The Loop Region of the Helix-Loop-Helix Protein Id1 Is Critical for Its Dominant Negative Activity

SUSAN PESCE AND ROBERT BENEZRA*

*Cell Biology Program, Memorial Sloan-Kettering Cancer Center,
1275 York Avenue, New York, New York 10021*

Received 14 July 1993/Returned for modification 26 August 1993/Accepted 2 September 1993

Id1, a helix-loop-helix (HLH) protein which lacks a DNA binding domain, has been shown to negatively regulate other members of the HLH family by direct protein-protein interactions, both in vitro and in vivo. In this study, we report the results of site-directed mutagenesis experiments aimed at defining the regions of Id1 which are important for its activity. We have found that the HLH domain of Id1 is necessary and nearly sufficient for its activity. In addition, we show that two amino acid residues at the amino terminus of the Id1 loop are critical for its activity, perhaps by specifying the correct dimerization partners. In this regard, replacing the first four amino acids of the loops of the basic HLH proteins E12 and E47 with the corresponding amino acids of Id1 confers Id1 dimerization specificity. These studies point to the loop region as an important structural and functional element of the Id subfamily of HLH proteins.

Three classes of transcription factors which utilize a common motif have now been identified. The helix-loop-helix (HLH), leucine zipper, and POU domain proteins all contain a cluster of basic amino acids which contact DNA adjacent to a protein-protein dimerization domain (9; for reviews, see references 23 and 25). The DNA binding affinities and sequence preferences of the leucine zipper and HLH proteins have been shown to be altered by the choice of dimerization partners (2, 8), thus expanding the potential functional repertoire of these proteins. In addition, all three classes of proteins have pairing partners which lack or have altered DNA binding domains, and heterodimeric complexes with these proteins are unable to bind DNA (1a, 17, 22).

The Id proteins are a class of HLH proteins which lack a DNA binding domain. Three members of the family have been identified thus far in the mouse (Id1, Id2, and Id3 [also called HLH462.3]) (1a, 3, 21). All three proteins have highly homologous HLH domains (81% identity in all three) but diverge nearly completely outside of this region. In addition, all three proteins have been shown to be capable of inhibiting basic HLH (bHLH) proteins from binding DNA in vitro. Furthermore, mRNA levels from all three genes decrease during differentiation in many different cell types in tissue culture and/or in the embryo, leading to the suggestion that the Id family may be controlling cell fate decisions in multiple lineages. Indeed, overexpression of Id1 has been shown to inhibit tissue-specific gene expression in cultured skeletal muscle cells, cardiac muscle cells, myeloid cells, B cells, T cells, pancreatic β cells, and osteoblasts (1a, 4, 10, 11, 14, 16, 18, 19, 26).

Although it is likely that the ubiquitously expressed E2A proteins are targets for Id1 in muscle cells (10), a detailed picture of the rules governing Id-bHLH interactions is not yet available. It is not known, for example, whether members of the Id family have different binding preferences for members of the E2A family and whether all members of the Id family bind poorly to all tissue-specific bHLH proteins (such as has been shown for Id1-MyoD and Id2-MyoD) (1a,

21). In addition, it is not known whether all bHLH proteins with leucine zippers are resistant to Id inhibition (as has been shown for some b-zip proteins) (21). And finally, is it not yet clear whether it is only helix-helix interactions which specify dimerization partners or whether other sequence elements play a role.

Mutagenesis experiments on MyoD and E2A proteins have established that the HLH domain is required for dimerization and that the basic region mediates DNA binding (5, 24). Disruption or deletion of the basic region prevents DNA binding but has no effect on the dimerization potential of these mutants. Disruption of either the first or second helix with proline residues, on the other hand, abolishes the ability of these molecules to dimerize. It has also been demonstrated that major changes in the loop sequence of MyoD (including replacement with the loops of c-Myc or achaete-scute) have no apparent effect on the dimerization or transactivation potential of MyoD, suggesting that the loop may be a relatively neutral structural element (5). However, the fact that subgroups within the HLH family (the muscle regulatory genes, the *myc* family, the Id genes, etc.) have loops that are more highly related to one another than to loops of members in the rest of the family (see Fig. 2 in reference 1a) suggests that this may be an oversimplified view.

