MOLECULAR AND CELLULAR BIOLOGY, Feb. 1993, p. 792–800
0270-7306/93/020792-09$02.00/0

Vol. 13, No. 2

A New Transcriptional-Activation Motif Restricted to a Class of Helix-Loop-Helix Proteins Is Functionally Conserved in Both Yeast and Mammalian Cells

MELANIE W QUONG, MARK E. MASSARI, RONALD ZWART, AND CORNELIS MURRE*

Department of Biology, 0116, and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093*

Received 29 September 1992/Returned for modification 28 October 1992/Accepted 4 November 1992

Previous studies demonstrated that the amino-terminal portions of E2A and E2-2 are crucial for transactivation. Subsequent findings showed that the same amino-terminal region of E2A is involved in two different translocation events contributing to the induction of a pre-B-cell acute lymphoblastic leukemia and a pro-B-cell acute lymphoblastic leukemia. These results led us to focus on the amino-terminal region of E2A to better understand its normal role in transcriptional regulation and its aberrant involvement in the two leukemias. We report here the identification of two conserved boxes in the E2A amino-terminal domain that show extensive homology within the transactivation domains of E12, E47, E2-2, HEB, and daughterless, all members of the same class of helix-loop-helix proteins. Together, both boxes are crucial for transcriptional activation and have the potential to form a new activation motif, that of a loop adjacent to an amphipathic α-helix, designated the loop-helix (LH) motif. A minimal region containing the LH motif is sufficient for transcriptional activation. Point mutations in the amphipathic helix of the minimal region reduce its transactivation capabilities dramatically. The same constructs expressed in yeast cells show identical patterns of activation, suggesting that the LH motif and its target proteins are functionally conserved in yeast cells. We propose that the LH motif represents a novel transactivation domain that is distinct from the previously characterized acidic blob, proline-rich, and glutamine-rich activation motifs. In addition, the LH motif is the first activation motif restricted to one class of DNA binding proteins.

The helix-loop-helix (HLH) proteins are a growing family of proteins containing a conserved structural motif (HLH) that mediates DNA binding and dimerization (7, 18a, 19, 27). These proteins play a major role in the control of cell type differentiation by forming homo- and heterodimers with themselves and other HLH proteins (7, 18a, 19, 27). The members of one particular class of HLH proteins (class I) are closely related (70 to 90% identity in the HLH region), are ubiquitously expressed, and include E12 (7, 18a, 19, 21, 27), E47 (7, 9, 18a, 19, 21, 27), E2-2 (9), HEB (10), and daughterless (3, 6). E12 and E47 are encoded by the same gene (E2A) but arise through differential splicing (26). The E2A gene products are involved in B-cell-, muscle-, and pancreas-specific gene expression (2, 4, 14, 20). For example, the presence of the E2A gene products or other ubiquitous class I HLH molecules is necessary for the mediation of specific differentiation programs by the formation of heterodimers with tissue-specific HLH proteins, such as MyoD, myogenin, and members of the achaete-scute protein family.

In contrast to its normal role in regulating transcription, the E2A gene has been found to be translocated in two acute lymphoblastoid leukemias (ALL). In pro-B-cell ALL, containing a t(17;19) translocation, the E2A N-terminal domain is fused to a leucine zipper-like domain (11). In pre-B-cell ALL, containing a t(1;19) translocation, the E2A N-terminal domain is fused to a homeobox-like domain, termed pbx1 (13). The involvement of the E2A N-terminal domain in two translocation events is puzzling and suggests an important function for this portion of the protein.

Here, we identify a strikingly conserved region in the N-terminal regions of E2A, E2-2, HEB, and daughterless. We show that this conserved region is important for transcriptional activation of these proteins and has the potential to form a structure consisting of a loop adjacent to an amphipathic α-helix (designated the loop-helix [LH] motif). Several amino acid substitutions of conserved residues within the putative LH motif obliterate the transactivation capability of the LH domain in both yeast and mammalian cell lines. Thus, we describe a new activation motif that is restricted to one class of DNA binding proteins and that is functionally conserved in yeast, *Drosophila*, and mammalian cells.

MATERIALS AND METHODS

Cell culture and transfections. HeLa and COS cells were cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum. For transfections, cells were plated at a density of 10^6 cells per plate on 10-cm plates. At 12 to 24 h later, the cells were transfected by the DEAE-dextran method. The total amount of transfected DNA was kept constant by the addition of an appropriate amount of carrier vector DNA. At 40 to 48 h posttransfection, the cells were harvested and equal amounts of protein were assayed for chloramphenicol acetyltransferase (CAT) activity. After autoradiography, the spots were excised and quantitated by liquid scintillation counting.

Western blot (immunoblot) analysis. COS cells were used to quantitate protein levels, since higher levels of expression are achieved with these cells. COS cells cotransfected with the reporter and activator constructs were harvested at 40 to 48 h posttransfection, and each plate was divided in two for CAT activity determinations (see above) and for Western

* Corresponding author.

