LETTER

ACS chemical biology

Structure and Properties of a Re-engineered Homeodomain Protein–DNA Interface

Matthew D. Simon[†,∥], Morris E. Feldman[‡,∥], Daniel Rauh[**], Ann E. Maris[†], David E. Wemmer[†], and Kevan M. Shokat[†,§,*]

[†]Department of Chemistry, University of California, Berkeley, California 94720, [‡]Graduate Group in Biophysics, University of California, San Francisco, California 94158, [§]Howard Hughes Medical Institute and Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, **Present address: Max Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. [∥]These authors contributed equally to this work.

ABSTRACT The homeodomain (HD)–DNA interface has been conserved over 500 million years of evolution. Despite this conservation, we have successfully re-engineered the engrailed HD to specifically recognize an unnatural nucleotide using a phage display selection. Here we report the synthesis of novel nucleosides and the selection of mutant HDs that bind these nucleotides using phage display. The high-resolution crystal structure of one mutant in complex with modified and unmodified DNA demonstrates that, even with the substantial perturbation to the interface, this selected mutant retains a canonical HD structure. Dissection of the contributions due to each of the selected mutations reveals that the majority of the modification-specific binding is accomplished by a single mutation (I47G) but that the remaining mutations retune the stability of the HD. These results afford a detailed look at a re-engineered protein–DNA interaction and provide insight into the opportunities for re-engineering highly conserved interfaces.

The redesign of one member of a highly conserved and large protein family has been useful for the generation of tools to dissect individual protein function in complex biological systems. Protein kinases, for example, contain highly conserved ATP binding pockets and can be studied by introducing a large-to-small mutation at a conserved position in the active site, thereby creating an engineered kinase that is uniquely inhibitable using an enlarged inhibitor that is too bulky to fit into the active site of wild-type kinases (*1–3*). Similarly, an aspartate-to-asparagine mutation in the active site of a GTPase can switch the nucleotide specificity from GTP to another nucleotide, XTP, allowing the activity of an engineered GTPase to be tracked in the presence of other GTPases (*4–7*). In these cases, the high sequence conservation of the engineered protein's family allowed extension of an engineering solution from one protein to other members within the family. Although these examples illustrate the utility of engineering protein–small molecule interactions, many proteins exert their regulatory roles in the cell through protein–protein or protein–DNA interactions. Therefore we wondered whether highly conserved, extended interfaces could be re-engineered to incorporate new functionality.

Several extended interfaces have been engineered, including the interacting surfaces between human growth hormone and its receptor (*8*), zinc fingers and diverse DNA sequences (*9–11*), and the protein heterodimerization interface of an endonuclease (*12*). Perhaps the most dramatic example of an adaptive protein surface is the variable region of an antibody, which can recognize chemically diverse haptens, including those derived from the synthetic elaboration of biomolecules (*13*). Although these examples highlight the adaptability of protein interfaces, they also demonstrate that protein engineers have focused on interfaces that have a high degree of natural functional variability, and generally the amino acids at these interfaces are poorly conserved. In contrast, to take advantage of the frequent occurrence of highly conserved domains and their interacting surfaces, we are interested in how much adaptability and corresponding potential for re-engineering exists within highly conserved biomolecular interfaces. To address this question, we chose to focus on the homeodomain (HD)–DNA interaction.

The HD is a DNA-binding domain that has been conserved over 500 million years of evolution in both structure and function (*14*). The 60 amino acid HD is composed of three α helices (Figure 1, panel a), the C-terminal of which is referred to as the recognition helix as a result of base-specific contacts it makes in the major groove of DNA with the sequence TAATXX (Figure 1, panels a and b, in green) (*15–18*). These conserved contacts include a hydrophobic interaction of either isoleucine orvaline at

*Corresponding author, shokat@cmp.ucsf.edu.

Received for review August 28, 2006 and accepted November 21, 2006.