In the present study, we have made site-directed amino acid changes in an inhibitory HLH protein, Id1, and tested these mutant molecules for their ability to antagonize bHLH protein activity in vitro and in a transient transfection assay. Our results indicate that the HLH domain of Id1 is necessary and nearly sufficient for its activity and that there is a strong correlation between the ability of a given Id1 mutant to inhibit MyoD's transactivation potential and its ability to associate with E2A proteins in vitro. In addition, we have determined that a proline residue and a leucine residue at the amino terminus of the loop region of Id1 are critical for its ability to pair with E proteins. Replacement of the proline residue with other helix-breaking residues fails to restore full Id1 activity. Conversely, replacing the four amino-terminal residues of the loop region of E12 with the corresponding amino acids of Id1 prevents homodimerization but still allows heterodimerization with MyoD, a property character-

* Corresponding author. Electronic mail address: rbenezra@bigmac.mskcc.org.

istic of the Id family (21). These studies therefore point to a previously unrecognized structural parameter that may govern the functioning of HLH proteins.

MATERIALS AND METHODS

Cell lines and transfections. C3H10T1/2 cells (from the American Type Culture Collection) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% bovine fetal calf serum (HyClone). Transfections were performed as previously described (1a), with the exception that 5 μ g of the muscle creatine kinase-chloramphenicol acetyltransferase (MCK-CAT) reporter, 5 μ g of the MyoD expression vector (emc11S), and 20 μ g of the Id test expression vectors were used unless otherwise indicated. In all transfections, the total plasmid concentration was kept constant by adding the appropriate amounts of EMSV expression vector without an insert.

Mutant construction and in vitro assays of Id activity. All Id mutants indicated in the text were constructed by using an Amersham in vitro mutagenesis kit in a protocol recommended by the manufacturer. All mutant constructions were confirmed by the Sanger dideoxy sequencing method. Proteins were synthesized in vitro by using a rabbit reticulocyte lysate system (Promega), and coimmunoprecipitations were performed as previously described (1a). The molar concentrations of all proteins were determined by using a Molecular Dynamics phosphorimage analysis system, correcting for relative cysteine content. Electrophoretic mobility shift assays used to test the activity of Id mutants were performed as previously described (1a) at a molar ratio of E47 to Id of 1:2. In heterodimeric shifts between MyoD and E2A proteins, (see Fig. 4B), the ratio of MyoD to E2A protein was ~4:1. All DNA binding reactions were performed at room temperature for 10 min.

Glutaraldehyde cross-linking. E12 and the 1/2 Id loop→E12 loop mutant were synthesized in rabbit reticulocyte lysates in the presence of [³⁵S]methionine (800 Ci/mmol; Amersham). Glutaraldehyde at twice the desired concentration was prepared just prior to use in 0.2 M KCl-0.1% Nonidet P-40 from a 25% stock solution (ICN Biochemicals). One volume of the glutaraldehyde solution was added, and the mixture was allowed to sit at room temperature for 3 min. One volume of 2× sodium dodecyl sulfate (SDS) sample buffer (13) was added, and the samples were heated to 95°C for 1 min before being loaded onto a discontinuous SDS-10% polyacrylamide gel.

RESULTS

Assays of Id1 activity. Previous experiments have suggested that Id1 may antagonize the activity of MyoD in vivo by binding to and inhibiting the activity of the E proteins, bHLH proteins which form stable heterodimers with MyoD in active transcription complexes (12, 15). We employed three independent assays in this study to determine the effects of altering the Id1 protein structure: coimmunoprecipitation with E proteins in vitro, inhibition of E protein DNA binding, and inhibition of MyoD transactivation of the MCK enhancer in transient transfection experiments. In the latter two assays, titrations with wild-type Id1 were performed to establish the minimum dose required for maximal Id1 inhibition, and all mutations were tested at this dosage in order to detect any loss of activity. For inhibition of E protein binding to DNA, maximal inhibition was observed at a 1:2 molar ratio of E47 to Id1 and a 4:1 molar ratio of E12

