Structure Comparison of Two Conserved HNF-3/fkh Proteins HFH-1 and Genesis Indicates the Existence of Folding Differences in Their Complexes with a DNA Binding Sequence†,‡

Wanyun Sheng, Mark Rance,§ and Xiubei Liao*,||

Department of Biochemistry and Molecular Biology, College of Medicine, University of Illinois at Chicago, 1835 West Polk Street, Chicago, Illinois 60612-4316, and Department of Molecular Genetics, Biochemistry, and Microbiology, University of Cincinnati, 213 Albert Sabin Way, Cincinnati, Ohio 45267-0524

Received October 9, 2001

ABSTRACT: The hepatocyte nuclear factor 3 (HNF-3)/fork head (fkh) family contains a large number of transcription factors that recognize divergent DNA sequences via a winged helix binding motif. HNF-3/fkh proteins show a broad profile of DNA sequence-specificity in which one DNA sequence can be recognized by more than one HNF-3/fkh protein and each individual HNF-3/fkh protein has several DNA binding sequences. In this study, heteronuclear NMR methods were used to study the structures of the DNA binding domain of a conserved winged helix protein HFH-1 and its DNA complexes. The structural comparison of winged helix proteins HFH-1 and Genesis and their DNA complexes indicates that even two highly conserved HNF-3 family members can adopt different local structures when they contact an identical DNA binding sequence, while one of these two HNF-3 proteins seems to adopt only slightly different structures on different DNA binding sites.

Transcription factors play a central role in cell growth, development, and differentiation by binding to specific DNA control elements to regulate gene expression. To understand transcription control, the structures of many transcription factors and their DNA complexes have been determined by NMR spectroscopy and X-ray crystallography. Several highly conserved structural motifs were identified from these studies, and their DNA contact mechanisms were elucidated (*1*,*2*).

One well-studied DNA binding domain is the winged helix motif, which is contained in a large number of DNA binding proteins and functions as the DNA binding domain of many transcription factors (*3*). The winged helix DNA binding motif was initially identified in the gene products of the *Drosophila homeotic* forkhead (fkh[1]) and hepatocyte nuclear factor 3 (HNF-3) (*4*, *5*). Since the initial discoveries, many proteins containing winged helix motifs homologous to HNF-3/fkh have been identified in a wide range of organisms; the functions of these proteins are linked to gene regulation during embryogenesis and tumorigenesis and in differentiated cell states (*3*). The structural determinations of winged helix proteins have shown that the motif contains three or four α-helices, a three-stranded *β*-shee, and two flexible loops termed wings (*6*−*10*). These proteins bind to DNA as monomers and use helix 3 (H3) as the principal element to recognize the DNA major groove, with the two wings making additional DNA contacts.

Although the base-specific DNA contact residues are highly conserved in the HNF-3/fkh family members, each individual protein shows a distinct DNA binding profile. Previous PCR-mediated binding sequence selection experiments have shown that each individual winged helix protein recognizes more than one DNA sequence and that usually one DNA sequence can be bound by more than one winged helix protein (*11*, *12*). An important question then arises regarding how the cellular transcriptional machinery can target a particular winged helix protein−DNA complex for further assembly to achieve its proper function in vivo.

In this report, structural studies of HFH-1 and two HFH-1/DNA complexes are presented. The structures of HFH-1 are also compared with the structures of Genesis. HFH-1 and Genesis (originally named HFH-2) are two homologous winged helix proteins (*13*, *14*). PCR-mediated selection and gel shift experiments show that HFH-1 and Genesis have both shared and distinct DNA binding sequences (*11*). By comparing the structures of the two proteins free of and bound to DNA, the data show that these two highly conserved proteins adopt slightly different local conformations even when they bind to the same DNA sequence. The data also indicate that when HFH-1 binds to two different

† This work was funded by the NIH (GM 52034 to X.L. and GM 40089 to M.R.). The Bruker DRX600 was purchased with funds from the University of Illinois at Chicago and a grant from the NSF Academic Research Infrastructure Program (BIR 9601705). These studies made use of the National Magnetic Resonance Facility at Madison, which is supported by the National Institutes of Health and the National Science Foundation.

‡ The atomic coordinate file of the HFH-1 protein has been deposited in the Brookhaven Protein Data Bank (code 1KQ8).