Published online December 15, 2006

10.1021/cb6003756 CCC: $33.50

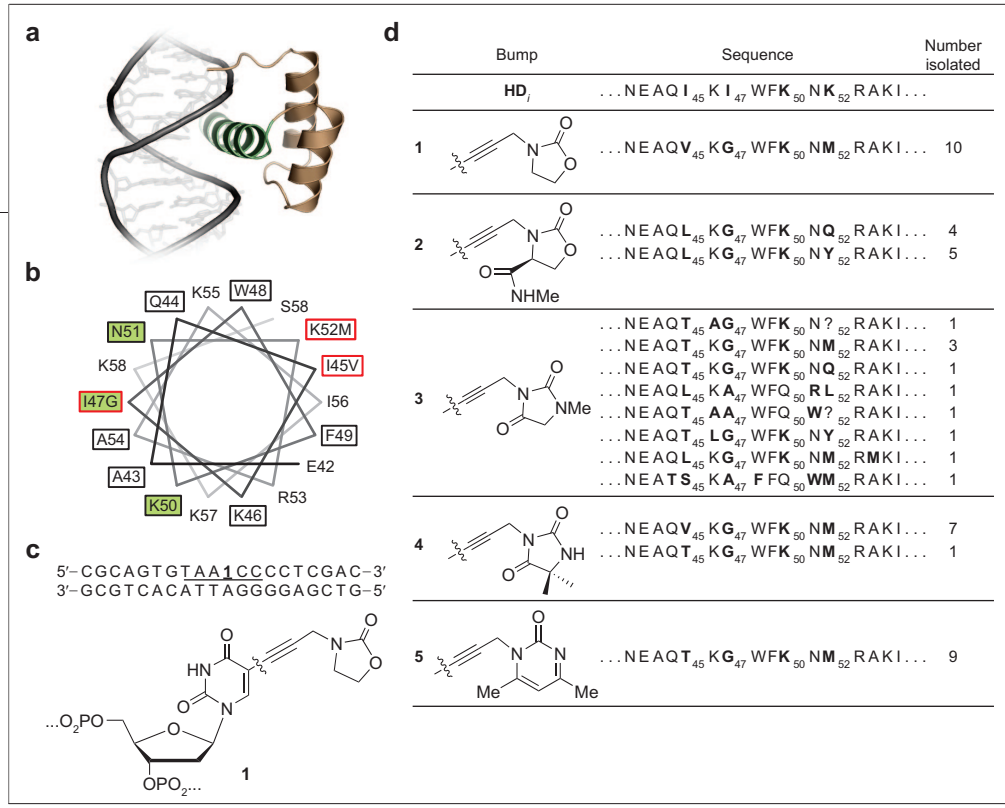

Figure 1. Sequences and schematic representation of the HD mutants recovered from phage selections. a) HD$_I$ is shown bound to TAATCC with its recognition helix (α3) shown in green. b) Helical wheel representation of the HD recognition helix with the three canonical base-specific contacts (Ile47, Lys50, and Asn51) shown in green, residues that were varied in the library presented in boxes, and mutations that comprise HD$_\Phi$, indicated in the sequence and shown in red. c) Example DNA sequence used in the phage selections with oxazolidinone-modified nucleoside 1 incorporated into the HD$_I$ binding site. d) Clones recovered from selections for binding to the modified DNA strands TAA1CC (1), TAA2CC (2), TAA3CC (3), TAA4CC (4), and TAA5CC (5). In two cases (indicated with a question mark) the sequencing led to ambiguous results at position 52.

position 47 with the C5 methyl group of thymidine 4, two invariant hydrogen bonds from Asn51 to adenosine 3, and contacts between either a glutamine or a lysine at position 50 that specify the last two base pairs in the recognition sequence (TAATTA or TAATCC, respectively) (*19*). Although these residues are nearly invariant in naturally occurring HDs, we recently reported the re-engineering of this interface using a phage display selection with a synthetic DNA oligomer bearing an unnatural nucleotide (*20*). This selected homeodomain (HD$_\Phi$) binds with specificity for DNA bearing an oxazolidinone appendage projecting into the major groove of the DNA on an alkynyl linker.

Here we report the design and synthesis of other novel nucleosides and the mutations recovered from phage display selections against DNA strands bearing these nucleotides. We also report extensive characterization of one mutant, HD$_\Phi$, including biochemical dissection of the roles for each mutation in binding studies, and we analyze the stability of these mutants using CD spectroscopy. Finally, we report the high-resolution crystal structure of HD$_\Phi$ bound to modified and unmodified DNA. Together, these results provide insight into the engineering of highly conserved interfaces.

