Asymmetric DNA Binding by A Homodimeric bHLH Protein†

Rachel L. Winston,‡,§ Jennifer A. Ehley,‡ Eldon E. Baird,|| Peter B. Dervan,*,|| and Joel M. Gottesfeld*,‡

Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, and Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received April 25, 2000; Revised Manuscript Received June 8, 2000

ABSTRACT: Protein−DNA interactions that lie outside of the core recognition sequence for the *Drosophila* bHLH transcription factor Deadpan (Dpn) were investigated using minor groove binding pyrrole−imidazole polyamides. Electrophoretic mobility shift assays and DNase I footprinting demonstrate that hairpin polyamides bound immediately upstream, but not immediately downstream of the Dpn homodimer selectively inhibit protein−DNA complex formation. Mutation of the Dpn consensus binding site from the asymmetric sequence 5′-CACGCG-3′ to the palindromic sequence 5′-CACGTG-3′ abolishes asymmetric inhibition. A Dpn mutant containing the unnatural amino acid norleucine in place of lysine at position 80 in the bHLH loop region is not inhibited by the polyamide, suggesting that the epsilon amino group at this position is responsible for DNA contacts outside the major groove. We conclude that the nonpalindromic Dpn recognition site imparts binding asymmetry by providing unique contacts to the basic region of each monomer in the bHLH homodimer.

Basic helix-loop-helix (bHLH)[1] transcription factors (*1*) recognize specific hexanucleotide DNA sequences in the major groove. Amino acids in the basic region make base-specific contacts, providing substantial affinity and specificity to the protein−DNA complex. The highly conserved HLH parallel four-helix bundle is responsible for homo- and heterodimerization of this class of proteins (*2*). Structural studies of six different bHLH domains show that the helical regions are nearly superimposable, while the loop regions differ in amino acid content and in length (ranging from 5 to 23 amino acids) and consequently display a large degree of structural variation (*3*−*6*). Although a minimum loop of five amino acids is thought to be necessary to correctly position helices 1 and 2 in the bHLH fold (*5*), longer loop regions may play more than a structural role, by contributing to DNA binding affinity and/or specificity through phosphate backbone (*4*−*6*) or base-specific interactions (*6*). Recently, we identified a single amino acid (Lys 80) in the loop region of the *Drosophila* bHLH transcription factor Deadpan (Dpn) that is essential for high affinity DNA binding (*7*). Biochemical data and homology modeling to other bHLH proteins with known 3D structures (*5*) suggested that this lysine side chain extends across the minor groove adjacent to the Dpn basic region binding site.

To investigate the significance and location of this peptide-loop/minor groove contact, synthetic DNA binding ligands were targeted to the minor groove of the sequences flanking the Dpn core recognition site. For this purpose, we used polyamides containing the aromatic amino acids *N*-methylpyrrole (Py) and *N*-methylimidazole (Im). Because the Py-Im polyamides have been shown to inhibit binding of a variety of eukaryotic transcription factors (*8*−*13*), these molecules can be used as molecular probes for uncovering the nuances of protein−DNA interactions. Hairpin polyamides bind predetermined DNA sequences with subnanomolar affinities comparable to DNA binding proteins (*14*) and recognize the minor groove of DNA through side-by-side amino acid pairings: an Im/Py pair targets G•C, a Py/Im pair targets C•G, and a Py/Py pair binds both A•T and T•A base pairs but not G•C or C•G (*14*−*19*). Using Py-Im polyamides to inhibit Dpn−DNA complex formation, an unexpected mode of protein−DNA recognition was revealed.

MATERIALS AND METHODS

Polyamide Synthesis. Py-Im polyamides **1** and **2** were synthesized by solid-phase methods (*20*) and their purity and identity were established by analytical HPLC, MALDI-TOF MS and ^{1}H NMR.

Solid-Phase Peptide Synthesis. The bHLH domain of Dpn [residues 39−102 from the wild-type sequence (*21*)] was synthesized by solid-phase peptide synthesis methods and characterized as described (*22*). Protein fractions were stored at −80 °C in small aliquots in the following buffer: 20 mM Tris, pH 6.3, 25 mM ammonium sulfate, 100 mM KCl, 1 mM EDTA, 10% glycerol, and 10 mM dithiothreitol.

† This work was supported by National Institutes of Health Grants GM47530 (to J.M.G.) and GM51747 (to P.B.D.). The Howard Hughes Medical Institute provided a predoctoral fellowship for E.E.B.

* To whom correspondence should be addressed. (J.M.G.) Phone: (858) 784-8913. Fax: (858) 784-8965. E-mail: joelg@scripps.edu. (P.B.D.) Phone: (626) 395-6002. Fax: (626) 683-8753. E-mail: dervan@cco.caltech.edu.

‡ The Scripps Research Institute.

§ Present address: Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720.

|| California Institute of Technology.

[1] Abbreviations: bHLH, basic helix-loop-helix; Dpn, Deadpan; EMSA, electrophoretic mobility shift assay; Py, *N*-methylpyrrole; Im, *N*-methylimidazole; bp, base pairs; HPLC, high-pressure liquid chromatography; MALDI-TOF MS, matrix-assisted laser desorption ionization-time-of-flight mass spectrometry; ^{1}H NMR, proton nuclear magnetic resonance; EDTA, ethylenediaminetetraacetic acid; K_d, apparent dissociation constant; WT, wild-type; Lys, lysine; Nle, norleucine; Orn, ornithine; ϵ-NH_2, epsilon amino group.

