pH-Dependent Enhancement of DNA Binding by the Ultrabithorax Homeodomain†

Likun Li,‡ Doris von Kessler,§ Philip A. Beachy,§ and Kathleen S. Matthews*,‡

Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77251, and Department of Molecular Biology and Genetics, Howard Hughes Medical Institute, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received March 11, 1996; Revised Manuscript Received May 8, 1996⊗

ABSTRACT: Ultrabithorax (Ubx) and Deformed (Dfd) proteins of *Drosophila melanogaster* contain homeodomains (HD) that are structurally similar and recognize similar DNA sequences, despite functionally distinct genetic regulatory roles for Ubx and Dfd. We report in the present study that Ubx-HD binding to a single optimal target site displayed significantly increased affinity and higher salt concentration dependence at lower pH, while Dfd-HD binding to DNA was unaffected by pH. Results from studies of chimeric Ubx−Dfd homeodomains showed that the N- and C-terminal regions of the Ubx-HD are required for this pH dependence. The increase in binding affinity at lower pH was greater for the Ubx optimal binding site than for other DNA binding sites, indicating that subtle sequence alterations in DNA binding sites may influence pH-dependent behavior. These data demonstrate enhanced DNA binding affinity at lower pH for the Ubx-HD *in vitro* and suggest the potential for significant discrimination of DNA binding sites *in vivo*.

Protein binding to specific target DNA sequences to activate or repress the expression of associated genes is a primary mechanism for gene regulation. Specificity and stability of protein−DNA complexes are determined primarily by the available functional groups on the macromolecular species, but these properties are also influenced by physical environment (ionic strength, ion type, pH, temperature, *etc*.) (Riggs *et al*., 1970; Record *et al.*, 1976, 1977; Barkley, 1981; Barkley *et al.*, 1981; Roe *et al.*, 1984, 1985; Roe & Record, 1985; McClure, 1980, 1985; Buc & McClure, 1985; Brenowitz & Jamison, 1993; Wong & Bateman, 1994) [see Record *et al*. (1991) for review].

The homeodomain is a motif of ∼60 amino acid residues (Figure 1A) that displays significant evolutionary conservation among the members of a large family of eukaryotic transcriptional regulatory proteins [reviewed in Scott *et al*. (1989), Hayashi and Scott (1990), and Gehring *et al*. (1994a,b)]. The homeodomain is capable of sequence-specific DNA recognition, binds to DNA as monomer, and contains a fold with significant similarity to the helix−turn−helix motif of prokaryotic DNA binding proteins (Hayashi & Scott, 1990; Kuziora & McGinnis, 1989, Hanes & Brent, 1989; Mann & Hogness, 1990; Gibson *et al*., 1990; Ekker *et al*., 1991, 1992, 1994; Florence *et al*., 1991; Damante & Lauro, 1991; Lin & McGinnis, 1992; Chan & Mann, 1993; Furukubo-Tokunaga *et al.*, 1993; Pomerantz & Sharp, 1994; Draganescu *et al*., 1995; Laughon & Scott, 1984; Scott *et al*., 1989; Pabo & Sauer, 1992; Laughon, 1991). Structural studies have shown that sequence-specific DNA contacts are made by the N-terminal arm of the homeodomain in the minor groove and by helix 3 at the C-terminus of the homeodomain in the major groove (Billeter *et al*., 1993; Li *et al.*, 1995; Otting *et al*., 1990; Qian *et al*., 1989, 1994; Wolberger *et al*., 1991; Kissinger *et al*., 1990).

Homeodomain-containing proteins include the proteins encoded by genes of the *Drosophila* homeotic complexes and by the clustered Hox genes of vertebrates (Akam, 1987; Duncan, 1987; Kaufman *et al*., 1990; McGinnis & Krumlauf, 1992; Scott *et al.*, 1989; Hayashi & Scott, 1990; Gehring *et al*., 1994a,b). HOM-C and Hox proteins in a variety of organisms are presumed to specify the developmental identities of spatial units along the anterior/posterior axis by regulating the expression of target genes in spatially and temporally precise patterns (Akam, 1987; Duncan, 1987; Kaufman *et al*., 1990; McGinnis & Krumlauf, 1992; Scott *et al*., 1989; Hayashi & Scott, 1990; Gehring *et al*., 1994a,b). A long-standing question regarding the biological specificities of HOM-C proteins has been the mechanism(s) by which individual HOM-C proteins address themselves to distinctive sets of regulatory targets. In *Drosophila melanogaster*, for example, Ultrabithorax (Ubx)[1] specifies the identities of parasegments five and six, comprising the posterior thorax and part of the first abdominal segment, while Deformed (Dfd) specifies mandibular and maxillary segment identities (Akam, 1987; Duncan, 1987; Kaufman *et al*., 1990; McGinnis & Krumlauf, 1992). Although functional specificity *in vivo* has been found to reside primarily in the homeodomain regions of the Ubx and Dfd proteins (Kuziora & McGinnis, 1989; Lin & McGinnis, 1992) and distinct DNA sequence preferences are observed for Ubx-HD and Dfd-HD, the differences in sequence-dependent binding affinities are small (Ekker *et al*., 1992, 1994). In the context of discrete biological function, this low level of discrimination in site recognition is surprising. We show here that changes in the physical environment can

† This work was supported by grants to K.S.M. from the NIH (GM22441) and the Robert A. Welch Foundation (C-576). P.A.B. is an investigator of the Howard Hughes Medical Institute. This work utilized facilities of the Keck Center for Computational Biology.

* Corresponding author.
‡ Rice University.
§ The Johns Hopkins University School of Medicine.

⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1996.

[1] Abbreviations: Ubx, Ultrabithorax; Dfd, Deformed; HD, homeodomain; TBE, Tris−boric acid−ethylenediaminetetraacetic acid (EDTA); DTT, dithiothreitol; BSA, bovine serum albumin.

