Interactions among Vertebrate Helix-Loop-Helix Proteins in Yeast Using the Two-hybrid System*

(Received for publication, December 4, 1992)

Jeff Stauning†, Michael Perry‡, Stephen J. Elledge§, and Eric N. Olson¶

From the †Department of Biochemistry and Molecular Biology, University of Texas M. D. Anderson Cancer Center and the §Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

The helix-loop-helix (HLH) motif is contained in a rapidly growing family of transcription factors and has been shown to mediate dimerization among heterologous HLH-containing proteins. E12 is a widely expressed HLH protein that preferentially forms heterodimers with cell type-specific HLH proteins such as MyoD, myogenin, and the achaete-scute gene products. As a first step toward screening for novel cell type-specific partners of E12, we used a modification of the two-hybrid assay for detection of protein-protein interactions in vivo to determine whether dimerization of HLH proteins with E12 can occur in yeast. Using the GAL4 DNA-binding domain fused to the E12 HLH motif and the GAL4 transcription activation domain fused to MyoD, we show that E12 and MyoD can efficiently dimerize in yeast and reconstruct a hybrid transcription factor that activates reporter genes linked to the GAL4 DNA-binding site. The GAL4 DNA-binding domain fused to E12 was used to screen a mouse T-cell cDNA library in which the cDNA was fused to the GAL4 activation domain. Several cDNA clones encoding proteins that interact with E12 were isolated, one of which corresponded to the HLH protein Id-2. Given the ability of E12 to dimerize preferentially with cell type-specific HLH proteins, this strategy should be useful for cloning novel partners for E12 from a variety of cell types.

A number of transcription factors that play important roles in the control of cell type-specific transcription share homology within a helix-loop-helix (HLH) motif (Murre et al., 1989a). The HLH motif is composed of two predicted amphipathic α-helices separated by an unstructured loop and mediates dimerization of heterologous HLH proteins (Murre et al., 1989b). Dimerization brings together the basic regions of HLH proteins, which lie immediately amino-terminal to the HLH motif, to form a composite DNA-binding domain. Members of the HLH family of transcription factors bind to a palindromic DNA consensus sequence CANNRC, known as an E-box, which is present in the control regions of numerous tissue-specific genes (Blackwell and Weintraub, 1990; Murre et al., 1989a).

Cell type-specific HLH proteins such as MyoD (Davis et al., 1987), myogenin (Edmondson and Olson, 1989; Wright et al., 1989), and products of the Drosophila achaete-scute complex dimerize preferentially with widely expressed HLH proteins of the E-protein class, which includes E12, E47, and HEB (Hu et al., 1992; Murre et al., 1989a; Brennan and Olson, 1990). Transcriptional activation by these types of HLH proteins can be prevented by HLH proteins lacking basic regions, which form heterodimers unable to bind the E-box consensus sequence. Among these inhibitory HLH proteins are Id-1 (Benezra et al., 1990) and Id-2 (Sun et al., 1991), and the extramacrochaetae gene product from Drosophila (Ellis et al., 1990; Garrell and Modelell, 1990). This class of HLH proteins has been implicated in repression of cell type-specific transcription through inactivation of E-proteins or their partners, which are required for activation of cell type-specific genes.

The presence of E-proteins in a wide range of cell types suggests the existence of cell type-specific partners for dimerization and transcriptional activation. To date, however, the only examples of tissue-specific HLH proteins in vertebrates are the myogenic HLH proteins from skeletal muscle (reviewed by Olson (1990)) and the mammalian achaete-scute homologs from certain types of neural cells (Johnson et al., 1990). Numerous attempts to clone cell type-specific HLH proteins by low stringency screening of cDNA libraries or polymerase chain reaction amplification have been unsuccessful, suggesting that if such proteins exist, they may have diverged significantly from known HLH proteins. Given the ability of several cell type-specific HLH proteins to dimerize preferentially with E-proteins, an alternate strategy to isolate such proteins would be through a functional assay based on dimerization. Previously, we have shown that interactions between heterologous HLH proteins can be detected in transfected tissue culture cells using a modification of the two-hybrid system for detection of protein-protein interactions in vivo (Chien et al., 1991; Fields and Song, 1989; Chakraborty et al., 1992). Here we demonstrate that vertebrate HLH proteins will dimerize in yeast and such interactions are sufficiently stable to allow screening of cDNA libraries for novel HLH proteins based on HLH-mediated dimerization. Using the GAL4 DNA-binding domain fused to the HLH motif of E12, we show that E12 and MyoD can efficiently dimerize in yeast and reconstitute a hybrid transcription factor that can activate expression of a reporter gene linked to the GAL4 DNA-binding site. We use the DNA-binding domain of GAL4 fused to E12 to screen a mouse T-cell cDNA library for potential dimerization partners for E12. One of the cDNAs, whose product reconstructed a functional transcrip-
tion factor when dimerized with the GAL4-E12 chimera, was found to encode Id-2. These results demonstrate that yeast can support HLH-mediated dimerization, and they suggest that the two-hybrid system may be a powerful strategy for isolation of novel partners for E12 from diverse cell types.