FIG. 1. Inhibition of E12 and E47 DNA binding activity by Id. E12 and E47 were mixed with Id1 in the molar ratios indicated, heated to 37°C for 30 min, and then incubated at room temperature for 10 min with a ³²P-labelled E-box-containing oligonucleotide derived from the MCK enhancer. Samples were then electrophoresed in a 6% nondenaturing polyacrylamide gel. Only complexes migrating behind the free probe are displayed. All proteins were synthesized in rabbit reticulocyte lysates and processed as previously described (1a).

to Id1 (Fig. 1). This difference in the sensitivities of E12 and E47 to Id1 inhibition was also observed when labelled proteins were used in the mobility shift assays in order to more carefully assess the molar protein input ratios (data not shown). In the transient transfection assay, maximal inhibition was obtained with a 1:3 molar ratio of MyoD to Id1 expression vector (data not shown).

The HLH domain of Id1 is necessary and nearly sufficient for activity. The first Id1 mutations generated were large-scale deletions aimed at defining the important domains of the protein. The HLH domain of Id1 (41 amino acids) is embedded between a 58-amino-acid N-terminal domain and a 49-amino-acid carboxy-terminal domain (see Fig. 2D). Deletion of residues 3 to 58 (referred to as ΔNH2) of Id1 had only a small effect on its ability to inhibit MyoD in the transient assay (Fig. 2A), its ability to coimmunoprecipitate with E12 (Fig. 2C; compare lanes 2 and 3 with 6 and 7), and its ability to inhibit E47 from binding DNA (Table 1). (All mutants constructed are summarized in Fig. 2D; note that amino acid residues 1 and 2 were kept constant in the deletion mutants in an attempt to prevent changes in protein stability defined by the N-end rule.) Deletion of the 47-amino-acid carboxy-terminal domain of Id1 (ΔCOOH) also had little effect on its activity in these assays (Fig. 2A and Table 1).

In contrast, deletion of the HLH domain of Id1 between residues 59 and 100 (ΔHLH) completely abolished its activity in the transient assay (Fig. 2A) and its ability to inhibit E47 from binding DNA (Table 1). In addition, disruption of amphipathic helix 2 by a valine-to-proline mutation at position 91 (PAH2) also abolished Id1 activity in all three assays (Fig. 2B; Fig. 2C, lanes 4 and 5; and Table 1). A similar loss in Id1 activity was observed when helix 1 was disrupted with a serine-to-proline mutation at amino acid residue 67 (PAH1 [Table 1]). These data suggest that the HLH domain of Id1 is the only region of the molecule that is absolutely required for its activity, as defined by the three assays employed in this study.

In order to determine whether the HLH domain of Id1 is sufficient to antagonize the activity of MyoD, we first overexpressed the Id1 HLH domain (plus amino acids 1 and 2) (abbreviated HLH-Id) in the transient assay. A typical result is shown in Fig. 2B. HLH-Id inhibited the expression of the MCK-CAT reporter gene with ~67% of the efficiency of wild-type Id1. Disruption of either helix of HLH-Id with a

FIG. 2. The HLH domain of Id1 is necessary and nearly sufficient for its activity. (A) C3H10T1/2 cells were transfected with a MyoD expression vector and an MCK-CAT reporter gene along with the indicated EMSV expression vectors as previously described (1a). EMSV is the vector without an insert, E:Id contains wild-type Id1, Δ NH₂ is a deletion mutant missing the entire N terminus from residue 3 up to the Id1 HLH domain, Δ HLH is missing the entire HLH domain, and Δ COOH is missing Id1 sequences carboxy terminal to the HLH domain. (B) Transfections were carried out as in for panel A. PAH2 is an Id1 mutant with a proline substitution in helix 2, HLH is the Id1 HLH domain alone, and HLH-PAH1 and HLH-PAH2 are HLH with proline substitutions in the first and second helices, respectively. EMSV and E:Id are as defined for panel A. (C) Wild-type Id1 (wt) and the Id1 mutants indicated were immunoprecipitated with an anti-E12 antiserum with and without added E12 as previously described (1a). (D) Schematic summary of mutants constructed (see text).

