Basic region helix-loop-helix (bHLH) transcription factors regulate key steps in early development by binding to regulatory DNA sites as heterodimers consisting of a tissue-specific factor and a widely expressed factor. We have examined the folding, dimerization, and DNA binding properties of the muscle-specific bHLH protein MyoD and its partner E47, to understand why these proteins preferentially associate in heterodimeric complexes with DNA. In the absence of DNA, the E47 bHLH domain forms a very stable homodimer, whereas MyoD is unfolded and monomeric. Fluorescence quenching experiments show that MyoD does not dimerize with E47 under dilute conditions in the absence of DNA. Residues in and around the loop of the E47 bHLH domain contribute to its markedly greater stability. An altered MyoD bHLH substituted with the loop segment from E47 folds in the absence of DNA, and it readily dimerizes with E47. In the presence of a specific DNA binding site, MyoD and E47 both form homodimeric complexes with DNA that have similar dissociation constants, despite the very different stabilities of these protein dimers off DNA. A 1:1 mixture of these bHLH domains forms almost exclusively heterodimeric complexes on DNA. Assembly of these bHLH-DNA complexes is apparently governed by the strength of each subunit’s interaction with the DNA and not by the strength of protein-protein interactions at the dimer interface. These findings suggest that preferential association of MyoD with E47 in DNA complexes results from more favorable DNA contacts made by one or both subunits of the heterodimer in comparison with either homodimeric complex.

At the heart of transcriptional initiation are numerous cooperative interactions between basal factors, transcriptional activators, and their promoter DNA sites. Many activating transcription factors bind to their DNA targets as heterodimers in which each subunit has a distinctive DNA binding preference and transcriptional activity. The strength of transcriptional activation is modulated by the choice of subunits forming these protein dimers, their cellular concentrations, and their affinities for DNA regulatory sites. Basic region helix-loop-helix (bHLH) transcription factors are developmental regulators of transcription in a variety of tissues, including muscle (1, 2), nervous tissue (3–7), and blood cell lineages (8, 9). In general, bHLH proteins function as heterodimers of a tissue-specific bHLH protein and one of the widely expressed bHLH proteins known as E-proteins. The well characterized bHLH protein MyoD is a master regulator of skeletal muscle development (2, 10, 11). MyoD binds to regulatory sequences upstream of muscle-specific genes and activates transcription. MyoD’s biological activity is contingent upon E-proteins, such as the alternative splice products of the E2A gene, E47 and E12, which physically associate with MyoD and bind to regulatory DNA sites as heterodimers (12–14). Dimerization of MyoD and other bHLH proteins is mediated by the helix-loop-helix motif, a four-helix bundle that positions the N-terminal basic region in the major groove of DNA (15, 16). The 60-residue bHLH domain of MyoD is necessary and sufficient for the myogenic conversion of cultured fibroblasts (1), implying that the delivery of this compact DNA-binding domain to E-box regulatory sites (CANNTG) can commit cells to a program of muscle development.

E-proteins promote MyoD’s transcriptional activity by recruiting it to DNA regulatory sites in heterodimeric DNA complexes that form in preference to either homodimeric complex (13, 14, 17, 18). The higher stability of these heterodimeric complexes might result from the strength of bHLH heterodimerization as well as from the strength of interactions with the DNA. Despite the similar crystal structures of the MyoD-DNA complex and the E47-DNA complex (15, 16) (Fig. 1), the protein fold of the MyoD bHLH is markedly less stable than that of E47 (19, 20). E-proteins could stabilize the folded conformation of the MyoD bHLH by forming stable heterodimers that deliver MyoD to DNA. Alternatively, the number or strength of the DNA contacts made by the MyoD-E47 heterodimer may be increased relative to those made by either homodimer. We have investigated the dimerization strengths and preferences of the MyoD and E47 bHLH domains to address the selective formation of MyoD-E47-DNA complexes.

**EXPERIMENTAL PROCEDURES**

**Protein Overexpression and Purification**—The MyoD and E47 bHLH domains were produced in Escherichia coli from the pSET71 expression vector (InVitrogen, Inc.) using the expression host BL21(DE3) pLysS (21). The MyoD bHLH peptide used for the crystal structure determination includes a Met-Glu-Leu sequence at its N terminus and the substitution of Cys with serine (15). We appended the sequence Gly-Gly-Cys to the C terminus of the MyoD peptide to specifically label it with fluorescent probes (see below). The E47 bHLH peptide is that used for the crystal structure (16) with the addition of a Gly-Gly-Cys sequence to its C terminus. All bHLH peptides were purified from E. coli cultures as described previously for the E47 bHLH (16). Mass spectrometry of the purified bHLH domains after modification of the cysteine confirmed that the bacterially produced peptides comprise the residues shown in Fig. 1 (see below). The purified peptides were quantitated by amino acid analysis, and from this a molar absorptivity at 280 nm was determined for MyoD (ε = 1434 M−1 cm−1). Because E47 lacks Tyr and Trp residues, E47 protein concentrations were routinely determined by a Bradford protein assay (Bio-Rad) that was calibrated with an E47 standard previously quantitated by amino acid analysis.
**Fig. 1.** A, the superimposed crystal structures of MyoD (thick lines; Ref. 16) and E47 (thin lines; Ref. 16) bHLH domains are very similar, except for the C terminus of helix 1 and the loop segments connecting helices 1 and 2. Tyr<sup>230</sup> of MyoD helix 2 (side chain shown) prevents the extension of helix 1 that is present in E47. The helix 1 extension increases the buried surface of the E47 bHLH domain and allows the formation of an intersubunit salt bridge, which probably contributes to the greater stability of E47 dimers. In addition, a network of hydrogen bonds that joins the E47 loop to helices 1 and 2 stabilizes the fold of each subunit (16). The DNA in the E47 complex is not shown here. B, sequences of the MyoD and E47 bHLH domains and of the chimeric MyoEL domain used in our studies are shown. The DNA-contacting residues (dark boxes) and buried residues (gray boxes) of the MyoD and E47 bHLH domains are highly conserved. We constructed a chimeric MyoEL peptide consisting of the MyoD bHLH domain with residues substituted from the structurally unique segments in and around the loop residues (sequence). We reasoned that substitution of these E47 residues into the MyoEL bHLH might prevent the extension of helix 1 and the formation of an intersubunit salt bridge seen in the E47 complex.