**MATERIALS AND METHODS**

**Yeast Strains and Expression Vectors**—Construction of the GAL(AD)-MyoD fusion vector was accomplished by first cloning the 1.35-kilobase pair SstI fragment from pBS/xlmf25 (Scales et al., 1990) into the bacterial expression vector pRX1 (Riezman and Pollard, 1989). The resulting plasmid (pRX/xlmf25) contained sequences encoding amino acids 2-286 of XMyoDa fused to the first 21 amino acids of TrpE and placed an EcoRI site immediately upstream of the SstI site, which can be complemented by the yeast LEU2 gene present on the GAL4(AD) plasmid.

**DNA Sequencing**—DNA sequencing was performed using double-stranded plasmid DNA with the Sequenase kit (United States Biochemical Corp.). Sequences were analyzed using the Beckman Microgene sequence analysis package.

**RESULTS**

**Detection of MyoD-E12 Interactions in Yeast**—To determine whether mammalian HLH proteins can heterodimerize in yeast, we created expression plasmids in which the bHLH region of human E12 was fused to the GAL4 DNA-binding domain (GAL4(DB)) and the complete Xenopus MyoD was fused to the GAL4 transcription activation domain (GAL4(AD)-MyoD). Potential interaction between the two hybrid proteins was then tested by expressing a yeast strain harboring lacZ under control of the GAL4 UAS; only transformants in which HIS3 is activated above a certain level will grow in the presence of 3-amino-1,2,4-triazole, an inhibitor of the HIS3 gene product. We added 25 mM 3-amino-1,2,4-triazole to SD-Trp-Leu-His plates since that amount gives approximately a 100-fold enrichment and reduces the total number of colonies that need to be screened for lacZ expression.

The structures of the GAL(DB)-E12 and GAL(AD)-MyoD expression plasmids are shown in Fig. 1A. GAL(DB)-E12 contains amino acids 508-654 of human E12 (Kamps et al., 1990) fused to the carboxy-terminal end of the GAL4 DNA-binding domain, encoded by residues 1-147 of GAL4. This portion of E12 contains the bHLH motif, which lies between residues 550-663, as well as the region immediately amino-terminal to the basic region, which has been shown to inhibit E12 homodimerization (Sun and Baltimore, 1991). Since MyoD is a substrate for PKC (Li et al., 1992), the failure of extracts of yeast transformed with the GAL(DB)-E12 expression vector. Antibody directed against GAL1-147 detected a protein of 35 kDa in GAL(DB)-E12 expressing cells, but not in untransformed yeast (Fig. 1B). This agrees with the predicted size of the GAL4-E12 fusion protein.

GAL(AD)-MyoD contains all of MyoD fused to the carboxyl-terminal end of the GAL4 activation domain, encoded by residues 768-881 of GAL4. When yeast were transformed with either GAL(DB)-E12 or GAL(AD)-MyoD alone, there was no activation of lacZ above background (Table I). However, when the two hybrids were coexpressed, we observed high levels of lacZ expression. Typically, yeast colonies expressing the two hybrids stained intensely blue within 2-5 min following exposure to the chromogenic substrate for β-galactosidase. Activation of transcription by the combined hybrids represented a specific protein-protein interaction because no activation was observed when GAL(DB) fused to PKC-α was coexpressed with GAL(AD)-MyoD (Table I). These results demonstrate that specific protein-protein interactions between E12 and MyoD are readily detectable in yeast. Since MyoD is a substrate for PKC (Li et al., 1992), the failure to detect activation with GAL(DB)-PKC plus GAL(AD)-MyoD suggests that this enzyme-substrate complex is relatively unstable in vivo.