proline substitution completely abolishes this inhibition (Fig. 2B, HLH-PAH1 and HLH-PAH2). No inhibition of E47 binding to DNA by HLH-Id, however, was observed in vitro (Table 1). Interestingly, however, HLH-Id can inhibit MyoD homodimers from binding DNA at about 60% of the efficiency of wild-type Id1 (data not shown), suggesting that MyoD may be a direct target for HLH-Id in the cell (see Discussion).

The proline and leucine residues at the amino terminus of the Id1 loop are critical for its dominant negative activity. We next sought to determine whether the structure of the Id1 loop played any role in its activity. Previous studies of MyoD indicated that its loop region could tolerate drastic changes with little effect on its activity (5) despite the high degree of sequence conservation between species in this domain. The loop region of Id1, however, appears not to be as flexible. As shown in Fig. 3A, replacing the Id1 loop with that of MyoD (MD loop→Id) completely abolishes Id1 activity in the transient assay. Similarly, MD loop→Id does not inhibit E47 from binding DNA in vitro (Fig. 3B) or associate with E47 in the coimmunoprecipitation assay (Table 1).

In order to determine whether single amino acid substitutions could account for this loss in Id1 activity, we systematically converted each residue of the Id1 loop to that of MyoD (Fig. 2D) and assayed these mutants in the DNA binding assay. As shown in Fig. 3C, the effects ranged from

virtually no loss of activity (R80-Q and S83-P; compare lanes 2, 3, 7, and 8) to intermediate loss of activity (T75-S, Δ Q78, V82-L, and K81-R; compare lanes 2, 3, 5, 6, 11, and 13) to complete loss of activity (P74-S and L76-N; compare lanes 2 through 4, and 12). Subsequently, three of these mutations were also tested in the coimmunoprecipitation assay, and as expected, the ability to associate with E47 correlated well with the ability of a given mutation to inhibit DNA binding, i.e., Δ Q78 and P77-A associate with E47 as well as wild-type Id1 and P74-S does not associate at all (Table 1). There is also good correspondence between the ability of these mutants to inhibit E proteins from binding DNA and their ability to inhibit MyoD-mediated transactivation of the MCK enhancer (Table 1). These data point to two residues in the Id1 loop, the proline at the amino terminus of the loop and the leucine at the third position, as being critical for Id1 activity.

Although deletion of a glutamine residue at position 78 of the Id1 loop resulted in only a small loss of activity, we have found that not all single amino acid deletions of the Id1 loop are as well tolerated. As shown in Table 1, deletion of S83 (a residue which can be changed to proline or glutamic acid without significant loss of activity) nearly completely inactivates the molecule in the transient and DNA binding assays.

The proline residue at position 74 may be essential for Id1 activity by breaking the α -helix at the helix 1-loop boundary. Replacement of the proline with other helix-breaking resi-

TABLE 1. Characteristics of Id proteins

Id1 protein	Inhibition ^a of:		Association with E12 ^b
	E47 DNA binding ^c	MyoD transactivation ^d	
Wild type	100	100	+++
ΔNH2	79	99 (6)	+++
ΔCOOH	89	100 (4)	+++
HLH-Id	0	67 (11)	
HLH-PAH2	ND	0 (4)	
PAH1	10	19 (6)	+
PAH2	0	2 (7)	—
ΔHLH	0	5 (7)	
MD loop→Id	0	10 (8)	—
P74-S	22	29 (13)	—
T75-S	75	97 (4)	
ΔQ78	90	98 (4)	+++
R80-Q	109	96 (4)	
S83-P	85	99 (6)	
V82-L	65	100 (7)	
L76-N	26	33 (7)	
K81-R	82	101 (4)	
L76-N	26	33 (7)	
K81-R	82	101 (4)	
S83-D	96	32 (8)	
P74-N	38	19 (8)	
P77-A	78	102 (4)	+++
S83-A	105	99 (4)	
S83-E	135	89 (4)	
P74-G	23	ND	
E47 loop→Id	8	ND	
ΔS	0	11 (4)	

^a Given as a percentage of wild-type value. ND, not determined.