**Generation of bHLH Mutants**—Some bHLH peptides were appended with the sequence -Gly-Gly-Cys-COOH to facilitate specific labeling with thiol-specific reagents. Nucleotides encoding these residues were introduced by polymerase chain reaction (PCR) using appropriately modified primers. PCR products were digested with NdeI and BamHI, inserted in the expression vector pRSET, and sequenced. In contrast to MyoD, the addition of the C-terminal Gly-Gly-Cys-COOH to the E47 bHLH greatly decreased the amount of peptide produced in E. coli. The E47 expression construct was then modified to include four additional E47 residues (-Val-Arg-Glu-Arg-) at the C terminus of the HLH segment, followed by the -Gly-Gly-Cys-COOH sequence. This peptide was produced in high levels in E. coli, and it bound DNA with the affinity of the wild-type E47 bHLH domain. A chimeric MyoD bHLH incorporating residues from the E47 loop and named MyoEL (Fig. 1) was generated by PCR in two steps, following published protocols (22). First, the C-terminal mutations were introduced by PCR with a mutagenic primer. This reaction product then served as template for a second PCR reaction using a mutagenic primer that extends to a unique StuI site. A StuI-BamHI fragment of the second PCR product was ligated into the complementary position of the wild-type MyoD expression construct. The resulting MyoEL bHLH peptide lacks a C-terminal Gly-Gly-Cys sequence, and it was purified by the methods used for the wild-type bHLH peptides.

**Fluorescent Labeling of bHLH Domains**—The C-terminal cysteines of MyoD and E47 were labeled with iodoacetamidofluorescein (IAF, Molecular Probes) to provide a sensitive monitor of bHLH dimerization by fluorescence spectroscopy. The C-terminal cysteines of the unlabeled dimers readily form an intermolecular disulfide bond, even in the presence of high concentrations of reducing agents. We therefore modified a protocol of Patel et al. (22) and reacted the peptides with IAF under denaturing conditions that favor the IAF-peptide reaction over disulfide bond formation. 4 mg of MyoD bHLH peptide was dissolved in 6 ml of 6 M guanidinium hydrochloride in 50 mM phosphate buffer (pH 7.4), 10% glycerol, and 1 mM DTT, yielding a peptide concentration of 90 μM at pH 5.5. The mixture was stirred at room temperature for 45 min, and then the reaction was adjusted to pH 8, and IAF was added to a calculated concentration of 2 mM. IAF is sparingly soluble, and it continued to dissolve during the course of the reaction. The mixture was stirred at 37 °C in the dark for 2 h. The labeling reaction was stopped by the addition of 6 mM DTT, and the solution was concentrated in a filtration device (Macrosep 3K, Filtron Technology Corporation). Excess IAF was removed by gel filtration of the labeled protein on a Sephadex G25 column, equilibrated in reaction buffer at pH 5.5. The E47 bHLH peptide has a cysteine in helix 1 (Cys<sup>363</sup>) in addition to the C-terminal cysteine. We substituted serine for Cys<sup>363</sup> to direct thiol-reactive labels to the C terminus, but the mutant peptide had an altered CD spectrum and a lower DNA binding affinity than wild-type E47. Thus, the cysteine-tagged wild-type E47 peptide was used in these studies, and it was labeled under partially denaturing conditions that favored reaction with the C-terminal cysteine in preference to the internal cysteine. The E47 bHLH peptide was diluted into 1.5 M guanidinium hydrochloride, 50 mM phosphate (pH 7.4), and 2 mM tri(2-carboxyethyl)phosphine (24). After adjusting the pH to 8, IAF was added to a concentration of 0.65 mM, and fresh tri(2-carboxyethyl)phosphine was added to a 4 mM final concentration. The reaction proceeded for 90 min at 37 °C in the dark before quenching with DTT and purifying the labeled E47CM peptide (fE47) as for fMyoD above. To control for the effect(s) of the C-terminal -Gly-Gly-Cys residues, the C-terminal cysteine of MyoD was carboxymethylated with iodoacetamide (Sigma) under the conditions used for IAF labeling, and the modified peptide was used in spectroscopic and electrophoretic mobility shift assay (EMSA) studies. Control experiments showed that the dimerization properties of the carboxymethylated MyoD bHLH peptide were indistinguishable from those of the MyoD bHLH domain used for the crystal structure determination (15). We were unable to selectively carboxymethylate the C-terminal cysteine of the E47 bHLH without modifying its internal cysteine (Cys<sup>363</sup>) under the partially denaturing conditions, indicating that both cysteines are accessible to this small compound. Therefore, we used an E47 bHLH peptide lacking the C-terminal cysteine for studies of the nonfluoresceinated peptide. Final purification of all C-terminally modified bHLH peptides was achieved by reverse-phase high performance liquid chromatography using a semipreparative C18 column (Waters).
It is very unlikely that the internal cysteine of fE47 (Cys363) was labeled, protein concentration determined by quantitative amino acid analysis. Scopistic experiments was 20 mM ammonium acetate (pH 7), 100 mM KCl, as wild-type E47.pared 9M stocks, and the results were analyzed by linear extrapolation presented here are the average of 10 wavelength scans performed at 25 °C, unless otherwise noted. The urea concentrations for denaturation studies were obtained by serial dilution of gravimetrically prepared 9 m stocks, and the results were analyzed by linear extrapolation to benign (no denaturant) conditions. Molar ellipticity values were calculated for the bHLH peptides, ignoring any nonnative residues at the peptide ends, and converted to α-helical content using the formula \[ \theta_{222} = [\text{helix}] \times 30,300 - 2860 \] (25).

Fluorescence Spectroscopy—Quenching of fluorescein-labeled bHLH dimers was monitored with an AB2 luminescence spectrometer (SLM/AMINCO) using an excitation wavelength of 492 nm, and recording emission at 520 nm in 1-cm methyl acrylate cuvettes. A bandwidth of 4 nm was used for both emission and excitation in most experiments; at low peptide concentrations the bandwidth was increased to a maximum of 10 nm. The emission values reported here are an average of three readings for each sample. Samples were prepared by mixing the fluorescein-labeled peptide with the same buffer as was used for CD experiments (above) and then incubating at 41 °C for 10 min to ensure the exchange of bHLH subunits. The same results were obtained for samples incubated 4 h or more at room temperature, although total fluorescence decreased at a rate of about 3%/h in this buffer. DNA complexes were formed by adding the DNA to the pre-equilibrated bHLH subunits, and then samples were incubated in the dark for an additional 120 min at room temperature prior to analysis.

