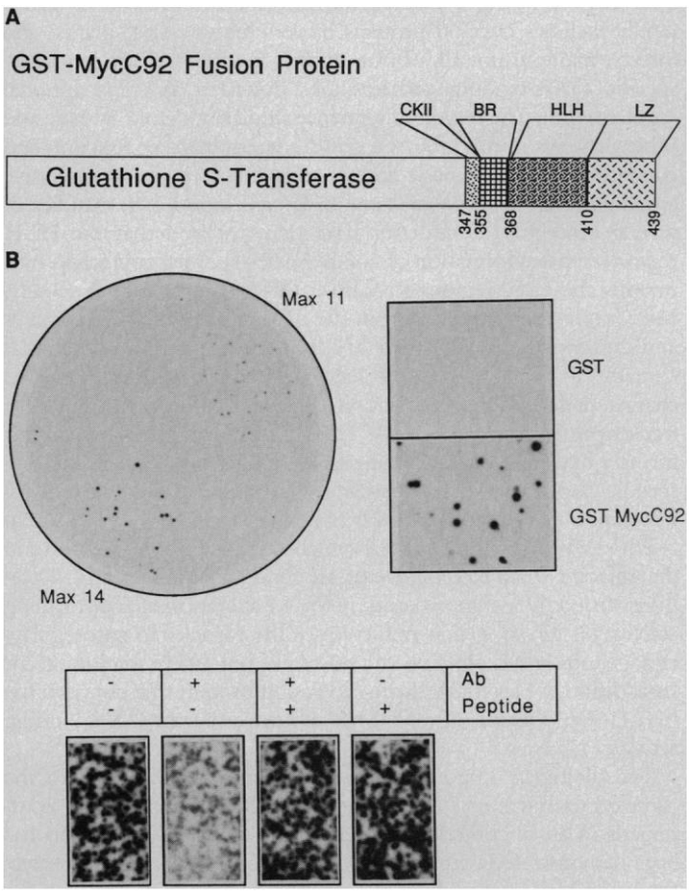


Fig. 1. Functional cloning of a Myc binding protein. (**A**) Diagram of the GST-MycC92 fusion protein used for iodination and screening. CKII, casein kinase II phosphorylation site; BR, basic region; HLH, helix-loop-helix; LZ, leucine zipper. (**B**) Plaques that express β-galactosidase fusion proteins were screened for their ability to react with ^{125}I-labeled GST-MycC92. Top left, secondary plating of five putative positives demonstrates the reactivity of two of the primary plaques, Max11 and Max14. Top right, as a negative control, GST was labeled to a similar specific activity and compared with GST-MycC92 for binding to Max14 plaques. Bottom, binding of GST-MycC92 to Max14 plaques was assayed with or without affinity purified carboxyl terminal–specific anti-Myc (Ab) or peptide immunogen (peptide).

Identification of a helix-loop-helix–zipper domain in Max. Nucleotide sequence analysis (*28*) of the inserts from both of the GST-MycC92 reactive λgt11 phages demonstrated that Max11 and Max14 encode the same protein as defined by the β-galactosidase open reading frame. Both appear to be partial, overlapping cDNA's. Max11 and Max14 encode 124 and 131 amino acids, respectively, between the junction with β-galactosidase and a TAA termination codon (*29*). Subsequent isolation of several overlapping cDNA's from a Manca (human Burkitt's lymphoma cell line) λgt10 library permitted deduction of an apparently complete open reading frame for Max that encodes 151 residues (Fig. 2). This is based on the

```
... CAG TGG CCG CTC CCT GGG CCG TAG GAA ATG AGC GAT AAC GAT GAC ATC GAG     24
                               ***     Met Ser Asp Asn Asp Asp Ile Glu      8

        GAA GAG CAA CCG AGG TTT CAA TCT GCG
        Glu Glu Gln Pro Arg Phe Gln Ser Ala

GTG GAG AGC GAC GCT GAC AAA CGG GCT CAT CAT AAT GCA CTG GAA CGA AAA CGT     78
Val Glu Ser Asp Ala Asp Lys Arg Ala His His Asn Ala Leu Glu Arg Lys Arg     26
                        +   +                                   +   +   +

AGG GAC CAC ATC AAA GAC AGC TTT CAC AGT TTG CGG GAC TCA GTC CCA TCA CTC    132
Arg Asp His Ile Lys Asp Ser Phe His Ser Leu Arg Asp Ser Val Pro Ser Leu     44
 +

CAA GGA GAG AAG GCA TCC CGG GCC CAA ATC CTA GAC AAA GCC ACA GAG TAT ATC    186
Gln Gly Glu Lys Ala Ser Arg Ala Gln Ile Leu Asp Lys Ala Thr Glu Tyr Ile     62

CAG TAT ATG CGA AGG AAA AAC CAC ACA CAC CAG CAA GAT ATT GAC GAC CTC AAG    240
Gln Tyr Met Arg Arg Lys Asn His Thr His Gln Gln Asp Ile Asp Asp Leu Lys     80

CGG CAG AAT GCT CTT CTG GAG CAG CAA GTC CGT GCA CTG GAG AAG GCG AGG TCA    294
Arg Gln Asn Ala Leu Leu Glu Gln Gln Val Arg Ala Leu Glu Lys Ala Arg Ser     98

AGT GCC CAA CTG CAG ACC AAC TAC CCC TCC TCA GAC AAC AGC CTC TAC ACC AAC    348
Ser Ala Gln Leu Gln Thr Asn Tyr Pro Ser Ser Asp Asn Ser Leu Tyr Thr Asn    116

GCC AAG GGC AGC ACC ATC TCT GCC TTC GAT GGG GGC TCA GAC TCC AGC TCA GAG    402
Ala Lys Gly Ser Thr Ile Ser Ala Phe Asp Gly Gly Ser Asp Ser Ser Ser Glu    134
                                    -               -               -

TCT GAG CCT GAA GAG CCC CAA AGC AGG AAG AAG CTC CGG ATG GAG GCC AGC TAA    453
Ser Glu Pro Glu Glu Pro Gln Ser Arg Lys Lys Leu Arg Met Glu Ala Ser ***    151
    -       -   -               +   +   +       +

GCC ACT CGG GGC AGG CCA GCA ATA AAA ...
```