S0006-2960(96)00604-6 CCC: $12.00

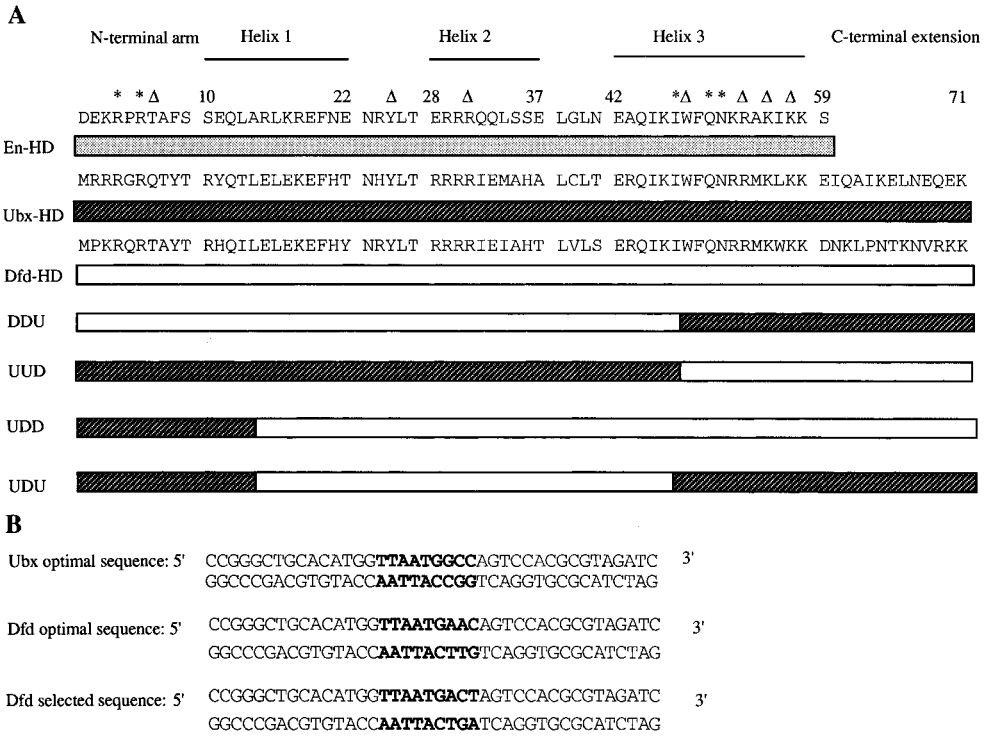

FIGURE 1: (A) Structures of Engrailed-HD (En-HD), Ultrabithorax-HD (Ubx-HD), and Deformed-HD (Dfd-HD) and four chimeric proteins derived from Ubx- and Dfd-HDs. Location of the three helices is based on the cocrystal structure of En-HD/DNA complex described by Kissinger *et al.* (1990). The asterisks (*) correspond to base contacts and triangles (Δ) to phosphate contacts assigned in the En-HD/DNA cocrystal. The amino acid sequences of the Ubx-HD and Dfd-HD are shown as hatched and open bars, respectively. The regions of Ubx and Dfd sequences contained in each of the four chimeras are shown schematically, and these are abbreviated as DDU, UUD, UDD, and UDU to indicate the segments corresponding to the parent HD. All homeodomains were constructed and purified as described in Ekker *et al.* (1991, 1992). Concentration of purified homeodomains was determined using molar absorptivity at 280 nm, and active concentration was determined by DNA binding activity assay under stoichiometric conditions. (B) DNA sequences used in this study. The binding-site sequences in each DNA are shown in boldface type. The 32-bp Dfd optimal sequence had 4 bp deleted from each end of the sequence shown.

significantly enhance Ubx-HD DNA binding affinity *in vitro* in a selective manner, suggesting potential mechanisms for generating site discrimination *in vivo*.

MATERIALS AND METHODS

Construction and Purification of Homeodomains. Ubx and Dfd homeodomains and four chimeric homeodomains derived from these proteins, DDU, UUD, UDD, and UDU (see Figure 1A for structures) were constructed and purified as described in Ekker *et al.* (1991, 1992). Concentration of purified homeodomains was determined using molar absorptivity for aromatic amino acids at 280 nm (Wetlaufer, 1962), calculated on the basis of the amino acid composition of each homeodomain (Ekker *et al.*, 1992). The values in units of M^{-1} cm^{-1} used were Ubx-HD, 9860; Dfd-HD, 15 460; DDU, 9860; UUD, 15 460; UDD, 14 040; and UDU, 8440. The active concentration of the homeodomains was determined by DNA binding activity assay under stoichiometric conditions. The percentage of activity measured for each proteins was as follows: Ubx-HD, ~100%; Dfd-HD, ~50%; DDU, ~100%; UUD, ~70%; UDD, ~40%; UDU, ~50%. The calculated active concentration based on percent activity was used in the experimental analyses.

Oligonucleotide Sequences for DNA Binding. Synthetic oligonucleotides of 40 and 32 bp were used for DNA binding measurements (see Figure 1B for sequences). The first contained the Ubx optimal DNA binding site sequence, TTAATGGCC; the second contained the Dfd optimal DNA binding site sequence, TTAATGAAC; and the third contained a sequence selected for high affinity for Dfd-HD, TTAATGACT (Ekker *et al.*, 1992). The 32-bp Dfd optimal binding DNA fragment was the same sequence as the 40-bp Dfd binding sequence with omission of 4 base pairs from both its ends. Similar binding affinities were measured for 40- and 32-bp fragments containing the Dfd optimal DNA sequence (data not shown). Equal molar amounts of synthesized single-stranded oligonucleotides (top and bottom strands) were annealed, purified by electrophoresis if necessary, and then end-labeled with [γ-^{32}P]ATP. The labeled oligonucleotide was separated from free [γ-^{32}P]ATP using a QIA quick-spin column (QIAGEN) or a NICK column (Pharmacia).

Equilibrium DNA Binding Measurements. Binding reactions were performed for 20 min at room temperature (22 °C) in 50 μL of buffer containing varying concentrations of active homeodomain polypeptide, constant labeled oligonucleotide (1×10^{-13} to 1×10^{-12} M, depending on the particular experiments), 1 mM dithiothreitol (DTT), 50 μg/mL bovine serum albumin (BSA), and 10% glycerol. Freshly thawed homeodomain polypeptide was used for each DNA binding experiment. The buffer type, pH, and ionic strength were varied as indicated in particular experiments. Loading dyes were not included in the reactions since it was reported that loading dyes significantly altered apparent thermodynamic equilibrium of the *lac* repressor−operator complex (Zhang *et al.*, 1993).

Two types of buffers, Tris (pK_a = 8.1 at 25 °C) and phosphate (pK_a = 6.9 at 25 °C), were selected for studies of

pH and ionic strength. In all experiments in which pH was varied, total ionic strength was kept constant (μ = 0.12 M, with 0.02 M contributed by Tris-HCl or potassium phosphate and 0.1 M by KCl). Individual ion concentration in the buffer at a particular pH was calculated according to the Henderson−Hasselbalch equation.