DNA Binding Assays—The DNA binding affinities of bHLH dimers were assessed by EMSAs (28), using a 21-base pair oligonucleotide containing either of two binding sites, 5′-CAGGTG-3′ or 5′-CAGCTG-3′, embedded in the sequence of the MyoD-responsive creatine kinase enhancer. In a typical DNA binding reaction, the bHLH peptide(s) were mixed with 0.002 pmol of DNA (0.1 μM final concentration) in 20 μl of binding buffer consisting of 20 mM Tris (pH 7.5), 100 mM KCl, 10% glycerol, 1 mM DTT, 0.5% Nonidet P-40, and 1.2 mg ml−1 sheared salmon sperm DNA. The protein concentration was at least 10-fold higher than the DNA concentration, even at the lowest protein concentration. Prior to mixing with the DNA, the bHLH proteins were incubated at 41 °C for 10 min to ensure subunit exchange. Longer incubation at room temperature gave the same results. The complete binding reaction was incubated for 60 min at 22 °C, and 5 μl was analyzed by electrophoresis through a 12% polyacrylamide gel in 90 mM Tris borate (pH 7.6), 2 mM EDTA (27). The positions of radiolabeled DNA in the gel were determined by phosphor image analysis (Fujix BAS1000, Fuji Co). The amounts of free and bound DNA were quantitated with MacBAS version 2.1 image analysis software, and the resulting binding isotherms were fitted to an equation that describes a monomer dissociation constant and cooperativity factor (see Equation 1) using the program Kaleidagraph (Abelbeck Software). At saturating peptide concentration, the percentage of DNA bound, was determined graphically from the binding isotherms. The DNA binding activity of each purified bHLH peptide was virtually 100%, as determined by the stoichiometric titration of a fixed amount of DNA with protein

2 H. R. Bosshardt and C. Berger, personal communication.

MyoD-E47 Dimerization

RESULTS

bHLH Folding and Dimerization in Solution—The bHLH domains of MyoD and E47 are predominately α-helical at peptide concentrations above 10 μM, as evidenced by the characteristic double minimum at 208 and 222 nm in their CD spectra (Fig. 2). The α-helical content of both peptides increases from approximately 50% helix to more than 80% α-helix upon the addition of stoichiometric amounts of a DNA binding site (Fig. 2), consistent with a DNA-induced folding of the NH2-terminal basic region segment that directly contacts DNA in crystal structures of these bHLH-DNA complexes (15, 16). In solution, the MyoD bHLH unfolds into unstructured monomers at lower peptide concentrations. This concentration-dependent equilibrium between the unfolded monomer and the α-helical oligomer fits a monomer-dimer equilibrium with a dissociation constant \( K_{d} = 9.4 \mu M \) (Fig. 3A and Table I). Analytical ultracentrifugation experiments using this MyoD peptide at concentrations above 15 μM reveal an additional dimer-tetramer equilibrium (not shown), as previously reported for the MyoD bHLH domain (20, 30). In contrast to MyoD, the E47 bHLH is a stable dimer over the full range of peptide concentrations accessible by CD spectroscopy and analytical ultracentrifugation (0.5–50 μM peptide data not shown).

The relative stabilities of E47 dimers and MyoD dimers were also determined from their helical stabilities in increasing concentrations of urea (Fig. 3C). Extrapolation to native conditions (28, 31) yields the Gibbs free energy of unfolding, which in turn provides the estimates of \( K_{d} = 20 \) nM for the E47 dimer and \( K_{d} = 1.4 \mu M \) for the MyoD dimer (Table I). The low stability of the

FIG. 2. The secondary structures of E47 (●), MyoD (●), and MyoEL (●) are almost identical. CD spectra of the peptides at 20 μM concentration show the characteristic double minimum typical of α-helical structures. The addition of a stoichiometric amount of an oligonucleotide containing the creatine kinase enhancer site to E47 (●) or MyoD (●) increases the α-helical content because the basic region folds upon binding to DNA.
MyoD-E47 Dimerization

**Table I**

| Dimer       | \( K_d \) in urea | \( \Delta G_{\text{obs}}^{\text{dimer}} \) | \([\text{urea}]_{50}^d \) |
|-------------|-------------------|---------------------------------|-----------------|
| E47 ND     | 0.020 ± 0.01      | 44.0 ± 1.5                     | 3.0             |
| MyoD 9.4 ± 2.2 | 1.4 ± 1.0       | 33.4 ± 3.1                     | 0.8             |
| MyoEL 1.5 ± 0.2 | 0.68 ± 0.07     | 35.2 ± 0.4                     | 1.75            |

\( a \) Dimerization constant determined from the concentration dependence of the CD signal.

\( b \) Dimerization constant determined from the free energy of unfolding, \( \Delta G_{\text{obs}}^{\text{dimer}} \).

\( c \) Free energy change associated with folding and dimerization of bHLH domains, determined from denaturation curve by the linear extrapolation method.

\( d \) Midpoint of the denaturation curve.

\* Not determined.

MyoD dimer prevented accurate measurement of the baseline slope (pre-unfolding) at low urea concentrations, and we have therefore used the slope from E47’s pre-unfolding transition in our estimate of \( K_o \) for MyoD. Nonetheless, this estimate of \( K_d \) for the MyoD dimer is in reasonable agreement with \( K_{\text{dimer}} \) determined directly (Fig. 3A and Table I). Despite their similar structures when bound to DNA (15, 16), it is evident that the E47 dimer is several orders of magnitude more stable than MyoD in the absence of DNA, corresponding to a free energy difference of approximately 10 kJ mol\(^{-1}\) for folding and dimerization.