* To whom correspondence should be addressed. E-mail: xiubei@uic.edu. Tel: (312) 996-7672. Fax: (312) 413-0364.

|| University of Illinois at Chicago.

§ University of Cincinnati.

[1] Abbreviations: HNF, hepatocyte nuclear factor; fkh, forkhead; HFH, hepatocyte nuclear factor/forkhead homologue; NMR, nuclear magnetic resonance; T1, longitudinal relaxation time constant; T2 transverse relaxation time constant; HSQC, heteronuclear single quantum correlation; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy.

10.1021/bi011908k CCC: $22.00
Published on Web 02/16/2002

DNA sequences it adopts similar structures. Thus, the data indicate the possibility that local structural differences between two homologous proteins in a complex with an identical DNA control sequence may be responsible for different in vivo activities of the two proteins.

MATERIALS AND METHODS

Protein Expression and Purification. The gene construction, protein expression, and purification of Genesis for NMR studies were described previously (*8, 9*). The gene encoding the DNA binding domain of HFH-1 (*11*), which in this report will henceforth be referred to only as HFH-1, was generated by PCR amplification from a rat genomic clone (*14*). The endonuclease recognition sites *Nde*I and *Xho*I were engineered for the cloning of HFH-1 into pET21b (Novagen). The expressed protein contains the functional DNA binding domain with an extra Met at the N terminus and a 6-His tag at the C terminus. The expression and purification of the HFH-1 protein were identical to those of the Genesis protein, as described previously (*8, 9*). Purified HFH-1 and Genesis were exchanged into NMR buffer (50 mM Na_2PO_4, pH 6.5, 100 mM NaCl, and 10 mM $Na_2S_2O_4$ in 10% D_2O/90% H_2O) and stored at 4 °C. Protein concentrations were between 0.7 and 1 mM, as measured by the BioRad protein assay using bovine albumin as the standard (BioRad Laboratories, Hercules, CA).

Protein−DNA Complex Formation. Two 17 base-pair double-stranded DNA sequences (HFH-2#12, 5′-GCTTAAAATAACAATAC-3′ and HFH-1#25, 5′-GCTACATAAACAACGCT-3′) were purchased from Oligos Etc. Inc., (Wilsonville, OR). HFH-2#12 is a strong binding site for Genesis and a moderate binding site for HFH-1, while HFH-1#25 is a strong binding site for HFH-1 and a weak binding site for Genesis (*11*). The formation of Genesis with HFH-2#12 was described previously (*8, 9*). The identical procedure was used to form the two HFH-1/DNA complexes. All three complexes were exchanged into NMR buffer (10 mM Tris-d_6, 20 mM NaCl, and 10 mM $Na_2S_2O_4$ in 10% D_2O/90% H_2O) by ultrafiltration.

NMR Spectroscopy. NMR spectra of the free proteins and their DNA complexes were recorded either on Bruker DMX 600 or DMX 750 or on Varian Inova 800 NMR spectrometers. The standard double and triple resonance experiments were used in the assignments of HFH-1 and its DNA complexes on single-labeled (^{15}N), double-labeled (^{15}N and ^{13}C) and triple-labeled (^{15}N, ^{13}C, and ^{2}H) proteins. (*15−20*). These experiments were performed on the free proteins at 17 °C and on the complexes at 25 or 30 °C. $\{^1H\}^{15}N$ heteronuclear NOE experiments were acquired at 17 °C on DNA-free HFH-1 according to a published pulse sequence (*21*). The $\{^1H\}^{15}N$ heteronuclear NOE values were measured from the differences between spectra acquired with and without saturation of protons, where saturation was achieved by a train of 120° flip-angle pulses separated by 5 ms for 3 s (*22*). The extraction of heteronuclear NOE values was described previously (*23*). The NMR data were processed and analyzed using either NMRPipe (NIH) (*24*) or the commercial software SYBYL (Tripos Inc.).

Secondary Structure Identification. The identification of protein secondary structures of HFH-1/DNA complexes was based on the analysis of backbone NOE connectivities and chemical shift information via the chemical shift index and backbone chemical shift pattern (*25−28*). The chemical shift index is a useful tool for identifying regions in a protein that form helices or β-strands. This method involves a comparison of the measured chemical shifts for each residue in the protein to a standard set of empirically derived values and ranges. A contiguous stretch of upfield-shifted $^{13}C^{\alpha}$ resonances indicates the presence of a helix, while a contiguous stretch of downfield shifts is characteristic for a β-strand.