Design and Synthesis of Derivatized Nucleosides. Our work re-engineering HD–DNA interactions has focused on the well-characterized Q50K mutant of the engrailed homeodomain (HD$_I$). The Q50K variant is naturally occurring in some HDs (*e.g.*, Bicoid) and alters the binding specificity from the palindromic consensus sequence of the wild-type HD, TAATTA, to the nonpalindromic sequence TAATCC, therefore simplifying biochemical analysis. To engineer HD–DNA interactions, we have focused on regions of the HD$_I$–DNA interface (*21*) where the HD makes highly conserved contacts to the DNA (Figure 1, panel a). The hydrophobic contact between Ile47 and the C5 methyl group of thymidine 4 in the recognition motif (TAATTA) was attractive because a large-to-small mutation at Ile47 might create sufficient space to accommodate prosthetic groups appended to the C5 position of T4. Furthermore, the Ile47 contact contributes substantially to DNA binding, but unlike Asn51 (another conserved residue that contacts the DNA), Ile47 is not completely essential ($K_{d,\ I47A}/K_{d,WT} \sim 20$; $K_{d,\ N51A}/K_{d,WT} > 1000$ (*22*)).

With this region of the protein–DNA interface in mind, we have designed synthetic appendages to the DNA that introduce diverse chemical functionality and steric demands into the interface yet are unlikely to compromise the basic structure of the B-form DNA or clash with the backbone of the helix-turn-helix motif, which is fundamental to the HD fold (*20*). The use of allyl or propargyl linkers appended to the C5 of thymidine bases allows the incorporation of diverse chemical functionality and the impact of these modifications has been systematically explored elsewhere (*23*). In these cases, the extended π-system can contribute to the base–base stacking within the DNA helix, thereby stabilizing the desired B-form DNA structure.

Because of the limited space between the major groove of the DNA and the protein backbone of the recognition helix, we chose to focus on flat five- and six-membered heterocycles connected to the base through a propargyl linker at C5 of thymidine (**1–5**). These alkynyl heterocycles were synthesized as shown in Supplementary Scheme 1. To install these appendages onto the nucleoside base, Sonogashira conditions were used to couple these terminal alkynes to 5′-dimethoxytrityl (DMT)-protected

2′-deoxy-5-iodouridine (see Supplementary Scheme 2) as we have reported previously (*20*). Installing the DMT-protecting group prior to the Sonogashira coupling proved advantageous both because this route allowed the use of a common starting material and also because this protection scheme allowed the reaction to be performed in THF. Using THF as solvent accelerated the reaction dramatically relative to the rate in DMF and simplified the workup of the desired nucleosides. These modified nucleosides were then incorporated into DNA oligomers (Figure 1, panel c) using solid-phase DNA synthesis as previously described (*20*).

Phage-Display Selections. Models and biochemical data (*20*) have demonstrated that a steric clash between Ile47 and alkynyl substituents prevents HD_I from binding to the modified DNA. Therefore, we have sought mutant HDs that bind specifically to the modified DNA strands, presumably including a large-to-small mutation at Ile47. To select for such mutants, we employed a phage selection with a HD displayed on the major coat protein (P8) of M13 phage. The phage display system has been used previously to select for mutants of the engrailed HD that bind to unmodified (*24, 25*) and oxazolidinone **1** modified DNA (*20*). The selections using unmodified DNA have validated this approach; the amino acids enriched in the selections largely recapitulated those found in naturally occurring HDs.

To generate the library of mutant HDs, we used a variant of Kunkel mutagenesis to introduce mutations focused around Ile47 (Figure 1, panel b, boxed residues), and to increase the proportion of functional HDs, the library was biased toward approximately four mutations per clone using split-and-pool DNA synthesis to construct the degenerate oligonucleotide used for mutagenesis (*20*). With this approach, we were able to obtain high coverage of the library using only modest numbers of unique transformants (1.8×10^7 Amp^r colonies).

Using this library and previously established selection conditions, we enriched for DNA-binding phage within 3–4 rounds of selection. To identify mutations enriched by these selections, we sequenced several clones from each enriched pool (Figure 1, panel d). The sequenced clones demonstrated common mutations derived from independent library members, and for TAA**1**CC a single clone dominated the selection.