^b The association of Id1 and Id1 mutants with E12 was assessed in the coimmunoprecipitation assay described in Materials and Methods. The background was considered the level of Id1 or mutant Id1 immunoprecipitated by the anti-E12 antiserum without added E12. Data show the approximate fold increase in the amount of Id1 immunoprecipitated with added E12: +++, 3.5- to 7-fold stimulation; +, 2- to 3.5-fold stimulation; —, no stimulation. Data are shown for immunoprecipitations performed.

^c Values were determined by phosphoimage analysis of the signal intensity of the shifted E47-DNA complex. The background was considered the signal generated from unprogrammed reticulocyte lysate in the region of the complex and was subtracted from all values. The signals from E47 alone and E47 plus Id1 were considered 100 and 0%, respectively, and all other values were generated by linear extrapolation. Typically, the E47-plus-Id1 signal was 5 to 10% of that of E47 alone.

^d Values were generated by determining the relative levels of CAT activity in the transient transfection assay. The ratio of acetylated to unacetylated [¹⁴C]chloramphenicol was determined by counting spots cut from thin-layer chromatographs. Typically, expression of MyoD resulted in 10 to 50% conversion of chloramphenicol to the acetylated form. The background was considered the percent conversion when the MyoD plasmid was omitted from the transfection (typically ~0.5% of the value when the MyoD expression plasmid was included). Id activities of 100 and 0% were taken as the percent conversion observed with the MyoD expression vector with or without the addition of the Id expression vector (E:Id), respectively. All other values were derived by linear extrapolation. The numbers in parentheses are the numbers of assays performed to generate each value.

dues such as asparagine or glycine, however, also failed to produce Id1 molecules with significant activity (Fig. 3C; compare lanes 2, 3, 15, and 21). Several other substitutions were tried (P74-M [data not shown] and P74-Q in the context of the entire E47 loop→Id [Fig. 3C, lane 22]), but these were also inactive. The transient transfection data for these mutants was in good agreement with the results of the DNA binding assay (Table 1). It seems possible, therefore, given the geometry of an α-imino acid, that P74 is specifying a tertiary structural constraint on the HLH domain of Id1.

The Id1 loop influences dimerization specificity. Since Id1 is

unique in the HLH family in that it does not homodimerize (21) but does heterodimerize weakly with MyoD (1a, 21), we asked whether the amino terminus of the Id1 loop could confer this specificity on other HLH proteins, namely, the E proteins. E12 dimer formation was assayed by glutaraldehyde cross-linking. As shown in Fig. 4A, when E12 is incubated with glutaraldehyde (at 0.025% or 0.05%; lanes 2 and 3, respectively), a new species with precisely twice the apparent molecular mass of E12 (57-kDa monomer to 114-kDa dimer) appears. That this band is in fact an E12 homodimer is supported by the observations that its formation is inhibited by heating to 37°C (which inhibits E12 binding to DNA [1]), is further inhibited by the addition of Id1, and is concentration dependent (data not shown). (Note that the 69-kDa band which also appears in this assay has not been defined but does not display any of these properties [data not shown].) When, however, the first four amino acids of the E12 loop are changed to the corresponding amino acids of Id1 (1/2 Id loop→E12), no dimer species is observed in this assay (Fig. 4A; compare lanes 2 and 3 with 5 and 6).

