Creating Temperature-sensitive Winged Helix Transcription Factors

AMINO ACIDS THAT STABILIZE THE DNA BINDING DOMAIN OF HNF3

Kimberly Stevens, Lisa Cirillo, and Kenneth S. Zaret

From the Cell and Developmental Biology Program, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Winged helix transcription factors contain two polypeptide loops, or “wings,” that make minor groove contacts with DNA from either side of a three-helix bundle that binds the DNA major groove. While wing 1 is stabilized by a β-sheet, parameters that stabilize wing 2 are unknown. Herein we identify two bulky aromatic residues in wing 2 that stabilize the loop structure and, thereby, the entire protein’s DNA binding and transcriptional stimulatory activity by interacting with other residues in the three-helix bundle. Mutations of these wing 2 residues create proteins that are temperature-sensitive for transcriptional activity. Aromatic and/or hydrophobic residues are highly conserved among the 150 known winged helix proteins, suggesting conserved function. We suggest that the winged helix structure evolved by the acquisition of aromatic and/or hydrophobic residues in distal polypeptide sequences that helped stabilize the association of a protein loop (wing 2) with the three-helix bundle, thereby enhancing DNA binding.

The winged helix structural motif for DNA binding was defined only 7 years ago (1), yet over 150 genes for winged helix factors have been discovered (2). These proteins have recently been renamed “Fox” transcription factors, for Fôrkhed lex (3). Most of the known Fox factors have critical functions in signalizing or development, and at least one of the factors, Foxa1 (HNF3β) (3), binds to its target sequences on nucleosomes (4, 5). The chromatin binding properties of HNF3 are underscored by the structural resemblance of linker histone to the winged helix motif (6, 7). Because of the importance of winged helix factors in diverse areas of biological regulation, we have been investigating the unique features of the winged helix and how they may have evolved.

Winged helix/Fox proteins comprise a subset of the so-called wHTH proteins. All wHTH proteins contain a classic helix-turn-helix DNA binding motif within a three-helix bundle, braced against an antiparallel β-sheet (8) (see Fig. 1A). The β-sheet consists of at least two β-strands that occur C-terminal to the HTH segment (9–11), and in many wHTH proteins, the sheet contains one or two β-strands that precede the HTH (1, 6, 12–16) (see Fig. 1C). In the Fox protein subset of wHTH proteins, extension of a loop between β-strands creates a wing-like structure that can be used to bind DNA, as with wing 1 of the transcription factor HNF3 (1) (Fig. 1A). The presence of a wing 2 loop, C-terminal to the β sheet, is also critical for the Fox/winged helix proteins to bind DNA as a monomer (1, 3) (Fig. 1, A and C). Although wing 2 has neither helical nor sheet characteristics, it is well organized in the HNF3 crystal (1).

Wing 2 appears to be uniquely associated with the 150 winged helix factors related to HNF3 and does not occur in other wHTH proteins. Also, amino acid sequences within the wHTH domain, upstream of wing 2, are highly conserved within the winged helix proteins (2), so the presence of wing 2 is just one of several features that distinguish Fox/winged helix factors from other wHTH proteins. For example, although linker histone is a wHTH protein whose crystal structure has only a 1.3-Å root mean square deviation between equivalent α-carbon coordinates of the wHTH domain of HNF3 (1, 6), sequence identity with HNF3 over this domain is a minimal 13%, and the linker histone lacks wing 2. Thus, it appears that the presence of wing 2 and upstream amino acid changes have evolved together to create the winged helix subdomain of wHTH proteins. This would predict that at least some of the amino acids upstream of wing 2 might interact with wing 2 in specific ways.

Since the presence of wing 2 is a distinguishing feature of the Fox/winged helix wHTH factors, and because the wing 2 sequence is essential for DNA binding (3), we have investigated the structural role of this segment of the protein. We find that the structural stability of wing 2 and the wing’s function in DNA binding is highly sensitive to the presence of wing 2 hydrophobic and aromatic amino acids that are on the opposite face of the protein from that which contacts DNA. Our findings provide insight into the maintenance of the wing 2 loop structure and the creation of temperature-sensitive Fox transcription factors and suggest a mechanism by which the wing 2 loop may have been established in evolution.