As we expected, the sequences largely contained Lys50, consistent with the known role for this residue specifying the last two base pairs of the DNA binding site used in the selection (*i.e.*, TAA**X**CC). Although it was encouraging that most of the clones contained a small residue at position 47, this bias was programmed into the original library (50% alanine; 50% glycine). Nonetheless, the bias toward Gly47 (see Figure 1, panel d) was expected on the basis of previous data demonstrating that I47A has low binding affinity and specificity. From the selected clones, we found a preponderance of mutations at two other positions: Ile45 and Lys52. These mutations are on the back side of the recognition helix from Ile47. Closer examination of the sequences from the selected clones demonstrates that these clones are clustered into two groups: those with consensus mutations (Lys50, Gly47, I45V/T/L, and K52M) and clones that do not have these mutations but are also missing residues identified as particularly important for HD function (*e.g.*, Trp48 and Asn51), suggesting that this second population is composed of residual clones that do not bind the desired DNA.

Having identified consensus mutations in selection against various modified DNA strands, we chose to focus on one mutant in particular, HD_Φ (I45V, I47G, Q50K, and K52M), because it completely dominated the pool enriched using oxazolidinone **1** and furthermore was the only clone to give a strongly positive result in a phage enzyme-linked immunosorbent assay (*20*). Replacement of Ile45 by valine is not entirely surprising given that isoleucine and valine are equally likely to occur in natural HDs at position 45 and this substitution is quite conservative. The replacement of Lys52 with methionine, however, is quite surprising; out of 129 human HDs, 103 have arginine at position 52, 10 have lysine, 7 have alanine, but none have methionine (*26*).

TABLE 1. Affinity for DNA[a] and thermal stability of engrailed HD mutants

Mutant	K_d (nM)					
	$TAATCC_{20}$	$TAA1CC_{20}$	$TAATGC_{20}$	$TAATCG_{20}$	T_m (°C)	ΔH (kcal/mol)
HD_I	2.1 ± 0.3	17.8 ± 3.5	8.8 ± 0.2	6.0 ± 1.0	52.6 ± 0.7	−30 ± 2
HD_Φ	5.8 ± 0.6	1.5 ± 0.1	4.5 ± 0.6	18.0 ± 1.5	53.9 ± 0.5	−36 ± 2
HD_I I47G	4.2 ± 0.6	2.3 ± 0.3			47.3 ± 0.5	−29 ± 1
HD_I I47G, I45V	2.3 ± 0.1	3.1 ± 0.3			59.4 ± 0.4	−36 ± 2
HD_I I47G, K52M	4.6 ± 1.0	2.3 ± 0.3			43.1 ± 0.5	−26 ± 1

[a] HD binding to 5′-CGCAGTGTAAXXXCCTCGAC and its complement was measured by EMSA.

Biochemical Analysis of a Selected Mutant. When given the choice in a competition electrophoretic mobility shift assay (EMSA), HD_Φ binds 5-fold more tightly to $TAA1CC_{20}$ than unmodified DNA, whereas HD_I binds specifically to unmodified DNA (Supplementary

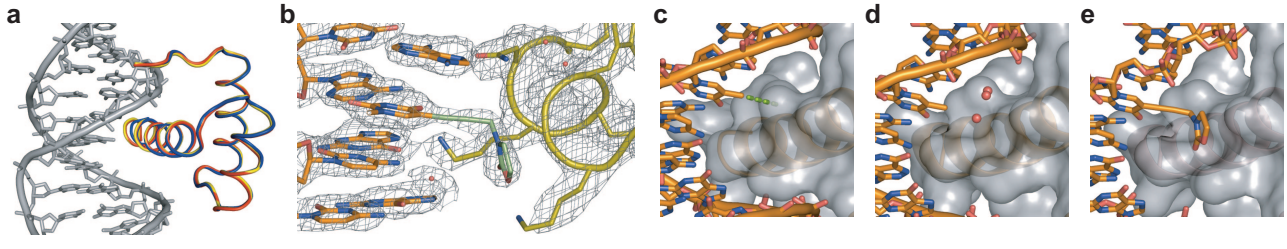

Figure 2. Structure of HD_Φ bound to TAATCC and TAA1CC. a) Ribbon overlay of HD–DNA complexes (HD_i–TAATCC (*15*) in blue, HD_Φ–TAA1CC in red, and HD_Φ–TAATCC in yellow). b) Electron density in the vicinity of the alkynyl nucleoside. $2F_o - F_c$ electron density from a simulated-annealing composite omit map is contoured at 1σ. c) HD_i–TAATCC. The hydrophobic contact between Ile47 and the C5 methyl of thymidine in the recognition sequence TAATCC is indicated by green dashes. d) HD_Φ–TAATCC. Three waters shown in red occupy the cavity created by the I47G mutation in HD_Φ. For clarity, only these three waters are displayed. e) HD_Φ–TAA1CC. The oxazolidinone substituent packs against the cavity created by I47G and displaces the three waters seen in panel d.