Fig. 2. Nucleotide and amino acid sequences of Max. The Max open reading frame, as generated from overlapping Manca cell cDNA's (human), encodes a 151–amino acid polypeptide. The 9–amino acid insertion found in several PCR clones is shown above the inverted triangle. Helix I and helix II of the bHLH homology region are underlined, while the hydrophobic heptad repeat, which extends from helix II into the zipper region, are in bold face and underscored. Basic and acidic regions are identified by their charge (+ or −), and termination codons are marked by asterisks.

assignment of an AUG in relatively good context for translational initiation (*30*) that was preceded by an upstream, in-frame termination codon. This amino acid sequence probably represents the complete Max open reading frame, because antibodies to Max (anti-Max) were used to immunoprecipitate a cellular protein that comigrated with, and produced an identical tryptic peptide map as, the in vitro translation product of the Max cDNA (*16*). Sequencing of Max-specific polymerase chain reaction (PCR) products from Manca cDNA's revealed a putative variant form of Max that differed only by a 9–amino acid insertion amino-terminal to the basic region (Fig. 2, inverted triangle). In the experiments described below, we utilized a Max cDNA that lacked this insertion.

A computer search of the protein database (PIR Version 25) revealed sequence similarity between a segment of the Max open reading frame and the bHLH proteins, including members of the Myc family (Fig. 3A). The Max sequence in this region represents a nearly exact match with the HLH consensus (Fig. 3A, bottom) (*31*). The similarity among Max and the bHLH proteins also extends in the amino-terminal direction into a basic region of Max. The Max sequence just carboxyl terminal to helix II contains a series of hydrophobic amino acid residues, three of which are leucines, spaced seven residues apart (Fig. 3A, boxed residues). Helical wheel analysis (*12*) of this region suggests that the amphipathic helix II may extend into and beyond the three leucines. These leucines and the other nonpolar residues might form a hydrophobic face similar to that in the leucine zipper proteins.

The Max polypeptide sequence is hydrophilic in nature. More than one-third of its residues are charged, and the most abundant amino acid is serine (14 percent). Max contains no cysteines. Despite a predicted molecular size of 17,200 daltons, Max, like Myc, exhibits aberrant electrophoretic mobility in SDS-polyacrylamide gels (see Fig. 5A). The organization of the Max cDNA protein coding sequence and the relative extents of the basic, HLH, zipper, and carboxyl-terminal regions are depicted in Fig. 3B. The similarity of Max with other bHLH proteins is limited to the Max bHLH region, and the sequences of baboon and human Max do not correspond to those of any previously identified protein.

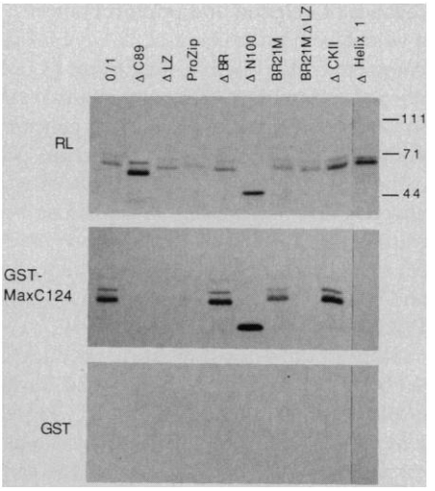
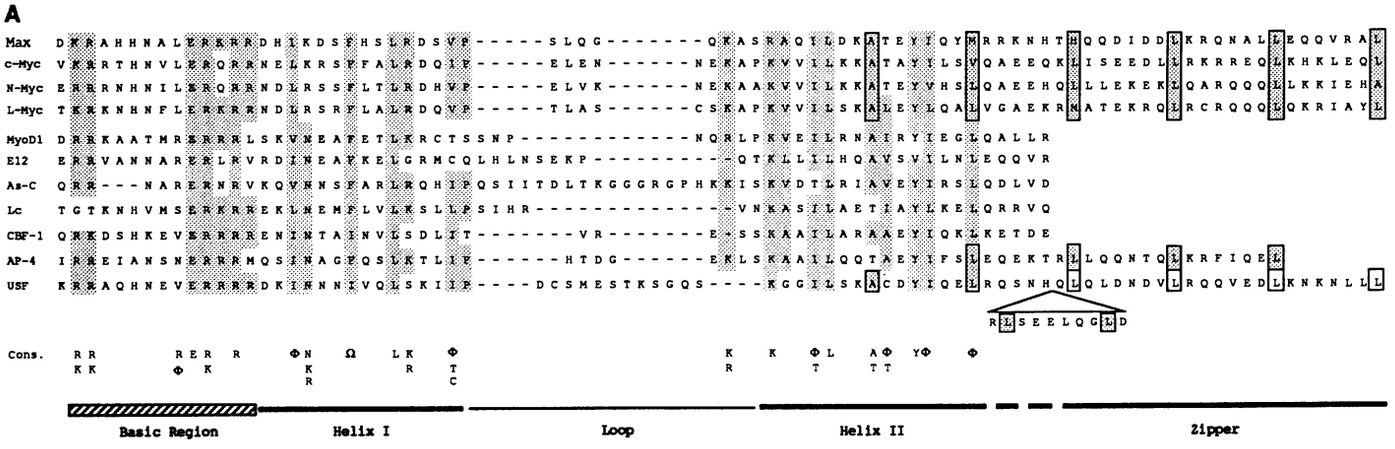


Fig. 4. Structural requirements for Myc-Max association. Wild-type (0/1) or mutant forms of the c-Myc protein were assayed for their ability to associate with the HLH-Zip motif of Max. After in vitro translation, programmed reticulocyte lysate (RL) was subjected directly to SDS-PAGE analysis (1 μl), or sample (10 μl) was purified on GST or GST-MaxC124 affinity columns and the bound material was subjected to SDS-PAGE (*34*). Mutations, ΔN100, deletion of the amino-terminal 100 amino acids of c-Myc (this mutation removes the two alternative initiation codons that normally are translated to produce the p64-p67 doublet); ΔC89, deletion of the carboxyl-terminal 89 amino acids; ΔLZ, deletion of the leucine zipper; ΔBR, basic region deletion; ΔHelix 1, helix I deletion; ΔCKII, casein kinase II phosphorylation site deletion (*36*); ProZip, proline was substituted for leucine at position 2 of the zipper region; BR21M, the basic region of c-Myc was replaced with that of MyoD. Migration of the molecular size markers are indicated.