EXPERIMENTAL PROCEDURES

Molecular Modeling and Sequence Comparisons—The hnf3g.pdb crystal structure coordinate file was kindly supplied by S. Burley (1). Structure was modeled with Swiss-PdbViewer version 3.1 (17) (Fig. 1A) and Esmond version 2.6 (R. Sayle) (Fig. 1B) running on a Micron 450 Mhz PC. The data base of winged helix factors and their sequence relationships was from the Pomona College site on the World Wide Web (maintained by W. Knöchel, K. Kaestner, and D. Martinez). Site-directed Mutagenesis and Protein Purification—The full-length mouse HNF3α cDNA in the Escherichia coli expression vector pET28b (Novagen) was used for mutagenesis (4). Site-directed mutagenesis employed overlapped extension polymerase chain reaction (18) as modified (19). DNA sequencing of the entire segments that were cloned from polymerase chain reaction products verified the sequence changes. Recombinant proteins were overexpressed in and purified from inclusion bodies in E. coli as described (20), which involved stepwise dilution from 6 M urea during the final purification.

Transient Transfections—For transient transfections, the wild type and HNF3α mutant cDNAs were inserted into the vector pR, down-stream of the Rous sarcoma virus promoter (21). The designated nanogram amounts of HNF3 expression plasmid were electroporated into mouse 3T3 cells (22) with 40 μg of pAN6-eGαs and 1 μg of pRTL1.
FIG. 1. Structural parameters governing HNF3 wing 2 stability. A and B, identical orientations of the HNF3-DNA crystal structure (1); DNA chain is in back. Amino acid segments are colored according to temperature factors in the crystal (red-orange-yellow; high; green-blue, low). A, backbone model. DNA strands are yellow; side chains are shown in magenta for selected surface hydrophobic residues. B, space-fill model. The patch of surface hydrophobic residues originally identified by Clark et al. (1) are labeled; other relevant features are noted. C, sequence of HNF3α and positions of conserved hydrophobic residues in winged helix proteins.

Figure 1 shows the structural parameters governing HNF3 wing 2 stability. A and B, identical orientations of the HNF3-DNA crystal structure (1); DNA chain is in back. Amino acid segments are colored according to temperature factors in the crystal (red-orange-yellow; high; green-blue, low). A, backbone model. DNA strands are yellow; side chains are shown in magenta for selected surface hydrophobic residues. B, space-fill model. The patch of surface hydrophobic residues originally identified by Clark et al. (1) are labeled; other relevant features are noted. C, sequence of HNF3α and positions of conserved hydrophobic residues in winged helix proteins.

pAN6-eG2 contains two HNF3 binding sites upstream of an albumin promoter fragment, the latter consisting of base pairs –56 to +8 with respect to the transcription start, fused to the neomycin resistance gene (23). pRT1 contains the Rous sarcoma virus promoter fused to the thymidine kinase gene and served as an internal control. Two days after transfection, total cellular RNA was isolated and subjected to primer extension analysis of neomycin and thymidine kinase RNAs in the same reaction (21). Gels were exposed to BAS2500 phosphor imager plates (Fuji) for quantitation. Signals from the HNF3 target plasmid were normalized to the signals from the control plasmid in each lane and estimated from the negative inverse of the slopes.

Footprinting reactions were as described (20) and employed 2.5 fmol of a 428-base pair end-labeled DNA probe of the albumin gene enhancer (5) in a 20-μl reaction volume with the designated concentrations of HNF3. The data in Fig. 3 depict the albumin enhancer eH binding site for HNF3 (21). The probe concentration was about 2-fold below the KD of wild type HNF3 for the eH site (20).

Aromatic residues, Phe254 and Tyr259, are near the distal bend of the loop. A van der Waals image shows that they contact the aromatic and hydrophobic residues near the distal bend of the loop. Two aromatic residues, Phe254 and Tyr259, are near the distal bend of the loop. A van der Waals image shows that they contact the aromatic and hydrophobic residues near the distal bend of the loop.

Aromatic residues, Phe254 and Tyr259, are near the distal bend of the loop. A van der Waals image shows that they contact the aromatic and hydrophobic residues near the distal bend of the loop. Two aromatic residues, Phe254 and Tyr259, are near the distal bend of the loop. A van der Waals image shows that they contact the aromatic and hydrophobic residues near the distal bend of the loop.

RESULTS

Protease Assays—Purified HNF3α proteins (4 μg) were suspended in a solution containing 20 mM Hepes (pH 6.5), 140 mM KCl, 1 mM MgCl2 5 mM dithiothreitol, 0.1% Nonidet P-40, and 18% glycerol in a total volume of 35 μL. The designated amounts of Staphylococcus aureus V8 protease, trypsin, proteinase K, or thermolysin (Sigma) were added, and the reactions proceeded for 30 min at either 37 or 12 °C. Digests were terminated with the addition of 15 μL of 3× SDS gel sample buffer (24). Samples were boiled, run on 12 or 15% polyacrylamide gels with SDS, and stained with Coomassie Brilliant Blue. Gels were scanned with a Umax device and imaged with Picture Publisher 7.0 (Micrografia). Digestion data was quantitated with Fuji ImageGauge software.

