Characterization of the DNA Binding Properties of the bHLH Domain of Deadpan to Single and Tandem Sites

Rachel L. Winston,* David P. Millar, Joel M. Gottesfeld, and Stephen B. H. Kent‡

Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037

Received December 3, 1998; Revised Manuscript Received February 17, 1999

ABSTRACT: The basic helix−loop−helix domain of the *Drosophila* transcription factor Deadpan (Dpn) was prepared by total chemical protein synthesis in order to characterize its DNA binding properties. Circular dichroism spectroscopy was used to correlate structural changes in Dpn with physiologically relevant monovalent (KCl) and divalent ($MgCl_2$) cation concentrations. In addition, we have used electrophoretic mobility shift assay (EMSA) and fluorescence anisotropy methods to determine equilibrium dissociation constants for the interaction of Dpn with two biologically relevant promoters involved in neural development and sex determination pathways. In this study, we have optimized DNA binding conditions for Dpn, and we have found a markedly higher DNA binding affinity for Dpn than reported for other bHLH domain transcription factors. Dpn binds as a homodimer (K_d = 2.6 nM) to double-stranded oligonucleotides containing the binding site CACGCG. In addition, we found that Dpn bound with the same affinity to a single or tandem binding site, indicating no cooperativity between adjacent DNA-bound Dpn dimers. DNA binding was also monitored as a function of physiologically relevant KCl and $MgCl_2$ concentrations, and we found that this activity was significantly different in the presence and absence of the nonspecific competitor poly(dI-dC). Moreover, Dpn displayed moderate sequence selectivity, exhibiting a 100-fold higher binding affinity for specific DNA than for poly(dI-dC). This study constitutes the first detailed biophysical characterization of the DNA binding properties of a bHLH protein.

Basic helix−loop−helix (bHLH)[1] proteins comprise a family of transcriptional regulators involved in a number of developmental pathways, including neurogenesis, myogenesis, sex determination, and cell differentiation (*1*). Mutagenesis studies have shown that the HLH regions of these proteins are essential for dimerization (*2*), while the basic region makes base-specific DNA contacts to a consensus hexanucleotide sequence (*3*). Recently, crystal structures of several bHLH proteins have been solved (*4*−*8*), and the X-ray crystallographic data confirmed initial predictions that the HLH domain forms two amphipathic α-helices that dimerize to form a parallel four-helix bundle (*9*−*11*).

The *Drosophila* bHLH protein, Deadpan (Dpn), belongs to the Hairy-related class of bHLH transcription factors, which include Hairy and members of the Enhancer-of-split complex (*12*−*17*). These proteins have been shown to act as transcriptional repressors (*15*, *18*−*25*), and to bind to a consensus hexanucleotide DNA sequence that is distinct from the E-box motif (CANNTG) recognized by other bHLH transcription factors. Specifically, the *Drosophila* E(spl) m8 protein (*18*) and mammalian HES-1 and HES-5 proteins (*15*, *16*) were shown to bind hexamer sequences CAC(G/A)AG (called the N-box) in DNase I footprint analyses. Hairy, a protein with 82% identity to Dpn, was shown to preferentially bind the sequence CACGCG (*23*, *24*). Genetic studies have indicated that Dpn, along with other factors, plays an early role in directing nervous system formation by repressing expression of genes promoting nonneuronal fates (*26*). Furthermore, loss of Dpn function in *Drosophila* results in weak motor activities, lethargic behavior, and death (*14*). While the molecular basis for these genetic findings has not been determined, a potential Dpn binding site exists within the promoter of the proneural achaete−scute gene complex (AS−C) (*27*). Moreover, ectopic expression of *dpn* causes reduced AS−C activities, suggesting a negative interaction between *dpn* and the AS−C genes (*14*).

Genetic studies have also shown that *dpn* plays a role in *Drosophila* sex determination, by acting as a denominator element in determining the X-chromosome-to-autosome ratio (*28*). Dpn appears to function as a negative regulator of *sex-lethal (sxl)* (*28*), the master regulatory gene in *Drosophila* sex determination, by binding a unique site within the *sxl* promoter (*20*). This site contains two Dpn binding sites, in which the center of each site is separated by approximately one helical turn of DNA. The significance of this tandem binding site in the regulation of *sxl* transcription has yet to be elucidated.

In this study, total chemical synthesis was used to prepare the bHLH region of Dpn, and electrophoretic mobility shift assay (EMSA) and fluorescence anisotropy techniques were used to measure the dissociation constants (K_d) for Dpn binding to sites within the AS−C (single site) and *sxl*

* To whom correspondence should be addressed.

‡ Present address: Gryphon Sciences, South San Francisco, CA 94080.

[1] Abbreviations: bHLH, basic helix−loop−helix domain; Dpn, Deadpan; EMSA, electrophoretic mobility shift assay; SPPS, solid-phase peptide synthesis; RP-HPLC, reversed-phase high-pressure liquid chromatography; TFA, trifluoroacetic acid; ESI-MS, electrospray ionization mass spectrometry; EDTA, ethylenediaminetetraacetic acid; BSA, bovine serum albumin.

10.1021/bi982856a CCC: $18.00
Published on Web 03/30/1999

(tandem site) promoters. Optimized DNA binding conditions were found, yielding a K_d value 2 orders of magnitude lower than values reported for other bHLH proteins. Dpn displayed a significantly different monovalent and divalent ion dependence in the presence or absence of the nonspecific competitor poly(dI-dC). Moreover, the bHLH domain of Dpn showed approximately 100-fold binding selectivity for specific DNA over poly(dI-dC). The binding of the bHLH domain of Dpn to single and tandem sites was also investigated in order to evaluate the potential significance, if any, the tandem site has for transcriptional regulation of *sxl*.

MATERIALS AND METHODS

Synthesis and Purification of the bHLH Region of Dpn. Dpn [amino acids 39−102 (*14*)] was manually synthesized using stepwise solid-phase peptide synthesis (SPPS), according to published in situ neutralization Boc chemistry protocols (*29*). The isostere aminobutyric acid was substituted for cysteine 59, and the N-terminus was acetylated. The crude synthetic product polypeptide was purified by semipreparative (1 × 25 cm) reversed-phase high-pressure liquid chromatography (RP-HPLC) using a gradient of 30−50% buffer B (90% acetonitrile, 0.1% TFA in H_2O) over 40 min at a flow rate of 3 mL/min. Fractions were analyzed by electrospray ionization mass spectrometry (PE Sciex API-III), and those with the correct mass were pooled and lyophilized. The overall yield of pure synthetic protein from a single synthesis was 50 mg (ca. 10% based on crude product).

Protein Folding. Pure lyophilized material was dissolved in 6 M guanidine hydrochloride, 100 mM sodium acetate, pH 7.0, and 10 mM dithiothreitol (DTT), and diluted with water to 1 M guanidine. Chromatography with a Pharmacia PD-10 column was used to exchange the protein into the following storage buffer for EMSA and fluorescence anisotropy experiments: 20 mM Tris, pH 6.3, 25 mM ammonium sulfate, 100 mM KCl, 1 mM EDTA, 10% glycerol, and 10 mM DTT. Protein fractions were collected, pooled, and stored at −80 °C in small aliquots. Protein used for CD studies was exchanged into 20 mM Hepes, pH 7.6, 1 mM EDTA. All protein concentrations were determined by amino acid analysis by AAA labs (Seattle, WA).