**Screening of a T-cell cDNA Library with GAL4-E12 “Bait”**—We next sought to determine whether GAL(DB)-E12 could be used as “bait” to screen for mammalian cDNAAs in a yeast cDNA expression library. The yeast strain used for these

---

2 T. Durfee, K. Becherer, P. Chen, S. Yeh, Y. Yang, A. Kilburn, W. Lee, and S. Elledge, submitted for publication.

3 J. Staudinger, S. J. Elledge, and E. N. Olson, manuscript in preparation.
Fig. 1. Structures of GAL(DB)-E12 and GAL(AD)-MyoD expression plasmids. A: left, GAL(DB)-E12 contains the basic HLH region of E12 fused to the DNA-binding domain of GAL4 and it carries the TRP1 selectable marker; right, GAL(AD)-MyoD contains the complete coding region of XMyoDa fused to the GAL4 transcriptional activation domain and carries the LEU2 selectable marker. The expression of GAL4 fusion proteins is controlled by the ADH promoter and the ADH terminator of the yeast ADH gene.

Table I
Activation of lacZ Reporter Gene by GAL4 Chimeras

| Expression plasmid     | β-Galactosidase activity |
|------------------------|-------------------------|
| None                   | <0.01                   |
| GAL(DB)-PKC            | <0.01                   |
| GAL(DB)-E12            | <0.01                   |
| GAL(AD)-Id2           | <0.01                   |
| GAL(AD)-MyoD          | <0.01                   |
| GAL(DB)-PKC + GAL(AD)-Id2 | <0.01           |
| GAL(DB)-PKC + GAL(AD)-MyoD | <0.01       |
| GAL(DB)-E12 + GAL(AD)-Id2 | 41              |
| GAL(DB)-E12 + GAL(AD)-MyoD | 100            |

Experiments also contains the HIS3 gene linked to GAL4 UAS. For this screen, we used a mouse T-cell cDNA library in which cDNA was fused to GAL4(AD). Screening of 500,000 transformants by cotransformation with GAL(DB)-E12 for cDNAs that could rescue activation of the lacZ and HIS markers linked to the GAL4 UAS yielded seven colonies. When stained for lacZ expression, two colonies showed strong expression within 5 min and the additional five colonies showed expression after incubation overnight.

We chose to focus our attention on the two clones (designated E12-bp1 and E12-bp2) that strongly activated lacZ expression. To determine whether activation of lacZ was dependent on an interaction between the polypeptides encoded by these clones and the E12 portion of GAL(DB)-E12, we coexpressed these plasmids with the GAL4(DB)-PKC construct. Clone E12-bp2 showed activity in this transformation, whereas clone E12-bp1 did not. These results suggest that clone E12-bp2 encodes a protein that interacts with GAL4(DB) or that it activates transcription of the reporter gene more directly, perhaps by binding to DNA sequences in the upstream region. Because this clone did not fulfill the criteria for specificity, we did not characterize it in greater detail.

E12-bp1 Encodes Id-2—To further investigate the nature of the interaction between the protein encoded by clone E12-bp1 and GAL4(DB)-E12 we determined the nucleotide sequence of the cDNA in the region of the junction with GAL4(AD) (Fig. 2). Comparison of the partial nucleotide sequence of clone E12-bp1 with the GenBank data base showed that it corresponded to mouse Id-2 (Sun et al., 1991) and contained the sequence beginning in exon 1 and extending through the stop codon to nucleotide 887 in the 3'-untranslated region fused in-frame with the GAL4(AD). This region of Id-2 contains all of the HLH motif. We conclude that E12-bp1 encodes Id-2 since it was expressed in yeast and that the stability of the interactions is sufficient to support screening of cDNA libraries.

In the previously published sequence of mouse Id-2, there were two in-frame AUG codons within exon 1, the more 5' of which was predicted to represent the site for translation initiation (Sun et al., 1991). Clone E12-bp1 began 43 nucleotides downstream of a C that was reported in the Id-2 sequence described previously (Sun et al., 1991). Clone E12-bp1 encodes Id-2 by the original cDNA clone E12-bp1 (Sun et al., 1991) and contains all of the sequence corresponding to the region fused in-frame with GAL4(AD).