Figure 1, panel a). The preference of HD_Φ for TAA**1**CC_{20} was seen previously using HD_Φ fused to maltose binding protein (*20*). Untagged HD_Φ also prefers binding to TAA**1**CC_{20} over TAA**2**CC_{20}, another oxazolidinone-bearing strand (Supplementary Figure 1, panel b).

Next we investigated the role of the two mutations in the hydrophobic core of the HD by mutating these residues back to the wild-type sequence, either individually (HD_i I47G, I45V and HD_i I47G, K52M) or together (HD_i I47G). In EMSAs using either modified or unmodified DNA strands, we were surprised to find that HD_i I47G functioned with specificity and affinity comparable to that of HD_Φ (Table 1 and Supplementary Figure 2). Although the conditions presented also suggest that the selectivity of HDs lacking the hydrophobic mutations is somewhat depressed, the importance of this result is mitigated by modest variability of the selectivities observed under different conditions. Nonetheless, under all conditions tested, HD_Φ is the tightest and most selective HD, demonstrating that HD_Φ binds with the desired specificity for TAA**1**CC_{20}.

Structural Analysis of the Selected Mutant in Complex with DNA. To understand the structural basis for this re-engineered HD–DNA interaction, we solved X-ray crystal structures of HD_Φ in complex with both modified and unmodified DNA to 2.2 and 1.9 Å, respectively (Supplementary Figure 3 and Supplementary Table 1). Analysis of the structures reveals that, despite the perturbation of the mutations in the HD and the modification to the DNA, HD_Φ still adopts a canonical HD structure as judged by the nearly identical ribbon representations of Cα positions for the HD_i–TAATCC, HD_Φ–TAATCC, and HD_Φ–TAA**1**CC structures (Figure 2, panel a). The oxazolidinone modification is clearly visible in the electron density (Figure 2, panel b) and projects into the cavity created by the I47G mutation (Figure 2, panel e). The average B factor of the oxazolidinone modification (34 Å^2) is similar to those of the base to which it is connected (33 Å^2) and also of side chains Lys50 (35 Å^2) and Asn51 (37 Å^2), suggesting that the oxazolidinone is well ordered within the context of the re-engineered interface. In the complex of HD_Φ with unmodified DNA, the cavity created by the I47G mutation (Figure 2, compare panels c and d) is occupied by three ordered water molecules (Figure 2, panel d). In both structures of HD_Φ, with the exception of the waters immediately surrounding the modification and cavity, the waters at the protein–DNA interface are similar to those found in the unengineered HD_i–TAATCC interface.

In the structure of HD_i bound to TAATCC, residue Lys50 contacts the last two bases through two alternate conformations allowing it to bind specifically to both base pairs (*15*). However, in the electron density for HD_Φ bound to either TAA**1**CC (Supplementary Figure 4) or TAATCC, K50 only occupies one of these conformations, with no apparent structural basis for specifying the final base TAATCC. To test if HD_Φ has decreased specificity for the last C=G base pair (TAATCC) compared to the previous C=G base pair (TAATCC), HD_Φ and HD_i binding to TAATCG and TAATGC were assayed by EMSA (Supplementary Figure 5, panels a and b). We found that HD_Φ retains specificity for the final base pair despite the single orientation of Lys50 observed in the crystal structure. Furthermore, we found that HD_Φ binding to its preferred DNA target is competed off approximately the same concentration of salmon sperm DNA as was found for HD_i (Supplementary Figure 5, panel c).

We next turned our attention to the two hydrophobic mutations I45V and K52M found in HD_Φ. Examination of the region of HD_Φ around Val45 revealed only minor changes relative to HD_i, such as changes in electron density best explained by Ser35 adopting a second conformation. It is possible that this second conformation requires the slightly reduced steric volume of the I45V mutation, but the importance of this second conformation is unclear. Examination of the environment surrounding the K52M mutation reveals that this mutation could relieve electrostatic repulsion caused by three lysine residues (Lys17, Lys52, and Lys55) in close proximity to one another. The engrailed HD is a member of a small subset of HDs that have basic amino acids at both position 17 and position 52. Most HDs possess a salt bridge between Glu17 and Arg52 in the HD consensus sequence (*27*). The high density of positive charge caused by the presence of Lys17, Lys52, and Lys55 is destabilizing and can be relieved by K52A and K52E mutations, which stabilize the engrailed HD as previously demonstrated (*28*). This result suggests that K52M may impact protein stability and led us to wonder more generally about the effects of the HD_Φ mutations on HD stability.