10.1021/bi000947d CCC: $19.00 © 2000 American Chemical Society
Published on Web 07/15/2000

Electrophoretic Mobility Shift Assays. Sets of complementary oligonucleotides were synthesized (Genosys Biotechnologies, Inc.) for use as double-stranded probes in gel shift assays. Sequences are given in Figure 1B. In addition, the probes contained a four base pair 5′ overhang (5′-TCGA-3′) for subsequent cloning. Additional sets of double stranded oligonucleotides were synthesized containing binding sites for polyamide **1** immediately upstream or downstream of a symmetrical Dpn recognition sequence (5′-CACGTG-3′, where the mutated base is underlined). Complementary single-stranded oligonucleotides were labeled with [γ-^{32}P]-ATP, annealed, gel purified, and resuspended in EMSA buffer (20 mM Hepes, pH 7.6, 100 mM KCl, 1 mM EDTA, 5% glycerol, and 1 mM dithiothreitol). Gel shift assays involving protein titrations were performed as described (*22*), with the following changes for experiments containing polyamide: radiolabeled probe, at a final concentration of 400 pM, was preincubated with polyamide for 15 min, Dpn was added subsequently, and the binding reaction was equilibrated for 30 min before loading the reaction on a 10% nondenaturing polyacrylamide gel. Phosphorimage analysis was used to quantitate the extent of DNA binding by Dpn. For protein titrations, apparent dissociation constants were obtained by fitting plots of fraction DNA bound versus Dpn concentration with the Hill equation using Kaleidagraph software. Hill coefficients of ~2 (ranging from 1.6 to 2.5) were obtained for Dpn binding to each of the double-stranded oligonucleotides. The mean and standard deviation for a given number of determinations (n) are given in the text. For polyamide inhibition experiments, Dpn concentrations that result in approximately 65% of the DNA bound were used. The fraction of DNA remaining bound at each polyamide concentration (normalized to 1.0 for the fraction bound at the particular Dpn concentration used in the experiment) is plotted versus polyamide concentration. Since the concentration of radiolabeled DNA used in these experiments is close to the dissociation constant for the polyamides binding to DNA, true K_i values cannot be obtained under these experimental conditions.

DNase I Footprinting. Unlabeled, gel purified, double-stranded oligonucleotides used in EMSA were phosphorylated with ATP using T4 polynucleotide kinase and cloned into the pBluscript II KS (±) phagemid DNA (Stratagene) digested with the restriction enzyme *Xho*I. The identities of the clones were verified by DNA sequencing. A 328 base pair *Pvu*II/*Xba*I restriction fragment containing the cloned sequence was end-labeled with Sequenase at the 3′ end of the *Xba*I site and gel purified. DNase I digestions were carried out in a 100 μL reaction volume containing approximately 50 pM labeled DNA in EMSA buffer. The DNA was equilibrated with polyamide for 15 min, followed by addition of Dpn and incubation for 30 min at ambient temperature. DNase I digestion was allowed to proceed for 30 s in the presence of 2 mM $CaCl_2$ and 5 mM $MgCl_2$ with 0.3 units of DNase I (Roche Molecular Biochemicals). Reactions were stopped by the addition of 110 μL of stop buffer (1.0% SDS, 0.5 M NaCl, 50 mM Tris, pH 7.6, and 25 mM EDTA), extracted with phenol/chloroform, and precipitated with ethanol using glycogen (20 μg) as a carrier. Samples were denatured by boiling for 5 min in 95% formamide and subjected to electrophoresis on a 6% denaturing polyacrylamide gel containing 7.6 M urea, 89 mM Tris borate, pH 8.3, and 2 mM EDTA. Dried gels were exposed to Kodak Biomax film with DuPont Cronex Lightening Plus intensifying screens at −80 °C. Phosphorimage analysis was used to determine polyamide equilibrium dissociation constants (as described in ref *15*).

RESULTS AND DISCUSSION

Oligonucleotide Design. To probe the DNA binding properties of Dpn, a series of synthetic double-stranded oligonucleotides were constructed based on the high affinity *Drosophila* neural promoter, achaete-scute (denoted WT, Figure 1B). The constructs shown in Figure 1B contain the core recognition sequence 5′-CACGCG-3′ for Dpn, and a binding site for polyamide **1** (Figure 1A). On the basis of the pairing rules for polyamide recognition (see above), polyamide **1** binds the nonpalindromic sequence 5′-WGCWCGW-3′ (where W = A or T) in an N-terminal to C-terminal orientation (*23*). The polyamide **1** site, 5′-AGCTCGT-3′, was placed immediately upstream (U-1) or immediately downstream (D-1) of the Dpn recognition sequence (Figure 1B). Polyamide **1** binds this site with an equilibrium dissociation constant (K_d) of 0.4 nM (as determined by quantitative DNase I footprinting (*15*) using a 328 bp restriction fragment; data not shown). Mismatch polyamide **2** (Figure 1A) contains the same number of Py and Im rings as polyamide **1** but differs in ring sequence, lowering its affinity for the polyamide **1** match site by at least 2 orders of magnitude (*14*). An additional set of oligonucleotides was synthesized with a match binding site for polyamide **2** (5′-AGCTGCT-3′) located immediately upstream (U-2) or immediately downstream (D-2) from the Dpn recognition sequence (Figure 1B). Polyamide **2** binds this match sequence with K_d = 0.1 nM (*9*).

Dpn Binding Affinities. A protein titration was performed on each of the DNA oligonucleotides to determine if mutation of sequences flanking the Dpn binding site influences protein binding affinity. Electrophoretic mobility shift assay (EMSA) was used to determine the K_ds for the binding reaction of the synthetic 64 amino acid bHLH domain of Dpn [residues 39−102 (*22*)] with each of the oligonucleotides shown in Figure 1B. The K_d for Dpn binding to the D-1 oligonucleotide (2.6 ± 0.2 nM, n = 3) is indistinguishable from the K_d obtained for Dpn binding to the WT recognition sequence [2.6 nM (*22*)]; however, we observe a reproducible 50% decrease in binding affinity to the U-1 oligonucleotide (K_d = 3.5 ± 0.2 nM, n = 6). Similarly, we obtained K_ds of 3.8 and 2.6 nM for Dpn binding to the U-2 and D-2 oligonucleotides, respectively. These results indicate that mutation of the sequence upstream of the protein binding site decreases Dpn affinity, whereas changes to the sequence downstream are without affect.