Fig. 2. Derived amino acid sequence of the junction between GAL(AD) and Id-2. The amino acid sequence of the cDNA insert in clone E12-bp1 was determined by DNA sequencing using a primer upstream of the junction of GAL(AD) and the cDNA insert. Nucleotide sequence and deduced open reading frame for a portion of the sequence are shown, and the regions corresponding to GAL(AD) and Id-2 are indicated. The predicted initiating methionine for translation of Id-2 is indicated with an asterisk, and the position of a C that was reported in the Id-2 sequence described previously (32) is shown with an arrowhead below the sequence.
tides upstream of the more 3' AUG (110 nucleotides down-
stream of the more 5' AUG). In sequencing the 5' region of
the Id-2 cDNA insert near the junction with GAL4, we de-
tected a discrepancy with the published cDNA sequence in
which a single dC present in the previously published sequence
was missing from our cDNA (see Fig. 2). We are certain that
this dC cannot be present in our cDNA because it would shift
the reading frame, resulting in translation termination of a
protein lacking an HLH motif. We suggest therefore that the
more 3' AUG of Id-2 may represent the true translational
initiation site. Clone E12-bp1 would therefore begin within
the 5'-untranslated region of the Id-2 mRNA.

**DISCUSSION**

There is considerable evidence for the involvement of cell
type-specific bHLH proteins in the control of tissue-specific
gene expression in vertebrates. (i) E12 and several related E-
proteins that are widely expressed dimerize preferentially with
cell type-specific HLH proteins such as MyoD and myogenin
(Chakraborthy et al., 1991; Murre et al., 1989a, 1989b). (ii) Id,
a negative regulator of HLH protein DNA binding, is ex-
pressed in many undifferentiated cell types and is down-
regulated upon activation of differentiation, which has been
proposed to release E-proteins to dimerize with cell type-
specific partners (Benezra et al., 1990). (iii) Many tissue-
specific genes contain E-boxes in their control regions that
are essential for cell type-specific transcription (Murre et al.,
1989a; Nelson et al., 1990; Olson, 1990; Sartorelli et al., 1992).
(iv) The involvement of numerous cell type-specific bHLH
proteins in cell fate specification in Drosophila suggests that
these proteins constitute a conserved mechanism for regula-
tion of cell type-specific transcription (Alonso and Cabrera,
1988; Caudy et al., 1988; Thiese et al., 1988).

Although HLH proteins contain several conserved amino
acids within the two amphipathic α-helices of the HLH motif,
the degree of nucleotide sequence homology between the genes
encoding different subclasses of HLH proteins is in most
cases not sufficient to allow cross-hybridization. Therefore,
to facilitate identification of novel HLH proteins independent
of sequence homology, we have used a modification of the
two-hybrid assay, which is based only on the ability of proteins
to dimerize with E12. There are several advantages to the
two-hybrid system as a method for isolating novel cell type-
specific partners for E12. It is a functional assay, based only
on the ability of proteins to dimerize with E12, not on nucleo-
tide sequence homology. The assay also does not depend on
DNA binding activity of heterodimers formed between E12 and
its putative partners. The assay is also highly sensitive
and specific, and in our hands has not yielded significant
background. The relative strength of protein-protein inter-
actions can also be assessed by quantitative determination of
lacZ activity in yeast extracts.

Using the E12 HLH motif linked to the GAL4 DNA-
binding domain as "bait," we showed that E12 and MyoD can
dimerize in yeast and we identified Id-2 from a mouse T-cell
cDNA library. The ability of E12 to dimerize with MyoD and
Id in yeast indicates that HLH-mediated dimerization does
not require post-translational modifications or enzymatic ac-
tivities unique to higher eukaryotic cells. The use of GAL(DB)-E12 to screen yeast expression libraries for dimer-
ization partners should therefore be a useful strategy for
identification of novel HLH proteins from a variety of cell
types.

In addition to serving as an interface for dimerization and
DNA binding, there is evidence that the bHLH motif interacts
with coregulators to induce muscle-specific transcription
(Brennan et al., 1991, Davis et al., 1990). The specificity of
interaction of these putative coregulators has been mapped to
the amino acids in the center of the basic domains of
myogenin and MyoD. The two-hybrid system may also be
useful for cloning such proteins.

**Acknowledgments**—We are grateful to M. Pisigna for technical
assistance and K. Tucker for preparation of the manuscript.