DNase Footprinting—Footprinting reactions were as described (20) and employed 2.5 fmol of a 428-base pair end-labeled DNA probe of the albumin gene enhancer (5) in a 20-μl reaction volume with the designated concentrations of HNF3. The data in Fig. 3 depict the albumin enhancer eH binding site for HNF3 (21). The probe concentration was about 2-fold below the KD of wild type HNF3 for the eH site (20). Footprinting gels were scanned with a Fuji BAS2500 phosphor imager, and the percentage of occupancy of the eH HNF3 site was determined relative to an internal control band in the same lane (out of the image shown in Fig. 3). Average and S.D. values of percentage of occupancy were determined from two or three independent experiments at each protein concentration. Scatchard plots were drawn from linear regression analysis of the relevant data points, and apparent KD values were estimated from the negative inverse of the slopes.

Modeling and Mutagenesis of Wing 2—In order to understand how the wing 2 segment is stably incorporated into the tertiary structure of the winged helix domain, we employed molecular modeling with the crystal structure coordinates of HNF3γ (1). Note that the numbering system we use is for Foxa1(HNF3α), which has six conservative substitutions over the 110-amino acid sequence of the winged helix domain of HNF3γ; none of these differences are close to the residues studied here. As seen in Fig. 1A, in which the DNA interacting surface is away from view, the polypeptide backbone of wing 2 consists of an elongated loop at the C terminus, beneath the three-helix bundle. Two aromatic residues, Phe254 and Tyr259, are near the distal bend of the loop. A van der Waals image shows that they contact the aromatic and hydrophobic residues in wing 2. All of these residues are part of a larger, interacting set of aromatic and hydrophobic amino acids that occur in a depression on the surface of HNF3 (Fig. 1B), as first noted by Clark et al. (1). It would seem that these surface residues would be important for protein-protein interactions. Although these residues are well spread out in the primary sequence (Fig. 1C), aromatic and hydrophobic amino acids are highly conserved at these posi-

Footprinting reactions were as described (20) and employed 2.5 fmol of a 428-base pair end-labeled DNA probe of the albumin gene enhancer (5) in a 20-μl reaction volume with the designated concentrations of HNF3. The data in Fig. 3 depict the albumin enhancer eH binding site for HNF3 (21). The probe concentration was about 2-fold below the KD of wild type HNF3 for the eH site (20). Footprinting gels were scanned with a Fuji BAS2500 phosphor imager, and the percentage of occupancy of the eH HNF3 site was determined relative to an internal control band in the same lane (out of the image shown in Fig. 3). Average and S.D. values of percentage of occupancy were determined from two or three independent experiments at each protein concentration. Scatchard plots were drawn from linear regression analysis of the relevant data points, and apparent KD values were estimated from the negative inverse of the slopes.

Modeling and Mutagenesis of Wing 2—In order to understand how the wing 2 segment is stably incorporated into the tertiary structure of the winged helix domain, we employed molecular modeling with the crystal structure coordinates of HNF3γ (1). Note that the numbering system we use is for Foxa1(HNF3α), which has six conservative substitutions over the 110-amino acid sequence of the winged helix domain of HNF3γ; none of these differences are close to the residues studied here. As seen in Fig. 1A, in which the DNA interacting surface is away from view, the polypeptide backbone of wing 2 consists of an elongated loop at the C terminus, beneath the three-helix bundle. Two aromatic residues, Phe254 and Tyr259, are near the distal bend of the loop. A van der Waals image shows that they contact the aromatic and hydrophobic residues in wing 2. All of these residues are part of a larger, interacting set of aromatic and hydrophobic amino acids that occur in a depression on the surface of HNF3 (Fig. 1B), as first noted by Clark et al. (1). It would seem that these surface residues would be important for protein-protein interactions. Although these residues are well spread out in the primary sequence (Fig. 1C), aromatic and hydrophobic amino acids are highly conserved at these posi-
Temperature-sensitive Winged Helix Transcription Factors

Fig. 2. Temperature-sensitive HNF3 proteins in cells. The designated WT and mutant HNF3 constructs were expressed in 3T3 fibroblasts for 2 days. A and B, RNA analysis. The data in B represent the mean and S.D. of two transfection experiments at each temperature. The black bars denote experiments at 37 and 39 °C, for which the data were similar. C and D, Western blot analysis of HNF3 proteins in transfected cells at 30 °C. The designated amounts of whole cell lysates were analyzed. An asterisk denotes a nonspecific band, which serves as a loading control. The data in D represent the mean and S.D. of two transfection experiments at each temperature. E, the bars show the percentage of transcription relative to wild type divided by the percentage of steady state level of transfected wild type protein and thus indicate the actual transcriptional stimulatory activity for each mutant HNF3.