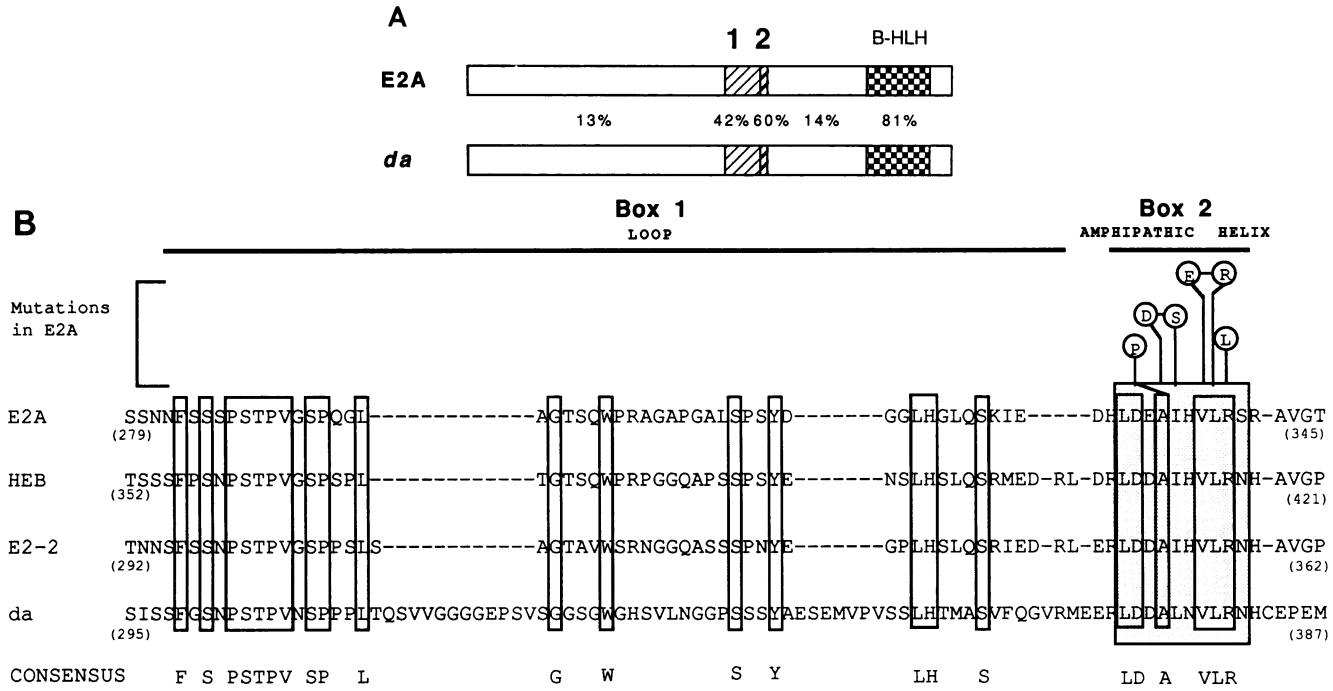


FIG. 1. Identification of a new motif conserved in class I HLH proteins. (A) Homology between E2A and daughterless (*da*). Shown are the percentages of amino acid homology between E2A and daughterless for box 1 and box 2 (1 and 2, respectively), the HLH (B-HLH), and the regions surrounding them. (B) Comparison of the amino acid sequences (single-letter code) of the region containing box 1 and box 2 of four class I HLH proteins: E2A, HEB, E2-2, and daughterless (da). Solid lines mark the loop and the potential amphipathic helix. Boxed amino acids signify conserved residues. The numbers below the sequences indicate the amino acid position of each protein. Single and double mutations made in the E2A protein are designated as circled letters above the sequences. Letters connected by lines indicate that multiple amino acid substitutions were made within the same molecule. Letters not connected by lines represent single-residue substitutions.

blot analysis. Fifty micrograms of each extract was fractionated on a 5 to 15% polyacrylamide gradient gel, transferred to Immobilon-P (Millipore), and probed with a polyclonal anti-GAL4 (amino acids 1 to 147) antibody (kindly provided by M. Ptashne). The blots were then probed with ^{125}I-conjugated protein A and visualized by autoradiography.

Generation of deletions and mutants. 5′ and 3′ deletions in E2A and E2-2 were generated by use of appropriate restriction enzyme sites and by digestion with Bal 31 enzyme. The deletions were treated with Klenow or T4 DNA polymerase as needed to maintain the correct reading frame and were subsequently fused in frame to GAL4 (1 to 147) in expression vector plasmid pBXG1. All constructs were verified by sequencing (in accordance with manufacturer instructions [U.S. Biochemicals Corp.]).

Mutations were generated by site-directed mutagenesis with mismatched oligonucleotides (in accordance with manufacturer instructions [Bio-Rad Laboratories]) and verified by sequencing (as described above). The oligonucleotides used to generate the mutations were as follows: 5′-CTG GAC GAG CCC ATC CAC G-3′ (for 334P); 5′-CGA GGA CAG GCA CGT GCT CC-3′ (for 334D/335S); 5′-CGT GCT CCT CAG CCA CGC C-3′ (for 339L); and 5′-ATC CAC GAG CGC CGC AGC C-3′ (for 337E/338R). The mutant nomenclature defines the position and the new residue. For example, mutant 334P means that a proline residue was substituted for the wild-type residue at position 334.

Yeast expression plasmids. Both the wild-type and mutant E2A sequences containing the LH region were excised as *Eco*RI-*Bam*HI fragments from the pBXG1 parent constructs. These were ligated into pMA424 (2μm containing *HIS3* and the *ADH1* promoter [16]) cut with *Eco*RI and *Bam*HI. This step created an in-frame fusion with GAL4 (1 to 147).

Media and β-galactosidase assay. Yeast strain GGY::171 (*Δgal4 Δgal80 his3 leu2 URA3::GAL1-lacZ* [8]) was transformed with various GAL4-E2A expression plasmids by the lithium acetate method (12). Individual transformants were grown to saturation in 5 ml of SC-His supplemented with 2% glucose (24a). Cultures were diluted 1:25 into SC-His supplemented with 2% galactose, 2% ethanol, and 3% glycerol and grown to an optical density at 600 nm of ~1. β-Galactosidase activity was determined by lysing cells with glass beads as described previously (24). Activity was normalized to the protein concentration as determined by the Bio-Rad protein assay.

RESULTS

Deletions in E2A and E2-2 define a conserved region crucial for transcriptional activation. Previous studies demonstrated that the N-terminal portion (amino acids 1 to 426 and 1 to 451 of E2A and E2-2, respectively) is crucial for transcriptional activation (9). Using dot matrix analysis, we noted that within the N-terminal portion, two stretches of amino acids that are extremely well conserved in class I HLH proteins are present (Fig. 1A). Within this sequence we noted two striking features: a stretch of highly conserved proline and serine residues (Fig. 1B) and a region of sequence similarity containing virtually identical hydrophilic and hydrophobic residues (Fig. 1B). To determine whether these two conserved boxes play an important role in the transactivation