Circular Dichroism (CD) Spectroscopy. CD spectra of Dpn were measured in 20 mM Hepes, pH 7.6, 1 mM EDTA at a concentration of 33 μM (KCl titrations) or 3 μM (for Dpn−DNA interaction studies) using an AVIV 202SF instrument. CD spectra were recorded from 190 to 260 nm at 25 °C using a quartz cuvette with a 0.1 cm path length. The results are presented as a plot of mean molar ellipticity per residue ($[\theta]$, in deg cm^2 dmol^{-1}) versus wavelength in 0.5 nm increments. The percent helicity in each sample was estimated from $[\theta]_{222}$ (*30*).

Sedimentation Equilibrium. Sedimentation equilibrium was carried out on a 10 μM solution of Dpn in 20 mM Hepes, pH 7.6, 1 mM EDTA, 0.5 mM DTT buffer with or without 100 mM KCl at 25 °C in a Beckman Optima XL-I analytical ultracentrifuge. The protein was dialyzed for 18 h against the aforementioned buffer to obtain an accurately "matched buffer" for reference. The sample was spun at a rotor speed of 30 000 rpm for 24 h, and the absorbance of protein was monitored at a wavelength of 230 nm as a function of the cell radius. Equilibrium was assessed by overlaying five consecutive scans taken 1 h apart. A partial specific volume ($\bar{\nu}$) was calculated to be 0.75 mL/g based on amino acid composition, and the data were fitted to an ideal single species model or an association model using nonlinear least-squares analysis with the Origin software provided by Beckman. The residuals were used to assess the "goodness of fit."

Electrophoretic Mobility Shift Assays (EMSA). The oligonucleotides used for gel shift assays were purchased from Genosys Biotechnologies, Inc. The sequence of the top strand single site oligonucleotide from the AS−C proximal promoter (*27*) is 5′-CGTACGCCGGCACGCGACAGGGC-3′, where the underlined sequence is the Dpn binding site. The sequence of the top strand tandem site oligonucleotide from the *sex lethal* promoter (*20*) is 5′-AGCCCACGCGACTGGCACGCGCACC-3′. Complementary single-stranded oligonucleotides were labeled with [γ-^{32}P]ATP, annealed, gel purified, and resuspended in assay buffer (20 mM Hepes, pH 7.6, 100 mM KCl, 1 mM DTT, 1 mM EDTA, and 5% glycerol). For protein titrations, serial dilutions of a protein stock solution (107 μM) were incubated in assay buffer with labeled oligonucleotide (ca. 40 pM) and 2.5 μg/μL bovine serum albumin (BSA) for 15 min in 20 μL aliquots. Samples were loaded onto a 10% nondenaturing polyacrylamide BioRad minigel (0.75 mm × 6.5 cm) and run for 25 min at 150 V/cm with 88 mM Tris−borate buffer, pH 8.3. Radioactivity in a given band was quantitated with a Molecular Dynamics PhosphorImager system using ImageQuant software. Quantitated data were plotted as fraction bound versus protein concentration, and fit with the Hill equation: fraction bound $= [\text{protein}]^n/([\text{protein}]^n + K_d^n)$, where n is the Hill coefficient and K_d is the apparent dissociation constant.

Fluorescence Anisotropy Assays. Fluorescence anisotropy titration measurements were carried out using a Spectronic 8100 Series 2 spectrofluorometer. The sequences of the top strand oligonucleotides used in the single and tandem site binding titrations, respectively, are 5′-F-TCGCCACGCGACC-3′ and 5′-F-TGCCCACGCGACTGGCACGCGCACC-3′, where F denotes fluorescein and the underlined sequence is the Dpn binding site. The sequences of the top strand oligonucleotides for the mis 1 and mis 2 sequences used to determine the intermediate dimer-bound species to the tandem site are 5′-F-TGCCGACCCGACTGGCACGCGCACC-3′ and 5′-F-TGCCCACGCGACTGGACGCCGCACC, respectively. A 1.5-fold excess of unlabeled, complementary strand was annealed to each strand to ensure fully hybridized labeled DNA. In all titrations, either a 1.0 or 0.5 μL aliquot of protein solution (400 nM, 1 μM, or 10 μM) was added to a 5 nM solution of DNA (200 μL) containing 20 mM Hepes, pH 7.6, 100 mM KCl, 1 mM EDTA, and 0.13 μg/μL BSA. The BSA was important for obtaining reproducible data at these low DNA concentrations, probably because it prevented the DNA and/or protein from adhering nonspecifically to the cuvette. Parallel and perpendicular polarization components of the fluorescein emission were measured in a T format by excitation at 480 nm and detection at 520 nm. The anisotropy was measured 30 times at each titration point with an integration time of 2.5 s for each intensity measurement; the resulting values were averaged and plotted against protein concentration.

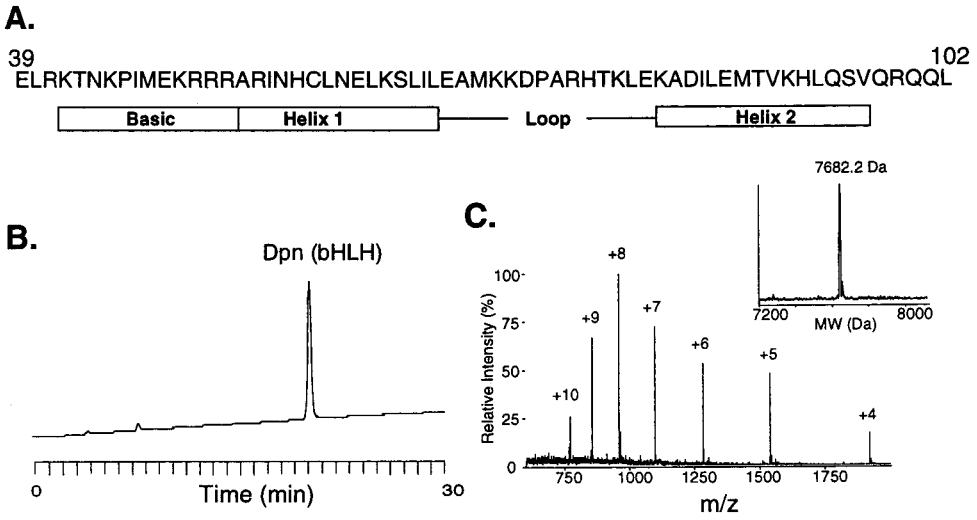

FIGURE 1: (A) Amino acid sequence of the bHLH domain of Dpn. (B) Analytical reversed-phase HPLC of purified synthetic bHLH Dpn, using a gradient of 0−72% buffer B over 30 min (see Materials and Methods). (C) Electrospray mass spectrum of the HPLC-purified product. The observed mass was 7682.3 ± 0.5 Da; the calculated mass was 7682.2 Da (average isotopic composition).