**DNA-mediated Folding of bHLH Dimers**—If the dimerization strength of the bHLH domain contributes substantially to DNA binding affinity, then the stable E47 dimer should bind to DNA with higher affinity than the unstable MyoD dimer. We measured the DNA binding affinities of the E47 bHLH and the MyoD bHLH by electrophoretic mobility shift assay (Figs. 4 and 5). The E47 bHLH binds to the -CAGGTG sequence of the creatine kinase enhancer with an apparent affinity (\( K_{\text{obs}} \)) of 18 nM (Fig. 5, Table II). MyoD bHLH binds to DNA with almost the same affinity (\( K_{\text{obs}} = 28 \) nM) as E47. Thus, in the presence of DNA, MyoD dimers form at concentrations that are 200-fold lower than \( K_{\text{dimer}} \) in the absence of a binding site (Tables I and II). The exclusive formation of MyoD-E47-DNA complexes from a mixture of these subunits (Fig. 4) results from the marginally higher affinity of the heterodimeric complex (\( K_{\text{obs}} = 14 \) nM) in comparison with either homodimeric complex.

The DNA binding reaction could follow either of two principal mechanisms. If only preformed dimers bind to DNA, and the concentration of monomers is negligible at the protein concentrations used in the binding assay, the binding isotherm could be represented by a simple relationship describing a hyperbolic curve. However, the binding curves for E47 dimers and for MyoD dimers are sigmoidal, suggestive of the cooperative binding of two bHLH monomers to DNA. We therefore fit the data to Equation 1 (32), which describes the dissociation constant (\( K_{\text{diss}} \)) of a monomer interacting with DNA and a cooperativity factor (\( a \)) that accounts for the much higher affinity of dimeric protein-DNA complexes.

\[
\frac{B}{B_{\text{max}}} = \frac{[P]}{K_{\text{diss}}} + \frac{[P]^2}{aK_{\text{diss}}} \\
1 + \frac{2[P]}{K_{\text{diss}}} + \frac{[P]}{aK_{\text{diss}}} 
\]  
(Eq. 1)

The DNA binding data of 4–7 independent experiments were fit to this relationship using a least squares fitting procedure (33), and the resulting curves are plotted in Fig. 5. The average values of the dissociation constants (\( K_{\text{diss}} \)) for E47 and MyoD are identical, and similar cooperativities (\( a \)) are observed for both homodimer-DNA complexes (Table II). The values obtained for the homodimer-DNA complexes were input as constants into an expanded form of Equation 1 to determine the \( a \) value (cooperativity) of MyoD-E47-DNA complex formation. This expanded equation includes terms for both monomers binding to DNA and for the heterodimer-DNA complex. Homodimer-DNA complexes are not observed under these conditions (Fig. 4). We set \( K_{\text{diss}} \) of the MyoD-E47-DNA complex equal to that obtained for the homodimer-DNA complexes and derived the cooperativity value (\( a \)) that best fits the EMSA data.
The cooperativity of heterodimer-DNA complex formation is nearly 10-fold higher than that of either homodimer-DNA complex (Table II).

The similar $K_{\text{obs}}$ for DNA complex formation by MyoD and E47, despite their markedly different dimerization strengths off DNA (Figs. 3 and 4), shows that the DNA binding affinities are dominated by each subunit’s interaction with DNA and not by the strength of protein-protein interactions at the dimer interface. Thus, the folding of the basic region of MyoD and its interaction with DNA are necessary for stable MyoD dimers to form at these protein concentrations. To further test this model, we oxidized the COOH-terminal cysteines of MyoD to form a stable disulfide-linked MyoD homodimer. The DNA binding isotherm of the covalently linked MyoD dimers is more hyperbolic (less sigmoidal) in comparison with those for the unlinked bHLH proteins (Table II), as expected if protein-DNA contacts predominate over protein-protein contacts in the stabilization of bHLH-DNA complexes.

The MyoD-E47 Heterodimer Is Favorable Only in the DNA Complex—The myogenic activity of MyoD and other muscle-specific bHLH transcription factors requires widely expressed E-proteins such as E47. MyoD freely associates with E47 on DNA, forming high affinity heterodimeric complexes (15, 18) (Figs. 3 and 4). Although the MyoD bHLH and the E47 bHLH bind to DNA with comparable affinities as homodimers (Table I), a stoichiometric mixture of these peptides preferentially binds to DNA as the MyoD-E47 heterodimer (open arrow). The almost exclusive formation of the heterodimeric DNA complex is striking because only the E47 subunit is folded at this concentration in the absence of DNA (Fig. 3). Protein concentrations for E47 and the heterodimer were 2, 3, 7.5, 10, 15, 17.5, 20, 25, 30, and 40 nM.

MyoD and E47 peptides should yield a 1:2:1 ratio of MyoD-MyoD, MyoD-E47, and E47-E47 in complex with DNA. However, the observed DNA binding equilibrium greatly favors the heterodimeric complex (Fig. 5). This finding suggests either that the MyoD-E47 dimer is exceptionally stable or that one or both subunits are positioned for tighter DNA interactions in the heterodimer than they are in either homodimer. To address these two possibilities, we investigated the bHLH heterodimer formation in the absence of DNA at peptide concentrations approximating the $K_{\text{dimer}}$ for DNA binding.

bHLH dimerization in solution was monitored with fluorescently labeled bHLH domains. Juxtaposition of two fluorescent groups in the bHLH dimer quenches the fluorescence (see “Experimental Procedures”), providing a convenient and sensitive assay of protein dimerization. Similar methods have been used to study the dimerization of basic region leucine zipper domains (23, 34). The emission of fMyoD was proportional to peptide concentration in the range of 1 nM to 1 μM peptide, indicating no change in aggregation state (data not shown).

The unlabeled MyoD bHLH is likewise unfolded and monomeric at these concentrations (Fig. 3A). At concentrations above 1 μM, the strong absorbance of fMyoD prevented an accurate measurement of $K_{\text{dimer}}$. Unlike fMyoD, the fluoresceinated E47 bHLH peptide (fE47) dimerizes at low nanomolar concentrations. fE47 fluorescence was significantly quenched at peptide concentrations higher than about 5 nM, although the weak fluorescence at subnanomolar concentrations precluded an accurate determination of $K_{\text{dimer}}$ from the fluorescence quenching data.