Asymmetric Polyamide Inhibition Of Dpn. The ability of polyamide **1** to inhibit Dpn (39−102) binding was measured by quantitative EMSA. Increasing concentrations of polyamide were incubated with each radiolabeled DNA duplex followed by addition of a constant amount of Dpn (39−102) (Figure 2A). The faster migrating band is free DNA, and the slower migrating band, which decreases in intensity with increasing polyamide concentration, is the Dpn-bound DNA complex. The radioactivity in each band was quantitated, and the fraction bound (normalized to the fraction bound in

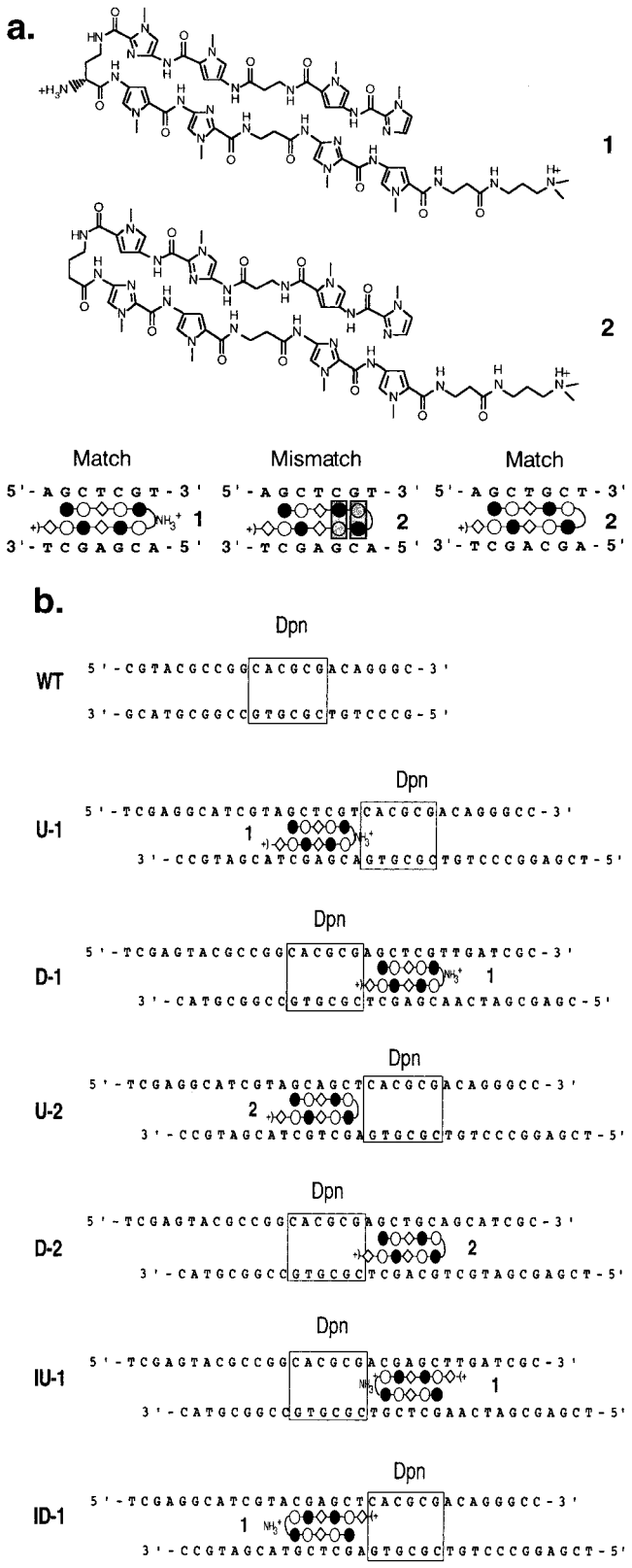

FIGURE 1: Polyamide structures and binding sites. (A) Structures and binding models for polyamides ImPy-β-PyIm-D-PyIm-β-ImPy-β-Dp (**1**) and ImPy-β-ImPy-γ-ImPy-β-ImPy-β-Dp (**2**), where Im = imidazole, Py = pyrrole, β = β-alanine, D = (*R*)-2,4-diaminobutyric acid, γ = γ-aminobutyric acid, and Dp = dimethylaminopropylamide. Shaded and unshaded circles, Im and Py rings, respectively; curved line, γ; curved line with NH_3^+, D; diamonds, β; and semicircle with plus sign, Dp. Note that β·β pairs recognize A·T and T·A base pairs (*17*). Mismatches are highlighted in gray boxes. (B) Oligonucleotide sequences with polyamide binding sites and Dpn recognition sites. Polyamide binding models are shown and Dpn binding sites boxed.

the absence of polyamide) was plotted as a function of polyamide concentration (Figure 2B). Dpn binding to the U-1 and D-1 oligonucleotides was unaffected by the addition of high concentrations (8.6 nM) of the mismatch polyamide **2** (Figure 2B). However, Dpn binding to the U-1 probe was inhibited by 50% at a concentration of 0.6 nM polyamide **1**, and nearly complete inhibition was achieved by 1.4 nM. The observed concentration of polyamide required for 50% inhibition is similar to the K_d for polyamide **1** binding to the U-1 sequence (0.4 nM, see above), suggesting that binding of the polyamide is directly responsible for the inhibition of Dpn. [Note, however, that no techniques are currently available to measure polyamide binding affinities for the double-stranded oligonucleotides utilized in the Dpn inhibition experiments; consequently, binding affinities were determined by quantitative DNase I footprinting (*15*) with the same DNA sequences within longer restriction fragments.] In contrast to the result with the U-1 oligonucleotide, polyamide **1** was not able to significantly inhibit Dpn binding to the D-1 oligonucleotide (Figure 2, panels A and B). With this probe, only 56% inhibition was achieved using 8.6 nM polyamide **1**, and increasing the concentration to 140 nM did not have any further inhibitory affect on Dpn−DNA complex formation. This differential inhibition observed for the upstream and downstream sequences cannot be due to differences in polyamide binding affinities since polyamide **1** binds the D-1 sequence with an affinity (K_d = 0.2 nM, data not shown) that is comparable to the affinity for the U-1 sequence.