RESULTS

Resonance Assignments. The backbone and side chain assignments of HFH-1 and the protein backbone assignments of the two HFH-1/DNA complexes were made through a combination of standard double- and triple-resonance experiments. The HSQC spectra of HFH-1 and its two DNA complexes with HFH-1#25 and HFH-2#12 are shown in Figure 1, with corresponding residue assignments indicated. The HSQC spectra of the DNA-free HFH-1 and its two DNA complexes show characteristic backbone chemical shift patterns of HNF-3/fkh homologous proteins, such as the ^1H and ^{15}N chemical shift values of highly conserved residues L25, S54, W46, and W77. These conserved chemical shift values indicate that both DNA-free HFH-1 and its DNA complexes adopt the compact winged helix motif (*9, 29*).

Structure Calculation of DNA-Free HFH-1. A total of 1201 interproton and 94 dihedral angle restraints were used in the structural calculation. The distance constraints contain 158 intraresidue, 283 sequential, 304 medium-range, and 318 long-range correlations derived from ^{13}C- and ^{15}N-edited 3D NOE spectra (*30, 31*). The dihedral angle restraints φ and ψ were deduced from the HAHN data (*32*), C′, C^{α}, and H^{α} chemical shift index values, and NOE patterns characteristic of secondary structures (*33*). For a residue in an α-helix, the φ and ψ angles were defined as $-60 \pm 20°$and $-50 \pm 50°$, respectively; for a residue in a β-strand structure, the angles were refined as $-120 \pm 40°$ and $-130 \pm 50°$, respectively (*34*). The HFH-1 structures were calculated using DYANA (*35*). At the final stage, 200 conformers were calculated, and the 20 conformers with the lowest target functions were selected to represent the structure shown in Figure 2 (A conformers, B ribbon). Good agreement with the experimental restraints is reflected in the low DYANA target functions (0.90 ± 0.12 Å^2) obtained for the family of structures. For the well-defined regions of the structure (Ile 8-Pro 37, Gly 45-Leu 64, Tyr 76-Asn 82), the RMSDs to the averaged structure are 0.60 ± 0.18 Å and 1.43 ± 0.21 Å for the backbone atoms and for all non-hydrogen atoms, respectively. A total of 74.4% of the residues are in the most favored region of the Ramachandran plot (*36*), 19.8% are in the additionally allowed region, 5.6% are in the generously allowed region, and 0.2% are in the disallowed region. The structural statistics are summarized in Table 1.

Structure Differences between DNA-Free HFH-1 and Genesis. The structures of several HNF-3/fkh proteins were determined previously by NMR spectroscopy and X-ray crystallography (*7−10*). In these structures, DNA-free Genesis, AFX, and FREAC-11 all contain four α-helices and three β-strands, as illustrated in the secondary structure of DNA-free Genesis (Figure 3A) (*8−10*). HNF-3γ contains