The fluoresceinated bHLH peptides are particularly good probes of MyoD-E47 heterodimer formation in solution. Heterodimers consisting of labeled and unlabeled bHLH subunits fluoresce more strongly than the labeled homodimers. Thus, heterodimer formation can be monitored by titrating a fixed concentration of a fluoresceinated bHLH domain with increasing concentrations of an unlabeled bHLH subunit. We fixed the concentration of fE47 at 50 nM, approximately the $K_{\text{dimer}}$ for DNA binding (Table II), and varied the amount of unlabeled E47 or unlabeled MyoD. The unlabeled E47 bHLH readily dimerizes with fE47, as evidenced by the increase in fluorescence as the equilibrium shifts from fE47 dimers to E47-fE47 heterodimers with increasing concentrations of competitor (Fig. 6A). The addition of DNA, at a concentration equal to the highest concentration of unlabeled competitor protein, has no measurable effect on E47 dimerization with fE47 (Fig. 6B). We conclude that the E47 bHLH is mainly dimeric at peptide concentrations that approximate the $K_{\text{dimer}}$ for DNA. In contrast to E47, MyoD has little effect on the fluorescent emission of fE47 over this range of peptide concentration in the absence of DNA. At higher MyoD concentrations, approaching the $K_{\text{dimer}}$ of the MyoD homodimer, the quenching of the fE47 dimer is relieved. Upon the addition of the -CAGGTG- binding site, MyoD readily associates with fE47 (Fig. 6B). The addition of nonspecific DNA to the mixture of MyoD and fE47 also promoted heterodimer formation, albeit less efficiently than the specific DNA. Thus, in the absence of a DNA binding site MyoD does not stably associate with E47. The predominance of the MyoD-E47 heterodimer in complexes with DNA (Fig. 4) is therefore not reflective of dimerization preferences in solution. Rather, the interaction with DNA promotes bHLH dimerization in general, and more specifically it favors the association of MyoD with E47.

The Loop Is a Key Determinant of bHLH Dimer Stability—To identify residues that contribute to E47’s greater dimerization strength, we attempted to stabilize MyoD dimers by substituting a limited number of residues from the structurally unique bHLH domains of E47.
segments of the E47 into the MyoD bHLH (Fig. 1B). A comparison of the crystal structures of the E47 and MyoD homodimers shows localized differences at the helix 1-loop junction near the dimer interface (Fig. 1A) (15, 16). Helix 1 of E47 is one turn longer than the corresponding helix of MyoD, and the E47 loop has several stabilizing interactions that link residues in the loop to nearby residues of helices 1 and 2. We therefore substituted the helix 1 extension and the loop of E47 into the bHLH domain of MyoD. In addition, the residues of helix 2 that interact with the loop in the E47 crystal structure were substituted with the corresponding E47 residues (Fig. 1B). We predicted that substitution of these 18 residues would stabilize the fold of the resulting MyoEL protein (MyoD with E47 loop) and thereby indirectly promote dimerization.

The CD spectrum of MyoEL shows that its secondary structure is similar to that of MyoD and E47, corresponding to 52% α-helix content in the absence of DNA (Fig. 2). However, the MyoD dimer is significantly more stable than the MyoD dimer (Fig. 3; Table I). The 6-fold difference in $K_{\text{dimer}}$ (Table I) conferred by substituting the loop and nearby residues of E47 into the MyoD bHLH corresponds to a free energy difference of 1.8 kJ mol$^{-1}$ for the wild-type and altered MyoD domains. MyoEL forms tetramers at very high peptide concentrations, albeit less efficiently than wild-type MyoD, and MyoEL binds to DNA with an apparent affinity similar to that of E47 or MyoD (data not shown). The enhanced dimerization of MyoEL in comparison with MyoD suggests that residues within the E47 loop region stabilize the fold of the bHLH domain (Table I). Because wild-type MyoD does not strongly associate with E47 in the absence of DNA, we expected that MyoEL would likewise fail to stably dimerize with E47 under these conditions. Surprisingly, MyoEL dimerizes strongly with E47 in the absence of DNA (Fig. 6A). Similar concentrations of E47 and MyoEL disrupt half of the dE47 dimers in solution, showing that the stability of the MyoEL–E47 heterodimer is similar to that of the E47 homodimer. Therefore, substitution of the E47 loop region into the MyoD bHLH markedly enhances dimerization with E47.

**DISCUSSION**

Transcriptionally active bHLH complexes consist of two protein subunits bound to a regulatory DNA site. Although many bHLH proteins can bind to DNA as homodimers, they regulate cellular development by the selective formation of DNA-bound bHLH heterodimers. Consistent with this in vivo function, heterodimeric DNA complexes predominate over homodimeric complexes when appropriate bHLH partners are mixed with DNA sites in vitro (Fig. 4; Refs. 1, 13, 14, and 18). A full description of transcriptional regulation by bHLH proteins re-

---

**TABLE II**

| Dimer               | $K_{\text{d}}$ a | $\alpha$ (cooperativity) a | $aK_{\text{d}}$ a | $K_{\text{obs}}$ b |
|---------------------|------------------|----------------------------|-------------------|-------------------|
| E47                 | 160 ± 97         | 0.013 ± 0.015               | 2.08 ± 1.46       | 17.8              |
| MyoD                | 160 ± 120        | 0.029 ± 0.042               | 4.64 ± 5.04       | 27.5              |
| E47–MyoD            | 0.0021 ± 0.0003  | 0.26 ± 0.14                 | 4.16 ± 0.74       | 8.0               |
| Cross-linked MyoD   | 16 ± 5.3         | 0.26 ± 0.14                 | 4.16 ± 0.74       | 8.0               |

a Parameters obtained from linear least squares fitting.
b Protein concentration at which half of the DNA is bound. It was determined graphically from the binding isotherm.