Analysis of Fluorescence Anisotropy Data. Binding curves generated from the increase in fluorescence anisotropy values as a function of protein concentration were fit to a model of dimer ⇌ dimer−DNA (one-state) or dimer ⇌ dimer−DNA ⇌ tetramer−DNA (two-state) equilibrium, using a numerically based algorithm (BIOEQS) (*31*, *32*). BIOEQS solves directly for ΔG values of the complexes based on protein concentration, anisotropy, and quantum yields. The one-state binding model contained a fixed anisotropy value for the free DNA species (empirically determined), and floating values for the ΔG and anisotropy of the fully bound protein−DNA species. The two-state binding model contained fixed anisotropy values for free DNA and dimer-bound DNA (empirically determined), and floating values for the ΔG and anisotropy of the tetramer-bound DNA species, as well as the ΔG value for the dimer-bound DNA species. Rigorous confidence interval testing was used to estimate the uncertainty in each floated parameter. The reported uncertainties in the ΔG values correspond to a 95% confidence interval for an F-statistical distribution. The equilibrium dissociation constants were calculated from the ΔG values using the relation $\Delta G = -RT \ln K_a$.

DNA Binding Specificity. Multiple protein titrations were performed in the presence of varying concentrations of the nonspecific competitor poly(dI-dC) (0, 20, 200, 300, 400, and 600 ng) using EMSA. The apparent K_d value for these protein titrations was determined for each poly(dI-dC) concentration as before (see EMSA section). The concentration of poly(dI-dC) was converted to base pair concentration (in μM) using 660 μg/μmol as a molar conversion factor and 20 μL as a final reaction volume. The K_d (μM) was plotted as a function of poly(dI-dC) concentration (μM), and the inverse slope was used to calculate the ratio of specific to nonspecific binding. The following equilibrium equations were established to derive this ratio:

$$A \overset{K_1}{\rightleftharpoons} S + P + N \overset{K_2}{\rightleftharpoons} B \qquad (1)$$

where A is the specific protein−DNA complex, S is the specific DNA, P is the protein, N is the nonspecific DNA, and B is the nonspecific protein−DNA complex. When the fraction of bound specific DNA (F_b) = 0.5, we make the following assumptions:

$$K_1 = [P] \qquad (2)$$

$$K_{1(\text{apparent})} = K_1 + B \qquad (3)$$

$$B = [P][N]/K_2 \qquad (4)$$

By substituting [P] in eq 4 for K_1 from eq 2, and replacing B from eq 3 with eq 4, the following equation is obtained:

$$K_{1(\text{apparent})} = K_1 + (K_1/K_2)[N] \qquad (5)$$

where K_1 is the intercept and K_1/K_2 is the slope, or ratio of specific to nonspecific binding. $K_{1(\text{apparent})}$ is the apparent K_d value measured for specific DNA binding, while [N] represents the concentration of poly(dI-dC).

RESULTS

Chemical Synthesis and Characterization. The bHLH domain of Dpn (Figure 1A), residues 39−102 from the wild-type sequence (*14*), was synthesized by highly optimized, manual stepwise solid-phase peptide synthesis (*29*). The crude synthetic product was purified by RP-HPLC (Figure 1B) and analyzed by ESI-MS (Figure 1C). The observed mass (7682.3 ± 0.5 Da) was identical within experimental uncertainty to the calculated mass (7682.2 Da, based on the average isotopic composition). The purified polypeptide chain was renatured from 6 M guanidine hydrochloride by size-exclusion chromatography into assay buffer (pH 7.6) for further characterization (see Materials and Methods).

CD Spectroscopy. To characterize the structural properties of Dpn, UV−CD spectroscopy was performed under a physiologically relevant range of mono- and divalent cation concentrations. Increasing amounts of KCl were added to a 33 μM solution of Dpn in 20 mM Hepes buffer, pH 7.6, at 25 °C. The corresponding spectra for each KCl concentration are shown in Figure 2A. Based on the molar ellipticity value at 222 nm (θ_{222}), the percent helicity of Dpn was calculated (*30*) to be 13% in the absence of KCl. In addition, there was a significant amount of random coil as evidenced by

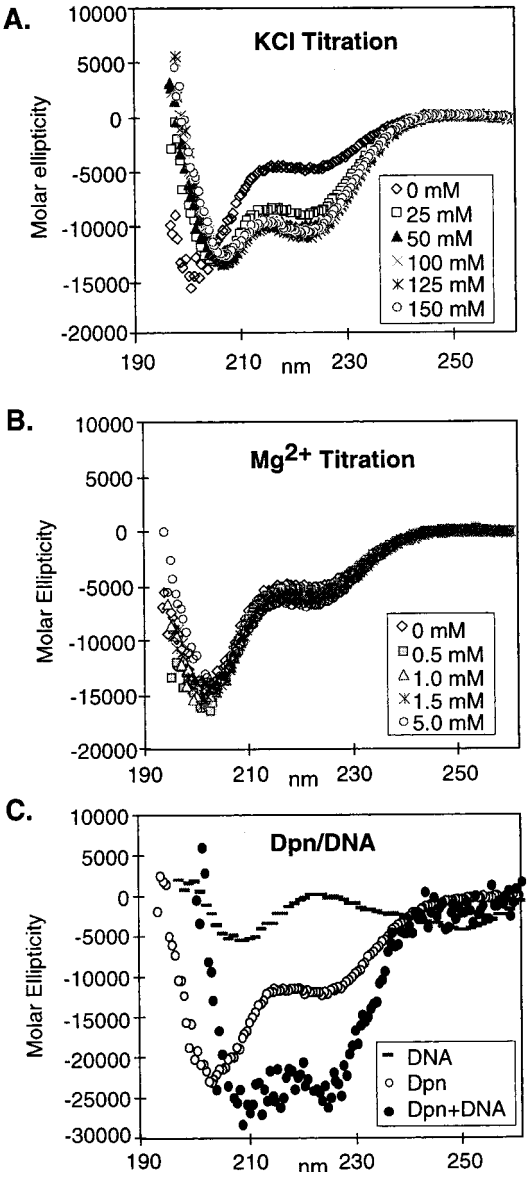

FIGURE 2: (A) UV−CD spectra for Dpn in 20 mM Hepes, pH 7.6, 1 mM EDTA buffer containing the indicated KCl concentrations. (B) UV−CD spectra for Dpn in 20 mM Hepes, pH 7.6, 1 mM EDTA buffer containing the indicated $MgCl_2$ concentrations. No significant increase in helicity was recorded. (C) UV−CD spectra for DNA alone (−), Dpn alone (○), and Dpn in the presence of equimolar 24 bp oligonucleotide from the AS−C promoter (●). DNA contributions to the spectrum have been subtracted.

the large degree of molar ellipticity at 200 nm compared to 222 nm. However, addition of KCl significantly increased the helical content of Dpn. At 25 mM KCl, the percent helicity doubled to 26%, and the minimum molar ellipticity value shifted from 200 to 205 nm, indicating some loss of random coil. Moreover, the ratio of molar ellipticity values at 222 and 200 nm ($\theta_{222}/\theta_{200}$) increased from 0.22 in the absence of KCl to 0.74 with the addition of 50 mM KCl, suggesting a coiled-coil interaction. Increasing the KCl concentration to 150 mM had little effect on the molar ellipticity values, edging the percent helicity up to 28%, and $\theta_{222}/\theta_{200}$ to 0.80.