Specific interaction of Max with Myc family proteins. The presence of putative bHLH and leucine zipper domains in Max suggests that Max interacted with a similar region of GST-MycC92 in the functional library screening. To further investigate the potential of this region of Max to associate with Myc, we developed an affinity chromatography assay in which the Max protein, linked to a solid support, was used to test for binding of full-length, wild-type Myc and a series of mutant Myc proteins.



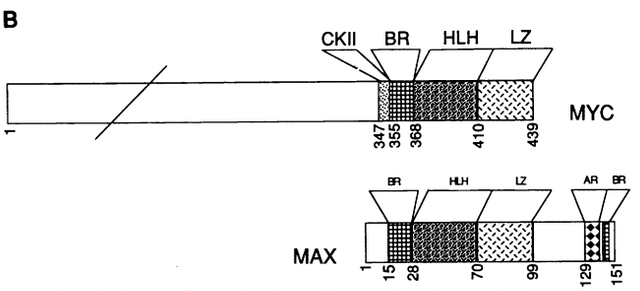


Fig. 3. Structure of the Max protein and its sequence similarity shared with other bHLH-Zip proteins. (**A**) Regions of sequence similarity shared with other bHLH transcription factors. The Max bHLH-Zip region is compared and contrasted to that of other bHLH proteins found in humans (MyoD, E12, AP-4, USF, c-Myc, L-Myc, and N-Myc), insects (As-C), plants (Lc), and yeast (CBF-1). Shaded regions identify residues that fit the consensus as derived for the known bHLH family (*31*) (ϕ = L, I, V, M; Ω = F, L, I, Y). Boxes denote the heptad repeat of hydrophobic residues, which extends from helix II into the putative leucine zipper. 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**) Schematic representation of Myc and Max proteins as aligned by their regions of sequence similarity (stippled boxes). Abbreviations are used to designate the casein kinase II phosphorylation site(CKII), basic region (BR), helix-loop-helix (HLH), leucine zipper (LZ), and acidic region (AR). Numbering corresponds to their respective amino acid sequences.

A set of deletion and point mutations (*32*) were introduced into a wild-type human c-*myc* cDNA (pHLmyc 0/1) that contained the complete c-Myc open reading frame (*33*). RNA's prepared from the normal and mutated clones by in vitro transcription were translated in a rabbit reticulocyte lysate to generate c-Myc proteins labeled with [^{35}S]methionine (*34*). A fusion protein that contained the carboxyl-terminal 124 amino acids of Max (GST-MaxC124) (*35*) was coupled to glutathione-Sepharose beads. The labeled in vitro translation products (Fig. 4, top) were incubated with GST-MaxC124 or GST resin, washed under low stringency conditions, and the bound material was eluted with SDS and analyzed by SDS-PAGE (*34*).

None of the c-*myc* translation products bound to GST alone (Fig. 4, bottom), while GST-MaxC124 resin retained the wild-type c-Myc proteins, p64 and p67 (0/1; Fig. 4, middle). The ability of c-Myc protein to interact with Max was dependent on an intact carboxyl terminus, as deletion of the carboxyl-terminal 89–amino acid residues (ΔC89) completely abolished binding to Max, while deletion of 100 residues at the amino terminus (ΔN100) had no effect (Fig. 4, middle). To ascertain what regions within the carboxyl-terminal domain were required for the binding, we examined a series of mutations. Neither deletion of the Myc basic region (ΔBR), its substitution with the MyoD basic region (BR21M), or deletion of one of the CKII phosphorylation sites (CKII, just amino terminal to the basic region) (*36*) had any effect on association with Max. In contrast, binding to Max was inhibited by deletion of either c-Myc helix I or the leucine zipper, as well as by substitution of a helix-disrupting proline residue for the second leucine in the zipper. These results suggest that full-length c-Myc interacts with the carboxyl-terminal region of Max and that this association is mediated by the c-Myc HLH-Zip domain.

Because N-Myc and L-Myc also have bHLH-Zip regions at their carboxyl termini (*7, 12,*), we assessed their ability to bind Max. [^{35}S]Methionine-labeled in vitro translation products, generated from N-*myc* and L-*myc* cDNA's, bound to GST-MaxC124 resin with the same efficiency as the c-Myc protein (Fig. 5A). In vitro translated, full-length Max protein also bound to the Max-containing resin suggesting that Max may homooligomerize. Neither the Myc family proteins nor Max bound GST alone. To test the possibility that any protein that contains an HLH or leucine zipper motif might associate with Max, we obtained cDNA's that encode other transcription factors and determined the ability of their in vitro translation products to bind to GST-MaxC124 resin. Categories of transcription factors examined included MyoD, E12, Tal, and myogenin, all of which possess bHLH domains (*8, 37, 38*); Fos and Jun, each of which contain a leucine zipper (*12*); and AP-4 (*39*) and USF (*40*), which contain adjacent HLH and leucine zipper regions. Although none of these proteins bound either GST or GST-MaxC124, specific interaction between c-Myc and GST-MaxC124 was again observed (Fig. 5B). This assay is a rather stringent test of association, because relatively low amounts of labeled protein compete for binding with a large excess of Max homodimers (or homo-oligomers). Furthermore, the reticulocyte lysate may contain competitors or inhibitors of binding. Therefore, the inability of specific proteins to interact with Max in the assay may not be a reflection of the in vivo situation.

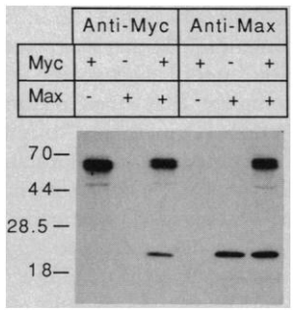


Fig. 6. Post-translational association of full-length Myc and Max. After separate in vitro translations, c-Myc and Max lysates were mixed, incubated for 30 minutes at 30°C, and immunoprecipitated with the indicated antibodies under low stringency conditions (*41*). Anti-Myc was specific for the carboxyl-terminal 12 amino acids of human c-Myc (*27*); anti-Max was raised against the GST-MaxC124 fusion protein. Immunoprecipitated [^{35}S]methionine labeled proteins were resolved by SDS-PAGE.