The CD spectrum of MyoEL shows that its secondary structure is similar to that of MyoD and E47, corresponding to 52% α-helix content in the absence of DNA (Fig. 2). However, the MyoD dimer is significantly more stable than the MyoD dimer (Fig. 3; Table I). The 6-fold difference in $K_{\text{dimer}}$ (Table I) conferred by substituting the loop and nearby residues of E47 into the MyoD bHLH corresponds to a free energy difference of 1.8 kJ mol$^{-1}$ for the wild-type and altered MyoD domains. MyoEL forms tetramers at very high peptide concentrations, albeit less efficiently than wild-type MyoD, and MyoEL binds to DNA with an apparent affinity similar to that of E47 or MyoD (data not shown). The enhanced dimerization of MyoEL in comparison with MyoD suggests that residues within the E47 loop region stabilize the fold of the bHLH domain (Table I). Because wild-type MyoD does not strongly associate with E47 in the absence of DNA, we expected that MyoEL would likewise fail to stably dimerize with E47 under these conditions. Surprisingly, MyoEL dimerizes strongly with E47 in the absence of DNA (Fig. 6A). Similar concentrations of E47 and MyoEL disrupt half of the dE47 dimers in solution, showing that the stability of the MyoEL–E47 heterodimer is similar to that of the E47 homodimer. Therefore, substitution of the E47 loop region into the MyoD bHLH markedly enhances dimerization with E47.
quires knowledge of the mechanism of bHLH subunit association during DNA binding. The association of particular DNA-bound dimers could result either from more favorable protein-protein interactions or from better protein-DNA contacts in the preferred complex. Because bHLH proteins show a wide range of dimerization strengths in the absence of DNA (19, 20), protein-protein interactions have been widely considered to be the more likely of these two possibilities. However, the DNA binding constants of most bHLH complexes are in the 1-30 nM range, and it is difficult to monitor these protein-protein interactions at the concentrations that are relevant for DNA binding. We therefore used a fluorescence assay to study the oligomerization of MyoD and E47 at concentrations approximating their dissociation constants to reveal the nature of the regulatory complex formation.

Although the structures of the bHLH domains of MyoD and E47 are very similar in the DNA complexes (Fig. 1), in the absence of DNA only E47 is stably folded and dimeric at low concentrations (Fig. 3). The instability of the MyoD bHLH in the absence of DNA is reminiscent of the behavior of basic region leucine zipper (bZIP) transcription factors. The bZIP protein c-Fos does not form homodimers, because of the repulsive interactions of several acidic residues in its leucine zipper. However, c-Fos readily dimerizes with its partner c-Jun, forming intersubunit salt bridges between basic residues in the leucine zipper of c-Jun and the corresponding acidic residues of c-Fos (35). We expected that MyoD and its dimerization partner E47 might likewise associate in a very stable heterodimer, and this could explain the selective binding of MyoD-E47 dimers to DNA. However, we find that the MyoD bHLH fails to stably associate with E47 in the absence of DNA (Fig. 5). The high concentrations of MyoD required for its interaction with fE47 (Fig. 5A) approach the dimerization constant of MyoD homodimers, indicating that in the absence of DNA there is no preference for the formation of MyoD-E47 heterodimers. The addition of DNA allows the formation of MyoD-E47-DNA complexes. MyoD likewise fails to dimerize with the closely related E12 protein in the absence of DNA (36). Unlike E47, E12 binds to DNA exclusively in heterodimeric complexes with MyoD or E47, because of an inhibitory domain located N-terminal to the basic region that prevents the formation of E12 homodimers (18, 36, 37). Our results suggest that the dimerization preferences in solution do not explain the DNA binding preferences of MyoD and its E-protein partners. Instead, the DNA binding equilibrium favors the MyoD-E47 heterodimer because of more favorable contacts made by one or both subunits to the DNA. This model is confirmed by the DNA binding affinities of the MyoD and E47 bHLH domains in homo- and heterodimeric complexes.

The binding isotherms for DNA complex formation are sigmoidal, which suggests two possible mechanisms of binding. Either bHLH domains can bind to DNA only as dimers and the equilibria for dimerization and DNA binding are coupled, or monomers bind initially to DNA and a second monomer binds with high cooperativity to form the stable dimeric complex. The fact that MyoD and E47 homodimers bind to DNA with comparable affinities (Table II), despite their vastly different dimerization constants off DNA (Table I), makes the mechanism of dimers binding to DNA in a coupled equilibrium seem unlikely. For coupled equilibria, the overall equilibrium binding constant of DNA complex formation is the product of the dimer dissociation constant and the dissociation constant for the dimer binding to DNA. Because the dimerization constants of E47 and MyoD differ by a factor of about 100, their very similar overall DNA binding constants can only be reached if the MyoD dimers bind to DNA 100-fold better than E47 dimers.
Considering the similar DNA contacts made by these proteins (15, 16), this possibility seems unlikely. In addition, the model of coupled equilibria predicts that covalently linked MyoD dimers should form an exceptionally stable protein-DNA complex. This is not the case, as shown by the data of Fig. 5 and Table II.

We therefore favor the model in which bHLH monomers bind transiently to DNA and a second subunit is cooperatively recruited into the complex. We have fitted the DNA binding isotherms to an equation describing a cooperative binding mechanism (Equation 1) and obtained values for the microscopic binding constants (Table II). The data are reasonably well described by this relationship. However, a nonrandom deviation of the calculated curve is evident in the transition region of some of the binding isotherms (Figs. 5, B and D). We attribute this deviation to the fact that monomeric bHLH-DNA complexes are not detected by our EMSA assays. The failure to quantitate binding intermediates in highly cooperative processes results in precisely the deflection observed here (38). Despite this shortcoming of EMSA, the calculated DNA binding affinities of the bHLH dimers (Table I) agree with previously reported values (18, 39), although we do not observe significantly different DNA binding affinities for MyoD and E47, as reported by Sun and Baltimore (18). Our results demonstrate that the dimerization strengths of MyoD and E47 have little influence on their DNA binding affinities. Although the HLH dimerization motif is required for DNA binding (40) and mutations in this segment can decrease DNA binding affinity, the primary role of the HLH interface might be to position each subunit of the dimer for optimal interactions with its DNA half-site. This positioning effect could be the basis for the observed cooperativity of binding to DNA.

Our results using bHLH domains of MyoD and E47 appear to conflict with previous reports that in vitro translated full-length MyoD is co-immunoprecipitated with epitope-tagged E47 protein in the absence of DNA (12, 14). We find that in vitro translated full-length MyoD and epitope-tagged full-length E47 do indeed coprecipitate in roughly equal amounts over a wide range of protein concentrations (0.5–100 μM concentration of each protein; data not shown). However, this protein-protein interaction appears to be nonspecific because it cannot be blocked by a 50-fold excess of the E47 bHLH studied here or the larger E47 fragment encoded by the partial cDNA originally isolated (12). These latter E47 fragments readily dimerize with full-length E47 or MyoD in complexes with DNA, leading us to conclude that the contents of immune precipitates with epitope-tagged E47 do not reflect the true binding equilibrium in solution.