In a separate titration, CD spectroscopy was used to assay the effect of $MgCl_2$ on the helical content of Dpn. In this assay, a physiologically relevant concentration of $MgCl_2$ was added to a 33 μM Dpn solution ranging from 0.5 to 5 mM;

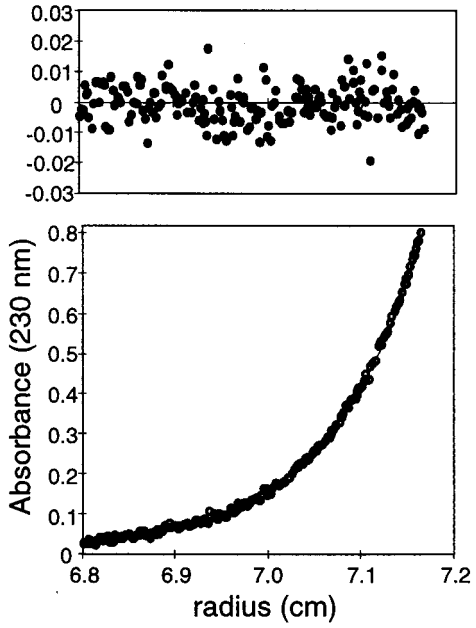

FIGURE 3: Sedimentation equilibrium analysis of Dpn (10 μM) in 20 mM Hepes, pH 7.6, 1 mM EDTA, 100 mM KCl. The fitted curve was obtained using a single ideal species model. The residuals are shown in the top panel.

however, no significant change in θ_{222} was recorded (Figure 2B). In addition, $\theta_{222}/\theta_{200}$ only increased from 0.22 to 0.28 at maximum magnesium concentration. Upon the addition of 25 mM KCl, the helicity increased to 26% as before (data not shown). These results suggest that $MgCl_2$ alone does not significantly alter the structure of Dpn.

To characterize the structural properties of Dpn alone and in the presence of DNA, CD spectroscopy was performed in the absence and presence of a DNA oligonucleotide containing a single Dpn binding site. The spectra were recorded at 100 mM KCl and a Dpn concentration of 3 μM (Figure 2C). The percent helicities of Dpn alone and in the presence of an equimolar amount of DNA were calculated to be 28% and 63%, respectively. In the absence of DNA, there was still a significant amount of random coil, as indicated by the large degree of molar ellipticity at 205 nm compared to 222 nm. In the presence of DNA, the 205 nm minimum was shifted approximately 3 nm to 208 nm, and the ratio between the two minima was approximately 1. This change in the shape of the protein-DNA spectrum indicates that the protein had adopted a fully formed, α-helical coiled-coil structure in the presence of DNA (*34*−*38*). Moreover, the percent helicities obtained from these measurements are in agreement with expected values based on the number of amino acids predicted to lie in helical regions of the bHLH structural motif (see Figure 1A).

Sedimentation Equilibrium. To characterize the oligomerization state of the bHLH domain of Dpn in solution, analytical ultracentrifugation was performed on a 10 μM solution of Dpn in 20 mM Hepes, pH 7.6, 100 mM KCl, and 1 mM EDTA. The solution was centrifuged at 30 000 rpm for 24 h in a sedimentation equilibrium experiment, where the absorbance at 230 nm was collected as a function of the cell radius (Figure 3). The data were fitted to both a single species model, which assumes that there is only one component in the solution giving rise to the absorbance, and an association model that assumes Dpn is in equilibrium

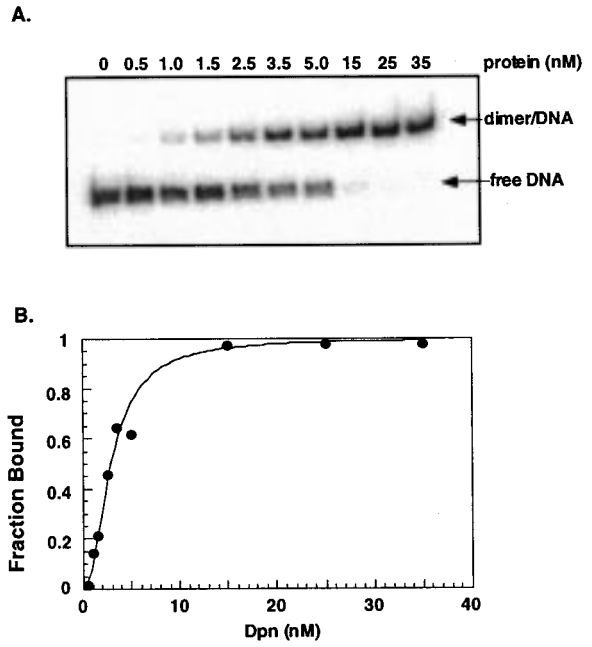

FIGURE 4: (A) EMSA of Dpn binding a 24 bp oligonucleotide containing a single binding site from the AS−C promoter. 40 pM radiolabeled duplex DNA was equilibrated with increasing concentrations of Dpn from left to right as shown. (B) Graphical representation of one EMSA trial. A K_d of 2.5 ± 0.3 nM was obtained with a Hill coefficient of 2.

between monomer and dimer. When the data were fit to the single species model, the best fit curve corresponded to a molecular mass of 13 555 ± 70 Da. The observed molecular mass was slightly lower than the calculated molecular mass of a dimer, presumably due to errors in estimating $\bar{v}$. Using the association model where the monomer molecular mass was fixed (7682 Da), an equally good fit was obtained; however, consistent with the single species model, this analysis revealed that the dimer was the predominant species in solution. Models incorporating higher order species gave an insignificant K_a for these higher molecular mass species. To assess the effects of KCl on dimerization, an additional sedimentation equilibrium experiment was performed in the absence of KCl (data not shown). Under these conditions, an association model was the best fit to the data, with a clear equilibrium between monomer and dimer. Taken together, these results suggest that Dpn is a dimer at high concentrations (10 μM), and that KCl promotes dimerization.

K_d Determinations. The DNA binding activity of synthetic Dpn was measured by quantitative EMSA, using a 24 base pair synthetic oligonucleotide containing a single binding site of sequence CACGCG from the AS−C promoter. The results of a protein titration performed in the presence of a constant amount of radiolabeled DNA (40 pM) and 100 mM KCl are shown in Figure 4A. The faster migrating band is uncomplexed DNA, and the slower migrating band, which increases in intensity with increasing concentrations of Dpn, is the Dpn-bound DNA complex. The radioactivity in each band was quantitated, and the fraction bound was plotted as a function of protein concentration (Figure 4B). The data were fit to the Hill equation with the Kaleidagraph software program. From this equation, the apparent dissociation constant (K_d) was calculated to be 2.5 ± 0.3 nM, and the stoichiometry of binding was determined to be a dimer (Hill coefficient = 2).