The experiments to determine effect of salt concentration on binding were carried out at two selected pHs, pH 7.5, at which most DNA binding experiments were performed previously, and pH 6.0, the optimal pH for the DNA binding of the Ubx-HD found in the present study. Potassium chloride in differing concentrations was used to vary ionic strength. Dependence of equilibrium constant on salt concentration was analyzed by log K_a *vs* log [M^+] plot (Record *et al.*, 1976), where [M^+] is the monovalent cation concentration contributed by KCl and K_a is the observed equilibrium association constant. The number of ionic interactions, *Z*, formed between homeodomain polypeptide and phosphate charges on the target DNA was calculated from the experimental value of $Z\psi$ according to the formula d log K_a/d log [M^+]= $-Z\psi$, where ψ is the number of counterions thermodynamically bound per phosphate. For double-helical DNA, ψ = 0.88 (Record *et al.*, 1976).

Retardation gels contained 8% polyacrylamide (19:1 acrylamide:bisacrylamide), 0.5× TBE (0.045 M Tris-borate, 0.001 M EDTA), and 3% glycerol. Electrophoresis of gels in 0.5 × TBE yielded results identical to those obtained in buffer similar to the reaction mixtures. Gel conditions were therefore kept constant while reaction conditions were altered. The gels were preelectrophoresed to constant current at 100 V with recirculation of 0.5× TBE buffer. Fifteen microliters of each sample was loaded onto the gels running at 300 V. Once tracking dyes (loaded side by side with experimental samples) had separated from each other, the voltage was reduced to 150 V, and the gel was electrophoresed at room temperature for 1.5 h [conditions modified from Carey (1988) and Liu and Matthews (1993)]. Dried gels were exposed to a FUJI phosphorimaging plate for 3 h (or overnight), and homeodomain−DNA complexes were quantified by scanning and analyzing the image generated using a Fuji Imaging Analyzer BAS1000 (Fuji Photo Film Co., Ltd, Japan) (Li & Matthews, 1995). The DNA binding data were then analyzed using the program Igor, version 2.0, to generate fits to the binding equation, R = [protein]/(K_d + [protein]), where R is the fractional saturation, K_d is the equilibrium dissociation constant, and [protein] is the protein concentration. The maximum value of R was not fixed in the fitting process to ensure that saturation was reached. The indicated equation represents the simplest binding process with no linked equilibria and was sufficient to fit the data. Experimental errors are expressed as standard deviations.

RESULTS

pH Dependence of Ubx and Dfd Homeodomain Binding. To monitor the effect of pH on Ubx-HD/DNA binding, the equilibrium association constant (K_a) was measured by gel retardation assays over the pH range 5.0−10.0 in Tris binding buffer (TBB) and over the pH range 5.0−7.5 in phosphate binding buffer (PBB) at constant ionic strength (see Materials and Methods for detail). In both TBB and PBB systems, Ubx-HD binding to DNA containing its optimal site, TTAATGGCC, is maximal not at neutral pH but at pH 6.0

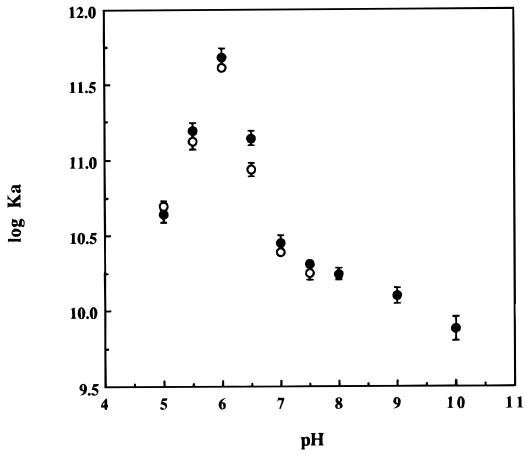

FIGURE 2: pH dependence of Ubx-HD/DNA binding. Ubx-HD binding to ^{32}P-labeled 40-bp dsDNA with the optimal site sequence for Ubx-HD (TTAATGGCC) was measured over the pH range 5.0−10.0 in Tris binding buffer (TBB, ●) and pH 5.0−7.5 in phosphate binding buffer (PBB, ○) at constant ionic strength (μ = 0.12 M) (see Materials and Methods for detail). The least-squares slope of the log K_a *vs* pH plot from pH 5.0 to 6.0 is 1.0 ± 0.1; from pH 6.0 to 7.0, −1.2 ± 0.1; and from pH 7.0 to 10.0, −0.17 ± 0.04.

(Figure 2). In the pH range 5.0−7.0, a dramatic effect on Ubx-HD binding is found. From pH 5.0 to 6.0, log K_a increases essentially linearly with pH, with a least-squares slope of the log K_a *vs* pH plot of 1.0 ± 0.1; from pH 6.0 to 7.0, log K_a decreases almost linearly with pH, with a least-squares slope of −1.2 ± 0.1. In the pH range from 7 to 10, the change of log K_a with pH diminished, with a slope of −0.17 ± 0.04 over this pH range.

To determine whether pH dependence of DNA binding is a characteristic in common with other *Drosophila* homeodomains, we examined the pH dependence of Dfd-HD using its optimal DNA binding site, TTAATGAAC, the Ubx optimal DNA binding site, TTAATGGCC, or a DNA binding site, TTAATGACT, to which Dfd-HD bound with high affinity (Ekker *et al.*, 1992) (for sequence comparison of Ubx-HD and Dfd-HD and their binding sites, see Figure 1). In contrast to the pH dependence of Ubx-HD/DNA interaction, only a small change for Dfd-HD DNA binding affinity was observed between pH 7.5 and 6.0 using any of the DNA sequences examined (Figure 3B, Table 1).