The different dimerization behaviors of E47 and MyoD in the absence of DNA could provide the basis for an additional regulatory function. The enhanced dimerization strength of E47 in comparison with MyoD is due to localized structural differences in the helix-loop-helix segment, as demonstrated with the hybrid protein MyoEL. Helix 1 of the E47 HLH is one turn longer than helix 1 of MyoD or the analogous helix of the bHLH-zipper proteins Max (41) and USF (42). Helix 1 of the MyoD HLH segment is capped by a Tyr from helix 2 (Fig. 1), whereas the corresponding Val in E47 is smaller, and it allows for the extension of helix 1 by one additional helical turn. This extension of helix 1 in the E47 HLH increases the buried surface at the dimer interface, and it facilitates an intersubunit salt bridge between His386 at the COOH terminus of helix 1 and Glu386 in helix 2 of the opposing subunit. Residues in the loop of the E47 bHLH domain form a network of hydrogen bonds that bridges between helices 1 and 2 within each subunit. This unique hydrogen bond network stabilizes the HLH fold and, thus, probably contributes indirectly to the stability of E47 dimers. The substitution of E47's loop into the MyoD bHLH domain confirmed this hypothesis. The resulting MyoEL dimer is significantly more stable than the wild-type MyoD dimer (Fig. 3; Table I), and, in contrast to MyoD, it readily dimerizes with E47 in the absence of DNA (Fig. 6C).

MyoD tolerates a variety of insertions and nonconservative substitutions in its loop segment (17), suggesting that the loop contributes little to the proper functioning of wild-type MyoD. In contrast, residues in the loop of the inhibitory HLH protein Id1 are important for its dimerization specificity (43). Id1 and related inhibitory HLH proteins lack a functional basic region and thus fail to bind to DNA (44–46). These inhibitory proteins strongly associate with E47 and E12, whereas dimerization with MyoD is less favorable (44, 45, 47). Id proteins probably block myogenic differentiation by sequestering E-proteins as inactive heterodimers, thereby depriving MyoD of a dimerization partner. The weak interaction between Id and MyoD might be explained by the relative instability of MyoD's HLH domain. We find that MyoD must interact with DNA to stably fold and dimerize at low peptide concentrations. Id proteins do not bind to DNA and are therefore disfavored as dimerization partners for MyoD. In contrast, the E47 HLH is a folded dimer at nanomolar peptide concentrations that can readily exchange with other folded bHLH domains (Figs. 3 and 5), including the Id protein. This could be an “Achilles heel” of regulated transcriptional responses, in which stably folded E-proteins promote the folding and DNA binding activities of their unstable partner MyoD. Id exploits the dimerization strength of E47 by sequestering it in inactive E47-Id heterodimers.

Despite the substitution of the loop and several flanking residues from E47 into MyoD, the resulting MyoEL dimer is not as stable as E47, implying that additional residues contribute to the exceptional stability of E47 dimers. One likely candidate is Cys363 of E47, which is a serine in MyoD and MyoEL (Fig. 1). Cysteine 363 is buried in the HLH interface of the E47 dimer, and its substitution with the more polar residue serine is probably disfavored in this hydrophobic environment. Although it is located in the protein interior (16), Cys363 is prone to oxidation during long term storage of the E47 bHLH peptide. It has been suggested that formation of this intersubunit disulfide bond promotes, and may be required for, the DNA binding activity of E47 dimers (48). However, a disulfide link at this position is not compatible with the crystal structure of the E47 dimer. In our experience, oxidation of the E47 bHLH domain decreases its ability to bind DNA, and this deficiency can be overcome by reducing the peptide with 10 mM dithiothreitol (not shown). Under mildly reducing conditions, the E47 and MyoD subunits readily exchange, and they bind to DNA as both homodimers and heterodimers (Figs. 4 and 5). These results suggest that disulfide-linked dimers of E47 adopt a non-native structure that does not bind to DNA. Substitution of Cys363 in E47 with the more polar residue serine alters the dimerization activity and the secondary structure of the bHLH domain (see “Experimental Procedures”). Consistent with this finding, it has been previously reported that an E47 bHLH peptide with the C363S substitution unfolds at significantly higher concentrations (20) than the wild-type E47 domain reported here.

Our finding that the stabilities of various bHLH dimers in solution have little influence on the affinities of the corresponding bHLH-DNA complexes presents an apparent paradox. Although dimerization is required for the stabilization of DNA complexes, the DNA binding equilibrium does not reflect the

---

3 T. Keith Blackwell, personal communication.
amounts of preformed bHLH dimers in solution. One explanation for this behavior is a model in which the rate of association of bHLH proteins with DNA is limited by the rate of folding of the basic region segment and its interaction with DNA, irrespective of the oligomerization state of bHLH subunits. In this model, the initial binding of DNA effectively involves monomers, and dimerization is subsequently required to stabilize the protein-DNA complex once it is formed. Evidence for the transient association of basic region leucine zipper monomers with DNA, prior to stabilization of dimeric bZIP-DNA complexes, has been presented (49, 50). As suggested for the bZIP proteins, electrostatic interactions with the DNA backbone may initially recruit bHLH subunits to the DNA, where a new equilibrium of specifically bound dimers is established. Our finding that specific and nonspecific DNA sequences promote the dimerization of MyoD with E47 supports this model of electrostatic recruitment for bHLH transcription factors. The nonspecific interaction of each bHLH subunit with DNA might shield unfavorable electrostatic interactions of the protein dimer interface, which destabilize MyoD dimers in the absence of DNA. In such a model, the relative stabilities of various bHLH-DNA complexes are determined by their rates of dissociation from DNA, and these rates depend on the particular DNA contacts made by each dimeric combination of subunits. This mechanism for transcriptional regulation does not require a strict control over the cellular levels of various bHLH subunits, because the equilibrium is determined by the DNA rather than the strength of protein-protein interactions.

Acknowledgments—We thank Carl Pabo for the bacterial expression system for the MyoD bHLH; Steve Harrison and Don Wiley for access to their CD spectrometer; Hans Rudolf Bosshard and Christine Berger for providing unpublished data; Keith Blackwell and Andrew Lassar for critical reading of the manuscript; and members of the Ellenberger, Blackwell, and Lassar laboratories for thoughtful suggestions and comments.