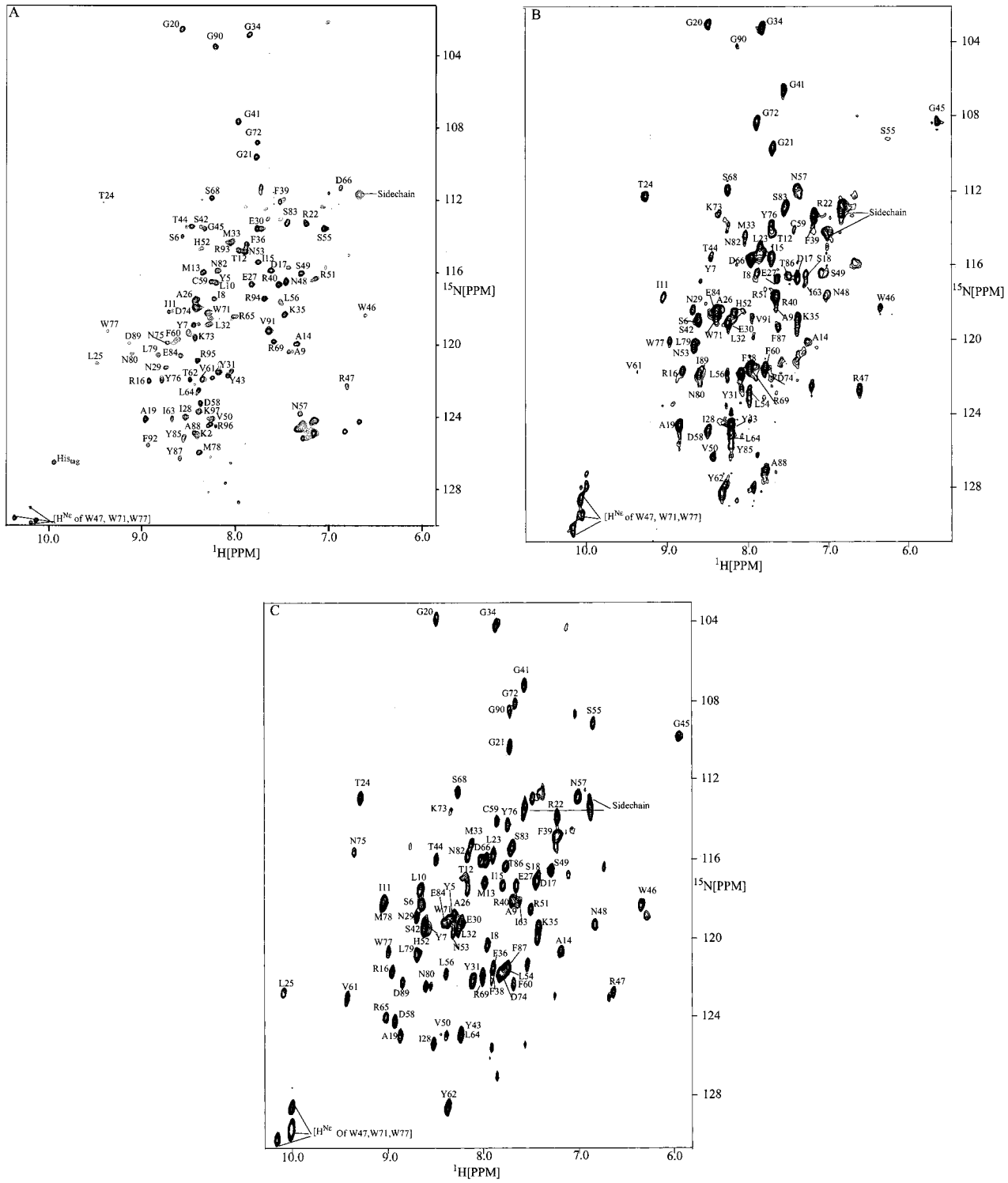

FIGURE 1: HSQC spectra of HFH-1 free and bound with DNA. (A) DNA-free HFH-1 at 17 °C and pH 6.5. (B) HFH-1/HFH-2#12−DNA complex at 25 °C and pH 6.5. (C) HFH-1/HFH-1#25−DNA complex at 30 °C and pH 6.5. The spectra were recorded at two different temperatures because of the much-reduced stability of DNA-free HFH-1 at 25 °C and the poor spectral quality of the HFH-1−DNA complexes at 17 °C. However, the temperature causes minimum backbone chemical shift perturbations in the HFH-1−DNA complexes as well as in the Genesis−DNA complex in the preliminary stage of data acquisition.

three α-helices and a helical loop between H2 and H3 (*7*). The NMR data and structural calculations both indicate that DNA-free HFH-1 contains only three helices, which are H1, H2, and H3 (Figure 3, B and C). These three helices and the three *β*-strands are located at the conserved positions, as observed in other HNF-3/fkh homologous proteins. The sequence (F39−Y43), which folds into helix 4 (H4) (Y39−K43) in Genesis and two other HNF-3/fkh proteins (*8*, *9*), adopts a helical loop structure in HFH-1, as indicated in the structures and NMR data (Figure 3C). It is likely that the existence of two glycines at positions 42 and 46 in the sequence break the formation of the helix. HFH-1 also contains two wing sequences at the conserved positions observed for other HNF-3/fkh homologues. A conclusion, therefore, is that DNA-free HFH-1 adopts the HNF-3/fkh-type winged helix motif.