Stimulations among the vast majority of the Fox/winged helix proteins (2), suggesting that they are critical for the proteins' structure or function or both.

As might be expected, the residues in the HNF3-DNA co-crystal with the highest temperature factors (B-factors) are at the distal loop bends of the wing domains (Fig. 1, A and B; yellow, orange, and red). Interestingly, Phe254 and Tyr259 in wing 2 separate the high B-factor residues at the distal loop bend from the low B-factor residues (green and blue) in the remainder of the wing. This suggests that Phe254 and Tyr259 stabilize the association of the wing 2 loop with the wHTH domain, perhaps by associating with Tyr173 and Met181 in the cluster of aromatic and hydrophobic residues. Without this interaction, temperature sensitivity of the entire loop would be predicted to increase, with detrimental consequences for winged helix domain function.

We sought to test these possibilities by creating site-directed mutants of HNF3Δo with alanine substitutions at either Tyr173 and Met181 (YM), Phe254 and Tyr259 (FY), or all four residues (YMFY). We anticipated that the FY mutation would primarily affect wing 2 stability and that the YM and YMFY mutations would affect both wing 2 stability and the stability of the HLH globular domain. If the YM and FY mutations were found to have similar effects on the protein, it would suggest that the primary role of the YM and FY amino acids is to maintain wing 2 structure. If YM and FY have different effects, and if YMFY is more destabilizing than either pair of changes, it would suggest that the YM and FY residues have additional or different roles from each other in maintaining winged helix domain stability.

Wing 2 Amino Acids That Are Essential for Transcriptional Stimulation—To test whether the YM and FY mutations affect HNF3 stability and/or function, we transiently transfected HNF3 expression plasmids into mouse 3T3 fibroblast cells, which lack HNF3. Transcriptional stimulation was monitored from another plasmid that contains two HNF3 target sites juxtaposed to a minimal albumin gene promoter. An HNF3 nonresponsive promoter in a third plasmid served as an internal transfection control. We cultured the transfected cells at 30 °C and 37–39 °C in order to determine if the temperature was a critical determinant for HNF3 stability or activity. We anticipated that the 30 °C temperature would be permissive for a temperature-sensitive function. RNA was isolated after 2 days and subjected to primer extension analysis (Fig. 2A). Two different levels of plasmid DNA were transfected, which showed that there was a linear transcriptional response over the range of HNF3 levels that were quantitated. The HNF3 target gene expression was normalized to the internal control signal, and this value was expressed as the percentage of activity of wild type (WT)1 HNF3, using data from the lower level of transfected HNF3 plasmid at 30 °C and the higher level of plasmid at 37–39 °C. Data from two independent transfection experiments were quantitated with a phosphor imager as shown in Fig. 2B. The values at 37 and 39 °C were very similar and were grouped together (see small error bars in Fig. 2B) To quantitate transfected protein levels, we performed Western blot analysis. Two different amounts of whole cell lysates were applied to SDS gels, and several nonspecific bands in each lane of the resulting chemiluminescence film served as a loading control (Fig. 2C, bands with asterisk). HNF3 bands were quantitated by scanning films of different exposure levels and from different amounts of loaded protein. Data from two independent transfection experiments were quantitated and are shown in Fig. 2D. As will be documented below, we found that the mutants exhibited temperature sensitivity for transcriptional stimulation that was not reflected by changes in steady-state protein levels.

At 30 °C, the transfected HNF3-YM mutant resulted in wild type levels of transcription, whereas the wing 2 loop mutants FY and YMFY reduced transcription by about half (Fig. 2, gray bars). By contrast, when the cells were transfected at 37–39 °C, the YM mutant gave about 30% of wild type transcription, the FY mutant gave about 10% of wild type transcription, and the YMFY mutant gave no detectable transcription (Fig. 2, black bars). In Western blots of titrations of whole cell lysates from cells at 30 °C, the abundance of the HNF3 mutant proteins generally reflected their transcriptional output. The YM mutant was at the wild type level of protein, and the FY and YMFY mutants were 67 and 44% of the wild type level, respectively (Fig. 2D, gray bars). Therefore, when the transcriptional data at 30 °C are normalized to the protein expression data, the YM and YMFY mutant proteins exhibited wild type transcriptional stimulation activity, and the FY mutant