N-Terminal Arm and Helix 3 plus the C-Terminal Extension of Ubx Homeodomain Are Required for pH Differential. To determine which subregion(s) of Ubx-HD might contribute to pH-dependent DNA binding, we have examined four chimeric homeodomains derived from Ubx- and Dfd-HDs (see Figure 1A for structures). Slightly increased DNA binding affinity for the Ubx optimal sequence was found at pH 6.0 relative to pH 7.5 for three chimeric homeodomains, DDU (1.8-fold), UUD (2.4-fold), and UDD (1.6-fold) (Figure 3C−E, Table 1). Each of these chimeric proteins contained one, but not both, of two regions: the Ubx N-terminal arm or helix 3 with C-terminal extension (Figure 1A). The pH dependence was maximally recovered with the chimeric protein UDU (11-fold) (Figure 3F, Table 1), which contained both of the segments that appear to contribute to pH-dependent behavior. Although a very small degree of apparent cooperativity is indicated by the shape of some of binding isotherms (Figure 3C,E,F), there was no evidence of additional bands or protein aggregation in the experiments. Therefore, it is unclear whether these small shape differences

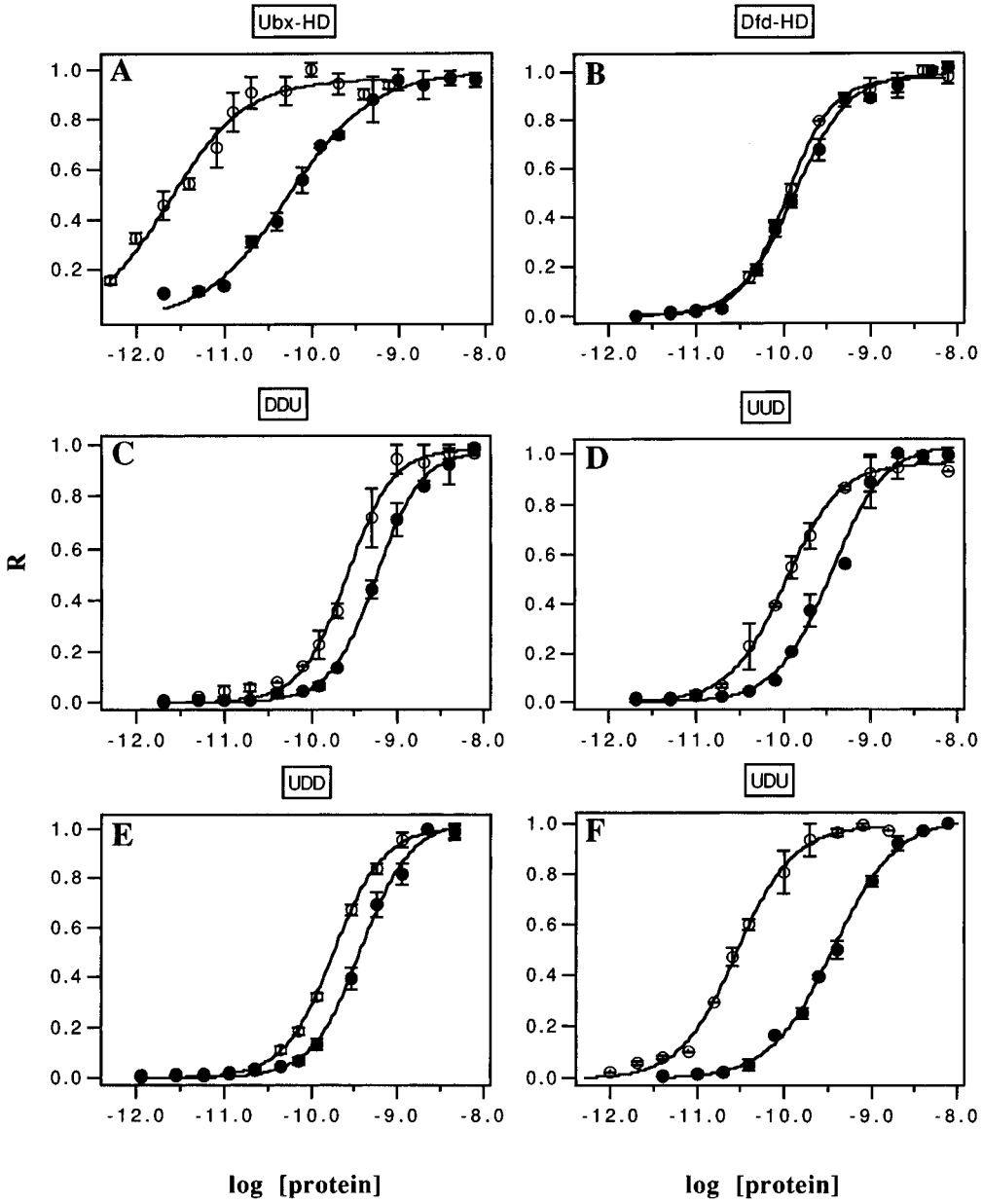

FIGURE 3: Comparison of pH dependence of Ubx, Dfd, and four chimeric homeodomains using the Ubx optimal binding site sequence, TTAATGGCC. Proteins used for each experiment are indicated in each panel. Measurements of equilibrium association constant (K_a) were carried out at pH 7.5 (●) and pH 6.0 (○) in PBB with constant ionic strength ($\mu = 0.12$) (see Materials and Methods for detail). This figure was derived from replicates within a single experiment, and the error bars indicate the standard deviation for these replicates. Values reported in Table 1 are the average of all replicates for multiple experiments.

Table 1: pH Dependence of DNA Binding by Ubx-HD, Dfd-HD, and Ubx/Dfd Chimeric Homeodomains[a]

homeodomains or chimeric proteins	Ubx-optimal: TTAATGGCC K_a (M^{-1}) × 10^{-9}			Dfd-optimal: TTAATGAAC K_a (M^{-1}) × 10^{-9}			Dfd-selected: TTAATGACT K_a (M^{-1}) × 10^{-9}		
	pH 7.5	pH 6.0	pH 6.0/pH 7.5	pH 7.5	pH 6.0	pH 6.0/pH 7.5	pH 7.5	pH 6.0	pH 6.0/pH 7.5
Ubx-HD	17 ± 4.1	400 ± 16	24	3.5 ± 1.8	25 ± 14	7.1	17 ± 2.3	71 ± 6.8	4.2
Dfd-HD	5.9 ± 1.8	8.1 ± 2.6	1.4	13 ± 9.8	18 ± 8.3	1.4	9.4 ± 0.9	10 ± 2.6	1.1
DDU	2.5 ± 0.6	4.6 ± 1.2	1.8	2.4 ± 0.9	3.9 ± 1.3	1.6	4.8 ± 1.8	11 ± 1.9	2.3
UUD	3.2 ± 0.7	7.8 ± 2.5	2.4	3.9 ± 1.6	6.3 ± 1.4	1.6	3.2 ± 1.6	7.0 ± 1.1	2.2
UDD	5.7 ± 3.3	9.1 ± 3.9	1.6	1.7 ± 0.2	3.7 ± 0.6	2.2	2.0 ± 0.4	3.9 ± 0.5	2.0
UDU	2.7 ± 0.2	29 ± 6.9	11	2.1 ± 1.0	6.9 ± 2.7	3.3	4.2 ± 0.9	16 ± 3.9	3.8