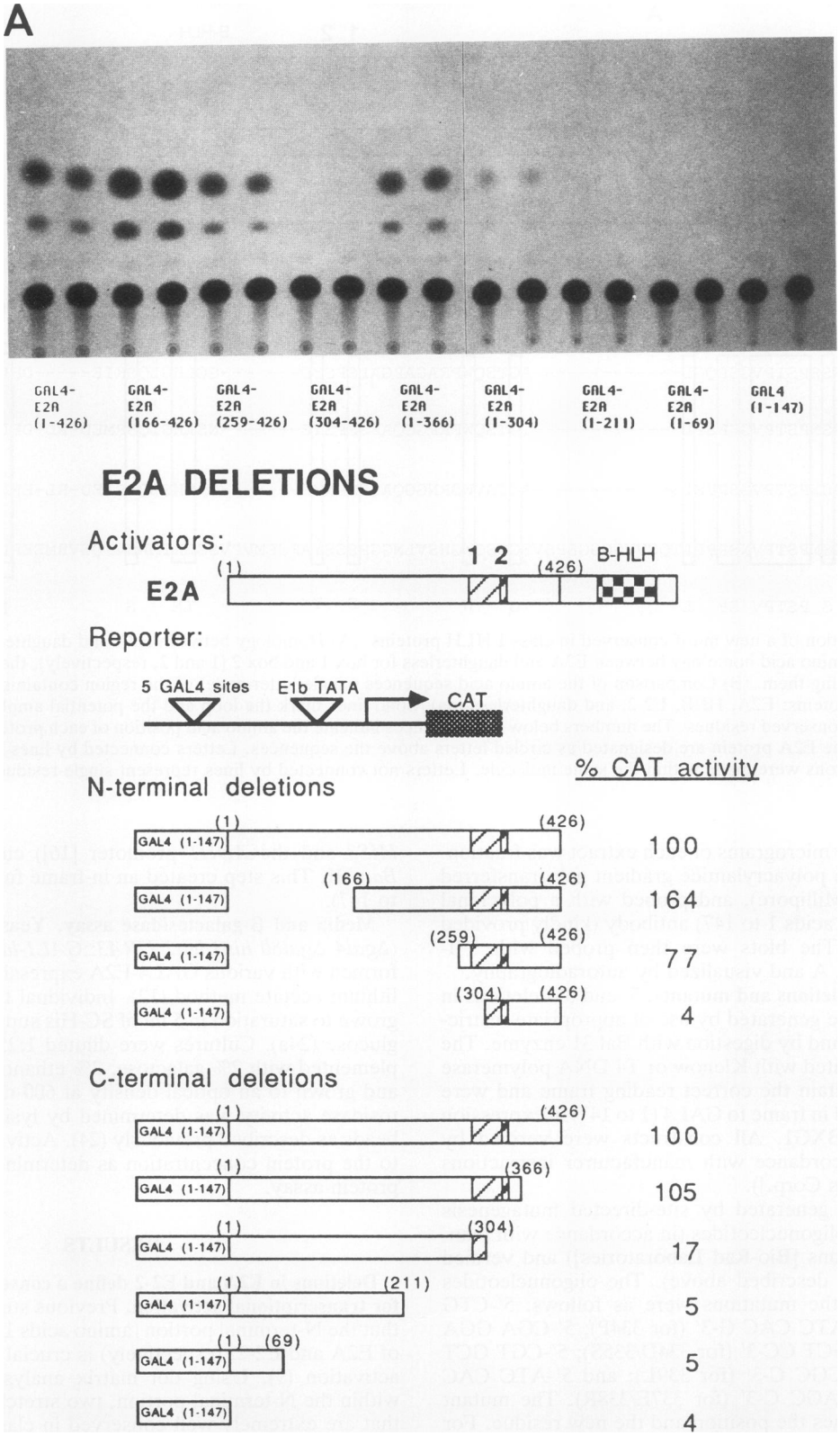
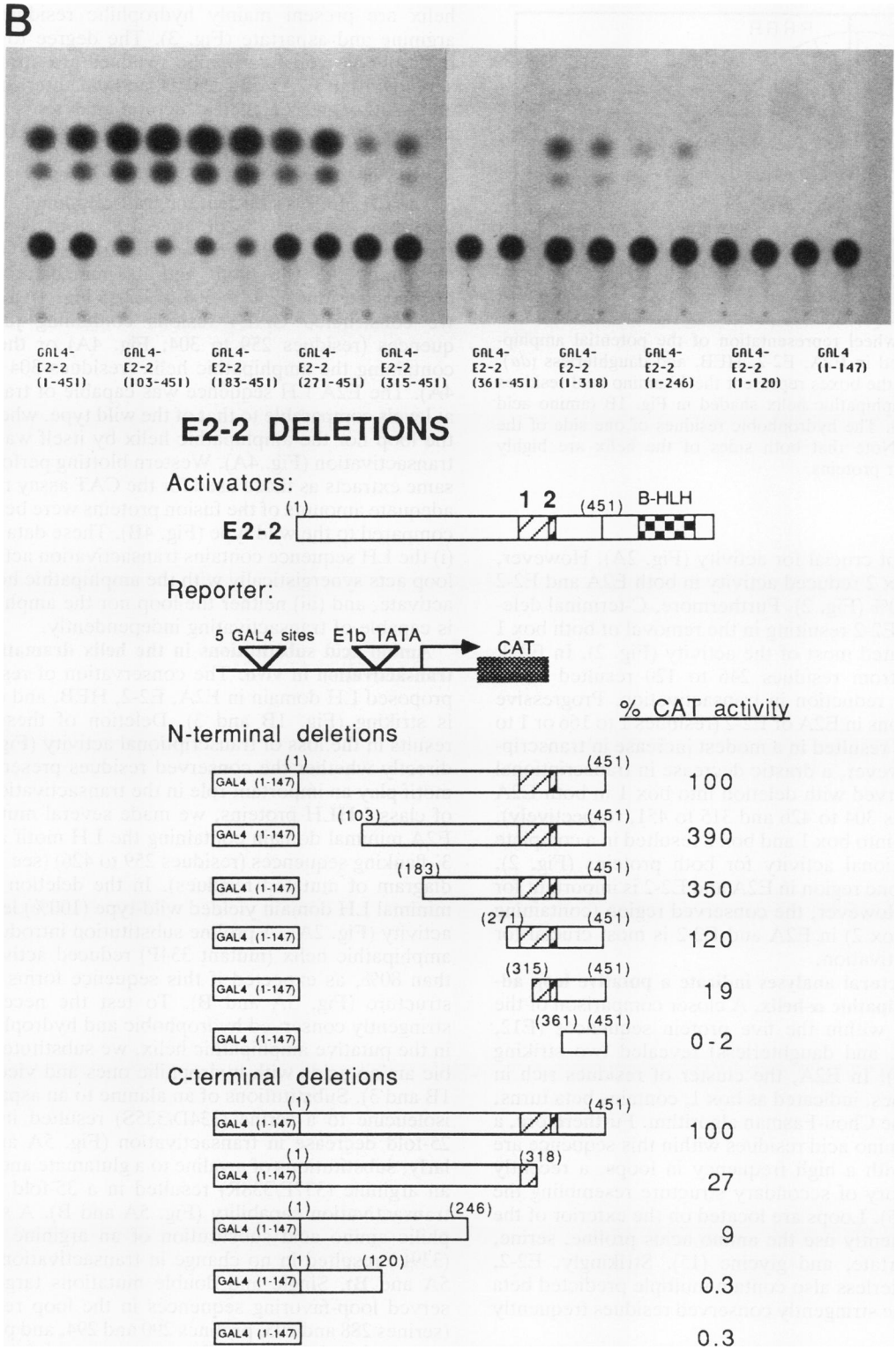