In order to determine if 1/2 Id loop→E12 could still heterodimerize with MyoD, we employed a gel mobility shift assay. As shown in Fig. 4B, 1/2 Id loop→E12 does not bind DNA by itself nor does MyoD (at this exposure length) (Fig. 4B, lanes 3 and 4). However, when 1/2 Id loop→E12 is mixed with MyoD, a protein-DNA complex which is presumably a complex of both molecules is now observed (Fig. 4B, lane 5). A similar result is observed with 1/2 Id loop→E47 and MyoD (Fig. 4B, lanes 7 to 12). The data shown in Fig. 4B also indicate that the 1/2 Id loop→E-MyoD complexes bind DNA less avidly than do the E-MyoD complexes (compare lanes 5 and 6 and lanes 10 and 12). This could reflect the fact that Id1 binds MyoD poorly compared with the E proteins (1a, 21). These experiments demonstrate, therefore, that the first four amino acids of the Id1 loop can confer upon a heterologous HLH protein a dimerization specificity similar to that of Id1, i.e., the ability to form weak heterodimers with MyoD but not homodimers.

DISCUSSION

In this study, we have identified regions of Id1 which are important for its dominant negative regulation of the bHLH protein family. Previous studies have indicated that in muscle cells, Id1 (which lacks a DNA binding domain adjacent to its HLH domain) acts by preferentially forming heterodimeric complexes with the ubiquitously expressed E proteins, thereby preventing them from binding DNA alone or in association with MyoD (1a, 10, 21). Since dimerization between members of the bHLH family has been shown to be mediated by the HLH domain (5, 24), we suspected that the HLH domain of Id1 would be similarly important for its activity.

We have determined that, as expected, the presence of the HLH domain of Id1 is absolutely required for its activity. Deletion of the HLH domain abolishes its ability to associate with E proteins, its ability to inhibit E proteins from binding DNA, and its ability to inhibit MyoD-dependent transactivation of the MCK enhancer. Disruption of either amphipathic α-helix with a proline residue also almost completely inhibits Id1 activity in these three assays. On the other hand, deletion of all sequences amino terminal to the HLH domain or all sequences carboxy terminal to the HLH domain has little effect on Id1 activity.

Expression of the Id1 HLH domain by itself (HLH-Id1) partially inhibits MyoD-dependent transactivation of the

In 21 of the 23 Id mutations analyzed in this study, there is a qualitative and quantitative correlation between the ability of a given mutant to inhibit MyoD-dependent transactivation of the MCK enhancer and the ability to inhibit E proteins from binding DNA *in vitro*. This strongly supports the idea that E proteins are cellular targets for Id inhibition. The only exceptions are HLH-Id (see above) and S83-D. The latter mutation inhibits E proteins from binding DNA *in vitro* but does not inhibit MyoD *in vivo*. While this might indicate that antagonizing the activity of E proteins in the cell is not

We have also shown here that E12 is inhibited from binding DNA at much lower Id1 concentrations than E47. This is probably not simply a result of differences in the stabilities of E12 and E47 homodimers, as the two molecules have been reported to have similar association constants (20). An alternate explanation might be that E12 forms higher-order multimers than E47 and that a single molecule

FIG. 4. E12 with an Id1 loop substitution does not homodimerize but heterodimerizes with MyoD. (A) E12 and 1/2 Id loop→E12 (see the text) were tested in a glutaraldehyde cross-linking assay. E12 synthesized in a reticulocyte lysate was incubated at 25°C with buffer alone (–) (0.2 M KCl plus 0.2% Nonidet P-40) or glutaraldehyde at a final concentration of 0.025% (+) or 0.05% (++). Molecular mass standards (in kilodaltons) are indicated on the left. The position of E12 and the E12 dimer are indicated. A 69-kDa band present in the unprogrammed reticulocyte lysate reaction is also visible. (B) Gel mobility shift assay with the indicated proteins and the MCK enhancer oligonucleotide were performed as for Fig. 1. All samples were heated to 37°C prior to the addition of the labelled probe.

of Id1 is sufficient to inhibit either complex from binding DNA. The observation that E12-DNA complexes run much more slowly in gel mobility shift assays than E47-DNA complexes despite the similar molecular weights of the two proteins is consistent with this possibility. Regardless of the mechanism, however, the result suggests that Id1 may selectively inhibit some members of the E protein family under certain conditions of intracellular concentration.