[a] The equilibrium association constant was measured in PBB (see Materials and Methods for details). In all experiments, ionic strength was kept constant ($\mu = 0.12$ M with 0.02 M contributed by PBB and 0.1 M by KCl). Experimental errors are expressed as standard deviations ($n \geq 4$).

are significant. A similar pattern of behavior was observed for Ubx-HD binding to either the Dfd optimal target DNA sequence or a sequence selected by Dfd-HD binding (Table 1). These results indicate that both N- and C-terminal regions of Ubx-HD contribute to pH-dependent DNA binding, and this property cannot be assigned to a single, specific amino acid residue but appears to derive from more dispersed structural influences.

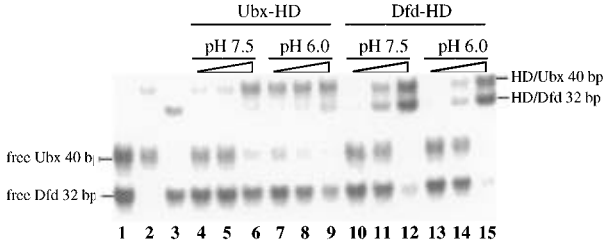

FIGURE 4: Discrimination in DNA binding by Ubx-HD. DNAs are present throughout at 1×10^{-10} M. Lane 1, free Ubx 40-bp DNA and free Dfd 32-bp DNA; lane 2, Ubx 40 bp DNA and Ubx-HD (1×10^{-10} M, pH 7.5); lane 3, Dfd 32-bp DNA and Dfd-HD (1×10^{-10} M, pH 7.5); lanes 4−15, Ubx 40 bp and Dfd 32-bp DNA [lanes 4−9, Ubx-HD (lanes 4−6, pH 7.5; lanes 7−9, pH 6.0); lanes 10−15, Dfd-HD (lanes 10−12, pH 7.5; lanes 13−15, pH 6.0)]. Lanes 4, 7, 10, and 13, protein concentration is 5×10^{-11} M; lanes 5, 8, 11, and 14, protein concentration is 1×10^{-10} M; lanes 6, 9, 12, and 15, protein concentration is 5×10^{-10} M. The affinities of 32-bp DNA and 40-bp Dfd-DNA sequences are similar for both Ubx-HD and Dfd-HD (data not shown); therefore, no influence of DNA length between 40 and 32 bp is involved in the binding reactions.

DNA Binding Site Sequences Contribute to pH Dependence of Ubx Homeodomain Binding. To explore whether protein structure is the only determinant for pH dependence of DNA binding, we have compared the binding of Ubx-HD, Dfd-HD, and the chimeric HDs to three different DNA sequences (see Figure 1B for comparison of DNA sequences). Interestingly, the difference in DNA binding affinity for Ubx-HD between pH 7.5 and 6.0 was 24-fold for its optimal DNA sequence, but this differential was reduced to ≤7-fold when the other two DNA binding site sequences were used (Figure 3A, Table 1). Similar effects were observed for the chimeric homeodomain UDU (Figure 3F, Table 1). These results demonstrate that pH dependence is determined not only by the differences between Ubx- and Dfd-HD structures but also by DNA sequence differences.

Since DNA binding affinity for Ubx-HD is greater when it interacts with its optimal DNA sequence, especially at pH 6.0, we tested the possibility that Ubx-HD could select its optimal DNA binding sequence when both Ubx and Dfd target sequences are present. To distinguish complexes formed from these two different DNAs, we employed mixtures of a 40-bp fragment containing the Ubx optimal site and a 32-bp fragment containing the Dfd optimal site for the binding reactions. No differences in binding affinity were observed for either homeodomain binding to fragments of 32 or 40 bp; thus, there is no length dependence for binding within this size range. As expected, Ubx-HD binds to its optimal DNA site preferentially at low protein concentrations and binds to the Dfd optimal site only at higher protein concentrations that exceed the concentration of the Ubx optimal DNA sequence (Figure 4). At pH 7.5, significant DNA binding is observed only at higher protein concentration with minimal Dfd DNA binding even at this protein concentration. Discrimination between DNA sequences is not observed for Dfd-HD at either pH (Figure 4).

pH Dependence of Ionic Interactions for Ubx and Dfd Homeodomains with DNA. The equilibrium association constants for the binding of Ubx- and Dfd-HDs and their respective target DNAs were determined as a function of salt concentration to measure the number of ionic interactions in the complexes formed (Record *et al.*, 1976). Within the salt concentration range studied (0.1−0.25 M), linear log−log plots of K_a *vs* salt concentration were observed (Figure 5). Presuming the absence of anion release over this range based on the linear responses (Record *et al.*, 1976; Barkley *et al.*, 1981), the slope of the log−log plots in Figure 5 yields the number of counterions ($Z\psi$) released from DNA upon formation of the protein−DNA complex, and this value in turn can be used to calculate the number of ionic interactions (Z). The slope is much steeper at pH 6.0 than at pH 7.5 for Ubx homeodomain (Figure 5A), indicating more counterions are released at pH 6.0. Approximately 8−10 monovalent counterions are released at pH 6.0 compared to ∼4 at pH 7.5. These results demonstrate that the effect of salt concentration on DNA binding by Ubx-HD is pH-dependent, and more counterions are released at pH 6.0 than at pH 7.5. The number of ionic interactions, Z, formed between the homeodomain and phosphates in the DNA backbone can be calculated from the experimental value of $Z\psi$ with the appropriate value of ψ (for double-helical DNA, $\psi = 0.88$; Record *et al.*, 1976). For Ubx-HD, the values for Z were ∼11 at pH 6.0 and ∼5 at pH 7.5. In contrast to the pH dependence of ionic interactions found for Ubx-HD, the number of ionic interactions for Dfd-HD binding to DNA was found to be pH-independent (Figure 5B). The number of monovalent counterions released is ∼4, and the number of ionic interactions, Z, was ∼5 at both pH 6.0 and 7.5.