FIG. 2. The LH motif in E2A and E2-2 is crucial for transcriptional activation. The transcriptional activities of the wild type and N-terminal and C-terminal deletions of E2A (A) and E2-2 (B) were examined. The wild-type and truncated versions were fused to the GAL4 DNA binding domain. The structures of the reporter and activator plasmids are diagrammed. Ten micrograms of the CAT reporter plasmid containing five GAL4 binding sites and the adenovirus E1b TATA box was cotransfected along with 5 μg of expression plasmid encoding the different activator proteins of E2A and E2-2. Results of independent duplicates are presented. Plasmids encoding the full-length amino-terminal proteins (amino acids 1 to 426 for E2A and 1 to 451 for E2-2) were considered to yield wild-type levels of CAT activity and



were assigned the value of 100. The activities of the deletions were normalized accordingly. Relative CAT activities (averages of the independent duplicates presented) as tested in HeLa cells are stated to the right of the diagrams. The relative CAT activities for N-terminal deletions from 166 to 426 and 259 to 426 were averages of duplicates from a separate identical experiment. All assays presented were repeated at least two times, with independent duplicates yielding comparable results.

ability of E2A and E2-2, we made a series of deletions (Fig. 2). Deletions generated from both the 5′ and the 3′ ends were subsequently fused in frame to the GAL4 DNA binding domain (Fig. 2). The expression plasmids together with a CAT reporter plasmid containing five GAL4 DNA binding sites upstream of the E1b TATA were cotransfected into HeLa cells and subsequently measured for CAT activity. Progressive C-terminal deletions showed that residues 366 to

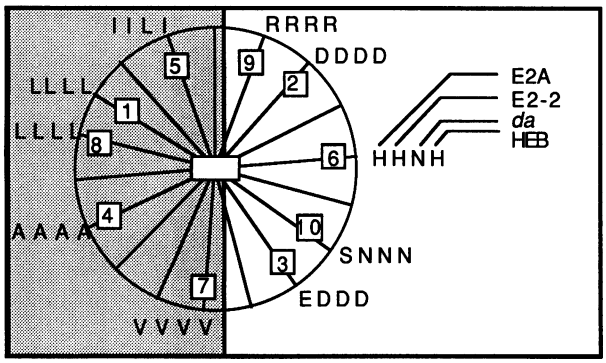


FIG. 3. Helical wheel representation of the potential amphipathic helix conserved in E2A, E2-2, HEB, and daughterless (*da*). Numbers 1 to 10 in the boxes represent the 10 amino acid residues of the proposed amphipathic helix shaded in Fig. 1B (amino acid residues 331 to 340). The hydrophobic residues of one side of the helix are shaded. Note that both sides of the helix are highly conserved in all four proteins.

426 of E2A are not crucial for activity (Fig. 2A). However, the deletion of box 2 reduced activity in both E2A and E2-2 by greater than 70% (Fig. 2). Furthermore, C-terminal deletions in E2A and E2-2 resulting in the removal of both box 1 and box 2 eliminated most of the activity (Fig. 2). In E2-2, further deletion from residues 246 to 120 resulted in an additional 30-fold reduction in transactivation. Progressive N-terminal deletions in E2A or E2-2 (residues 1 to 166 or 1 to 183, respectively) resulted in a modest increase in transcription (Fig. 2). However, a drastic decrease in transcriptional activity was observed with deletion into box 1 in both E2A and E2-2 (residues 304 to 426 and 315 to 451, respectively). Further deletions into box 1 and box 2 resulted in a complete loss of transcriptional activity for both proteins (Fig. 2). Thus, more than one region in E2A and E2-2 is important for transactivation. However, the conserved region (containing both box 1 and box 2) in E2A and E2-2 is most crucial for transcriptional activation.

Secondary structural analyses indicate a putative loop adjacent to an amphipathic α-helix. A closer comparison of the conserved boxes within the five protein sequences (E12, E47, E2-2, HEB, and daughterless) revealed two striking features (Fig. 1B). In E2A, the cluster of residues rich in serines and prolines, indicated as box 1, contains beta turns, as predicted by the Chou-Fasman algorithm. Furthermore, a majority of the amino acid residues within this sequence are residues found with a high frequency in loops, a recently recognized category of secondary structure resembling the Greek letter Ω (15). Loops are located on the exterior of the protein and frequently use the amino acids proline, serine, asparagine, aspartate, and glycine (15). Strikingly, E2-2, HEB, and daughterless also contain multiple predicted beta turns and the same stringently conserved residues frequently found in loops.

A closer examination of the second conserved box, indicated as box 2, revealed a number of absolutely conserved hydrophobic and hydrophilic residues (Fig. 1B). In each of the gene products analyzed, Chou-Fasman analysis predicted this stretch of amino acids to adopt an α-helical structure. This prediction prompted us to plot box 2 on a helical wheel (Fig. 3). Five highly conserved hydrophobic residues (valine, alanine, leucine, leucine, and isoleucine) are present on one side of the helix. On the other side of the helix are present mainly hydrophilic residues, including arginine and aspartate (Fig. 3). The degree to which both hydrophobic and hydrophilic residues are stringently conserved within E2A, E2-2, HEB, and daughterless is remarkable. Although a detailed structural analysis is required, we propose that as a working model, this conserved motif might be a loop joined to an amphipathic helix, designated the LH motif.