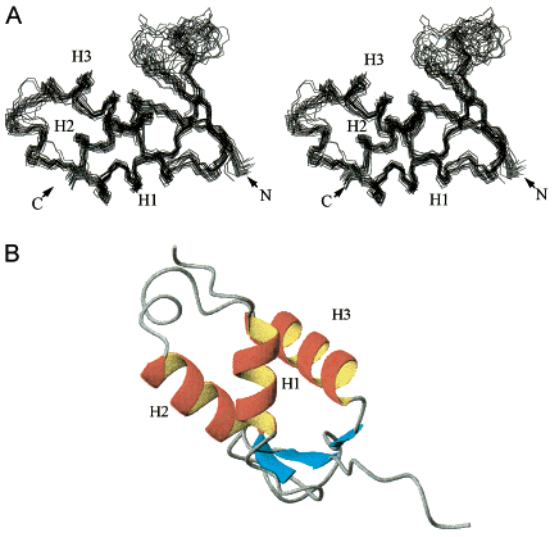

FIGURE 2: NMR structure of DNA-free HFH-1. (A) Stereo superposition of 20 selected conformers with the lowest target functions from the final DYANA calculations. (B) Ribbon diagram showing the NMR-derived tertiary structure used in this study. The average structure was generated and analyzed by MolMol (*43*).

Table 1: Structural Statistics of the Best 20 DYNYA Conformers

(A) NMR-derived restraints	
total interproton restraints	1201
intraresidue ($\|i-j\|=0$)	158
sequential ($\|i-j\|=1$)	283
medium-range ($1<\|i-j\|<5$)	304
long-range ($\|i-j\|>4$)	318
hydrogen bonds	44
dihedral angles (φ,ψ)	94
(B) residual violations[a]	
DYANA target function (Å^2)	0.90 ± 0.12
upper limit	
sum (Å)	3.90 ± 0.50
maximum (Å)	0.21 ± 0.05
van der Waals	
sum (Å)	3.60 ± 0.40
maximum (Å)	0.22 ± 0.06
(C) average RMSD to mean structure[b] (Å)	
backbone atoms N, C$^\alpha$, C′ (Å)	0.60 ± 0.18
all heavy atoms (Å)	1.43 ± 0.21
(D) Ramachandran plot (% residues)	
residues in most favored regions	74.4
residues in additional allowed regions	19.8
residues in generous allowed regions	5.6
residues in disallowed regions	0.2

[a] In part (B), values are means (standard deviations). [b] Residues 8−37, 45−64, and 76−81.

Structural Differences of Genesis and HFH-1 on DNA Binding Sites. The structure of Genesis complexed with DNA site HFH-2#12 was determined previously (*9*). When bound to the DNA, the compact HTH motif including H4 (Y39−K43) is unperturbed, while a small helix is induced at the N terminus of wing 2 (N82−F86) (Figure 4A). Although Genesis also binds to HFH-1#25, the assignment of Genesis complexed with HFH-1#25 was not pursued because of the weak interaction of the two molecules that results in fast exchange between the free and bound states and poor spectral quality under the NMR conditions employed (Sheng and Liao, unpublished results). DNA-free HFH-1 does not contain the corresponding H4. One important question is how protein−DNA interactions modify the structure of HFH-1. Secondary structures of HFH-1 on two DNA sequences HFH-1#25 and HFH-2#12 were deduced from backbone NOE patterns and chemical shift indices. The $^{13}C^\alpha$ chemical shift (Figure 4C) and backbone NOE pattern (Figure 4D) indicate that the sequence between H2 and H3 (F39−T45) and the sequence at the N terminus of wing 2 (N83−F87) in HFH-1 are not dramatically perturbed and still adopt extended conformations when HFH-1 binds to a strong DNA binding site such as HFH-1#25, which seems to indicate that the homologous HFH-1 and Genesis proteins adopt different local structures on the two DNA sequences. HFH-1#25 and HFH-2#12 are two different sequences. Thus, it is important to compare the structures of HFH-1 and Genesis when they are bound to an identical sequence, such as HFH-2#12. Because of a relatively weak interaction between the HFH-1 protein and the HFH-2#12 DNA site, the experiments used for backbone NOE assignments could not generate enough information. Thus, the $^{13}C^\alpha$ chemical shift index alone is used to deduce the secondary structure of HFH-1 on the DNA binding site (Figure 4B). As expected, the HTH motif in HFH-1 is retained. The $^{13}C^\alpha$ chemical shift indices of only a small number of residues of HFH-1 show differences in the two different DNA complexes, while the patterns of the overall chemical shift indices of the two complexes are similar. For example, the residues (F37−W47) between H2 and H3 and at the N terminus (N83−Y98) of wing 2 still show negative $^{13}C^\alpha$ chemical shift indices characteristic of extended structures in both of the HFH-1/DNA complexes. However, the HFH-1/HFF-2#12 and Genesis/HFH-2#12 DNA complexes show rather different $^{13}C^\alpha$ chemical shift index patterns. These data seem to indicate that the highly homologous HFH-1 and Genesis proteins adopt different local conformations on an identical DNA binding site, while HFH-1 adopts similar conformations on two different DNA sequences.