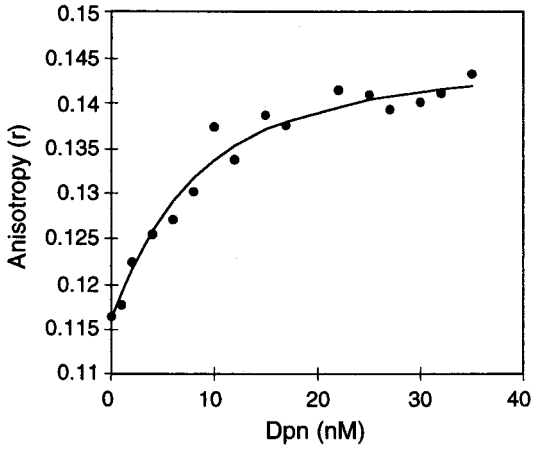

FIGURE 5: Fluorescence anisotropy titration of Dpn binding a single DNA site from the AS−C promoter. A fluorescein-labeled 13 bp oligonucleotide containing a single Dpn binding sequence at a concentration of 5 nM was titrated with increasing concentrations of Dpn. The data were fit (solid line) to a one-state binding model (dimer ⇌ dimer/DNA) using the program BIOEQS. From this analysis, a ΔG value of −11.1 (+0.2, −0.4) kcal/mol was determined, corresponding to a K_d of 5.1 (+2.5, −2.4) nM.

K_d measurements of the DNA binding activity of Dpn were also determined by fluorescence anisotropy. Because the entire experiment takes place in solution under equilibrium conditions, this method is thought to more accurately reflect protein behavior in solution (*40*, *41*). In this assay, a fluorescein-conjugated DNA oligonucleotide at a concentration of 5 nM was titrated with small increments of Dpn, and the resulting increase in the fluorescence anisotropy of the fluorophore was determined. Multiple measurements were taken to ensure that equilibrium had been reached before additional protein was added. The anisotropy data obtained from this method were plotted as a function of protein concentration and analyzed with BIOEQS, a numerically based equilibrium binding program (Figure 5). Using this program, a $\Delta G = -11.1$ (+0.2, −0.4) kcal/mol for Dpn binding was determined by fixing the fluorescence anisotropy of free DNA ($r = 0.1164$) and fitting the data to a simple one-state equilibrium model (free dimer ⇌ dimer−DNA). The quoted uncertainties correspond to 95% confidence intervals from an F-statistical distribution. The ΔG value obtained from this experiment corresponds to a K_d of 5.1 (+2.5, −2.4) nM, consistent with the results obtained from EMSA. An associating equilibrium model (monomer ⇌ dimer ⇌ dimer−DNA) also provided an equally good fit to the data, and yielded a similar binding affinity ($\Delta G = -11.3$ kcal/mol, data not shown), consistent with the hypothesis that Dpn exists as a monomer at nanomolar concentrations. A similar K_d was obtained in a separate titration using a DNA concentration of 2.5 nM, confirming that the K_d obtained by this method is independent of DNA concentration.

Effects of KCl on Dpn−DNA Binding. To further investigate the DNA binding characteristics of Dpn, the effects of potassium chloride concentration on Dpn binding activity were measured by EMSA. In this assay, radiolabeled DNA was equilibrated with a constant amount of protein in buffer containing increasing concentrations of KCl, in the presence or absence of poly(dI-dC) (20 ng). Aliquots of these reaction mixtures were subjected to electrophoresis on a native polyacrylamide gel, and the fraction bound was determined by phosphorimage analysis. The radioactivity in each band

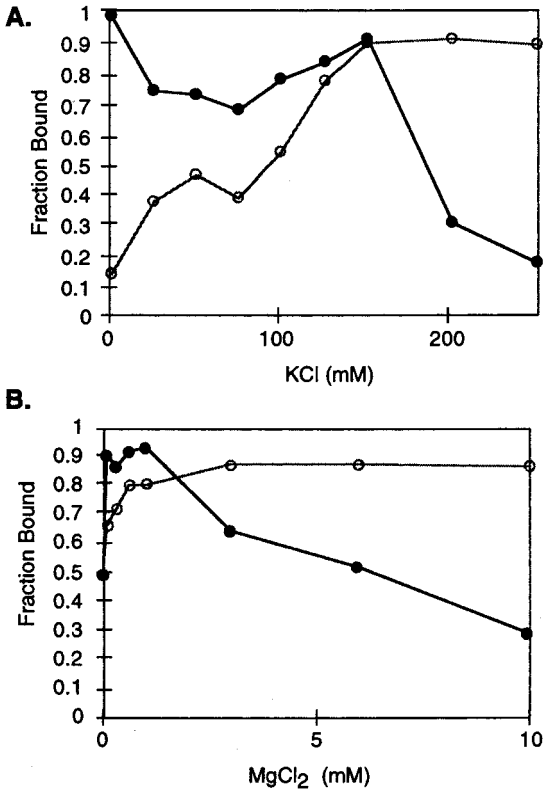

FIGURE 6: Radiolabeled duplex DNA (40 pM) was equilibrated with 100 or 10 nM Dpn in the presence or absence of poly(dI-dC), respectively. (A) Graphical representation of EMSA results from Dpn−DNA interactions at increasing concentrations of KCl, in the presence (○) or absence (●) of 20 ng of poly(dI-dC). (B) Graphical representation of EMSA results from Dpn−DNA interactions in the presence of 100 mM KCl and increasing concentrations of $MgCl_2$, in the presence (○) or absence (●) of 20 ng of poly(dI-dC).

was quantitated, and the fraction bound was plotted as a function of KCl concentration (Figure 6A). Dpn binding behavior showed a different KCl dependence with and without poly(dI-dC). In the presence of poly(dI-dC), the DNA binding activity of Dpn increased with increasing KCl concentration, where maximal activity was maintained from 150 to 250 mM KCl. This behavior suggests that higher KCl concentrations prevent Dpn from binding nonspecifically to poly(dI-dC), increasing the apparent affinity for specific DNA. In the absence of poly(dI-dC), Dpn binding activity was somewhat erratic at KCl concentrations below 100 mM, and fell sharply at KCl concentrations above 150 mM. From this plot, the optimal KCl concentration for DNA binding activity was determined to be between 100 and 150 mM. Protein titrations performed at 100 or 150 mM KCl yielded identical dissociation constants (data not shown).

Effects of $MgCl_2$ on Dpn−DNA Binding. The effect of magnesium ions on the binding activity of Dpn was also examined by EMSA. Radiolabeled DNA was equilibrated with a constant amount of protein in buffer containing 100 mM KCl and increasing concentrations of $MgCl_2$, in the presence or absence of poly(dI-dC). Aliquots of these reactions were subjected to electrophoresis as before, and the results are shown in Figure 6B. Dpn displayed a markedly different behavior in the presence and absence of poly(dI-dC). In the presence of poly(dI-dC), the DNA binding activity of Dpn increased with increasing magnesium concentration, reaching maximal activity in 3−10 mM $MgCl_2$. However, in the absence of poly(dI-dC), Dpn binding activity peaked at 1.0 mM $MgCl_2$, and fell sharply at higher concentrations of $MgCl_2$. This behavior suggests that magnesium ions compete electrostatically with Dpn for DNA sites in the absence of nonspecific competitor.