REFERENCES
1. Davis, R. L., and Weintraub, H. (1992) Science 256, 1027–1030
2. Weintraub, H., Davis, R. L., Tapscott, S., Thayer, M., Krause, M., Benzeria, R., Blackwell, T. K., Turner, D., Bugg, R., Hollenberg, S., Zhuang, Y., and Lassar, A. (1991) Science 251, 761–766
3. Garrell, J., and Modolell, J. (1990) Cell 61, 39–48
4. Cabrera, C. V., and Alonso, M. C. (1991) EMBO J. 10, 2965–2973
5. Caudy, M., Va¨ssin, H., Brand, M., Tuma, R., Jan, L. Y., and Jan, Y. N. (1988) Cell 55, 1061–1067
6. Hinz, U., Giebel, B., and Campos-Ortega, J. A. (1994) Cell 76, 77–87
7. Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N., and Weintraub, H. (1995) Science 268, 836–844
8. Aphił, P. D., Nakahara, K., Orkin, S. H., and Kirsch, I. R. (1992) EMBO J. 11, 4073–4081
9. Hsu, H. L., Cheng, T. J., Chen, Q., and Baier, R. (1991) Mol. Cell. Biol. 11, 3037–3042
10. Weintraub, H. (1993) Cell 75, 1241–1244
11. Olson, E. N., and Klein, W. H. (1994) Genes Dev. 8, 1–8
12. Murre, C., McCaw, P. S., Vassin, H., Caudy, M., Jan, L. Y., Jan, N. Y., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., Weintraub, H., and Baltimore, D. (1989) Cell 58, 537–544
13. Blackwell, T. K., and Weintraub, H. (1990) Science 250, 1104–1110
14. Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991) Cell 66, 305–315
15. Ma, P. C. M., Rouald, M. A., Weintraub, H., and Pabo, C. O. (1994) Cell 77, 451–459
16. Ellenberger, T., Pass, D., Arnaud, M., and Harrison, S. C. (1994) Genes Dev. 8, 970–980
17. Davis, R. L., Cheng, P. F., Lassar, A. B., and Weintraub, H. (1990) Cell 69, 725–746
18. Sun, X. H., and Baltimore, D. (1991) Cell 64, 459–470
19. Anthony-Cahill, S. J., Benfield, P. A., Fairman, R., Wassermann, Z. R., Brenner, S. L., Stafford, W. F. I., Ailenbach, C., Hubbell, W. L., and DeGrado, W. F. (1992) Science 255, 979–982
20. Fairman, R., Beran-Steed, R. K. Anthony-Cahill, S. J., Lear, J. D., Stafford, W. F. I., DeGrado, W. F., Benfield, P. A., and Brenner, S. L. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10429–10433
21. Studier, F. W., and Moffat, B. A. (1986) J. Mol. Biol. 189, 133–130
22. Sarkar, G., and Sommer, S. S. (1990) Biotechniques 8, 404–407
23. Patel, I. R., Curran, T., and Kerppola, T. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 7760–7764
24. Burns, J. A., Butler, J. C., Moran, J., and Whitesides, G. M. (1991) J. Org. Chem. 56, 2648–2650
25. Chen, Y., Yang, Y. T., and Martinez, H. M. (1972) Biochemistry 11, 4120–4131
26. Carey, J. (1991) Methods Enzymol. 208, 133–117
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., p. 8.23, Cold Spring Laboratory, Cold Spring Harbor, NY
28. Wендт, H. Berger, C., Baici, A., Thomas, R. M., and Bosshard, H. R. (1995) Biochemistry 34, 4097–4107
29. Laue, T. M., Bhaivrav, D. S., Ridgeway, T. M., and Pelletier, S. L. (1992) Analytical Ultra centrifugation in Biochemistry and Polymer Science (Haring, S. E., Rowe, A. J., and Horton, J. C., eds), Royal Society of Chemistry, Cambridge
30. Starosvasnik, M. A., Blackwell, T. K., Laue, T. M., Weintraub, H., and Klevit, R. E. (1995) Biochemistry 34, 6762–6767
31. Pace, C. N. (1986) Methods Enzymol. 131, 266–280
32. Segel, I. H. (1975) Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-state Enzyme Systems, John Wiley and Sons, Inc., New York
33. Marquardt, D. W. (1963) J. Soc. Ind. Appl. Math. 11, 431–441
34. Wendt, H.; Baici, A., and Bosshard, H. R. (1994) J. Am. Chem. Soc. 116, 6973–6974
35. O'Shea, E. K., Ruthkowski, R., and Kim, P. S. (1992) Cell 68, 699–708
36. Maleki, S. J., Royer, C. A., and Hurlburt, B. K. (1997) Biochemistry 36, 6702–6767
37. Shirakata, M., and Paterson, B. R. (1995) EMBO J. 14, 1766–1772
38. Senear, D. F., and Borenwitz, M. (1991) J. Biol. Chem. 266, 13661–13671
39. Meierhans, D., el-Aris, C., Neuenenschwardner, M., Sieber, M., Stackhouse, J. F., and Allemann, R. K. (1995) Biochemistry 34, 11026–11036
40. Murre, C., McCaw, P. S., and Baltimore, D. (1989) Cell 56, 777–783
41. Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) Nature 363, 39–45
42. Ferre-D'Amare, A. R., Pognoone, P., Roeder, R. G., and Burley, S. K. (1994) EMBO J. 13, 180–189
43. Pease, S., and Benzeria, R. (1993) Mol. Cell. Biol. 13, 7874–7880
44. Benzeria, R., Davis, R. L., Lockshorn, D., Turner, D. L., and Weintraub, H. (1990) Cell 61, 49–59
45. Sun, X. H., Copeland, N. G., Jen, Y., Weintraub, H., and Benezra, R. (1992) Genes Dev. 6, 1466–1479
46. Benezra, R. (1994) Cell 79, 1057–1067
47. Metallo, S. J., and Schepartz, A. (1997) Nat. Struct. Biol. 4, 115–117
48. Park, C., Campbell, J. L., and Goddard, W. A. (1996) J. Am. Chem. Soc. 118, 6287–6291