The LH motif is sufficient for transcriptional activation. To determine whether the putative LH motif is sufficient for transcriptional activation, we constructed a GAL4 fusion containing the LH motif and its immediate flanking sequences (residues 259 to 366 of E2A; Fig. 4A). In addition, we constructed GAL4 fusions containing just loop sequences (residues 259 to 304; Fig. 4A) or the sequences containing the amphipathic helix (residues 304 to 366; Fig. 4A). The E2A LH sequence was capable of transactivation at levels comparable to that of the wild type, whereas neither the loop nor the amphipathic helix by itself was capable of transactivation (Fig. 4A). Western blotting performed on the same extracts as those used for the CAT assay revealed that adequate amounts of the fusion proteins were being made, as compared to the wild type (Fig. 4B). These data suggest that (i) the LH sequence contains transactivation activity, (ii) the loop acts synergistically with the amphipathic helix to transactivate, and (iii) neither the loop nor the amphipathic helix is capable of transactivating independently.

Amino acid substitutions in the helix dramatically reduce transactivation in vivo. The conservation of residues in the proposed LH domain in E2A, E2-2, HEB, and daughterless is striking (Fig. 1B and 3). Deletion of these sequences results in the loss of transcriptional activity (Fig. 2). To test directly whether the conserved residues present in the LH motif play an important role in the transactivation capability of class I HLH proteins, we made several mutations in an E2A minimal domain containing the LH motif and adjacent 3′-flanking sequences (residues 259 to 426) (see Fig. 1B for a diagram of mutated residues). In the deletion assays, this minimal LH domain yielded wild-type (100%) levels of CAT activity (Fig. 2A). A proline substitution introduced into the amphipathic helix (mutant 334P) reduced activity by more than 80%, as expected if this sequence forms an α-helical structure (Fig. 5A and B). To test the necessity of the stringently conserved hydrophobic and hydrophilic residues in the putative amphipathic helix, we substituted hydrophobic amino acids with hydrophilic ones and vice versa (Fig. 1B and 3). Substitutions of an alanine to an aspartate and an isoleucine to a serine (334D/335S) resulted in a dramatic 25-fold decrease in transactivation (Fig. 5A and B). Similarly, substitutions of a valine to a glutamate and a leucine to an arginine (337E/338R) resulted in a 35-fold reduction in transactivation capability (Fig. 5A and B). A single hydrophilic amino acid substitution of an arginine to a leucine (339L) resulted in no change in transactivation levels (Fig. 5A and B). Single and double mutations targeted to conserved loop-favoring sequences in the loop region of E2A (serines 288 and 293, prolines 290 and 294, and phenylalanine 283) had little or no effect on transactivation (data not shown). To determine whether the differences in the activities of the various mutants could have been due to variations in protein levels, we performed a Western blot analysis. The various GAL4 derivatives were of the expected size and were expressed at comparable levels (data not shown). Thus, the ability of E2A to transactivate is highly sensitive to mutations that affect either the amphipathy or the secondary structure of its amphipathic helix, whereas mutations within

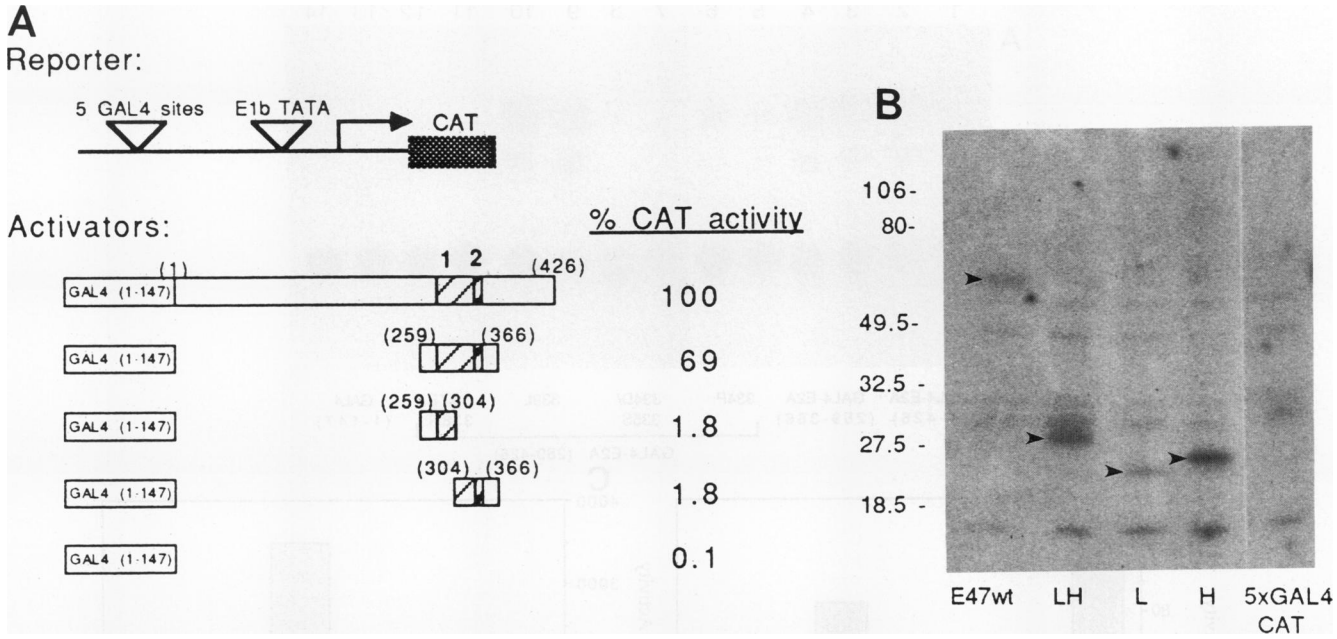


FIG. 4. The LH motif is sufficient for transcriptional activation. (A) The transactivation abilities of the LH motif, the loop alone, and the amphipathic helix alone were compared with those of wild-type E2A. The corresponding GAL4 fusion proteins are diagrammed with their relative CAT activities. (B) Western blot showing levels of expression of the LH domain (LH)-, loop (L)-, and amphipathic helix (H)-containing GAL4 fusion proteins. Arrowheads indicate the correctly sized protein products. E47wt is the GAL4 fusion to the wild-type E2A (residues 1 to 426). 5×GAL4CAT is the GAL4 DNA-binding domain (residues 1 to 147) alone. The transactivation assay in COS cells was performed as stated in Materials and Methods with 5 μg of the reporter construct and 0.5 μg of the activator constructs. The transfections were repeated at least two times and were similarly performed in HeLa cells, which yielded comparable patterns of transcriptional activation.

the loop do not have an impact on the transactivation capability of E2A.