To demonstrate that polyamide **1** was indeed binding the designed target site, DNase I footprinting was employed, and Dpn and polyamide contacts on the U-1 and D-1 probes were visualized. Singly end-labeled restriction fragments containing the U-1 or D-1 sequences were equilibrated with polyamide alone, Dpn alone, or polyamide followed by addition of Dpn. As expected, both protein and polyamide bound their respective target sites as evidenced by the distinct footprints (Figure 3, panels A and B). On the U-1 probe, polyamide **1** protected the site 5′-CGTAGCTCGTCACGCGA-3′ (where the underlined sequence corresponds to the cognate binding site), while Dpn specifically protected the site 5′-CTCGTCACGCGACAG-3′ (Figure 3, panels A and C). Nonspecific mismatch sites (located in the adjacent vector sequence; see Figure 3 caption) are also occupied by polyamide **1** since the polyamide concentrations used in this experiment are 10−100-times the apparent dissociation constant for polyamide **1** binding to its match site (adjacent to the Dpn site). The Dpn footprint, characterized by a marked DNase I hypersensitive site, changes to the polyamide footprint in the presence of polyamide **1** and U-1 DNA (Figure 3A, lanes 3−6). This result directly demonstrates that occupancy of the DNA by polyamide **1** displaces Dpn from its DNA site. On the D-1 probe, polyamide **1** protected a smaller site, 5′-CGAGCTCGT-3′, while Dpn protected its binding site 5′-GCCGGCACGCGAGCTCG-3′ (Figure 3, panels B and C). In contrast to the U-1 DNA, increasing concentrations of polyamide **1** did not alter the Dpn footprint on the D-1 probe (Figure 3B, lanes 3−6). In this experiment, the polyamide and protein footprints overlap, making it impossible to distinguish whether the polyamide and protein co-occupy the DNA or the protein is bound alone. Regard-