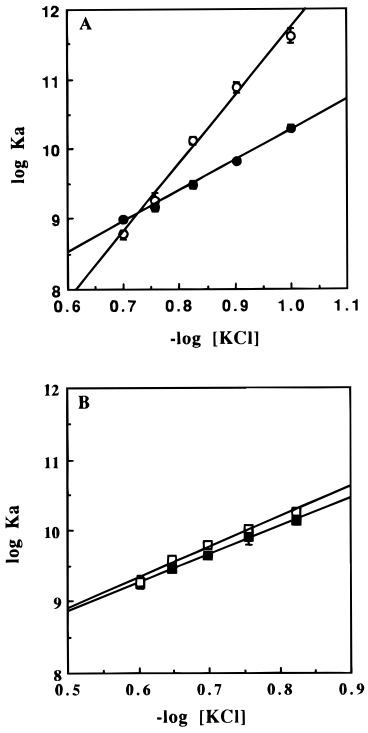

FIGURE 5: Ionic strength dependence of HD/DNA interactions. (A) Ubx-HD binding to 40-bp DNA containing Ubx optimal sequence in PBB at pH 7.5 (●) and pH 6.0 (○) and (B) Dfd-HD binding to 40 bp DNA containing Dfd optimal sequence in PBB at pH 7.5 (■) and pH 6.0 (□). KCl in differing concentrations was used to vary ionic strength, and plots of log K_a *vs* log [M^+] were used to analyze the data (see Materials and Methods for detail). The slopes of the plots were used to determine the number of counterions released, $Z\psi$, from which the number of ionic interactions, Z, was calculated (see Materials and Methods). For Ubx-HD the values for Z were ∼11 at pH 6.0 and ∼5 at pH 7.5. In contrast, Z for Dfd-HD was found to be pH-independent, with a value of ∼5 at both pH 6.0 and 7.5.

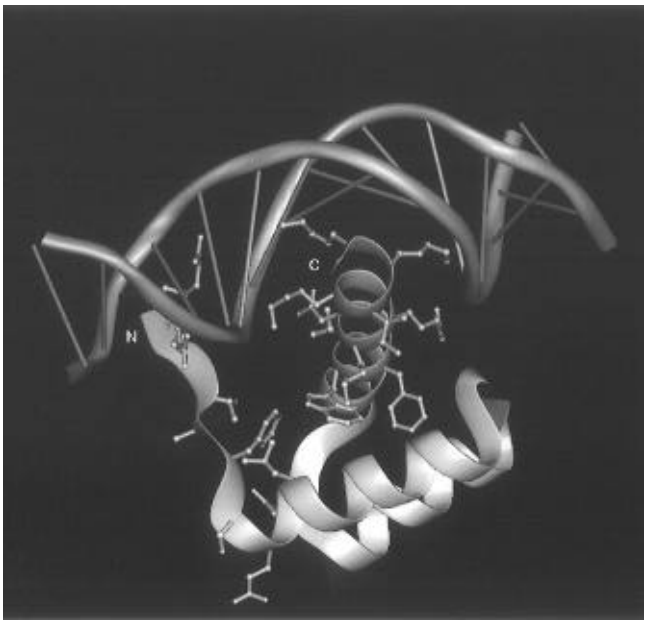

FIGURE 6: X-ray crystallographic structure of En-HD complexed with DNA. This figure was generated from PDB file 1HDD (Kissinger *et al.*, 1990) using the program QUANTA. Side chains are shown only for residues in the N-terminal region (amino acids 3−13) and C-terminal helix 3 (amino acids 47−59) to facilitate identification of these regions. Note the separation in the structure of these two regions that contribute to pH-dependent behavior.

DISCUSSION

Subtle Differences in Protein Structure Can Cause Differential Responses to Physical Environment. The results reported demonstrate that the DNA binding properties of Ubx and Dfd homeodomains are differentially influenced by alterations in physical environment. DNA binding of Ubx-HD is highly pH-dependent, with 24-fold increased affinity at pH 6.0, while that of Dfd-HD is pH-independent. Furthermore, the number of ionic interactions for Ubx-HD/DNA binding derived from change of K_a with monovalent cation concentration exhibits strong pH dependence, while the number of these interactions is unaffected by pH for Dfd-HD. The differences in the pH dependence of these two homeodomains must arise from their structural differences, in spite of the fact that the amino acid sequences of these two homeodomains are 72% identical and the recognition helices of the two homeodomains are almost identical with the exception of residue 56 (see Figure 1A). The discrepancy in response to environmental factors, however, mirrors differences observed in the *in vivo* functions of these two proteins (Beachy *et al.*, 1985, 1988; Krasnow *et al.*, 1989; Bergson & McGinnis, 1990; Regulski *et al.*, 1991). It is interesting to note that not only these two homeodomains but many proteins in this family share a high degree of identity and homology (Scott *et al.*, 1989). These conserved elements nonetheless result in distinct physiological effects, reflecting the complexity and delicacy of these genetic regulatory processes.

The homeodomain/DNA structures determined to date have revealed that the homeodomain has a similar molecular fold and interacts with DNA sequences in a conserved manner (Figure 6), with a bidentate contact between an invariant Asn side chain and an adenine base in the major groove providing a fixed frame of reference for all of these complexes (Qian *et al.*, 1989, 1994; Otting *et al.*, 1990; Billeter *et al.*, 1993; Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; Li *et al.*, 1995). Biochemical studies of DNA sequence preferences indicated that the optimal binding site generally includes a core TAAT motif (the underlined A participates in the bidentate contact with the invariant Asn), with some variation in preference for flanking sequences (Scott *et al*., 1989; Florence *et al*., 1991; Hanes & Brent, 1989; Hoey & Levine, 1988; Laughon, 1991; Treisman *et al*., 1989; Walter *et al*., 1994; Gehring *et al*., 1994a,b; Hayashi & Scott, 1990; Ekker *et al*., 1991, 1992, 1994; Kuziora & McGinnis, 1989). Despite these differences in flanking sequence preference, however, and particularly for HOM-C and other closely related homeodomains, sequence-dependent differences in affinity are not large when optimal DNA sites for different homeodomains are compared (Ekker *et al*., 1991, 1992, 1994; Kuziora & McGinnis, 1989; Desplan *et al.*, 1988; Hayashi & Scott, 1990; Hoey & Levine, 1988; Walter *et al.*, 1994; Pomerantz & Sharp, 1994). The question thus arises to what extent can DNA sequence discrimination account for the target specificity clearly evident from the distinctive biological properties of HOM-C proteins? Ubx and Dfd are two relatively well-studied homeodomains. Despite the fact that the optimal binding sites for Ubx-HD and Dfd-HD differ by only two base differences at position 7 and 8 in the 9-bp binding site sequences (TTAATGGCC for Ubx-HD and TTAATGAAC for Dfd-HD) (Ekker *et al.*, 1992), recent examination of these proteins by hydroxyl radical footprinting and missing nucleoside experiments revealed striking differences in contacts with DNA between the Ubx and Dfd homeodomains using the optimal sequence for either Ubx or Dfd binding (Draganescu *et al.*, 1995). These observations suggest that subtle differences in protein structure can result in different side chain−nucleoside interactions in the protein/DNA complexes. The data presented here suggest that small differences in homeodomain structures are enough to generate differential responses of protein−DNA interactions to physical environment.