The LH motif has a common cellular target. The extraordinary conservation of the LH domain suggests that E12, E47, E2-2, HEB, and daughterless might interact with a common cellular target. To determine whether the LH sequence present in E2A, E2-2, and daughterless has a common target, we tested whether the overexpression of E47, E2-2, and daughterless could inhibit the transactivation activity of the GAL4-E2A minimal LH domain (Fig. 6). This inhibition or "squelching" would result from the full-length protein E47, E2-2, or daughterless competing for and thereby titrating away a common target from the LH activating region. All three HLH proteins competed efficiently for the transactivation capability of the E2A LH motif (Fig. 6), whereas the control parent plasmid, pJ3Ω, failed to compete (data not shown). These experiments suggest that the E2A LH target can interact with other class I HLH proteins, including daughterless and E2-2.

The LH-interacting target is functionally conserved in yeast and mammalian cells. Recent studies have shown that both E47 (9) and daughterless/lethal of scute heterodimers can activate transcription from their cognate sequences in yeast cells (2). Thus, we tested whether the LH domain could transactivate in yeast cells. Indeed, the wild-type LH domain strongly activated the β-galactosidase reporter gene (Fig. 5C). This result raises the question of whether the target interacting with the LH motif is conserved in yeast cells as well. To test this idea, we examined whether amino acid substitutions within the proposed helix had a similar impact on transactivation in yeast cells and in mammalian cells. A proline substitution introduced into the helix reduced activity to approximately 30%. Substitutions of a conserved alanine to an aspartate and an isoleucine to a serine resulted in a fourfold reduction in transactivation capability (Fig. 5C). Similarly, substitutions of a valine to a glutamate and a leucine to an arginine reduced transactivation by a factor of five (Fig. 5C). Interestingly, as in mammalian cells, substitution of the conserved arginine within the helix to a leucine did not reduce the transactivation capability of the LH domain (Fig. 5C). A Western blot revealed that mutant protein levels were comparable to those in the wild type (data not shown). These data correlate extremely well with the results obtained for mammalian cells (Fig. 5B and C). Thus, mutations in the putative α-helix have similar effects in both yeast and mammalian cells. These data and the results described above suggest that the molecule(s) interacting with the LH domain is likely to be conserved in yeast, *Drosophila*, and mammalian cells.

DISCUSSION

In this study, we identified a transactivation domain, the LH motif, present in a class of HLH proteins. Although it is clear from our data that the LH motif is responsible for a majority of the transcriptional activation of E2A and E2-2, it is not the sole determinant required for full transcriptional activation. Our deletion studies (Fig. 2) revealed a modest increase in transcriptional activation in both E2A and E2-2 which was eradicated upon further deletion. Furthermore, in E2-2, subsequent deletion from the C terminus of residues 246 to 120 resulted in an additional 30-fold reduction in transactivation. Similarly, when the E2A minimal LH domain (residues 259 to 366) was fused to GAL4, it accounted