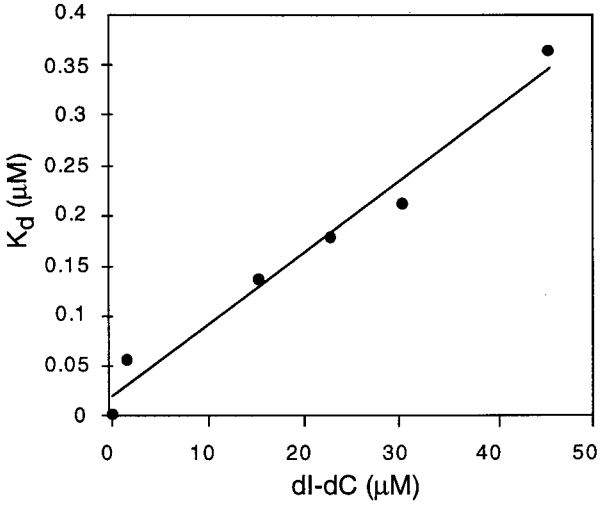

FIGURE 7: Graphical representation of EMSA results from Dpn−DNA interactions performed with increasing amounts of poly(dI-dC) in the binding reactions. The poly(dI-dC) concentration was plotted as a function of the K_d determined for each protein titration. The ratio of specific to nonspecific binding was calculated to be 140-fold from the inverse slope of the fitted line.

DNA Binding Specificity. The specificity of DNA binding was characterized by comparing the K_d of Dpn for specific DNA in the presence of increasing amounts of poly(dI-dC) using EMSA. The dependence of the K_d on poly(dI-dC) concentration is shown in Figure 7. In the absence of nonspecific competitor, the dissociation constant was calculated to be 2.5 nM. However, as increasing amounts of poly(dI-dC) were added to the binding reaction, the apparent K_d for Dpn increased in a linear fashion. This binding behavior suggests that Dpn bound nonspecifically to the poly(dI-dC). The slope of the plot in Figure 7 represents the ratio of specific DNA binding to nonspecific binding [poly(dI-dC)] (see Materials and Methods). From the inverse of this slope, the specificity ratio was calculated to be 140, indicating that Dpn bound specific DNA approximately 2 orders of magnitude better than poly(dI-dC).

Significance of the Tandem Binding Site. It had been proposed that the unique spacing between the two binding sites may permit protein−protein interactions between Dpn dimers (S. Burley, personal communication). To assess the significance of a naturally occurring tandem repeat found in the *sxl* promoter, both EMSA and fluorescence anisotropy methods were used to determine the K_d for Dpn binding to this sequence. Figure 8A shows the results of the EMSA titration at 100 mM KCl. As protein concentration was increased, uncomplexed DNA (fast migrating band) shifted first to the dimer−DNA complex (middle band) and then to the tetramer−DNA complex (top band). In this case, the fraction bound was calculated by extracting the fraction of bound sites: the dimer−DNA complex counted for only one of two sites filled, while the tetramer−DNA complex counted for both sites filled. Plotting these data as before yielded a K_d of 2.6 ± 0.2 nM with a Hill coefficient of 2 (Figure 8B).

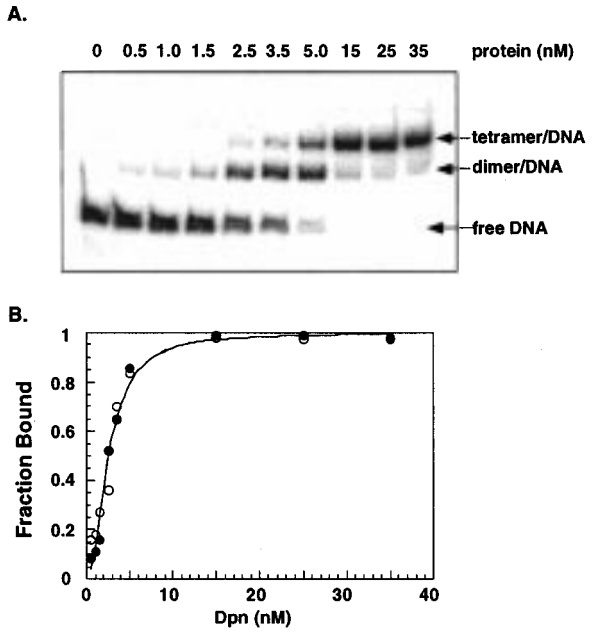

FIGURE 8: Binding of the Dpn bHLH domain to a tandem DNA site is noncooperative. (A) EMSA of Dpn binding an oligonucleotide containing a tandem binding motif. 40 pM radiolabeled duplex DNA was equilibrated with increasing concentrations of Dpn as shown. (B) Graphical representation of EMSA trials, where Dpn binding to an oligonucleotide containing the tandem site (●) or a tandem site with a five base pair insertion (○) is shown. A K_d of 2.6 ± 0.2 nM was obtained for the tandem site, and a K_d of 2.6 ± 0.1 nM was obtained for the "insert" site. Both curves yielded a Hill coefficient of 2.

From these results, we infer a sequential binding model, where binding of one Dpn dimer does not influence the binding of a second Dpn dimer.

To corroborate these results, EMSA was used to assay Dpn binding to an oligonucleotide containing an extra five bases between the tandem binding sites (termed "insert"). If protein−protein interactions occur in the wild-type tandem binding site, we hypothesized that insertion of five extra base pairs would disrupt any such interactions. The results of the protein titration are shown in Figure 8B. Dpn bound the insert site with a K_d of 2.6 ± 0.1 nM and a Hill coefficient of 2, statistically identical to the results obtained for the wild-type tandem site. Because an identical dissociation constant was obtained for the single, tandem, and "insert" oligonucleotides, binding to the tandem site was judged to be noncooperative.

We also used fluorescence anisotropy to determine the mechanism of Dpn binding and to quantitate the relative binding energies to the tandem site. The fluorescence anisotropy data were collected at a DNA concentration of 5 nM, and the results are shown in Figure 9. The data were fitted to a two-state binding model (free dimer ⇌ dimer−DNA ⇌ tetramer−DNA), which assumes that one Dpn dimer molecule binds a single site, followed by a second dimer binding the adjacent site. The ΔG values for the two binding events were obtained from the BIOEQS program by fixing the following anisotropy values (r): free DNA ($r = 0.1306$) and dimer−DNA ($r = 0.1500$). We experimentally determined $r_{dimer-DNA}$ by assaying a mutated version of the tandem site that could only support binding of a single dimer (each of the binding sites was mutated in turn, so that one hexanucleotide sequence was no longer a consensus sequence

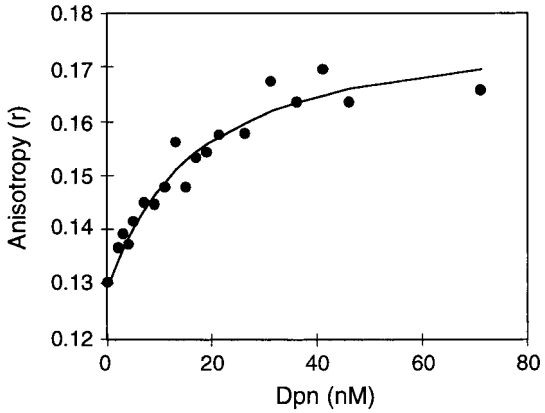

FIGURE 9: Fluorescence anisotropy titration of Dpn binding an oligonucleotide containing a tandem binding motif. The fluorescein-labeled 23 bp oligonucleotide (5 nM) was titrated with increasing concentrations of Dpn. The data were fit to a two-state binding model (−) using the program BIOEQS. ΔG values of −11.4 (+1.4, −0.9) kcal/mol for the dimer/DNA complex and −21.6 (+1.2, −1.0) kcal/mol for the tetramer/DNA were determined (i.e., approximately double the value for the single site), showing an absence of cooperative binding.

for Dpn). Because binding of Dpn to each "half" site was indistinguishable, we assumed that Dpn bound to either site within the tandem DNA sequence were equivalent species in our subsequent calculations. After performing a protein titration with the tandem site oligonucleotide, a ΔG value of −11.4 (+1.4, −0.9) kcal/mol was obtained for the dimer−DNA complex. The ΔG value calculated for the sum of the dimer−DNA complex and the tetramer−DNA complex was −21.6 (+1.2, −1.0) kcal/mol, within experimental uncertainty to double the ΔG value calculated for the dimer−DNA complex. Therefore, the fluorescence anisotropy data are also consistent with the noncooperative binding apparent in the EMSA experiments.