Analysis of the Stability of the Selected Mutant. To examine the stability of the HDs used in this study, we monitored their thermal denaturation by CD spectroscopy (Figure 3). Starting with HD_i, introducing the I47G mutation is destabilizing to the protein ($\Delta T_m = -5.3$ °C). In general, the replacement of an amino acid with glycine is destabilizing because the unfolded state of the protein is entropically stabilized by extra

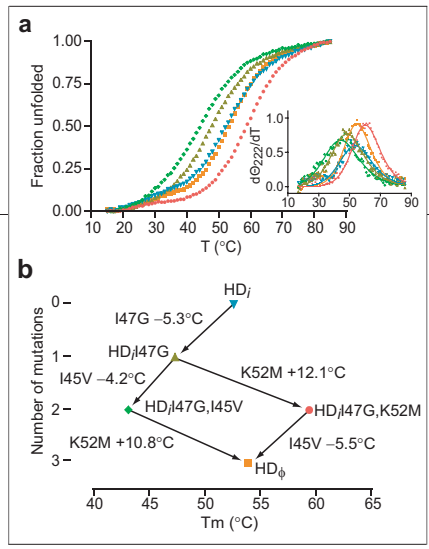

Figure 3. Analysis of the effect of HD_Φ mutations on HD stability. a) CD spectroscopy was used to monitor the thermal denaturation of HDs. The derivative of the CD curve was used to determine the melting temperature (inset) (*31*). b) Colors and symbols for each mutant along with the effects of mutations on the thermal stability of the mutant HDs.

conformations available to glycine (*29*), especially in an α helix (*30*). As predicted from the relief of repulsive charge–charge interactions, the K52M mutation is stabilizing, not only recovering the stability lost by I47G but further stabilizing HD_Φ V45I to a net 6.8 °C above HD_i. The dramatic stabilizing effect of the K52M mutation is consistent with the notion that selection favored HDs with stabilizing mutations, but more surprising, the other mutation, I45V, destabilizes the HD by 4–6 °C, depending on the sequence context, bringing the net thermal stability of HD_Φ close to that of HD_i (Figure 3, panel b, and Table 1). The phage selected mutations retuned the stability of the HD, leading to similar stabilities of HD_Φ with HD_i.

Despite the high conservation of HD–DNA contacts, particularly Gln/Lys50, Asn51, and Ile/Val47, using phage display and nucleoside chemistry, we were able to re-engineer the HD–DNA interface, completely disrupting one of the conserved contacts (Ile47-T4), and replace it with an elaborated nucleoside that packs against the cavity created by an I47G mutation. Although these modifications represent a dramatic perturbation, the high-resolution structure of this interface revealed that the selected mutant functions without affecting the HD fold. Furthermore, we found that the selected mutations tuned the stability of the selected HD to be similar to that of the starting HD. Analysis of the HD_Φ–TAA**1**CC interaction reported here demonstrates that even highly conserved interfaces, such as the HD–DNA interface, contain sufficient adaptability to allow the installation of novel function, in this case specific binding to modified DNA. This latent adaptability is of importance to protein engineers who wish to make tractable the enormity of the proteome by exploiting its conserved motifs and domains, including the HD.

METHODS

Synthesis. Detailed experimental procedures and characterization of the alkynes and nucleosides can be found in Supporting Information.

Phage Library Construction and Selection. The phage library used for these selections and the conditions for the selections have been published elsewhere (*20*).

Expression and Purification of Mutant HDs. Mutant HDs were expressed as MBP fusions, and the MBP affinity tag was cleaved using Factor Xa. A detailed description of the expression, purification, and characterization can be found in Supporting Information.

Electrophoretic Mobility Shift Assays. Binding of the HDs to DNA was determined essentially as described previously (*20, 24*), except by using HDs cleaved from the maltose binding protein tag (see above) and conditions described in Supporting Information.