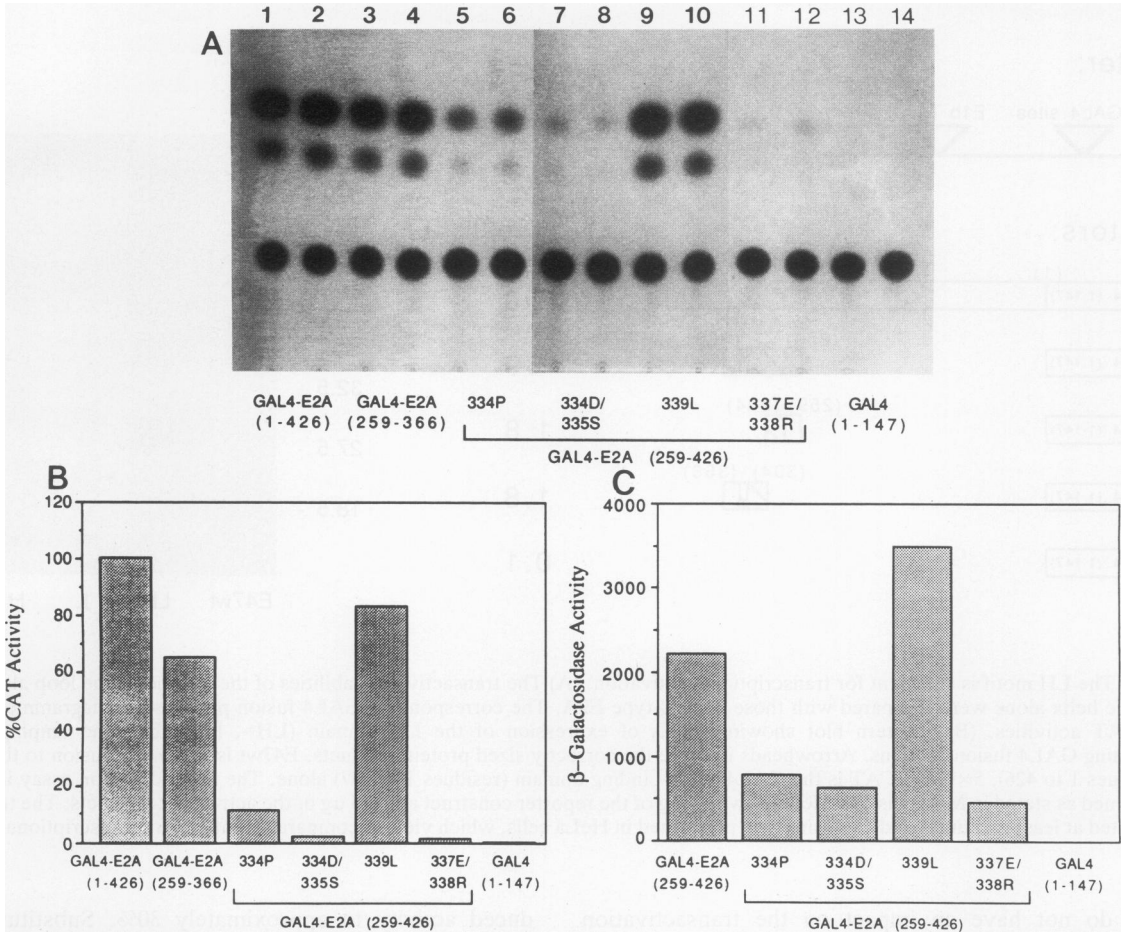


FIG. 5. Conserved hydrophobic and hydrophilic residues in the amphipathic helix are crucial for transcriptional activation in mammalian and in yeast cells. (A) Amino acid substitutions were made in the E2A minimal LH domain (residues 259 to 426) as shown and described in Fig. 1B and 2. Independent duplicates are presented. Lanes: 5 and 6, the alanine residue at position 334 was changed to a proline residue; 7 and 8, alanine and isoleucine residues at positions 334 and 335, respectively, were substituted with aspartate and serine residues; 9 and 10, the arginine residue at position 339 was changed to a leucine residue; 11 and 12, valine and leucine residues at positions 337 and 338, respectively, were substituted with glutamate and arginine residues. The mutated expression plasmids were cotransfected with a CAT reporter plasmid into HeLa cells, and CAT activity was assayed (as described in Materials and Methods) with 5 μg of the reporter construct and 2 μg of each activator construct. (B) The average relative values are shown in the histogram. The CAT activity for the E2A wild-type LH domain (residues 259 to 426) was determined in a separate experiment and is not shown here; however, the CAT activities for the full-length E2A protein (1 to 426) and the minimal LH domain (259 to 366) are shown here as a reference. In separate experiments, the E2A wild-type LH domain (259 to 426) showed CAT activity equal to that of the full-length E2A protein (1 to 426) (Fig. 2A). The same assay performed with COS cells yielded comparable transactivation results, and Western blotting performed on the same extracts revealed that equal amounts of proteins were expressed (data not shown). (C) The putative LH motif activates transcription in yeast cells. The indicator yeast strain GGY::171 (8), which contains a *GAL1-lacZ* reporter gene activated by UAS_G (four GAL4 binding sites) integrated at the *URA3* locus, was transformed with the indicated GAL4-E2A fusion constructs, harvested, and assayed for β-galactosidase activity as described in Materials and Methods. The averages for at least four independent transformants are shown. Standard deviations were <20%. β-Galactosidase activity is expressed as units per milligram of total protein.

for 69% of wild-type levels of activity. Obviously, there are likely additional regions within E2A and E2-2 that account for these differences. In this study, we have chosen to focus on the major determinant, the LH motif. The LH domain contains a putative loop adjacent to a potential amphipathic helix. The potential amphipathic helix consists of strikingly conserved hydrophobic and hydrophilic residues (Fig. 3). Most of the LH residues are conserved in E2A and the *Drosophila* gene product daughterless, two proteins that show little homology outside the HLH region. The hydrophobic residues of the putative helix appear to be critical for transactivation. Substitutions of the conserved alanine, isoleucine, valine, and leucine residues obliterate transactivation in both yeast and mammalian cells. Whether the conserved hydrophilic residues play a crucial role is not clear. Substitution of a conserved hydrophilic residue does not affect transactivation. Obviously, more extensive mutagenesis will be required to determine exactly the structure-function relationship of the LH motif. Nevertheless, the data presented here clearly demonstrate the importance of the conserved hydrophobic residues in LH-mediated transactivation.

Inspection of the primary amino acid sequence reveals that the putative loop region also contains a series of

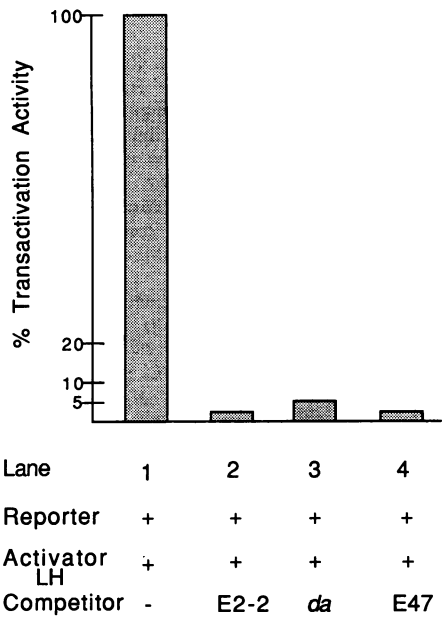


FIG. 6. E47, E2-2 and daughterless (*da*) squelch the transactivation ability of the E2A minimal LH domain. Five micrograms of the E2A minimal LH activator (residues 259 to 366), 10 μg of the CAT reporter plasmid, and 10 μg of each competitor plasmid (simian virus 40 expression plasmids encoding full-length E47, E2-2, and daughterless proteins) were cotransfected into HeLa cells and assayed for CAT activity as described in Materials and Methods. Relative CAT activities are shown in the histogram.

conserved residues. Although our mutations in the loop represented a nonrandom set, loop mutations that we introduced into the E2A loop region had no effect on E2A transactivation (unpublished results). However, the loop is required for function, as deletion of its sequences obliterates transactivation in both E2A and E2-2. It is conceivable that the loop domain folds into a malleable structure that is insensitive to single amino acid substitutions but is required to enhance the affinity of the LH target molecule for interaction with the proposed amphipathic helix.

The homology of the LH motifs is so striking that an obvious suggestion is that they have a common and conserved target. Competition experiments with the various class I HLH proteins indeed indicated the presence of such a conserved target molecule. For example, overexpression of daughterless in mammalian cells inhibits the transactivation capability of E2A. In addition, substitutions of residues within the putative helix that decrease transactivation in mammalian cells also reduce transactivation in yeast cells. Similarly, a substitution within the putative helix that does not affect transactivation in mammalian cells does not reduce transactivation in yeast cells. Thus, mutations in the LH affect function similarly in yeast and mammalian cells. These data are most easily explained by the presence of a universal target molecule.

We attempted to identify other proteins containing the LH motif; however, no homologies of significance were detected. This result implies that the LH domain is restricted to a class of DNA binding proteins. Thus, we have identified a new transactivation domain that is clearly distinct from activation domains such as the acidic blob (23, 25) and proline-rich (18) and glutamine-rich (5) domains. It is possible that different activation domains interact with separate target proteins. Recently, it was shown that the activating region of E1A interacts with a target distinct from that of an acidic activating region (17). Furthermore, a potential transcriptional adaptor was identified which potentiated the activity of particular acidic activation domains (1). It will be interesting to determine whether E2A, E2-2, and daughterless, members of the class I HLH proteins containing the conserved LH motif, are able to inhibit the activity of other classes of transactivators. It seems likely that the LH domain interacts with a target specific for this class of transactivators.