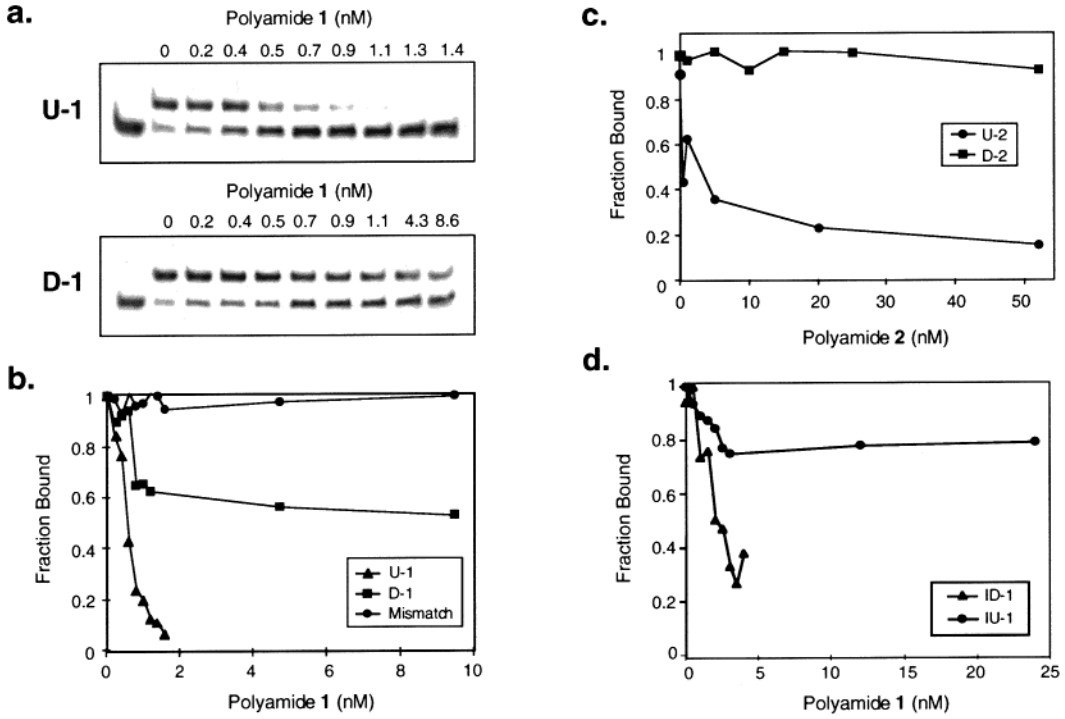

FIGURE 2: Differential inhibition of Dpn−DNA binding. (A) EMSA of Dpn and polyamide **1** bound to U-1 and D-1 probes (Figure 1B). Radiolabeled duplex DNA was equilibrated with increasing concentrations of polyamide **1** as shown, followed by addition of 5.4 or 2.7 nM Dpn, for U-1 and D-1 probes, respectively. In the absence of polyamide, approximately 65% of the DNA is complexed with Dpn at the protein concentrations used in these assays. (B) Graphical representation of EMSA titrations plotted as fraction bound, normalized to the fraction bound in the absence of polyamide, versus polyamide concentration. A mismatch polyamide **2** titration is shown (circles), and polyamide **1** titrations with U-1 and D-1 are indicated with triangles and squares, respectively. (C) Differential inhibition of Dpn−DNA binding by polyamide **2** as determined by EMSA. Radiolabeled duplex U-2 and D-2 probes (Figure 1B) were equilibrated with increasing concentrations of polyamide **2**, followed by addition of a constant amount of Dpn (4.3 and 3.6 nM for the U-2 and D-2 probes, respectively). Graphical representations of EMSA titrations are shown, with circles and squares representing the U-2 and D-2 probes, respectively. (D) Differential Dpn inhibition by polyamide **1** with IU-1 and ID-2 oligonucleotides. Radiolabeled duplex IU-2 and ID-2 probes (Figure 1B) were equilibrated with increasing concentrations of polyamide **1**, followed by addition of a constant amount of Dpn (2.5 and 5 nM for the IU-1 and ID-1 probes, respectively). Graphical representations of EMSA titrations are shown, with circles and triangles representing the IU-1 and ID-1 probes, respectively.

less, this result is consistent with the lack of inhibition observed in EMSA for D-1.

To ensure that the results obtained with polyamide **1** were not artifactual, additional inhibition experiments were performed with polyamide **2** (Figure 1A) and the U-2 and D-2 oligonucleotides (Figure 1B). Asymmetric inhibition of Dpn binding to these constructs was also observed (Figure 2C), reproducing the differential effects of polyamide **1** with the U-1 and D-1 oligonucleotides. Polyamide **2** was found to inhibit Dpn binding to the U-2 oligonucleotide with a with 50% inhibition at 0.5 nM polyamide, while no inhibition of Dpn binding to the D-2 oligonucleotide was observed at polyamide **2** concentrations as high as 50 nM (Figure 2C). As an additional control, we asked whether inversion of the Dpn binding site would cause inhibition downstream instead of upstream of the protein binding site. For this experiment, we synthesized oligonucleotides containing an inverted Dpn binding site with respect to the polyamide binding sites in the U-1 and D-1 oligonucleotides (IU-1 and ID-1, Figure 1B). In this case, the binding affinity of Dpn for ID-1 should be lower than that for IU-1 because the protein will now bind in a "reverse" orientation. Consistent with this hypothesis, we find that the K_ds for ID-1 and IU-1 are 3.7 and 2.5 nM, respectively. This result is consistent with the observation that the sequence upstream of the Dpn binding site (in the wild-type orientation) is important for binding affinity. Additionally, the predicted polyamide inhibition pattern was indeed observed with these oligonucleotides (Figure 2D). Polyamide **1** inhibited Dpn binding to the ID-1 oligonucleotide (50% inhibition at 2 nM) but had little effect on Dpn binding to the IU-1 oligonucleotide, even at a polyamide concentration of 25 nM.

The results with three separate sets of oligonucleotides firmly establish that mutations in the upstream flanking sequence lower Dpn binding affinity, while mutations downstream are without affect. Moreover, polyamides consistently inhibit Dpn binding when targeted 5′ of the core recognition sequence (in the wild-type orientation). Taken together, these observations strongly suggest that Dpn makes an asymmetric DNA contact upstream of its recognition sequence.

Non-Palindromic Binding Site Imparts Asymmetric DNA Binding. The differential polyamide inhibition of Dpn suggests that this protein binds DNA in an asymmetric fashion, despite the fact that it is active as a homodimer (*22*). It is conceivable that the nonpalindromic hexanucleotide binding site 5′-CACGCG-3′ imparts an asymmetry to Dpn by providing a unique chemical environment for each protein monomer. To test this hypothesis, a new set of oligonucleotides containing a single base change (C to T) in the Dpn binding site was constructed to create the palindromic site 5′-CACGTG-3′. As expected, protein titrations with oligonucleotides containing these sequences yielded slightly higher dissociation constants than the wild-type sequence (K_d = 5.5