CD. Thermal stability of engrailed HD mutants was measured by CD essentially as described previously (*19, 22*). Detailed conditions are described in Supporting Information. Melting temperature and enthalpy of denaturation were determined by fitting the derivative of the denaturation curves to the van't Hoff difference equation (*31*). Fitting the difference data obviates the need to fit baselines to the data, thus reducing the number of free parameters available to the fit. Thermal denaturation curves were numerically differentiated, smoothed over a 3 °C window, and fit to the van't Hoff difference equation using Levenberg–Marquardt least squares minimization (Figure 3, inset of panel a) using scripts written in Matlab (*32*).

Crystallization, Data Collection, and Structure Refinement. Crystals were grown in hanging drop essentially as described previously (*21*) except that a higher concentration of poly(ethylene glycol)-400 was used in the well solution. The HD_Φ–TAATCC structure was solved by molecular replacement using HD_i bound to TAATCC (*15*) as the initial model. The HD_Φ–TAATCC structure was then used as the initial model for HD_Φ bound to TAA**1**CC. Details of the crystallization, data collection, and structural refinement can be found in Supporting Information.

Accession Codes: Structure factors and final coordinates have been deposited in the RCSB PDB with ID codes 2HOT and 2HOS for HD_Φ bound to TAA**1**CC and TAATCC, respectively.

Acknowledgment: We thank L. Rice and C. Waddling for training and assistance with crystallography and structure refinement, the ALS 8.3.1 beamline, J. Chung and I. Gomez Pinto in the James Lab University of California, San Francisco (UCSF) for assistance with large-scale purification of the DNA strands, the Frankel lab (UCSF) for use of their DNA synthesizer, G. Weiss and K. Sato for assistance setting up the phage selections, C. Pabo for a critical discussion of this work, and R. Grant for providing unpublished structure factors for the Q50A HD mutant. M.D.S. was supported by fellowships from the National Science Foundation and the ACS Organic Division. M.E.F. was supported in part by National Institutes of Health (NIH) training grant number GM08284. Mass spectrometry studies were carried out at the UCSF Mass Spectrometry Facility supported by NIH grant number NCRR RR01614. Financial support was provided by the Volkswagen foundation.

Supporting Information Available: This material is free of charge *via* the Internet.

REFERENCES

1. Bishop, A. C., Shah, K., Liu, Y., Witucki, L., Kung, C., and Shokat, K. M. (1998) Design of allele-specific inhibitors to probe protein kinase signaling, *Curr. Biol. 8*, 257–266.
2. Bishop, A. C., Buzko, O., and Shokat, K. M. (2001) Magic bullets for protein kinases, *Trends Cell Biol. 11*, 167–172.
3. Bishop, A. C., Ubersax, J. A., Petsch, D. T., Matheos, D. P., Gray, N. S., Blethrow, J., Shimizu, E., Tsien, J. Z., Schultz, P. G., Rose, M. D., Wood, J. L., Morgan, D. O., and Shokat, K. M. (2000) A chemical switch for inhibitor-sensitive alleles of any protein kinase, *Nature 407*, 395–401.
4. Hwang, Y. W., and Miller, D. L. (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein, *J. Biol. Chem. 262*, 13081–13085.
5. Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., Zhang, C., and Shokat, K. M. (2000) Unnatural ligands for engineered proteins: new tools for chemical genetics, *Annu. Rev. Biophys. Biomol. Struct. 29*, 577–606.
6. Weijland, A., Parlato, G., and Parmeggiani, A. (1994) Elongation factor Tu D138N, a mutant with modified substrate specificity, as a tool to study energy consumption in protein biosynthesis, *Biochemistry 33*, 10711–10717.
7. Weijland, A., and Parmeggiani, A. (1993) Toward a model for the interaction between elongation factor Tu and the ribosome, *Science 259*, 1311–1314.
8. Atwell, S., Ultsch, M., De Vos, A. M., and Wells, J. A. (1997) Structural plasticity in a remodeled protein-protein interface, *Science 278*, 1125–1128.
9. Greisman, H. A., and Pabo, C. O. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites, *Science 275*, 657–661.

10. Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Design and selection of novel Cys2His2 zinc finger proteins, *Annu. Rev. Biochem. 70*, 313–340.
11. Segal, D. J., Dreier, B., Beerli, R. R., and Barbas, C. F., 3rd. (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences, *Proc. Natl. Acad. Sci. U.S.A. 96*, 2758–2763.
12. Kortemme, T., and Baker, D. (2004) Computational design of protein-protein interactions, *Curr. Opin. Chem. Biol. 8*, 91–97.
13. Allen, J. J., Lazerwith, S. E., and Shokat, K. M. (2005) Bio-orthogonal affinity purification of direct kinase substrates, *J. Am. Chem. Soc. 127*, 5288–5289.
14. Gehring, W. J., Affolter, M., and Burglin, T. (1994) Homeodomain proteins, *Annu. Rev. Biochem. 63*, 487–526.
15. Tucker-Kellogg, L., Rould, M. A., Chambers, K. A., Ades, S. E., Sauer, R. T., and Pabo, C. O. (1997) Engrailed (Gln50-Lys) homeodomain-DNA complex at 1.9 Å resolution: structural basis for enhanced affinity and altered specificity, *Structure 5*, 1047–1054.
16. Clarke, N. D., Kissinger, C. R., Desjarlais, J., Gilliland, G. L., and Pabo, C. O. (1994) Structural studies of the engrailed homeodomain, *Protein Sci. 3*, 1779–1787.
17. Fraenkel, E., Rould, M. A., Chambers, K. A., and Pabo, C. O. (1998) Engrailed homeodomain-DNA complex at 2.2 Å resolution: a detailed view of the interface and comparison with other engrailed structures, *J. Mol. Biol. 284*, 351–361.
18. Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B., and Pabo, C. O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions, *Cell 63*, 579–590.
19. Ades, S. E., and Sauer, R. T. (1994) Differential DNA-binding specificity of the engrailed homeodomain: the role of residue 50, *Biochemistry 33*, 9187–9194.
20. Simon, M. D., and Shokat, K. M. (2004) Adaptability at a protein-DNA interface: re-engineering the engrailed homeodomain to recognize an unnatural nucleotide, *J. Am. Chem. Soc. 126*, 8078–8079.
21. Grant, R. A., Rould, M. A., Klemm, J. D., and Pabo, C. O. (2000) Exploring the role of glutamine 50 in the homeodomain-DNA interface: crystal structure of engrailed (Gln50-Ala) complex at 2.0 Å, *Biochemistry 39*, 8187–8192.
22. Ades, S. E., and Sauer, R. T. (1995) Specificity of minor-groove and major-groove interactions in a homeodomain-DNA complex, *Biochemistry 34*, 14601–14608.
23. He, J., and Seela, F. (2002) Propynyl groups in duplex DNA: stability of base pairs incorporating 7-substituted 8-aza-7-deazapurines or 5-substituted pyrimidines, *Nucleic Acids Res. 30*, 5485–5496.
24. Simon, M. D., Sato, K., Weiss, G. A., and Shokat, K. M. (2004) A phage display selection of engrailed homeodomain mutants and the importance of residue Q50, *Nucleic Acids Res. 32*, 3623–3631.
25. Sato, K., Simon, M. D., Levin, A. M., Shokat, K. M., and Weiss, G. A. (2004) Dissecting the Engrailed homeodomain-DNA interaction by phage-displayed shotgun scanning, *Chem. Biol. 11*, 1017–1023.
26. Banerjee-Basu, S., and Baxevanis, A. D. (2001) Molecular evolution of the homeodomain family of transcription factors, *Nucleic Acids Res. 29*, 3258–3269.
27. Clarke, N. D. (1995) Covariation of residues in the homeodomain sequence family, *Protein Sci. 4*, 2269–2278.
28. Stollar, E. J., Mayor, U., Lovell, S. C., Federici, L., Freund, S. M., Fersht, A. R., and Luisi, B. F. (2003) Crystal structures of engrailed homeodomain mutants: implications for stability and dynamics, *J. Biol. Chem. 278*, 43699–43708.
29. Matthews, B. W., Nicholson, H., and Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663–6667.
30. Serrano, L., Neira, J. L., Sancho, J., and Fersht, A. R. (1992) Effect of alanine versus glycine in alpha-helices on protein stability, *Nature 356*, 453–455.
31. John, D. M., and Weeks, K. M. (2000) van't Hoff enthalpies without baselines, *Protein Sci. 9*, 1416–1419.
32. (2004) *Statistics Toolbox User's Guide*; The Mathworks: Natick, MA.