1 The abbreviation used is: WT, wild type.
was close to the wild type level, at 75% (Fig. 2E).

At 37 °C, the levels of all HNF3 proteins, including wild type, were lower than that at 30 °C (Fig. 2C), which we attribute to the lower transient transfection efficiency at the higher temperature. Still, the relative level of the YM mutant protein was the same as wild type, the FY mutant was about 90% of wild type, and the YMFY mutant was about 40% of wild type, which is very close to the relative levels seen at 37 °C (Fig. 2D). Although the YMFY protein is difficult to see in the reproduction in Fig. 2C, lane 17, it was evident on the original autoradiographs and scans thereof. If amino acids Phe254 and Tyr259 interacted primarily with Tyr173 and Met181 to maintain wing 2 domain stability, either set of mutations should be equally destabilizing. However, the YM mutations were neutral, and the FY mutations decreased the steady-state level of the protein, albeit only partially (Fig. 2D), in a non-temperature-sensitive fashion. These data, and the fact that the YM mutation in the YMFY mutant enhances the destabilizing effect of the FY mutation, suggest that although residues Tyr173 and Met181 may interact with Phe254 and Tyr259 to promote wing 2 stability in vivo, FY is a more critical determinant of overall steady state protein levels.

Interestingly, when the transcriptional data are normalized to the protein expression data at 37 °C and compared with that at 30 °C, the YM, FY, and YMFY proteins are markedly temperature-sensitive for transcriptional activation (Fig. 2E). The YM and FY proteins are 4- and 6-fold deficient at higher temperature, compared with wild type, while the YMFY mutant appears completely temperature-sensitive for activity. The mutants’ intrinsic temperature sensitivity for transcriptional activation suggests that the YM and FY residues confer far more temperature-sensitive changes in the protein (wild type, FY, and YM). V8 protease specifically recognizes certain residues in the YM and FY loops and exhibits no difference between WT and FY proteins and consequently, its activity, than was reflected in the steady-state levels of the proteins. On this basis, we next investigated structural aspects of the protein in vitro.

**V8 Protease as a Probe of Wing 2**—Because wing 2 lacks helical or sheet characteristics that are readily detected by physical techniques, we used enzymatic probes to compare the structure of wild type and mutant HNF3α. V8 protease in particular has a potential cleavage target in the distal loop of wing 2, in that Glu255 is exposed to solvent (Fig. 1, A and C); therefore, V8 could be a probe of wing 2 structure. In initial experiments, we found that the FY mutant, which contains Phe254 changed to an alanine in the residue that precedes Glu255, had an altered V8 protease digestion pattern compared with wild type HNF3α (Fig. 3A). Note that the starting material for these proteins, as purified from E. coli, migrates as a doublet, designated band a (Fig. 3A, lanes 2 and 5); the larger protein of the doublet may be due to improper translation termination. While the wild type protein yielded prominent protease-resistant fragments designated b–e (Fig. 2A, lanes 3 and 4), the FY variant yielded much lower levels of fragments c–e and instead had elevated levels of fragments designated x and y (Fig. 3A, lanes 6 and 7). Note that the x and y bands are also generated from the wild type protein and that part of the band intensity with fragment b is due to V8 protease (see Fig. 4B, lane 20). The YMFY mutant had the same altered V8 cleavage pattern as the FY mutant, while the YM mutant looked like wild type (data not shown and Fig. 4A below). We conclude that the Phe254 change to alanine upstream of the V8 target residue Glu255 caused a change in V8 cleavage specificity. Since this change altered the abundance of one of the four prominent V8 digest products (c, d, and e to x and y), wing 2 must be a prominent cleavage target for the enzyme.

The specificity of V8 for wing 2 is supported by digestion studies with trypsin, which lacks cleavage targets in the wing 2 loop and exhibits no difference between WT and FY proteins in the generation of 12 different proteolytic fragments (Fig. 3B, lanes 4, 5, 8, and 9, bands with black dots). A similar lack of difference in cleavage pattern was obtained when the YM and YMFY proteins were digested with trypsin and when the WT, YM, FY, and YMFY proteins were digested with thermolysin and proteinase K (data not shown). Thus, the YM and FY mutations do not induce conformational changes throughout the HNF3 protein. We conclude that V8 specifically recognizes
Temperature-sensitive Winged Helix Transcription Factors

FIG. 5. Wing 2 amino acids essential for DNA binding. A and B, portions of DNase footprinting gels are shown, along with the designated nanomolar concentrations of HNF3 and mutant proteins. A, lane G shows a G-sequence cleavage marker of the probe DNA. The DNA protected by HNF3 and the HNF3-induced hypersensitive site is marked by the bracket. C, summary of at least two independent footprinting experiments. Closed and open dots indicate the mean percentage of occupancy of the DNase footprints at 12 and 20 °C, respectively. Vertical bars indicate S.D. D, Scatchard plot analysis of the footprinting data.

wing 2 and that the amino acid changes studied here do not grossly affect HNF3 structure.