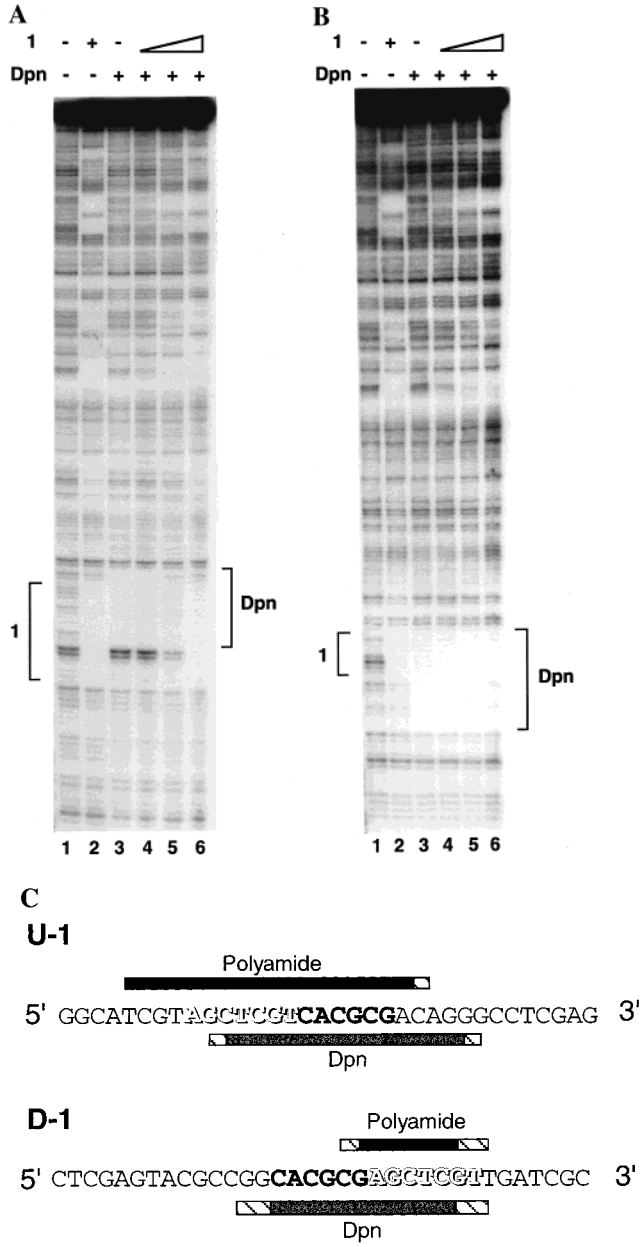

FIGURE 3: DNase I footprint analysis of polyamide inhibition. (A) DNase I footprint of the U-1 probe in the presence of: lane 1, DNA alone; lane 2, polyamide **1** (36 nM); lane 3, Dpn (36 nM); lanes 4−6, Dpn (36 nM) and polyamide **1** at 3.6, 18, or 36 nM, respectively. Specific footprints are indicated with brackets. Additional nonspecific sites of protection were observed in adjacent vector DNA (5′-TGCAGGA-3′, 5′-AGCATAA-3′, and 5′-AGCTAAC-3′) since these polyamide concentrations range from 10- to 100-fold higher than the K_d for polyamide **1** binding its match site. (B) DNase I footprint of the D-1 probe in the presence of lane 1, DNA alone; lane 2, polyamide **1** (36 nM); lane 3, Dpn (36 nM); lanes 4−6, Dpn (36 nM) and polyamide **1** at 3.6, 18, or 36 nM, respectively. Specific footprints are indicated with brackets. (C) Schematic representation of footprinting results. Only the top strand DNA sequence is shown. The Dpn recognition sequence is shown in bold; the polyamide binding site is highlighted; and, the footprints are indicated with solid boxes above and below the sequence, along with possible additional sites of protection in hatched boxes.

nM and 3.0 nM for U-1-Pal and D-1-Pal, respectively; data not shown). Polyamide titrations were performed on these oligonucleotides, and a graphical representation of the EMSA results is shown in Figure 4A. Strikingly, polyamide **1** displays a similar inhibition profile when the polyamide is targeted upstream or downstream of the Dpn binding site. These results suggest that changing the protein recognition sequence alters the protein contacts to sequences flanking the binding site. Additionally, these results support our hypothesis that the nonpalindromic binding site induces asymmetric Dpn binding.

Role of Dpn Lys 80 in Polyamide Inhibition. In a previous study, we showed that Lys 80, a residue which resides at the C-terminal end of the loop region, contributes −1.3 kcal/mol in DNA binding affinity, presumably via a peptide−DNA phosphate contact (*7*). Characterization of a mutant peptide containing the unnatural amino acid norleucine at position 80 (Dpn Nle 80) showed that the epsilon amino group of Lys 80 (Lys 80-ϵNH_2) is directly responsible for this binding affinity. To further investigate the role of this lysine in DNA binding, polyamide inhibition experiments were performed with Dpn Nle 80 (Figure 4, panels B and C). In the presence of a constant amount of protein and radiolabeled DNA (U-1), increasing the concentration of polyamide **1** had little effect on the Dpn Nle 80−DNA complex. Even at high concentrations (143 nM), polyamide **1** did not inhibit this mutant peptide from binding the U-1 probe. In contrast, a Dpn mutant peptide containing ornithine at position 80 (Dpn Orn 80), which retains a terminal amine but has a shortened alkyl chain, displayed 50% inhibition at a polyamide concentration of 0.7 nM (Figure 4, panels B and C). Therefore, polyamide **1** is only inhibitory to Dpn peptides that retain a terminal amine at position 80. Taken together, these results suggest that polyamide **1** prevents Lys 80-ϵNH_2 of WT-Dpn from contacting DNA.

If this conclusion is true, then the K_d for the WT-Dpn/U-1 DNA interaction in the presence of polyamide should be similar to the K_d for Dpn Nle 80−DNA alone. To test this hypothesis, EMSA was used to assay WT-Dpn activity in the presence of a saturating concentration of polyamide **1** (4 nM) and the U-1 or D-1 oligonucleotides (Figure 4D). WT Dpn bound D-1 with similar affinity in the presence (Figure 4D) or absence (see above) of a constant amount of polyamide **1**, corroborating previous results that demonstrated a lack of inhibition for the Dpn/D-1 DNA complex (Figures 2, panels A and B, and 3B). In contrast, WT Dpn bound U-1 DNA with a K_d of 20 ± 0.7 nM in the presence of polyamide **1**, an affinity comparable to that of the Dpn Nle 80−DNA complex alone (*7*). In addition, the DNA binding specificity of Dpn under these conditions is clearly compromised, as evidenced by the nonspecific band observed on the gel (indicated by an arrow). Interestingly, Dpn Nle 80 also binds DNA with a low degree of specificity (*7*). Thus, polyamide bound upstream of the Dpn recognition sequence induces binding behavior (affinity and specificity) in WT-Dpn that is similar to that of Dpn Nle 80. These results are consistent with a model where polyamide **1** inhibits Dpn−DNA interactions by interfering with the same contact made by Lys 80-ϵNH_2.