The Magnitude of pH Dependence Is Affected by DNA Binding Site Sequences. By using different DNA sequences, we have found that the magnitude of the pH differential in Ubx homeodomain binding can be altered by base sequence. When the optimal site for Ubx-HD was employed, the DNA binding affinity at pH 6.0 was 24-fold higher than that at pH 7.5, while this difference diminished to ≤7-fold when sequences that bound better to Dfd were used. It is apparent from these results that both protein and DNA structure contribute to the pH dependence observed for Ubx-HD. The increase in DNA binding affinity was maximal only when Ubx-HD and its optimal DNA binding sequence were employed in the reaction.

Low pH May Improve Contacts between Ubx-HD Protein and DNA. The experiments with chimeric proteins demonstrated that the N-terminal arm and helix 3 with C-terminal extension of Ubx-HD are required for differential DNA binding at low pH. Only the chimeric protein UDU, which contains both the N-terminal arm and helix 3 with C-terminal extension, showed pH-dependent differential binding approaching that of Ubx-HD. According to the homeodomain/DNA structures determined (Qian *et al.*, 1989, 1994; Otting *et al.*, 1990; Billeter *et al.*, 1993; Kissinger *et al.*, 1990; Wolberger *et al*., 1991), the N-terminal arm contacts DNA in the minor groove and helix 3 contacts DNA in the major

groove. Figure 6 shows the structure for the *engrailed* homeodomain oriented to view these contacts. The absence of information on the structure of a significant portion of the C-terminal region (residues past position 59) and the apparent physical separation of the regions that influence pH dependence of Ubx-HD binding preclude facile identification of amino acid residues that might contribute to the observed binding effects. The difference in affinity between Ubx-HD binding at pH 6.0 and 7.5 can be derived from the equivalent of ≤ 2 hydrogen bonds and may be the sum of small increases in the strength of multiple single interactions. The latter mechanism is consistent with the DNA sequence dependence of this differential binding, as optimal alignment may depend not only on the protein fold but also on the DNA conformation.

On the basis of identification of subdomain requirements for pH dependence of Ubx-HD binding, measurements of salt dependence of Ubx- and Dfd-HD/DNA binding, and examination of the structures determined for homeodomains, we postulate that low pH may generate a protein configuration that optimizes Ubx-HD contacts along the DNA backbone for increased contribution of ionic interactions to the overall affinity. The differences in behavior between Ubx-HD and Dfd-HD with respect to pH are consistent with the striking distinctions in DNA contacts revealed for these proteins by the hydroxyl radical footprinting and missing nucleoside experiments of Draganescu *et al*. (1995). This conformation with increased affinity and enhanced ionic contacts may be stabilized not only by pH but also by other interactions or environments.

Potential Mechanism for the Discrimination of DNA Binding Sites in Vivo. The data presented in this study indicate that marginal binding affinity differences derived from small differences in protein structures and DNA sequences can be amplified by differential responses of protein−DNA interactions to physical environment. In the case of Ubx-HD, the physical environment can be manipulated *in vitro* to generate significant differences in DNA binding. Although pH differentials are not known to be associated with normal developmental processes, cellular pH shifts in response to growth factors or hormones have been identified in other signal transduction pathways [*e.g.*, Sugimoto *et al.* (1992) and Isfort *et al.* (1995)]. Presuming the absence of such environmental alterations, the conformational state of Ubx-HD corresponding to that elicited by pH variation could nonetheless be generated *in vivo* by association with other proteins. For example, differences in DNA binding specificity, and therefore presumably affinity, for the homologous Hox proteins can be generated by association with Pbx proteins (Chang *et al.*, 1996). The data given here arise from protein binding to DNA containing a single binding site, whereas, in contrast, naturally occurring binding sites for the Ubx-encoded protein contain multiple clusters of individual binding site sequences to which cooperative binding of Ubx protein occurs (Beachy *et al.*, 1993). According to a multisite integrative model of homeotic protein action proposed by Beachy *et al.* (1993), even small differences in binding of homeotic proteins to individual sites can be summed to yield large overall differences in binding to multiple sites. Differential responses of homeodomain proteins, whether to protein−protein association or local environment, to generate states with increased sequence discrimination would further expand the potential distinctions in DNA binding. It will be of interest to determine whether other HOM-C and related homeodomain proteins can assume states with increased sequence discrimination or increased affinity and whether and how such states may be induced *in vivo*.

ACKNOWLEDGMENT

We thank Dr. Liskin Swint-Kruse for assistance in generating the depiction of the the X-ray crystallographic structure of *engrailed* homeodomain complexed with DNA.