The loop region of MyoD can tolerate drastic alterations in sequence with little effect on its activity (5). Replacing seven of the nine loop residues with alanine reduces MyoD activity by only about 50%. In addition, swapping the loop region of MyoD for that of the c-Myc or achaete-scute T4 protein, which changes the composition of the loop and increases the size by 1 or 12 residues, respectively, has little effect on MyoD activity. Our studies here indicate that the loop region of Id1 is not nearly that flexible. It was first observed that swapping the Id1 loop for that of MyoD led to a complete

loss of Id1 activity in all three assays. By changing each residue in the Id1 loop to the corresponding residue of MyoD, we determined that two residues were absolutely essential for Id1 activity: a proline residue at the amino terminus of the loop (P-74) and a leucine at the third position in the loop (L-76). Other residue changes produced effects ranging from partial loss of Id1 activity to no detectable change in activity (see Results).

Since P-74 would be expected to disrupt any α -helical structure at the beginning of the Id1 loop as well as impose a tertiary structural constraint, we attempted to determine which one of these parameters was more important in maintaining Id1 activity. Changing P-74 to helix-breaking residues (asparagine or glycine) resulted in little or no Id1 activity. All other substitutions at this position also failed to produce an active molecule. It seems reasonable, therefore, to postulate that the geometry imposed by the α -imino group at this position fixes the three-dimensional structure of the HLH domain in a way which is important for maintaining its activity.

Since Id1 exists as a monomer and binds tightly to the E protein family, we tested whether the Id1 loop was responsible in any way for these properties. Swapping the first four amino acid residues of the E12 and E47 loops for the corresponding amino acids of Id1 resulted in molecules that could no longer form homodimers (as assayed by cross-linking experiments) but could still form weak heterodimers with MyoD, a property characteristic of the Id1 family. We should point out that the inability of Id to homodimerize and its ability to heterodimerize with another member of the bHLH family are properties which are not necessarily linked. That is, although we show that the Id loop mutations which inactivate the protein cannot heterodimerize with the E proteins, we do not necessarily believe that this is due to the formation of stable homodimers. The issue of whether the loop mutants we describe or the MyoD loop has the property of promoting homodimers in a molecule which normally exists as a monomer is currently under investigation.

Our results indicate that in the case of the HLH protein Id1, the loop is not simply a nonfunctional spacer between the two amphipathic helices but may play an important role in specifying the correct dimerization partner. It will be of interest to determine whether the first four amino acids of the twist loop (which are identical to the corresponding amino acids of Id) play a similar role in its activity. It is clear, however, from inspection of the loop sequences of other HLH proteins that a proline at position 1 is not sufficient to prevent homodimer formation; i.e., Max, TFE3, and USF, for example, have prolines in this position yet do form stable homodimers. Other structural elements (i.e., the leucine zippers present in these proteins) may contribute to the stability of the homodimer. In addition, it is possible that the entire PTLF motif at the beginning of the loop may be required to facilitate the dimerization specificity we observe.

The crystal structure of the Max homodimer indicates that this bHLH protein forms a parallel four-helix bundle when bound to DNA (7). It has been proposed that the packing of the conserved hydrophobic residues of the HLH domain in the interior of this structure contributes to the stability of the homodimer. The suggestion, however, that Id can form homodimers because of the conservation of these hydrophobic residues is most probably incorrect, on the basis of previously published biochemical data (21) and this study. Rather, it would appear that other elements in the structure of the full-length Id1 protein override this stabilizing force,

and we propose that the PTLP subdomain in the loop is at least one of those elements.

ACKNOWLEDGMENTS

We thank Hal Weintraub and Keith Blackwell for their comments on the manuscript. R.B. thanks Jill, Noah, and Stella for their unwavering support.

This work was supported by NSF grant IBN-9118977 and NCI grant P30-CA-08748.