DNase I footprinting was performed with Dpn Nle 80 to corroborate the EMSA results. Under identical conditions used for WT-Dpn, Dpn Nle 80 did not yield full protection of the Dpn site, even at peptide concentrations as high as 4 μM. At these high polyamide concentrations, a partial footprint was observed at the Dpn site along with partial protection of additional sites along the DNA probe (data not shown). We measured the half-life for this complex by

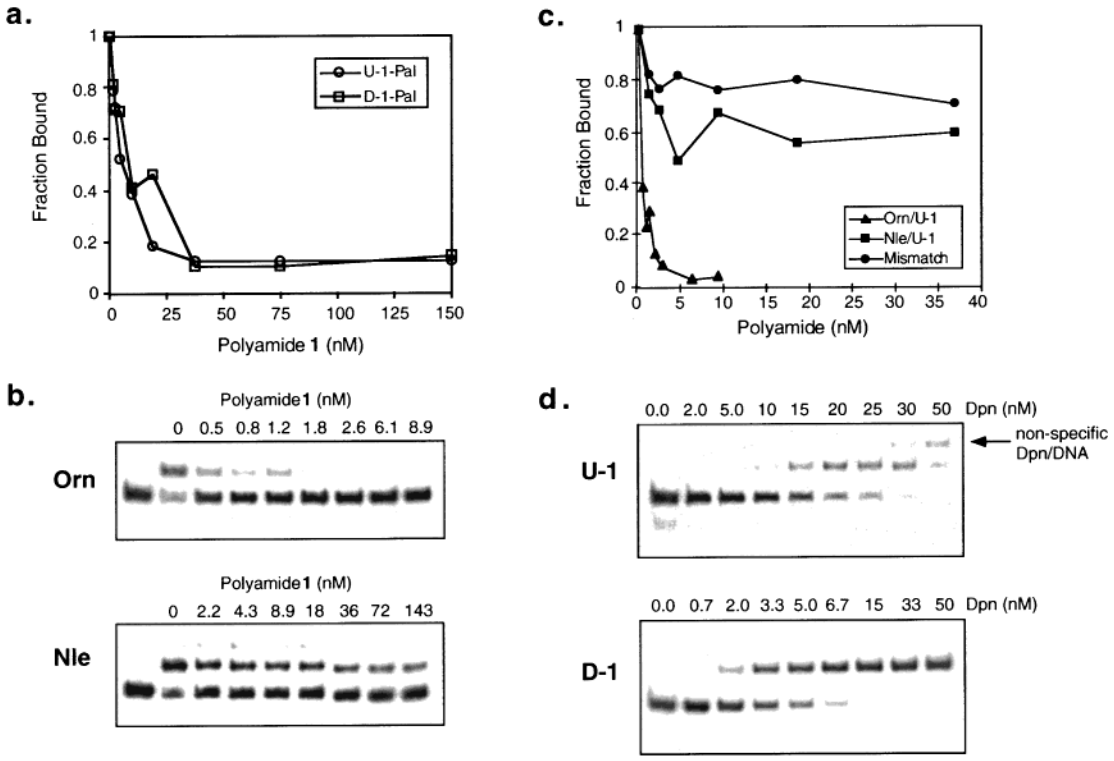

FIGURE 4: Effects of polyamides with altered DNA sequences and mutant Dpn peptides. (A) Equivalent polyamide **1** inhibition of Dpn−DNA complex formation with the palindromic Dpn binding site. Graphical representation of EMSA with U-1-Pal (open circles) and D-1-Pal (open squares) are plotted as fraction bound, normalized to the fraction bound in the absence of polyamide, versus polyamide concentration. Polyamide inhibition at 50% is achieved with both probes at 8.6 nM polyamide **1**. (B) Polyamide inhibition of Dpn mutants. EMSA with Dpn Nle 80 or Dpn Orn 80 and polyamide **1** bound to U-1. Radiolabeled duplex DNA was equilibrated with increasing concentrations of polyamide **1** as shown, followed by addition of 30 nM Dpn Nle 80 or 4.6 nM Dpn Orn 80. (C) Graphical representation of EMSA titrations plotted as fraction bound, normalized to the fraction bound in the absence of polyamide, versus polyamide concentration. A mismatch polyamide **2** titration is shown (circles), and polyamide **1** titrations with Dpn Orn 80 and Dpn Nle 80 are shown (triangles and squares, respectively). (D) EMSA of Dpn titration in the presence of saturating amounts of polyamide **1** (4 nM) in the presence of U-1 or D-1 probes. K_ds of 20 ± 0.7 and 3.5 ± 0.7 nM were obtained for the U-1 and D-1 probes, respectively, from plots of fraction DNA bound versus protein concentration (not shown). Nonspecific binding to the U-1 probe is indicated with an arrow.

EMSA to explore the possibility that the Dpn Nle 80−DNA complex was too transient to detect in DNase I footprinting assays. We found that the half-life was approximately 200 min for WT-Dpn, and less than 15 s for Dpn Nle 80 (data not shown). These results are consistent with the failure to observe full binding for Dpn Nle 80 under the conditions used for WT protein.

CONCLUSIONS

On the basis of our Py-Im polyamide inhibition results, we suggest that the nonpalindromic Dpn recognition site imparts binding asymmetry by providing unique contacts to the basic region of each monomer in the bHLH homodimer. These differential protein monomer−DNA interactions could uniquely position the loop region upstream of the Dpn consensus site, such that the side chain of Lys 80 stretches across the minor groove to make a specific phosphate contact (Figure 5). Alternatively, the nonpalindromic DNA site may impart an asymmetry to the DNA structure itself, so that only the phosphate backbone upstream of the binding site is accessible for interaction with the loop region of Dpn. In either model, placement of a polyamide (shown in yellow in Figure 5) in the minor groove upstream of the recognition sequence sterically blocks the crucial Lys 80 backbone contact or alters the DNA structure in a manner such that this contact is no longer possible. The Lys 80−DNA contact also requires an optimal DNA sequence context as mutations