Wing 2 Residues Imparting Thermal Stability—We hypothesized that the YM and FY mutations might alter the structure of wing 2 or its relation to the three helix bundle, apart from V8 enzyme cleavage specificity. If such were the case, the mutants might exhibit temperature sensitivity for V8 digestion. At 12 °C, increasing amounts of enzyme up to 0.4 μg/reaction yielded 30–60% loss of the starting full-length proteins (band a) (Fig. 4, A and C). At 37 °C, the FY and YMFY mutants, but not the wild type protein, were more protease sensitive than they were at 12 °C (Fig. 4, B and C). Specifically, at 37 °C the wild type protein consistently yielded full-length material in the presence of up to 0.1–0.4 μg of enzyme, but the full-length band a of the mutants was mostly degraded with 0.1 μg of protease (Fig. 4B). The YM mutant was measurably protease sensitive but still had some full-length protein at 0.4 μg of protease (Fig. 4B, lane 18). This correlates well with the transcriptional stimulatory activity of the transfected proteins, where the FY and YMFY mutants were more temperature sensitive than the YM mutant (Fig. 2E). We conclude that bulky aromatic residues at the FY positions contribute to the thermal stability of the protein. Furthermore, since these studies were conducted with purified protein, these residues must help maintain structural conformation in the wing 2 region in the absence of DNA.

Temperature-sensitive DNA Binding—If the FY residues and to some extent the YM residues are critical for the structure of wing 2, then alanine substitutions would be predicted to perturb DNA binding. In quantitative DNase footprinting assays with a DNA probe containing an HNF3 binding site, we found that the YM, FY, and YMFY mutants were indeed reproducibly deficient in site-specific binding to DNA at the standard reaction temperature of 20 °C, compared with wild type (Fig. 5A). We could visualize this in two ways, as seen in portions of the footprinting gels in Fig. 3. First, all three mutants were defective in protecting DNA at the HNF3 binding site; saturation at 2 μM protein was achieved only with the wild type protein (Fig. 5A, lanes 4, 7, 10, and 13, bracketed region). Second, all three mutants were impaired in generating the characteristic DNase hypersensitive cleavage within the HNF3 footprint. Hypersensitivity may be due to the ability of the winged helix factor to slightly open the minor groove on the DNA side that is opposite to that on which the protein binds (1, 25), thereby enhancing DNase access. Quantitative phosphor image analysis of multiple footprinting experiments was used to determine binding site saturation (Fig. 5C) and a general estimate of the apparent $K_d$ of the proteins for DNA (Fig. 5D). We note that in the experiments used for quantitation, the DNA substrate concentration was below the wild type HNF3 apparent $K_d$. The apparent $K_d$ of wild type mouse HNF3α for the eh1 binding site, under these conditions at 20 °C, was 0.42 nM ($r^2 = 0.94$), which agrees well with the 0.65 nM value determined previously with the rat HNF3α, which was a slightly less active protein (20).

The quantitation analysis showed that at 20 °C, the YM, FY, and YMFY mutants were increasingly deficient in DNA binding site saturation (Fig. 5C, dashed lines) and in apparent $K_d$ (Fig. 5D, slope of dashed lines relative to that of wild type HNF3). Specifically, while the YM mutant had a 2-fold reduction in affinity for the HNF3 target site, the FY mutant had a 2–4-fold lower affinity, and the YMFY mutant had a 10–20-fold lower affinity. However, when the same proteins were tested for binding site occupancy at the lower temperature of 12 °C (Fig. 5B), the YM mutant exhibited the wild type affinity (Fig. 5, C and D). Thus, the YM protein is temperature-sensitive for DNA binding in vitro, similar to the temperature sensitivity of YM function in vivo (Fig. 2E). In contrast, the FY and YMFY mutants were nearly as deficient in DNA binding at 12 °C as they were at 20 °C (Fig. 5, C and D). Since mutation of the FY residues leads to more severe phenotypes than mutation of the YM residues, in various functional assays, the Phe254 and Tyr259 residues must interact with residues in addition to YM to maintain the structural integrity of wing 2 in the winged helix domain. We suggest that improper refolding of the renatured FY protein in vitro, during purification, leads to a deficiency in DNA binding that is not compensated by a lower temperature. We conclude that the wing 2 residues FY, as well as the YM residues within the HTH domain, all of which are on the non-DNA binding surface of the protein, are critical to allow stable DNA binding by the winged helix protein.