**REFERENCES**

Alonso, M. C., and Cabrera, C. V. (1988) EMBO J. 7, 2583-2591.
Brennan, T. J., Davis R. L., Lockshon, D., Turner, D. L., and Weintraub, H.
(1990) Cell 61, 49-59
Blackwell, K. T., and Weintraub, H. (1990) Science 250, 1104-1110
Brennan, T. J., and Olson, E. N. (1990) Genes & Dev. 4, 582-595
Brennan, T. J., Chakraborthy, K., and Olson, E. N. (1991) Proc. Natl. Acad. Sci.
U. S. A. 88, 5675-5679
Caudy, M., Vassen, H., Brand, M., Tuma, R., Jan, L. Y., and Jan, Y. N. (1988)
Cell 55, 1061-1087
Chakraborthy, T., Brennan, T. J., Li, L., Edmondson, D., and Olson, E. N.
(1991) Mol. Cell. Biol. 11, 3633-3641
Chakraborthy, T., Martin, J. F., and Olson, E. N. (1992) J. Biol. Chem. 267,
17498-17501
Chien, C.-T., Bartel, P. L., Sternglanz, R., and Fields, S. (1991) Nature 340,
245-246
Davis, R. L., Weintraub, H., and Lassar, A. B. (1987) Cell 51, 987-1000
Davis, R. L., Cheng, P.-F., Lassar, A. B., and Weintraub, H. (1990) Cell 60,
733-746
Edmondson, D. G., and Olson, E. N. (1989) Genes & Dev. 3, 628-640
Ellis, E. D., Mulligan, J. T., Ramer, S. W., Stoutwook, M., and Davis, R. W.
(1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1751-1755
Ellis, H. M., Spann, D. R., and Posakony, J. W. (1990) Cell 61, 27-38
Fields, S., and Song, O. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9758-9762
Garrell, J., and Modolell, J. (1990) Cell 61, 99-45
Hoffman, C. S., and Winston, F. (1987) Gene (Amst.) 57, 267-272
Hu, J.-S., Olson, E. N., and Kingston, R. E. (1992) Mol. Cell. Biol. 12, 1031-1042
James, G., and Olson, E. N. (1992) J. Cell Biol. 116, 863-874
Johnson, J. E., Birren, S. J., and Anderson, D. J. (1989) Nature 340, 858-861
Kamps, M. P., Murre, C., Sun, X.-H., and Baltimore, D. (1990) Cell 60, 547-555
Li, L., Zhou, Z., James, G., Heller-Harrison, R., Czech, M. P., and Olson, E. N.
(1992) Cell 71, 1151-1154
Murre, C., McCaw, P. S., and Baltimore, D. (1986a) Cell 66, 777-783
Murre, C., McCaw, P. S., Vassen, H., Caudy, M., Jan, L. Y., Jan, Y. N.,
Cabrera, C. V., Buskin, J. N., Haushka, S. D., Lassar, A. B., Weintraub, H.
and Baltimore, D. (1989b) Cell 68, 537-544
Nelson, C., Shen, L.-P., Meister, A., Fodor, E., and Rutter, W. J. (1990) Genes &
Dev. 4, 1038-1043
Olsen, E. N. (1990) Genes & Dev. 4, 1454-1461
Rimm, D. L., and Pollard, T. D. (1989) Gene (Amst.) 75, 323-337
Rose, M. D., Winston, F., and Heizer, P. (1990) Methods in Yeast Genetics: A
Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor,
NY
Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A
Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor,
NY
Sartorelli, V., Hong, N. A., Bishopic, N. H., and Kedes, L. (1992) Proc. Natl.
Acad. Sci. U. S. A. 89, 4047-4051
Scales, J. B., Olson, E. N., and Perry, M. (1990) Mol. Cell. Biol. 10, 156-1524
Sun, X.-H., and Baltimore, D. (1991) Cell 64, 459-470
Sun, X.-H., Copeland, N. G., Jenkins, N. A., and Baltimore, D. (1991) Mol.
Cell. Biol. 11, 5063-5071
Thiese, B., Stroetzl, C., Gorostiza-Thiese, C., and Perin-Schmitt, F. (1988)
EMBO J. 7, 2175-2183
Wright, W. E., Sassoon, D. A., and Lin, V. K. (1989) Cell 56, 607-617