REFERENCES

Akam, M. (1987) *Development 101*, 1−22.
Barkley, M. D. (1981) *Biochemistry 20*, 3833−3842.
Barkley, M. D., Lewis, P. A., & Sullivan, G. E. (1981) *Biochemistry 20*, 3842−3851.
Beachy, P. A., Helfand, S. L., & Hogness, D. S. (1985) *Nature 313*, 545−551.
Beachy, P. A., Krasnow, M. A., Gavis, E. R., & Hogness, D. S. (1988) *Cell 55*, 1069−1081.
Beachy, P. A., Varkey, J., Young, K. E., von Kessler, D. P., Sun, B. I., & Ekker, S. C. (1993) *Mol. Cell. Biol. 13*, 6941−6956.
Bergson, C., & McGinnis, W. (1990) *EMBO J. 9*, 4287−4297.
Billeter, M., Qian, Y. Q., Otting, G., Müller, M., Gehring, W., & Wüthrich, K. (1993) *J. Mol. Biol. 234*, 1084−1097.
Brenowitz, M., & Jamison, E. (1993) *Biochemistry 32*, 8693−8701.
Buc, H., & McClure, W. R. (1985) *Biochemistry 24*, 2712−2723.
Carey, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 975−979.
Chan, S.-K., & Mann, R. S. (1993) *Genes Dev. 7*, 796−811.
Chang, C.-P., Brocchieri, L., Shen, W.-F., Largman, C., & Cleary, M. L. (1996) *Mol. Cell. Biol. 16*, 1734−1745.
Damante, G., & Di Lauro, R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 5388−5392.
Desplan, C., Theis, J., & O'Farrell, P. H. (1988) *Cell 54*, 1081−1090.
Draganescu, A., Levin, J. R., & Tullius, T. D. (1995) *J. Mol. Biol. 250*, 595−608.
Duncan, I. (1987) *Annu. Rev. Genet. 21*, 285−319.
Ekker, S. C., Young, K. E., von Kessler, D. P., & Beachy, P. A. (1991) *EMBO J. 10*, 1179−1186.
Ekker, S. C., von Kessler, D. P., & Beachy, P. A. (1992) *EMBO J. 11*, 4059−4072.
Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E., & Beachy, P. A. (1994) *EMBO J. 13*, 3551−3560.
Florence, B., Handrow, R., & Laughon, A. (1991) *Mol. Cell. Biol. 11*, 3613−3623.
Furukubo-Tokunaga, K., Flister, S., & Gehring, W. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6360−6364.
Gehring, W. J., Affolter, M., & Bürglin, T. (1994a) *Annu. Rev. Biochem. 63*, 487−526.
Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G., & Wüthrich, K. (1994b) *Cell 78*, 211−223.
Gibson, G., Schier, A., LeMotte, P., & Gehring, W. J. (1990) *Cell 62*, 1087−1103.
Hanes, S. D., & Brent, R. (1989) *Cell 57*, 1275−1283.
Hayashi, S., & Scott, M. P. (1990) *Cell 63*, 883−894.
Hoey, T., & Levine, M. (1988) *Nature 332*, 858−861.
Isfort, R. J., Stuard, S. B., Cody, D. B., Ridder, G. M., & LeBoeuf, R. A. (1995) *Eur. J. Biochem. 234*, 801−810.
Kaufman, T. C., Seeger, M. A., & Olsen, G. (1990) *Adv. Genet. 27*, 309−362.
Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B., & Pabo, C. O. (1990) *Cell 63*, 579−590.
Krasnow, M. A., Saffman, E. E., Kornfeld, K., & Hogness, D. S. (1989) *Cell 57*, 1031−1043.
Kuziora, M. A., & McGinnis, W. (1989) *Cell 59*, 563−571.
Laughon, A. (1991) *Biochemistry 30*, 11357−11367.
Laughon, A., & Scott, M. P. (1984) *Nature 310*, 25−31.
Li, L., & Matthews, K. S. (1995) *J. Biol. Chem. 270*, 10640−10649.
Li, T., Stark, M. R., Johnson, A. D., & Wolberger, C. (1995) *Science 270*, 262−269.

Lin, L., & McGinnis, W. (1992) *Genes Dev. 6*, 1071−1081.
Liu, Y.-C., & Matthews, K. S. (1993) *J. Biol. Chem. 268*, 23239−23249.
Mann, R. S., & Hogness, D. S. (1990) *Cell 60*, 597−610.
McClure, W. R. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5634−5638.
McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171−204.
McGinnis, W., & Krumlauf, R. (1992) *Cell 68*, 283−302.
Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gehring, W. J., & Wüthrich, K. (1990) *EMBO J. 9*, 3085−3092.
Pabo, C. O., & Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.
Pomerantz, J. L., & Sharp, P. A. (1994) *Biochemistry 33*, 10851−10858.
Qian, Y. Q., Billeter, M., Otting, G., Müller, M., Gehring, W. J., & Wüthrich, K. (1989) *Cell 59*, 573−580.
Qian, Y. Q., Furukubo-Tokunaga, K., Resendez-Perez, D., Müller, M., Gehring, W. J., & Wüthrich, K. (1994) *J. Mol. Biol. 238*, 333−345.
Record, M. T., Jr., Lohman, T. M., & deHaseth, P. (1976) *J. Mol. Biol. 107*, 145−158.
Record, M. T., Jr., deHaseth, P. L., & Lohman, T. M. (1977) *Biochemistry 16*, 4791−4796.
Record, M. T., Jr., Ha, J.-H., & Fisher, M. A. (1991) *Methods Enzymol. 208*, 291−343.
Regulski, M., Dessain, S., McGinnis, N., & McGinnis, W. (1991) *Genes Dev. 5*, 278−286.
Riggs, A. D., Suzuki, H., & Bourgeois, S. (1970) *J. Mol. Biol. 48*, 67−83.
Roe, J.-H., & Record, M. T., Jr. (1985) *Biochemistry 24*, 4721−4726.
Roe, J.-H., Burgess, R. R., & Record, M. T., Jr. (1984) *J. Mol. Biol. 176*, 495−521.
Roe, J.-H., Burgess, R. R., & Record, M. T., Jr. (1985) *J. Mol. Biol. 184*, 441−453.
Scott, M. P., Tamkun, J. W., & Hartzell, G. W., III (1989) *Biochim. Biophys. Acta 989*, 25−48.
Sugimoto, T., Kano, J., Fukase, M., & Fujita, T. (1992) *J. Cell. Physiol. 152*, 28−34.
Treisman, J., Gönczy, P., Vashishtha, M., Harris, E., & Desplan, C. (1989) *Cell 59*, 553−562.
Walter, J., Dever, C. A., & Biggin, M. D. (1994) *Genes Dev. 8*, 1678−1692.
Wetlaufer, D. B. (1962) *Adv. Protein Chem. 17*, 303−390.
Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D., & Pabo, C. O. (1991) *Cell 67*, 517−528.
Wong, J. M., & Bateman, E. (1994) *Nucleic Acids Res. 22*, 1890−1896.
Zhang, X., Duggan, L. J. L., & Gottlieb, P. A. (1993) *Anal. Biochem. 214*, 580−582.

BI9606049