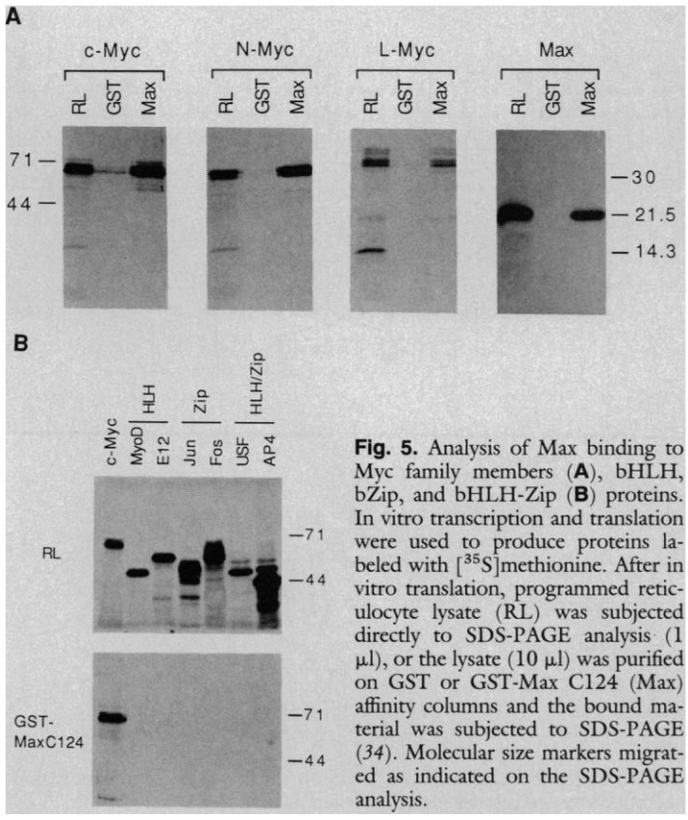


Fig. 5. Analysis of Max binding to Myc family members (**A**), bHLH, bZip, and bHLH-Zip (**B**) proteins. In vitro transcription and translation were used to produce proteins labeled with [^{35}S]methionine. After in vitro translation, programmed reticulocyte lysate (RL) was subjected directly to SDS-PAGE analysis (1 μl), or the lysate (10 μl) was purified on GST or GST-Max C124 (Max) affinity columns and the bound material was subjected to SDS-PAGE (*34*). Molecular size markers migrated as indicated on the SDS-PAGE analysis.

Formation of a Myc-Max complex with sequence-specific DNA-binding activity. Experiments in which bacterially expressed GST-MycC92 was used to select preferred DNA sequences from a pool of partially randomized oligonucleotides have shown that c-Myc has specific DNA-binding activity for the sequence CACGTG (*17*). Compared to other bHLH proteins used in this assay, in vitro translated c-Myc bound relatively poorly to an oligonucleotide that contained this sequence (CM-1). These results might be explained in terms of inefficient homodimerization, inefficient binding of homodimers to the CM-1 sequence, or both. Because Max is capable of specifically associating with Myc, we tested the possibility that the Myc-Max heterocomplex might exhibit increased binding to CM-1 compared to Myc alone.

For the DNA-binding assays, it was important to use full-length Myc and Max proteins in a soluble complex. Therefore, we first determined whether the full-length forms of both Max and c-Myc specifically associate in solution. The p64 and p67 c-Myc proteins and the p21 Max protein produced by in vitro translation of their respective cDNA's were recognized by their cognate antisera (Fig. 6) (*41*). Under low stringency immunoprecipitation conditions, anti-Myc failed to recognize Max and anti-Max failed to recognize c-Myc. However, when Max and c-Myc were combined after translation, each antiserum precipitated Max as well as c-Myc. The ability of a specific antiserum to precipitate the two proteins after mixing is best explained by formation of Myc-Max complexes that are stable under the immunoprecipitation conditions. This idea is supported by the results of blocking experiments, which demonstrate that coprecipitation of both proteins occurs only through the antigenic determinants of one of them (*16*). The Myc mutants that fail to bind to truncated Max in the affinity chromatography

experiments (Fig. 4) also did not associate with full-length in vitro translated Max in the coimmunoprecipitation assay (*16*).

Having established that full-length Max and c-Myc associate in solution, we next determined whether the Myc-Max complex could bind a specific DNA sequence in a gel retardation assay (*42*). Incubation of an unprogrammed reticulocyte translation lysate with the ^{32}P-labeled CM-1 oligonucleotide resulted in retardation of the probe (Fig. 7A). This binding appeared to be due to endogenous USF protein, which also recognizes the CM-1 sequence (*43*). When the translation lysates were programmed with Max RNA, no additional binding to the probe was detected, while lysates that contained c-Myc reproducibly showed a faint band of retarded probe (Fig. 7A). Retardation of the CM-1 probe was observed when reticulocyte lysate that contained both c-Myc and Max were used in the assay. That both c-Myc and Max proteins were bound to the retarded DNA probe was demonstrated by the ability of both anti-Myc and anti-Max to decrease the electrophoretic mobility of the bound probe. The specificity of this antibody effect on mobility of the probe was confirmed by the fact that it could be reversed for each antibody by addition of the cognate immunogen (Fig. 7A). The specificity of binding to CM-1 was verified in competition experiments in which a fivefold excess of unlabeled CM-1 was sufficient to compete for binding by the Myc-Max complex. By contrast, a 500-fold excess of an oligonucleotide (B1/B2) that contained a binding site for MyoD and E12 (*42, 44*) and differed by only three nucleotides from CM-1 was required to achieve a similar degree of competition (Fig. 7B).

Binding of Myc to immobilized Max was dependent on the integrity of the HLH and leucine zipper domains (Fig. 4). To ascertain whether the association of Myc and Max in a nucleoprotein complex required the same sequences, some of the c-Myc mutants were examined for their ability to bind to CM-1 in a complex with Max. Specific binding to the CM-1 probe by Max and c-Myc was abolished when c-Myc mutants that lacked the putative leucine zipper domain or basic region were used in place of wild-type c-Myc. By contrast, a c-Myc deletion mutant that did not directly affect the bHLH-Zip region (such as ΔN100, which lacks 100 amino-terminal residues of c-Myc), both associated with Max (Fig. 5) and bound to the CM-1 oligonucleotide. Therefore, loss of DNA binding correlates with the inability of c-Myc mutants to associate with Max in the binding assay. An exception to this is the basic region deletion mutant, which associated with Max but did not bind the CM-1 probe (Fig. 7), a result that suggests a requirement for the basic region of c-Myc in specific DNA binding but not in protein-protein interactions.