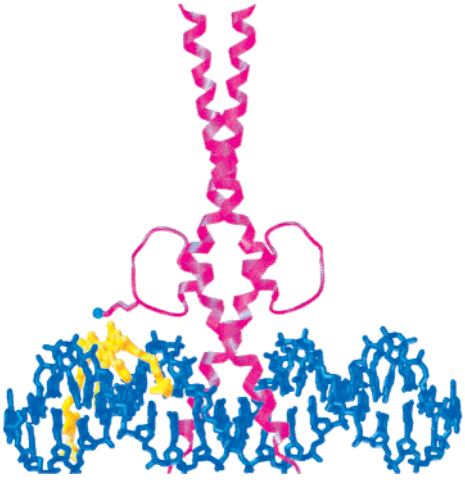

FIGURE 5: Model for Dpn−DNA recognition in the presence of polyamide. A model of Dpn, based on the Max crystal structure (*5*), is shown in ribbon form in pink. The alkyl chain of Lys 80 in one monomer is highlighted in pink and the epsilon amino group is shown in blue. A model of DNA-bound polyamide (*18*) is shown in yellow adjacent to the Dpn recognition sequence. X-fit (*28*) was used to dock the polyamide structure onto the Max-DNA structure and an AVS (*29*) representation is shown.

upstream from the Dpn recognition element clearly affect Dpn binding affinity, but mutations downstream are without effect. For bHLH proteins that bind palindromic sites [such as the Max protein (*5*)] the analogous lysine residue interacts

differently with the DNA backbone, making symmetric contacts with both sides of the core recognition sequence across the minor groove. In this case, a polyamide targeted *either* upstream or downstream of the recognition site will inhibit DNA binding (Figure 4A). Although there are examples of homodimeric proteins binding asymmetrically to DNA (*24*−*26*), there is also an example of a homodimer binding symmetrically to an asymmetric DNA site (*27*). Those examples show that it is not possible to know a priori how a given protein will interact with its target sequence. In contrast to the X-ray structural studies described above, the experiments presented here take advantage of a rapid chemical method employing sequence-specific DNA ligands to uncover molecular details of a protein−DNA interaction.

REFERENCES

1. Littlewood, T., and Evan, G. I. (1998) *Helix-loop-helix transcription factors*, Oxford University Press, New York.
2. Anthony-Cahill, S. J., Benfield, P. A., Fairman, R., Wasserman, Z. R., Brenner, S. L., Stafford, W. F., III, Altenbach, C., Hubbell, W. L., and DeGrado, W. F. (1992) *Science 255*, 979−983.
3. Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S. C. (1994) *Genes Dev. 8*, 970−980.
4. Ma, P. C. M., Rould, M. A., Weintraub, H., and Pabo, C. O. (1994) *Cell 77*, 451−459.
5. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature 363*, 38−45.
6. Ferré-D'Amaré, A. R., Pognonec, P., Roeder, R. G., and Burley, S. K. (1994) *EMBO J. 13*, 180−189.
7. Winston, R. L., and Gottesfeld, J. M. (2000) *Chem. Biol. 7*, 245−251.
8. Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E., and Dervan, P. B. (1997) *Nature 387*, 202−205.
9. Dickinson, L. A., Gulizia, R. J., Trauger, J. W., Baird, E. E., Mosier, D. E., Gottesfeld, J. M., and Dervan, P. B. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 12890−12895.
10. Dickinson, L. A., Trauger, J. W., Baird, E. E., Dervan, P. B., Graves, B. J., and Gottesfeld, J. M. (1999) *J. Biol. Chem. 274*, 12765−12773.
11. Dickinson, L. A., Trauger, J. W., Baird, E. E., Ghazal, P., Dervan, P. B., and Gottesfeld, J. M. (1999) *Biochemistry 38*, 10801−10807.
12. McBryant, S. J., Baird, E. E., Trauger, J. W., Dervan, P. B., and Gottesfeld, J. M. (1999) *J. Mol. Biol 286*, 973−981.
13. Bremer, R. E., Baird, E. E., and Dervan, P. B. (1998) *Chem. Biol. 5*, 119−133.
14. Trauger, J. W., Baird, E. E., and Dervan, P. B. (1996) *Nature 382*, 559−561.
15. White, S., Baird, E. E., and Dervan, P. B. (1997) *Chem. Biol. 4*, 569−578.
16. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) *Nature 391*, 468−471.
17. Turner, J. M., Swalley, S. E., Baird, E. E., and Dervan, P. B. (1998) *J. Am. Chem. Soc. 120*, 6219−6226.
18. Kielkopf, C. L., Baird, E. E., Dervan, P. B., and Rees, D. C. (1998) *Nat. Struct. Biol. 5*, 104−109.
19. Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., and Rees, D. C. (1998) *Science 282*, 111−115.
20. Baird, E. E., and Dervan, P. B. (1996) *J. Am. Chem. Soc. 118*, 6141−6146.
21. Bier, E., Vaessin, H., Younger-Shepherd, S., Jan, L. Y., and Jan, Y. N. (1992) *Genes Dev. 6*, 2137−2151.
22. Winston, R. L., Millar, D. P., Gottesfeld, J. M., and Kent, S. B. H. (1999) *Biochemistry 38*, 5138−5146.
23. Herman, D. M., Baird, E. E., and Dervan, P. B. (1998) *J. Am. Chem. Soc. 120*, 1382−1391.
24. Parraga, A., Bellsolell, L., Ferre-D'Amare, A. R., and Burley, S. K. (1998) *Structure 6*, 661−672.
25. Beamer, L. J., and Pabo, C. O. (1992) *J. Mol. Biol. 227*, 177−196.
26. Sarai, A., and Takeda, Y. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6513−6517.
27. Albright, R. A., and Matthews, B. W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3431−3436.
28. McRee, D. E. (1999) *J. Struct. Biol. 125*, 156−165.
29. Upson, C., Faulhabe, T., Kamins, D., Laidlaw, D., Schlegel, D., Vroom, J., Gurwitz, R., and Vandam, A. (1989) *IEEE Comput. Graph. Appl. 9*, 30−42.

BI000947D