DISCUSSION

Considering that the ~150 known Fox/winged helix proteins have extensive sequence similarity throughout the entire 110-amino acid domain and that the sequence conservation distinguishes the family from other wHTH proteins, there are likely
to be a variety of sequence determinants that impart the distinct regulatory activities of these proteins. We investigated the winged helix domain’s most striking structural characteristic, wing 2, by focusing on how the wing 2 loop is stabilized and thereby may help the protein bind DNA. Our data indicate that two aromatic residues, Phe\textsuperscript{254} and Tyr\textsuperscript{259}, near the loop bend of wing 2 contribute to the stability of wing 2 even in the absence of DNA, apparently by interacting with residues Tyr\textsuperscript{173} and Met\textsuperscript{181} and others in the three-helix bundle of the protein. We find that the stabilization of wing 2 structure is critical for DNA binding activity and therefore for the transcriptional function of the winged helix protein.

We speculate that hydrophobic and aromatic ring system interactions between Phe\textsuperscript{254} and Tyr\textsuperscript{259} residues in the wing 2 loop and Tyr\textsuperscript{173} and Met\textsuperscript{181} and other residues in the three-helix bundle (Fig. 1B) are critical to maintain wing 2 structural integrity. In support of this, of 144 winged helix proteins with a completely sequenced winged helix domain in the data base, 97, 91, and 94% have hydrophobic or aromatic residues at positions corresponding to 173, 181, and 254 in HNF3\textalpha, respectively; and 83% of the proteins have a hydrophobic or aromatic residue at position 259. Of the five winged helix proteins that have charged residues corresponding to position 181, four are deleted for five amino acids at the distal loop region, making it unlikely that they form a wing 2 structure. One winged helix protein, daf16b (26), that has a charged (lysine) residue at 181 and that has extensive wing 2 sequences has an oppositely charged glutamate at 173, an arginine at 254, and a threonine at 259. Thus, the putative wing 2 in this protein may be stabilized primarily by opposing charge interactions rather than by hydrophobic and aromatic residue interactions.

The YM mutation exhibited temperature-sensitive DNA binding \textit{in vitro}, in accord with the temperature sensitivity of transcriptional activation \textit{in vivo}. However, HNF3 proteins containing the FY mutation were deficient in DNA binding at both temperatures tested \textit{in vitro}. By all criteria tested, the FY mutations had more severe phenotypic defects than the YM mutations, indicating that the residues are more critical to maintain protein integrity. We note that the HNF3 proteins used for \textit{in vitro} footprinting were renatured from insoluble precipitates in \textit{E. coli}. We therefore suggest that improper folding of the winged helix domain in the FY and YMFY mutants \textit{in vitro} leads to deficiencies in their recovery of DNA binding activity, in comparison with the less severe YM mutation. By contrast, the normal folding pathways inside the cell, which may involve chaperone proteins, apparently allow reconstitution of DNA binding and transcriptional activity of FY mutants at the permissive temperature of 30 °C.

Considering the high degree of conservation of aromatic and hydrophobic residues at the respective Tyr, Met, Phe, and Tyr positions in all winged helix proteins, it seems likely that mutation of homologous amino acids in other winged helix family members will also create temperature-sensitive phenotypes \textit{in vivo}. Given the different ranges of temperature-sensitive transcriptional activation exhibited by the YM, FY, and YMFY mutations (Fig. 2, B and C), it may be possible to generate mutants of other winged helix factors that are temperature-sensitive over different ranges of activity. Each class of mutant could be useful in a different regulatory context.

Our findings suggest a model for the evolution of the winged helix domain from the more general wHTH structure (Fig. 6). We propose that a primordial wHTH protein may have acquired, by mutation, bulky aromatic and/or hydrophobic amino acids in unstructured sequences that were distal to the wHTH domain. The distal polypeptide segment then could associate with aromatic and/or hydrophobic residues in the three-helix bundle that themselves contribute to the stability or function of the wHTH domain, thus creating wing 2. Further mutations might have strengthened the association and conferred a functional advantage to the protein. For example, some wHTH proteins function as dimers, and the appearance of wing 2 would extend the DNA binding face of the protein so that the resultant winged helix domain could bind efficiently as a monomer. As noted in the Introduction, the deletion of wing 2 prevents HNF3 binding to DNA, despite the presence of sequences for wing 1 and the three-helix bundle. The deficiency in the FY mutants in inducing DNase hypersensitivity within the HNF3 footprint shows that the wing 2 may help the protein stabilize a conformational change in DNA that is essential for monomeric binding.