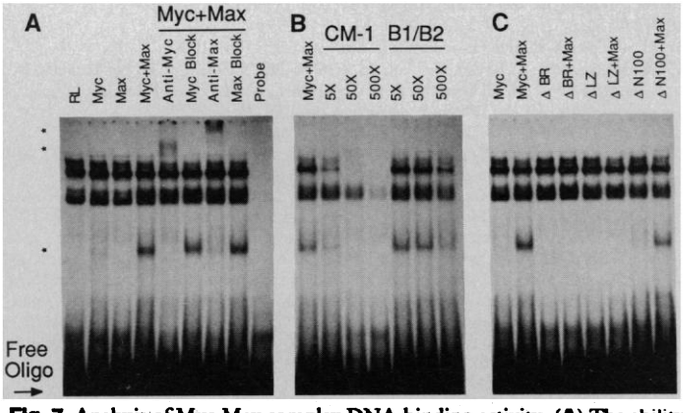


Fig. 7. Analysis of Myc-Max complex DNA-binding activity. (**A**) The ability of in vitro translated Myc and Max proteins to bind to the CM-1 oligonucleotide (CACGTG core consensus) was assessed by electrophoretic mobility shift assay. Post-translational mixes of Myc and Max were performed as in Fig. 6. Lysates were incubated with ^{32}P-labeled CM-1 prior to resolution in a 5 percent acrylamide gel. In experiments where antibodies were added, affinity purified antibody (1 μg) was added after formation of the nucleoprotein complex to minimize steric interference. To block the antibody effect, cognate immunogen (1 μg) was added. The positions of probe specifically bound and further retarded by antibody are indicated with asterisks. The arrow indicates free oligonucleotide. CM-1 oligonucleotide alone (probe) and unprogrammed reticulocyte lysate (RL) served as background controls. (**B**) The specificity of the Myc-Max shift was tested by competition with 5-, 50-, and 500-fold excess of unlabeled oligonucleotide. B1/B2 contains the 3′ MCK enhancer binding site for MyoD (CACCTG core consensus) and differs from CM-1 at only three positions. (**C**) Requirements for the formation of a nucleoprotein complex. Various c-Myc mutants (see Fig. 4 for abbreviations) were assayed for their ability to bind CM-1 in association with Max.

Implications for Myc function. Studies on the bHLH and bZip regions within a number of eukaryotic transcription factors have distinguished two essential yet separable functions for these domains: dimerization (HLH, Zip) and DNA-binding activity (basic region) (*9, 10*). Here we show that the bHLH-Zip domain of c-Myc is capable of specific interaction with a newly identified bHLH-Zip protein, Max. Our in vitro experiments are consonant with studies on the structure and properties of Myc (*1*) and may serve as a basis for understanding the mechanism of Myc function in vivo. Using anti-Myc, we have identified the Max protein in immunocomplexes from avian and human cells, a further indication that the Myc-Max association is likely to be biologically relevant (*16*).

Two regions within the c-Myc protein appear to be critical for c-Myc function as judged by assays for cotransformation, inhibition of differentiation, and suppression of endogenous Myc expression; these are (i) an approximately 40– to 60–amino acid segment centered about residue 120 and (ii) the 95–amino acid carboxyl-terminal region (*3–5*). Our results suggest that the carboxyl terminus mediates association with Max and formation of a sequence-specific DNA-binding complex. The mutations that negatively affected the ability of c-Myc to interact with Max and bind CM-1 (Fig. 4 and 7), such as deletion or disruption of the zipper, are either identical or very similar to those that abolish c-Myc activity in biological assays (*3–5*).

The dimerization function that resides within the c-Myc HLH-Zip domain appears to be independent of the basic region, which is likely to directly mediate DNA binding. The same c-Myc basic region deletion mutant that had no effect on association with Max completely abolished the ability of the Myc-Max complex to bind the CM-1 DNA probe (Fig. 7C). This basic region deletion mutation also abolished the ability of c-Myc to transform Rat 1 cells in collaboration with *bcr-abl*, while deletion of the adjacent upstream segment had no effect (*16*). Replacement of the c-Myc basic region with that of MyoD was likewise biologically inactive although, as expected, the chimeric protein was capable of association with Max (Fig. 4). Taken together, these results demonstrate that both the dimerization and DNA-binding activities that reside in the carboxyl-terminal bHLH-Zip domain are essential for important aspects of c-Myc activity.

A striking finding of our study is that Max interacts specifically with three members of the Myc family of proteins. Numerous attempts to demonstrate heterodimer formation between Myc and other bHLH, bZip, and bHLH-Zip proteins have not been successful (*10, 15, 16*). However, under our assay conditions Max is capable of associating with c-Myc, N-Myc, and L-Myc (Fig. 5A). Other proteins that contain related dimerization domains, including the bHLH-Zip proteins USF and AP-4, did not associate (Fig. 5B). Leucine zipper segments alone determine specificity in Fos-Jun association (*45*) and act to organize the two proteins in a parallel array (*13*). Max and the Myc proteins, however, all have HLH domains in addition to zipper regions, and our data show that the integrity of the HLH region is also important for heterodimer

formation (Fig. 4). If an initial interaction between parallel zipper regions is required for proper orientation, then the appropriate alignment of contiguous HLH regions might influence binding. In Max and the Myc family, the hydrophobic residues of the putative leucine zipper appear to maintain their heptad spacing well into helix II, possibly extending the coiled-coil interaction. By contrast, in USF (*40*) the heptad phasing is disrupted at the helix II–zipper boundary, and in AP-4 (*39*) the hydrophobic array does not extend as far into helix II (Fig. 3A). While it remains to be determined whether these differences are important for the apparent restricted specificities of binding, other factors are likely to influence association, including the size and composition of the loop region (*11*), the nature of specific residues within the helical segments (*46*), and the presence of other domains in the protein that may facilitate or block interaction. Although we have assumed that Myc and Max interact to form dimers, it is possible that they may also participate in higher order associations.

The fact that N-Myc and L-Myc as well as c-Myc specifically associate with Max suggests that Max may serve to integrate the functions of these three proteins that are differentially expressed during development, differentiation, and neoplasia (*1*). If so, Max might be expected to be expressed in at least as many cell types as are Myc family proteins. Initial experiments with Northern (RNA) blotting indicate that a 2.1-kb Max RNA is expressed in many cells and tissues at concentrations comparable to those of c-Myc. In addition, low stringency Southern (DNA) blot analysis suggests that Max is highly conserved as a single gene or a small family of genes in vertebrate genomic DNA, but is absent from invertebrates that also lack Myc homologs (*16*). These results are consistent with the possibility that Max, or a small number of Max-related proteins, interacts with Myc family proteins to mediate their specific biological functions. Whether Max can also be oncogenically activated poses an interesting biological question.