REFERENCES

- Benezra, R. Unpublished data.
- Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* 61:49-59.
- Blackwell, T. K., and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* 250:1104-1110.
- Christy, B. A., L. K. Sanders, L. F. Lau, N. G. Copeland, N. A. Jenkins, and D. Nathans. 1991. An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. *Proc. Natl. Acad. Sci. USA* 88:1815-1819.
- Cordle, S. R., E. Henderson, H. Masuoka, P. A. Weil, and R. Stein. 1991. Pancreatic β -cell-type-specific transcription of the insulin gene is mediated by basic helix-loop-helix DNA-binding proteins. *Mol. Cell. Biol.* 11:1734-1738.
- 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.
- Fairman, R., R. K. Beran-Steed, S. J. Anthony-Cahill, J. D. Lear, W. F. Stafford III, W. F. DeGrado, P. A. Benfield, and S. L. Brenner. Multiple oligomeric states regulate the DNA-binding of helix-loop-helix peptides. *Proc. Natl. Acad. Sci. USA*, in press.
- Ferre-D'Amare, A. R., G. C. Prendergast, E. B. Ziff, and S. K. Burley. 1993. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature (London)* 363:38-45.
- Halazonitis, T. D., K. Georgopoulos, M. E. Greenberg, and P. Leder. 1988. c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell* 55:917-924.
- Ingraham, H. A., S. E. Flynn, J. W. Voss, V. R. Albert, M. S. Kapiloff, L. Wilson, and M. G. Rosenfeld. 1990. The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. *Cell* 61:1021-1033.
- Jen, Y., H. Weintraub, and R. Benezra. 1992. Overexpression of Id protein inhibits the muscle differentiation program: in vivo association of Id with E2A protein. *Genes Dev.* 6:1466-1479.
- Kreider, B., R. Benezra, G. Roverea, and T. Kadesch. 1992. Inhibition of myeloid differentiation by the helix-loop-helix protein Id. *Science* 255:1700-1702.
- Lassar, A. B., R. L. Davis, W. E. Wright, T. Kadesch, C. Murre, A. Varonova, 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.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Murray, S. S., C. A. Glackin, K. A. Winters, D. Gazit, A. J. Kahn, and E. J. Murray. 1992. Expression of helix-loop-helix regulatory genes during differentiation of mouse osteoblastic cells. *J. Bone Miner. Res.* 7:1131-1138.
- Murre, C., P. S. McCaw, H. Vassin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, N. 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.
- Pongubala, J. M. R., and M. L. Atchison. 1991. Functional characterization of the developmentally controlled immunoglobulin kappa 3' enhancer: regulation by Id, a repressor of helix-loop-helix transcription factors. *Mol. Cell. Biol.* 11:1040-1047.
- Ron, D., and J. F. Habener. 1992. CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as dominant-negative inhibitor of gene transcription. *Genes Dev.* 6:439-453.
- Sartorelli, V., N. A. Hong, N. H. Bishopric, and L. Kedes. 1992. Myocardial activation of the human cardiac α -actin promoter by helix-loop-helix proteins. *Proc. Natl. Acad. Sci. USA* 89:4047-4051.
- Sawada, S., and D. R. Littman. 1993. A heterodimer of HEB and an E12-related protein interacts with the CD4 enhancer and regulates its activity in T-cell lines. *Mol. Cell. Biol.* 13:5620-5628.
- 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.
- Sun, X.-H., N. G. Copeland, N. A. Jenkins, and D. Baltimore. 1991. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol. Cell. Biol.* 11:5603-5611.
- Treacy, M. N., X. He, and M. G. Rosenfeld. 1991. I-POU: a POU-domain protein that inhibits neuron-specific gene activation. *Nature (London)* 350:577-584.
- Vinson, C. R., and K. C. Garcia. 1992. Molecular model for DNA recognition by the family of basic-helix-loop-helix-zipper proteins. *New Biol.* 4:396-403.
- Voronova, A., and D. Baltimore. 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci. USA* 87:4722-4726.
- Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, 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.
- Wilson, R. B., M. Kiledjian, C.-P. Shen, R. Benezra, P. Zwollo, S. M. Dymecki, S. V. Desiderio, and T. Kadesch. 1991. Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: implications for B-lymphoid-cell development. *Mol. Cell. Biol.* 11:6185-6191.