Comparison of DNA Binding-Induced Backbone Chemical Shift Perturbations. To further probe the structural differences between Genesis and HFH-1, average chemical shift changes caused by protein−DNA interactions were calculated for the two proteins. The average chemical shift change was defined as $(\Delta\delta^2_{HN} + \Delta\delta^2_N/25 + \Delta\delta^2_{C\alpha}/4)/3]^{1/2}$ (*37*). When Genesis interacts with HFH-2#12, chemical shift perturbations are rather small for most of the residues, particularly for those in the less structurally perturbed helix−turn−helix motif. The large chemical shift perturbations that are observed are mostly in residues located at the N terminus and in the two wings of the Genesis protein (Figure 5A). Because these segments undergo obvious structural changes, the large chemical shift perturbations are expected. The data shown in Figure 4 indicate that the DNA complex formation causes more severe chemical shift perturbations in HFH-1 than in Genesis. The average chemical shift changes are larger for HFH-1 when it contacts HFH-2#12 and HFH-1#25 than for Genesis when it binds to HFH-2#12 (Figure 5), which is true even for the residues in the HTH motif of HFH-1. However, HFH-1 shows similar chemical shift change patterns in both the HFH-2#12 and the HFH-1#25 complex, while HFH-1 and Genesis show rather different patterns, even when they bind to the identical DNA sequence (Figure 5). Therefore, the data further support the conclusion that HFH-1 undergoes similar structural transitions and adopt similar structures when it binds to two different DNA sequences.