Important questions raised by this work concern the way in which the properties of Myc and Max are altered through association. Our experiments demonstrate that complex formation generates sequence-specific DNA-binding activity for the CM-1 oligonucleotide under conditions where neither Myc nor Max alone bound significant amounts of probe (Fig. 7). This oligonucleotide contains the CACGTG consensus, which serves as a binding site for presumptive Myc homodimers (*17, 19*). That this is a weak binding site may be reflected by the low but detectable binding by in vitro translated Myc. No binding by Max alone could be detected, indicating that either Max does not recognize CM-1 or that it does not homodimerize under the conditions of the assay. A key point becomes whether the Myc-Max heterocomplex has a specificity for DNA binding that is distinct from that of either of the homodimers. By analogy with MyoD and the E2A proteins, each member of a Myc-Max complex might contribute half-site recognition in defining DNA-binding specificity (*21*). The Myc-Max complex can be used directly to select a putative new binding sequence with the method for preferential binding and amplification of random sequences (*21*).

Another major question concerns the function of the Myc-Max complex. It has been suggested that Myc may function in transcription, DNA replication, or both (*1*). The characteristics of the Myc-Max complex places these proteins in the same general class as bHLH transcription factors, but the results do not rule out other possible functions. The essential amino-terminal region of c-Myc has been shown to act as a transcriptional activation domain when linked to yeast or prokaryotic DNA-binding domains (*47*). However, introduction of c-*myc* alone into cells only induces variable, usually low, activation of different promoters (*48*). While the HLH-Zip region constitutes the very carboxyl terminus of all the Myc family members, the HLH-Zip region of Max is nearly 50 residues from its carboxyl terminus (Fig. 3B). This region, which probably extends past the dimerized regions of Myc and Max, contains additional acidic and basic patches (Fig. 2) that could interact with components of the transcriptional machinery or other factors (*49*). Whatever their function, the ability of these polypeptides to form multiprotein complexes suggests that the differential regulation of their relative concentrations could be an important determinant of Max-Myc family associations, consequent DNA-binding specificities, and ultimately, the influence of Myc on cell proliferation and behavior.

REFERENCES AND NOTES

1. L. J. Z. Penn, E. M. Laufer, H. Land, *Semin. Cancer Biol.* **1**, 69 (1990); B. Lüscher and R. N. Eisenman, *Genes Dev.* **4**, 2025 (1990).
2. H. Abrams, L. Rohrschneider, R. N. Eisenman, *Cell* **29**, 427 (1982); P. Donner, I. Greiser-Wilke, K. Moelling, *Nature* **296**, 262 (1982); S. R. Hann and R. N. Eisenman, *Mol. Cell. Biol.* **4**, 2486 (1984); R. A. Watt, A. R. Shatzman, M. Rosenberg, *ibid.* **5**, 448 (1985); G. Ramsay, L. Stanton, M. Schwab, J. M. Bishop, *ibid.* **6**, 4450 (1986); J. De Greve *et al.*, *ibid.* **8**, 4381 (1988); C. V. Dang, H. V. Dam, M. Buckmire, W. M. F. Lee, *ibid.* **9**, 2477 (1989); K. Saksela, T. P. Makela, G. Evan, K. Alitalo, *EMBO J.* **8**, 149 (1989).
3. J. Stone *et al.*, *Mol. Cell. Biol.* **7**, 1697 (1987); P. J. Enrietto, *Virology* **168**, 256 (1989); H. Nakajima, M. Ikeda, N. Tsuchida, S. Nishimura, Y. Taya, *Oncogene* **4**, 999 (1989); D. H. Crouch, C. Lang, D. A. F. Gillespie, *ibid.* **5**, 683 (1990).
4. S. O. Freytag, C. V. Dang, W. M. F. Lee, *Cell Growth Differ.* **1**, 339 (1990).
5. M. J. Smith, D. C. Charron-Prochownik, E. V. Prowchownik, *Mol. Cell. Biol.* **10**, 5333 (1990); L. J. Z. Penn *et al.*, *ibid.*, p. 4961.
6. N. Jones, *Cell* **61**, 9 (1990); E. N. Olson, *Genes Dev.* **4**, 1454 (1990).
7. C. Murre, P. S. McCaw, D. Baltimore, *Cell* **56**, 777 (1989).
8. S. J. Tapscott *et al.*, *Science* **242**, 405 (1988).
9. A. Voronova and D. Baltimore, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4722 (1990).
10. C. Murre *et al.*, *Cell* **58**, 537 (1989).
11. R. L. Davis, P.-F. Cheng, A. B. Lassar, H. Weintraub, *ibid.* **60**, 733 (1990).
12. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
13. E. K. O'Shea, R. H. Rutkowski, P. S. Kim, *ibid.* **243**, 538 (1989).
14. C. R. Vinson, P. B. Sigler, S. L. McKnight, *ibid.* **246**, 911 (1989).
15. T. Kouzarides and E. Ziff, *Nature* **336**, 646 (1988).
16. E. M. Blackwood, B. Lüscher, R. N. Eisenman, unpublished data.
17. T. K. Blackwell, L. Kretzner, E. M. Blackwood, R. N. Eisenman, H. Weintraub, *Science* **250**, 1149 (1990).
18. C. V. Dang, M. McGuire, M. Buckmire, W. M. F. Lee, *Nature* **337**, 664 (1989).
19. G. Prendergast and E. Ziff, *Science* **251**, 186 (1991).
20. C. V. Dang *et al.*, *Mol. Cell. Biol.* **11**, 954 (1991).
21. T. K. Blackwell and H. Weintraub, *Science* **250**, 1104 (1990).
22. Y. Nakabeppu, K. Ryder, D. Nathans, *Cell* **55**, 907 (1988); T. D. Halazonetis, K. Georgopoulos, M. E. Greenberg, P. Leder, *ibid.*, p. 917; F. J. Rauscher III *et al.*, *Science* **240**, 1010 (1988); P. Sassone-Corsi, W. W. Lamph, M. Kamps, I. M. Verma, *Cell* **54**, 553 (1988); R. Chiu *et al.*, *ibid.*, p. 541.
23. J. M. Sikela and W. E. Hahn, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3038 (1987); H. Singh, J. H. LeBowitz, A. S. Baldwin, Jr., P. A. Sharp, *Cell* **52**, 415 (1988); C. R. Vinson, K. L. LaMarco, P. F. Johnson, W. H. Landschulz, S. L. McKnight, *Genes Dev.* **2**, 801 (1988); H. Singh, R. G. Clerc, J. H. LeBowitz, *BioTechniques* **7**, 252 (1989); P. F. Macgregor, C. Abate, T. Curran, *Oncogene* **5**, 451 (1990).
24. GST-MycC92 fusion protein was expressed from a pGEX-2T plasmid (Pharmacia) that contained the 570–base pair Ava II–Eco RI fragment of a human c-*myc* cDNA clone (0/1) ligated into the Sma I–Eco RI cloning sites. Fusion protein was purified as described [D. B. Smith and K. S. Johnson, *Gene* **67**, 31 (1988)]. GST (50 μg) or GST-MycC92 (50 μg) were ^{125}I-labeled to high specific activity (72 μCi/μg) with Iodobeads (Pierce) as recommended by the manufacturer [M. A. K. Markwell, *Anal. Biochem.* **125**, 427 (1982)].
25. The λgt11 expression library was constructed from the baboon lymphoid cell line 594S as described [R. L. Idzerda *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4659 (1989)].
26. The 594S λgt11 library was plated in the Y1088 bacterial strain. As plaques became visible, β-galactosidase fusion protein expression was induced by overlaying the lawns with IPTG [isopropyl β-D-thiogalactopyranoside (10 mM)]-impregnated nitrocellulose filters (Amersham; Hybond C Extra). Transfer of released proteins was allowed to proceed overnight. Filters were marked, rinsed to remove bacterial debris, and blocked with 5 percent dry milk in HND buffer [20 mM Hepes, pH 7.2, 50 mM NaCl, 0.1 percent NP-40, and 5 mM dithiothreitol (DTT)] for 1 hour at 4°C. ^{125}I-labeled GST or GST-MycC92 (100 ng/ml, about 3 nM) was added to the filters in HND buffer supplemented with 1 percent dry milk. After a 4-hour incubation on a rotating platform at 4°C, filters were rapidly washed seven times with PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \cdot 7H_2O$, 1.4 mM KH_2PO_4, pH 7.3) that contained 0.2 percent Triton X-100 (room temperature). Filters wrapped in plastic were exposed to x-ray film for 3 hours to overnight [see (*23*) for related protocols].
27. S. R. Hann and R. N. Eisenman, *Mol. Cell. Biol.* **4**, 2486 (1984).
28. Sequence analysis of Max11 and Max14 clones, along with Max clones derived from a Manca λgt10 library, was performed by the dideoxy method [F. Sanger, S. Miklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977)]. The 513-nucleotide sequence presented (Fig. 2) was constructed from two overlapping