What selective forces could explain the persistence of a surface hydrophobic patch that apparently maintains wing 2 structure for the majority of winged helix factors? A hydrophobic patch within the wHTH motif may not only provide stability to the three-helix bundle; it may also serve as a site for interacting with another protein important for transcriptional regulation. The presence of two bulky hydrophobic residues in wing 2 with four such residues on the back of the wHTH domain expands the surface hydrophobic patch substantially (Fig. 1B). Thus, it seems likely that multiple selective forces maintain the hydrophobic patch within the wHTH domain, providing a base for distal strand recruitment into wing 2 to create a thermostable winged helix structure. Similar mechanisms may account for the creation and maintenance of other tertiary structures from primordial structural motifs.

\textbf{Acknowledgments}—We thank Stephen Burley for crystal structure coordinates; Gail Levine for technical assistance; Nicolas Guex for advice on using Swiss-Pdb viewer; Art Landy, Joanne Yeh, and Randy Strich for comments; and Elaine Butler and Sarah Costello-Berman for help in preparing the manuscript.

\textbf{REFERENCES}

1. Clark, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993) Nature \textbf{364}, 412–420
2. Kaufmann, E., and Knochel, W. (1996) Mech. Dev. \textbf{57}, 3–20
3. Lai, E., Prezioso, V. R., Smith, E., Litvin, O., Costa, R. H., and Darnell, J. E., Jr. (1990) *Genes Dev.* 4, 1427–1436
4. Cirillo, L. A., McPherson, C. E., Bossard, P., Stevens, K., Cherian, S., Shim, E.-Y., Clark, E. A., Burley, S. K., and Zaret, K. S. (1998) *EMBO J.* 17, 244–254
5. Shim, E. Y., Woodcock, C., and Zaret, K. S. (1998) *Genes Dev.* 12, 5–10
6. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993) *Nature* 362, 219–224
7. Cerf, C., Lippens, G., Ramakrishnan, V., Muyldermans, S., Segers, A., Wyns, L., Wodak, S. J., and Hallenga, K. (1994) *Biochemistry* 33, 11079–11086
8. Brennan, R. (1993) *Cell* 74, 773–776
9. Wilson, D. B., Ghysdael, J., and Owen, M. J. (1992) *Nucleic Acids Res.* 20, 699–704
10. Fogh, R., Oettleben, G., Ruterjans, H., Schnarr, M., Boelens, R., and Kaptein, R. (1994) *EMBO J.* 13, 3936–3944
11. Qiu, X., Verlinde, C., Zhang, S., Schmitt, M., Holmes, R., and Hol, W. (1995) *Structure* 3, 87–101
12. Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) *Science* 253, 1001–1007
13. Damberger, F., Pelto, J., Harrison, C., Nelson, H., and Wemmer, D. (1994) *Protein Sci.* 3, 1806–1821
14. Harrison, C., Bohm, A., and Nelson, H. (1994) *Science* 263, 224–226
15. Vuister, G., Kim, S.-J., Wu, C., and Bax, A. (1994) *Biochemistry* 33, 10–16
16. Donaldson, L. W., Petersen, J. M., Graves, B. J., and McIntosh, L. P. (1996) *EMBO J.* 15, 125–134
17. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* 18, 2714–2723
18. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene (Amst.)* 77, 51–59
19. Mikaelian, I., and Sergeant, A. (1992) *Nucleic Acids Res.* 20, 376
20. Zaret, K., and Stevens, K. (1995) *Protein Expression Purif.* 6, 821–825
21. Jackson, D. A., Rowader, K. E., Stevens, K., Jiang, C., Milos, P., and Zaret, K. S. (1993) *R. Soc. London, Ser. B, Biol. Sci.* 2491–2410
22. Zaret, K. S., Dipersio, C. M., Jackson, D. A., Montigny, W. J., and Weinstat, D. L. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 9076–9080
23. Milos, P. M., and Zaret, K. S. (1992) *Genes Dev.* 6, 99–1004
24. Laemmli, U. K. (1970) *Nature* 227, 680–685
25. Pierron, S., Helleqvist, M., Samuelsson, L., Enerbäck, S., and Carlsson, P. (1994) *EMBO J.* 13, 5002–5012
26. Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A., and Ruvkun, G. (1997) *Nature* 389, 994–999