DISCUSSION

A series of quantitative DNA binding studies were carried out on the chemically synthesized bHLH region of Dpn. Solid-phase synthesis methods (*29*) were employed to prepare the 64 amino acid polypeptide because the bHLH domain was easily accessible to these methods. Synthesis of Dpn was straightforward and convenient, yielding large quantities (ca. 50 mg) of pure protein (Figure 1) in a short period of time (1 week).

To characterize the structure of the synthetic protein domain, CD spectroscopy was used to assess the helical content of Dpn under a physiological range of KCl and $MgCl_2$ concentrations (Figure 2). We infer from these data and sedimentation equilibrium studies (Figure 3) that Dpn is monomeric at nanomolar concentrations, but predominantly dimeric at high concentrations (10−33 μM) and in the presence of KCl. [These conclusions are also consistent with the Hill coefficient ($n = 2$) determined by EMSA analysis (Figure 4), which suggests that at nanomolar concentrations Dpn exists as a monomer and forms dimers to bind DNA.] We hypothesize that K^+ cations induce conformational changes in the preformed dimer which promote dimerization of the HLH region, while the basic region and contiguous helix 1 are largely unfolded in the absence of DNA. Complete folding of Dpn is nucleated by

the presence of its cognate DNA site, presumably locking the random-coil basic region and helix 1 into an α-helical recognition motif. These results are consistent with reports of other bHLH proteins undergoing a random-coil to α-helix transition with addition of double-stranded DNA oligonucleotides (*9*, *42*, *43*).

To further investigate the DNA binding properties of Dpn, the equilibrium dissociation constant (K_d) was determined for an isolated single binding site found in the AS−C promoter. Previous reports have shown that the related protein, Hairy, binds this sequence (*23*, *24*), but the apparent dissociation constant was not measured. Traditional techniques used to quantitate protein−DNA binding interactions include EMSA, CD spectroscopy, and sedimentation equilibrium (SE). Interpretation of data obtained using CD and SE techniques can be difficult, as both require high concentrations of protein which may drive nonphysiologically relevant oligomeric interactions. In contrast, EMSA can be used to measure dissociation constants of protein−DNA complexes in the subnanomolar to submicromolar range; however, it has been criticized as a valid method for studying protein−DNA interactions because it is not an equilibrium assay (i.e., separation of the free DNA and protein-bound species is required) (*40*, *41*). More recently, fluorescence anisotropy has been used for detecting, characterizing, and quantitating protein−DNA interactions (*44*). This technique appears to provide true equilibrium conditions because titration and analysis take place in solution, and it is sensitive enough to measure protein−DNA interactions with dissociation constants in the nanomolar to micromolar range [for a review, see (*40*, *41*)].

In this work, the dissociation constant for Dpn bound to an isolated site was obtained from both EMSA and fluorescence anisotropy methods. To our knowledge this is the first report of EMSA and fluorescence anisotropy methods used in a single study to determine the functional properties of a DNA-binding protein. Although there has been speculation concerning the "equilibrium nature" of the EMSA (*40*, *41*), the studies reported here reveal similar values for the K_d of Dpn binding DNA from EMSA (2.5 ± 0.3 nM) and fluorescence anisotropy [5.1 (+2.5, −2.4) nM] methods (see Figures 4 and 5).[2]

Surprisingly few quantitative studies have been carried out on the DNA binding activity of bHLH proteins. The dissociation constant determined here for Dpn reflects a significantly higher affinity for DNA than found for other bHLH homodimer proteins including MASH-1 (K_d = 118 nM), MyoD (K_d = 148 nM), E12 (K_d = 155 nM) (*42*), E47N (36 nM), and MyoD (126 nM) (*45*). However, these proteins were assayed under very different buffer conditions. For example, Künne et al. used no KCl and fairly high concentrations of ammonium sulfate in their binding assays of MASH-1, MyoD, and E12 (*42*), while Sun and Baltimore used a low concentration of monovalent cation (50 mM NaCl) and 1 μg of poly(dI-dC) in their K_d determinations for E47N and MyoD (*45*). Our results indicate that the bHLH domain of Dpn displays significantly different binding activity depending on the concentration of mono- and divalent ions, and whether poly(dI-dC) is present in the binding reaction (Figure 6). During our investigation of Dpn binding behavior, we also noticed that even low concentrations of poly(dI-dC) could compete with specific DNA for Dpn binding (Figure 7). Our results demonstrate how sensitive the dissociation constant can be to poly(dI-dC) concentration and mono- and divalent ion concentrations added to DNA binding reactions.

Significance of the Tandem Binding Site. To investigate possible cooperative interactions between Dpn dimers bound to DNA, the K_d for a tandem binding site found in the *sxl* promoter was also determined by EMSA (Figure 8) and fluorescence anisotropy (Figure 9) methods. In this case, the EMSA results revealed an identical dissociation constant found for both the single and tandem binding sites (2.6 ± 0.2 nM). Insertion of five extra base pairs between the two sites also had no effect on Dpn binding (Figure 8B). Taken together, these results suggest that there is no difference between a single Dpn dimer binding one DNA site and two Dpn dimers binding two DNA sites in tandem. Fluorescence anisotropy was also used to measure Dpn binding the tandem sequence. From a *single* experiment, we were able to calculate a ΔG for the first binding event [−11.4 (+1.4, −0.9) kcal/mol] and a ΔG for the sum of both the first and second binding events [−21.6 (+1.2,−1.0) kcal/mol] by extracting multiple ΔG values from the information contained in the tandem site binding profile (Figure 9). This binding profile also showed no measurable difference between binding to one or two sites. If there had been a cooperative binding event, the sum of both binding events would be more than twice the value obtained for the first binding event. These results highlight the advantages of EMSA and fluorescence anisotropy: EMSA allowed us to directly detect the fractional population of the free and bound complexes (free DNA ⇌ dimer−DNA ⇌ tetramer−DNA), providing a simple quantitative assessment of cooperativity. On the other hand, fluorescence anisotropy allowed straightforward extraction of multiple equilibrium binding constants without the separation of free and bound species.

This paper describes the thorough characterization of the DNA binding properties of the bHLH domain of Dpn. No evidence for cooperative binding was detected between two Dpn dimers binding to adjacent sites in a single DNA molecule. It is possible that cooperative interactions could occur in the context of the full-length protein. The protein has several other domains that could be involved in protein−protein interactions and that could influence DNA binding (*14*). Alternatively, these domains may provide a recognition surface to contact other transcription factors. For example, it has been shown that Hairy-related proteins share a conserved C-terminal tetrapeptide sequence, WRPW, which interacts with the corepressor Groucho protein (*46*−*49*). Linking two tandem Dpn binding sites may provide a scaffold for recruiting multiple Groucho proteins to achieve effective transcriptional repression.