Manca cDNA clones.
29. In retrospect, it is not surprising to have cloned only two functional inserts from the 10^6 plaques screened. Size selection of the cDNA inserts along with the presence of an in-frame stop codon located two codons 5′ to the initiating AUG (see Fig. 2) limits the number of potentially functional points of *lacZ* fusion (that is, those that contain an intact HLH-Zip region) to 40. For comparison, screening of the 594S λgt11 library with a c-Myc carboxyl-terminal specific antiserum identified only 12 immunoreactive plaques.
30. M. Kozak, *J. Cell Biol.* **108**, 229 (1989).
31. M. Cai and R. W. Davis, *Cell* **61**, 437 (1990); R. Benezra, R. L. Davis, D. Lockshon, D. L. Turner, H. Weintraub, *ibid.*, p. 49.
32. Oligonucleotide-directed mutagenesis was used to generate a variety of mutant Myc proteins: ΔC89, deletion of the carboxyl-terminal 89 residues (deleted amino acids 351 to 439, added Arg-Arg-Thr-Ser); ΔBR, deletion of the basic region (deleted amino acids 353 to 367); ΔCKII, deletion of the casin kinase II phosphorylation site located 5′ of the basic region (deleted amino acids 346 to 354); ΔHelix 1, deletion of helix I (deleted amino acids 368 to 381); ΔLZ, deletion of the leucine zipper (deleted amino acids 416 to 439); BR21M, replace the basic region of Myc with that of MyoD (replaced Myc amino acids 347 to 367 with MyoD amino acids 102 to 122); BR21MΔLZ, double mutant that consists of BR21M and ΔLZ; ProZip, replacement of Leu (amino acid residue 420) with Pro (from A. J. Street). Deletion of sequences 5′ to the Pvu II site in the 0/1 cDNA resulted in ΔN100; translation from this construct initiates at the first internal methionine (amino acid 101). Numbering corresponds to the amino acid sequence of human c-Myc [R. Watt *et al., Nature* **303**, 725 (1983)]. Dideoxy sequencing and immunoprecipitation of in vitro translation products were used to confirm the identity of each construct.
33. S. R. Hann, M. W. King, D. L. Bentley, C. W. Anderson, R. N. Eisenman, *Cell* **52**, 185 (1988); A. J. Street, D. Bentley, R. N. Eisenman, unpublished data.
34. In vitro transcription and translation were performed under conditions recommended by the Promega Protocols and Applications Guide. [^{35}S]Methionine labeled proteins were produced from each of the following vectors: pVZ1-Max11/13/14, pBluescribe vectors that contained the mutant Myc constructs, pU313S (L-Myc) (from K. Alitalo), pNmycB (N-Myc) (from R. Bernards), pV2C11α (MyoD) (from A. Lassar) (*11*), E12R (E12) (from C. Murre) (7), pBS-B065 (myogenin) (from W. Wright) (*38*), tal-SP6pGem (Tal) (from R. Baer) (*37*), pJun7′8 (Jun) (from R. Turner), pSP65fos1B (Fos) (from T. Curran), Δ12.2 (USF) (from R. Roeder) (*40*), and T7βAP-4 (AP-4) (from Y.-F. Hu and R. Tjian) (*39*). Programmed reticulocyte lysate (1 μl) was subjected directly to SDS-PAGE, or lysates (20 μl) were diluted into HND buffer (400 μl) (*26*) that contained bovine serum albumin (BSA) (10 mg/ml). Half of this dilution was incubated with either GST or GST-Max124 beads [approximately 5 μg of fusion protein adsorbed to 10 μl of glutathione-Sepharose (Pharmacia)] for 1 hour at 4°C. The resin was then washed four times with PBS that contained NP-40 (0.1 percent) at room temperature. The bound proteins were eluted with SDS-containing sample buffer and subjected to SDS-PAGE and autoradiography.
35. GST-MaxC124 was constructed by insertion of the Ava II–Eco RI fragment of Max14 into the Sma I site of pGEX-3X expression vector (Pharmacia). The resulting fusion protein had the 124 carboxyl-terminal amino acids of Max in frame with GST sequences. Fusion protein was purified as described in (*24*).
36. B. Lüscher, E. A. Kuenzel, E. G. Krebs, R. N. Eisenman, *EMBO J.* **8**, 1111 (1989).
37. Q. Chen *et al., ibid.* **9**, 415 (1990).
38. W. E. Wright, D. A. Sassoon, V. K. Lin, *Cell* **56**, 607 (1989).
39. Y.-F. Hu, B. Lüscher, A. Admon, N. Mermod, R. Tjian, *Genes Dev.* **4**, 1741 (1990).
40. P. D. Gregor, M. Sawadogo, R. G. Roeder, *ibid.*, p. 1730.