The restricted presence of the LH motif in class I HLH proteins is interesting. The class I HLH proteins E12, E47, E2-2, HEB, and daughterless are known to play important roles in various, seemingly unrelated developmental pathways. For example, E12 and E47 are required for B-cell-, muscle-, and pancreas-specific gene expression. E2-2 is involved in immunoglobulin gene expression. daughterless is involved in at least two known developmental pathways: sex determination and neurogenesis. Thus, class I HLH proteins have pleiotropic roles. They form either homodimers or heterodimers with tissue-specific HLH proteins. For example, E2A forms heterodimers with MyoD and myogenin (14, 19, 20), and daughterless forms heterodimers with achaete-scute (19). In B cells, E2A products likely form homodimers (20). It is conceivable that one function of the class I HLH proteins is to interact with a common target to activate transcription of downstream genes in the various developmental pathways.

Understanding the normal role of the LH motif in activating transcription is of critical importance in deducing its aberrant role in the two B-cell ALLs. It is intriguing that in two independent pediatric ALLs, the E2A N-terminal region containing the LH motif is fused to a heterologous DNA binding domain (13, 22). In pre-B-cell ALL, containing a t(1;19) translocation, the E2A N-terminal portion, including the LH motif, is fused to a homeobox-like domain, designated pbx1 (13, 22). In pro-B-cell ALL, containing a t(17;19) translocation, the same E2A LH motif-containing portion is fused to a DNA binding and dimerization domain containing a basic leucine zipper-like region, designated HLF (11). It is likely that the translocated E2A LH domain, with its inappropriate heterologous DNA binding domain (i.e., pbx1 or HLF), is able to interact with components of the basic transcriptional machinery, thereby activating promoters that it normally does not activate. Why the E2A N-terminal transactivation domain, including the LH motif, is involved in two ALLs of the B-cell lineage is an intriguing but open question. Understanding the role of E2A in these leukemias awaits the isolation and characterization of molecules interacting with the LH domain.

ACKNOWLEDGMENTS

We thank Tom Kadesch for providing plasmids and Russell Doolittle for sequence analysis.

This work was supported by the National Institutes of Health, the Council for Tobacco Research, and the Searle Family Trust.

REFERENCES

1. **Berger, S. L., B. Pina, N. Silverman, G. A. Marcus, J. Agapite, J. L. Regier, S. J. Triezenberg, and L. Guarente.** 1992. Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. Cell **70**:251–265.
2. **Cabrera, C. V., and M. C. Alonso.** 1991. Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless*

gene products of *Drosophila*. EMBO J. **10**:2965–2973.
3. **Caudy, M., H. Vaessin, M. Brand, R. Tuma, L. Y. Jan, and Y. N. Jan.** 1988. *daughterless*, a gene essential for both neurogenesis and sex determination in *Drosophila*, has sequence similarities to *myc* and the *achaete-scute* complex. Cell **55**: 1061–1067.
4. **Cordle, S. R., E. Henderson, H. Masuoka, P. A. Weil, and R. Stein.** 1991. Pancreatic B-cell-type-specific transcription of the insulin gene is mediated by basic helix-loop-helix DNA-binding proteins. Mol. Cell. Biol. **11**:1734–1738.
5. **Courey, A. J., and R. Tjian.** 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell **55**:887–898.
6. **Cronmiller, C., P. Schedl, and T. W. Cline.** 1988. Molecular characterization of *daughterless*, a *Drosophila* sex determination gene with multiple roles in development. Genes Dev. **2**:1666–1676.
7. **Davis, R. L., P.-F. Cheng, A. B. Lassar, and H. Weintraub.** 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. Cell **60**:733–746.
8. **Gill, G., and M. Ptashne.** 1987. Mutants of GAL4 protein altered in an activation function. Cell **51**:121–126.
9. **Henthorn, P., M. Kiledjian, and T. Kadesch.** 1990. Two distinct transcription factors that bind the immunoglobulin enhancer μE5/κE2 motif. Science **247**:467–470.
10. **Hu, J.-S., E. N. Olson, and R. E. Kingston.** 1992. HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol. Cell. Biol. **12**:1031–1042.
11. **Inaba, T., W. M. Roberts, L. H. Shapiro, K. W. Jolly, S. C. Raimondi, S. D. Smith, and A. T. Look.** 1992. Fusion of the leucine zipper gene *HLF* to the E2A gene in human acute B-lineage leukemia. Science **257**:531–534.
12. **Ito, H., Y. Fakuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153**:163–168.
13. **Kamps, M. P., C. Murre, X. Sun, and D. Baltimore.** 1990. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. Cell **60**:547–555.
14. **Lassar, A. B., R. L. Davis, W. E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore, and H. Weintraub.** 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell **66**:305–315.
15. **Lesczynsky, J. F., and G. D. Rose.** 1986. Loops in globular proteins: a novel category of secondary structure. Science **234**:849–855.
16. **Ma, J., and M. Ptashne.** 1987. A new class of yeast transcriptional activators. Cell **51**:113–119.
17. **Martin, K. J., J. W. Lillie, and M. R. Green.** 1990. Evidence for interaction of different eukaryotic transcriptional activators with distinct cellular targets. Nature (London) **346**:147–152.
18. **Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. Cell **58**:741–753.

18a. **Murre, C., P. S. McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and Myc proteins. Cell **56**:777–783.

19. **Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58**:537–544.
20. **Murre, C., A. Voronova, and D. Baltimore.** 1991. B-cell- and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. Mol. Cell. Biol. **11**:1156–1160.
21. **Nelson, C., L. P. Shen, A. Meister, E. Fodor, and W. J. Rutter.** 1990. Pan: a transcriptional regulator that binds chymotrypsin, insulin and AP-4 enhancer motifs. Genes Dev. **4**:1035–1044.
22. **Nourse, J., J. D. Mellentin, N. Galili, J. Wilkinson, E. Stanbridge, S. D. Smith, and M. L. Cleary.** 1990. Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell **60**:535–545.
23. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335**:683–689.
24. **Rose, M., and D. Botstein.** 1983. Construction and use of gene fusions to lacZ (β-galactosidase) that are expressed in yeast. Methods Enzymol. **101**:167–180.

24a. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

25. **Struhl, K.** 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell **49**:295–297.
26. **Sun, X.-H., and D. Baltimore.** 1991. An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimers but not in E12 heterodimers. Cell **64**:459–470.
27. **Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar.** 1991. The *myoD* gene family: nodal point during specification of the muscle cell lineage. Science **251**:761–766.