ACKNOWLEDGMENT

We are grateful to Steve Burley for drawing our attention to this system as a fruitful target for study by total chemical

[2] The small difference may arise because the oligonucleotide used in the fluorescence anisotropy assay was a truncated form of the oligonucleotide used for the EMSA experiment. Preliminary data indicate that Dpn makes phosphate contacts to flanking bases that do not exist in the shorter oligonucleotide (unpublished results).

synthesis. In addition, we thank Jeff Kelly for many scientific discussions and for generously sharing his lab space and equipment. We also thank Catherine Royer for providing the BIOEQS program, Edwin Van der Schans for synthesis of the fluorescein-labeled oligonucleotides, Lynne Canne and Jay Levy for peptide synthesis, Hans Purkey for assistance with the analytical ultracentrifugation experiments, Mike Otto for assistance with the BIOEQS program, and Dan Treiber for helpful scientific discussions.

REFERENCES

1. Littlewood, T., and Evan, G. I. (1998) *Helix-loop-helix transcription factors*, Oxford University Press, London, England.
2. Murre, C., McCaw, P. S., and Baltimore, D. (1989) *Cell 56*, 777−783.
3. Davis, R. L., Cheng, P.-F., Lassar, A. B., and Weintraub, H. (1990) *Cell 60*, 733−746.
4. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature 363*, 38−45.
5. Ma, P. C. M., Rould, M. A., Weintraub, H., and Pabo, C. O. (1994) *Cell 77*, 451−459.
6. Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S. C. (1994) *Genes Dev. 8*, 970−980.
7. Shimizu, T., Toumoto, A., Ihara, K., Shimizu, M., Kyogoku, Y., Ogawa, N., Oshima, Y., and Hakoshima, T. (1997) *EMBO J. 16*, 4689−4697.
8. Párraga, A., Bellsolell, L., Ferré-D'Amaré, A. R., and Burley, S. K. (1998) *Structure 6*, 661−672.
9. 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.
10. Halazonetis, T. D., and Kandil, A. N. (1992) *Science 255*, 464−466.
11. Vinson, C. R., and Garcia, K. C. (1992) *New Biol. 4*, 396−403.
12. Klambt, C., Knust, E., Tietze, K., and Campos-Ortega, J. A. (1989) *EMBO J. 8*, 203−210.
13. Rushlow, C. A., Hogan, A., Pinchin, S. M., Howe, K. M., Lardelli, M., and Ish-Horowicz, D. (1989) *EMBO J. 8*, 3095−3103.
14. Bier, E., Vaessin, H., Younger-Shepherd, S., Jan, L. Y., and Jan, Y. N. (1992) *Genes Dev. 6*, 2137−2151.
15. Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992) *Genes Dev. 6*, 2620−2634.
16. Akazawa, A. C., Sasai, Y., Nakanishi, S., and Kageyama, R. (1992) *J. Biol. Chem. 267*, 21879−21885.
17. Delidakis, C., and Artavanis-Tsakonas, S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8731−8735.
18. Tietze, K., Oellers, N., and Knust, E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6152−6156.
19. Oellers, N., Dehio, M., and Knust, E. (1994) *Mol. Gen. Genet. 244*, 465−473.
20. Hoshijima, K., Kohyama, A., Watakabe, I., Inuoe, K., Sakamoto, H., and Shimura, Y. (1995) *Nucleic Acids Res. 23*, 3441−3448.
21. Jiménez, G., Pinchin, S. M., and Ish-Horowicz, D. (1996) *EMBO J. 15*, 7088−7098.
22. Barolo, S., and Levine, M. (1997) *EMBO J. 16*, 2883−2891.
23. Ohsako, S., Hyer, J., Panganiban, G., Oliver, I., and Caudy, M. (1994) *Genes Dev. 8*, 2743−2755.
24. Van Doren, M., Bailey, A. M., Esnayra, J., Ede, K., and Posakony, J. W. (1994) *Genes Dev. 8*, 2729−2742.
25. Chen, H., Thiagalingam, A., Chopra, H., Borges, M. W., Feder, J. N., Nelkin, B. D., Baylin, S. B., and Ball, D. W. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5355−5360.
26. Roark, M., Sturtevant, M. A., Emery, J., Vaessin, H., Grell, E., and Bier, E. (1995) *Genes Dev. 9*, 2384−2398.
27. Villares, R., and Cabrera, C. V. (1987) *Cell 50*, 415−424.
28. Younger-Shepherd, S., Vaessin, H., Bier, E., Jan, L. Y., and Jan, Y. N. (1992) *Cell 70*, 911−922.
29. Schnölzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B. H. (1992) *Int. J. Pept. Protein Res. 40*, 180−193.
30. Chen, Y., Yang, J. T., and Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
31. Royer, C. A. (1993) *Anal. Biochem. 210*, 91−97.
32. Royer, C. A., Smith, W. R., and Beechem, J. M. (1990) *Anal. Biochem. 191*, 287−294.
33. Straume, M., Frasier-Cadoret, S. G., and Johnson, M. L., (1991) in *Topics in Fluorescence Spectroscopy* (Lakowicz, J. R., Ed.) pp 177−240, Plenum Press, New York.
34. Lau, S. Y. M., Taneja, A. K., and Hodges, R. S. (1984) *J. Biol. Chem. 259*, 13253−13261.
35. Hodges, R. S., Semchuk, P. D., Taneja, A. K., Kay, C. M., Parker, J. M. R., and Mant, C. T. (1988) *Peptide Res. 1*, 19−30.
36. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992) *Biochemistry 31*, 5739−5746.
37. Zhu, B., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Protein Sci. 2*, 383−394.
38. Su, J. Y., Hodges, R. S., and Kay, C. M. (1994) *Biochemistry 33*, 15501−15510.
39. Atchley, W. R., and Fitch, W. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5172−5176.
40. Heyduk, T., Ma, Y., Tang, H., and Ebright, R. H. (1996) *Methods Enzymol. 274*, 492−503.
41. Hill, J. J., and Royer, C. A. (1997) *Methods Enzymol. 278*, 390−416.
42. Künne, A. G. E., Meierhans, D., and Allemann, R. K. (1996) *FEBS Lett. 391*, 79−83.
43. Meierhans, D., el-Ariss, C., Neuenschwander, M., Sieber, M., Stackhouse, J. F., and Alleman, R. K. (1995) *Biochemistry 34*, 11026−11036.
44. LeTilly, V., and Royer, C. A. (1993) *Biochemistry 32*, 7753−7758.
45. Sun, X.-H., and Baltimore, D. (1991) *Cell 64*, 459−470.
46. Paroush, Z., Finley, R. L., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R., and Ish-Horowicz, D. (1994) *Cell 79*, 805−815.
47. Fisher, A. L., Ohsako, S., and Caudy, M. (1996) *Mol. Cell. Biol. 16*, 2670−2677.
48. Dawson, S. R., Turner, D. L., Weintraub, H., and Parkhurst, S. M. (1995) *Mol. Cell. Biol. 15*, 6923−6931.
49. Wainwright, S. M., and Ish-Horowicz, D. (1992) *Mol. Cell. Biol. 12*, 2475−2483.

BI982856A