41. In vitro transcripts from the c-Myc and Max vectors were added to a Promega reticulocyte lysate translation mixture and incubated for 1 hour at 30°C. c-Myc and Max (2:1) were mixed after translation, thus compensating for differences in the translational efficiencies of the two RNA species; association was allowed to proceed for 30 minutes at 30°C, after which the lysates were diluted into PBS with 1 percent NP-40. Proteins were immunoprecipitated under these mild conditions with anti-Myc (5 μg of affinity purified) (*27*) or anti-Max (5 μl of polyclonal antisera to the GST-Max124 fusion protein) (*35*). Immunoprecipitations were blocked by the addition of the cognate immunogen. Antigen-antibody complexes were isolated on protein A–Sepharose beads (Sigma), and the pellets were washed five times with PBS that contained 1 percent NP-40. The [^{35}S]Methionine labeled samples were analyzed by SDS-PAGE under reducing conditions.
42. Max or c-Myc transcripts were translated in vitro with nonradioactive methionine. Post-translational mixes were performed as in (*41*), and the resulting lysates were analyzed for binding to the synthetic CM-1 oligonucleotide by the electrophoretic mobility shift assay [A. Revzin, *BioTechniques* **7**, 346 (1989)]. Final conditions within a 25-μl binding reaction were: 20 mM Hepes, pH 7.2, 50 mM KCl, 3 mM $MgCl_2$, 1 mM DTT, 1 mM EDTA, 8 percent glycerol, 25 ng of sheared salmon sperm DNA as a nonspecific competitor, 10 μl of programmed reticulocyte lysate, and 0.2 ng of ^{32}P-labeled CM-1 oligo (*17*). The DNA-binding reaction was allowed to proceed at room temperature for 10 minutes. For antibody experiments, affinity purified anti-Myc or anti-Max (1 μg) was added for 10 minutes after the formation of the nucleoprotein complex; the cognate immunogen (10 μg) blocked this supplemental shift. As competitors, double-stranded oligonucleotides were added at 1, 10, and 100 ng per reaction; the core sequence of the B1/B2 and CM-1 templates are 5′-CCCCCAACACCTGCTGCCTGA-3′ and 5′-CCCCCACCAC-GTGGTGCCTGA-3′, respectively (*17*). Protein-DNA complexes were resolved on a 5 percent acrylamide gel (50 mM tris base, 50 mM borate, 1 mM EDTA), and gels were dried prior to autoradiography.
43. The background bands in the gel retardation assays were due to endogenous USF binding factor activity, and binding of USF could be inhibited by the addition of the CM-1 probe. USF specifically binds to the CM-1 consensus [R. W. Carthew, L. A. Chodosh, P. A. Sharp, *Cell* **43**, 439 (1985); M. Sawadogo and R. Roeder, *ibid.*, p. 165; A. C. Lennard and J. M. Egly, *EMBO J.* **6**, 3027 (1987)] and is present in the reticulocyte lysates, as evidenced by the ability of antibodies to USF to alter the mobility of these bands (L. Kretzner, unpublished data). Antibodies to USF were provided by M. Sawadogo.
44. A. B. Lassar *et al., Cell* **58**, 823 (1989).
45. E. K. O'Shea, R. Rutkowski, W. F. Stafford III, P. S. Kim, *Science* **245**, 646 (1989).
46. J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, *ibid.* **250**, 1400 (1990).
47. K. Lech, K. Anderson, R. Brent, *Cell* **52**, 179 (1988); G. J. Kato, J. Barrett, M. Villa-Garcia, C. V. Dang, *Mol. Cell. Biol.* **10**, 5914 (1990).
48. R. E. Kingston, A. S. Baldwin, P. A. Sharp, *Nature* **312**, 280 (1984); R. Kaddurah-Daouk, J. M. Greene, A. S. Baldwin, R. E. Kingston, *Genes Dev.* **1**, 347 (1987); R. Onclercq, P. Gilardi, A. Lavenu, C. Cremisi, *J. Virol.* **62**, 4533 (1988).
49. M. Ptashne, *Nature* **335**, 683 (1988); P. B. Sigler, *ibid.* **333**, 210 (1988); P. J. Mitchell and R. Tjian, *Science* **245**, 371 (1989).
50. We thank T. St. John for the phage library and advice on its use; K. Alitalo, R. Baer, R. Bernards, T. Curran, L. Graf, M. Groudine, W. Hokanson, Y.-F. Hu, A. Lassar, T. Mäkelä, R. Roeder, M. Sawadogo, A. J. Street, R. Tjian, D. Turner, R. Turner, and J. Wallace for gifts of reagents; and T. K. Blackwell, J. Cooper, L. Kretzner, B. Lüscher, T. Palmer, and H. Weintraub for critical readings of the manuscript as well as advice and encouragement. Supported by the Molecular Training Program in Cancer Research at the University of Washington (NIH grant T32 CA09437)(E.M.B.) and by NIH grants RO1 CA20525 and PO1 CA28151 (R.N.E.).

9 January 1991; accepted 30 January 1991