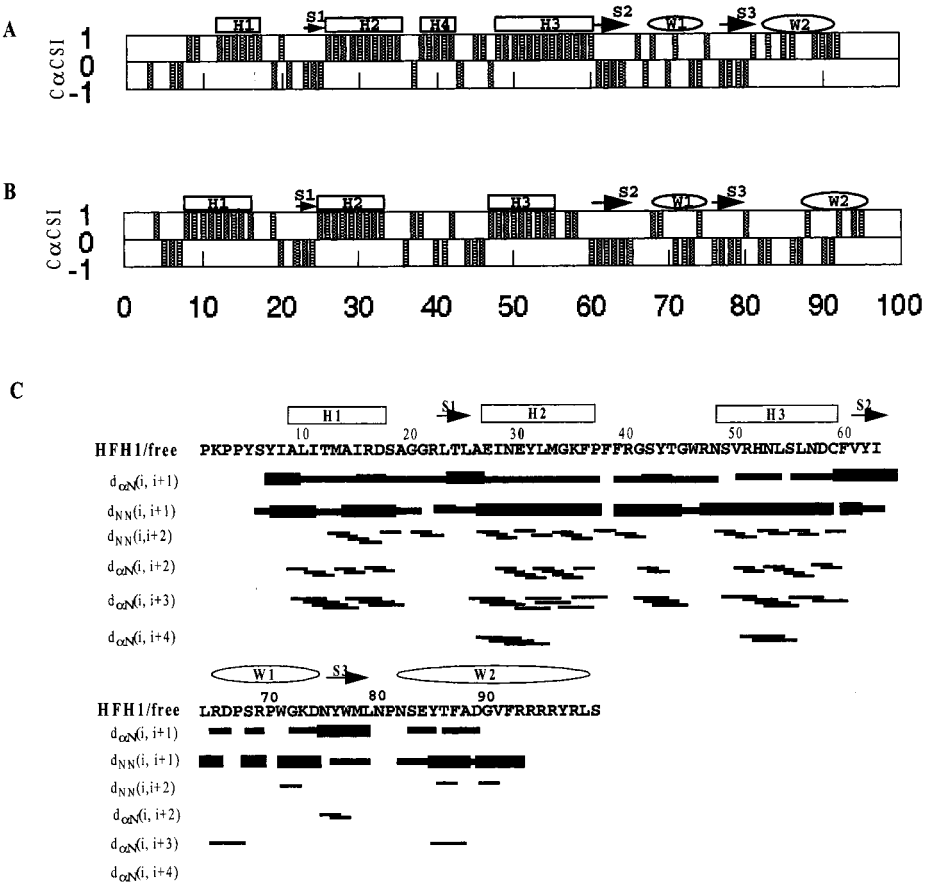

FIGURE 3: $^{13}C^{\alpha}$ chemical shift indices [$\Delta\delta(^{13}C^{\alpha})(PPM_{obs} - PPM_{random\ coil})$] plotted as a function of amino acid residues for (A) DNA-free Genesis and (B) DNA-free HFH-1. (C) Backbone NOE pattern of DNA-free HFH-1. The positions of the predicted secondary structural elements are shown on top of the corresponding plot.

Comparison of Heteronuclear NOE Data of DNA-Free HFH1 and DNA-Free Genesis. HFH-1 does not contain the corresponding helix H4, and the DNA interactions cause more chemical shift changes in HFH-1 than in Genesis. It is possible that the structure of DNA-free HFH-1 is more dynamic than that of DNA-free Genesis. To understand how the DNA binding properties are correlated to the dynamic properties of HFH-1 and Genesis, the backbone $\{^1H\}^{15}N$ heteronuclear NOE values of DNA-free Genesis and DNA-free HFH-1 were measured at 17 °C and compared (*23*). Similar to those of Genesis, the residues in the HTH motif of HFH-1 show heteronuclear NOE values around 0.8 (Figure 6), which are comparable to what was observed for DNA-free Genesis. Therefore, the data show that the HTH motif of HFH-1 is also well-formed and that H4 is dispensable in the HTH motif. However, the data also show a major difference in the dynamic behavior of the two proteins. The residues in wing 1 (E67−G74) of Genesis have relatively small heteronuclear NOE values, while in HFH-1, the corresponding residues (D67−D75) except G73 have NOE values of around 0.8. These results indicate that wing 1 of HFH-1 is more rigid than that of Genesis.

DISCUSSION

HNF-3/fkh homologous winged helix proteins are highly conserved DNA binding proteins. Although they have almost invariable residues in their primary DNA contact helices (H3), the family members still show a broad DNA binding sequence profile. By studying the DNA binding specificity of chimeric proteins, previous experiments have shown that the amino acid composition in the sequence between H2 and H3 influences the DNA binding specificity of the winged helix family members without making base-specific DNA contacts (*11*,*12*). A hypothesis was proposed that the folding differences of this divergent sequence in conserved family members are responsible for the observed differences in DNA binding specificity (*11*). Structural studies indicated that winged helix proteins, Genesis, AFX, and FREAC-11 (*8*−*10*), contain a helix (H4) between H2 and H3 and that HNF-3γ contains a helical conformation between H2 and H3. The present study has shown that HFH-1 also contains a helical structure between H2 and H3 (Figure 7). Because H3 is slightly shorter in HFH-1 than in HNF-3γ, the helical loop in HFH-1 is longer between H2 and H3. Because HFH-1, Genesis, and HNF-3 are the three proteins used in the original sequence-swap experiment, our data seems to support the conclusion of that experiment (*11*). Because of the insufficient number of NOEs detected, the structure of the sequence between H2 and H3 in HFH-1 is less well-defined. One possibility is that this sequence also adopts a short α-helix. However, the chemical shift indexes and *J*-coupling constants of this sequence of HFH-1 indicate otherwise. The chemical shift indexes of this sequence in Genesis show characteristic values of α-helix. It is difficult to rationalize the formation of an α-helix in the corresponding sequence of HFH-1 with conserved aromatic residues showing different patterns of chemical shift indexes. However, the different conformation in this sequence between H2 and H3 does not

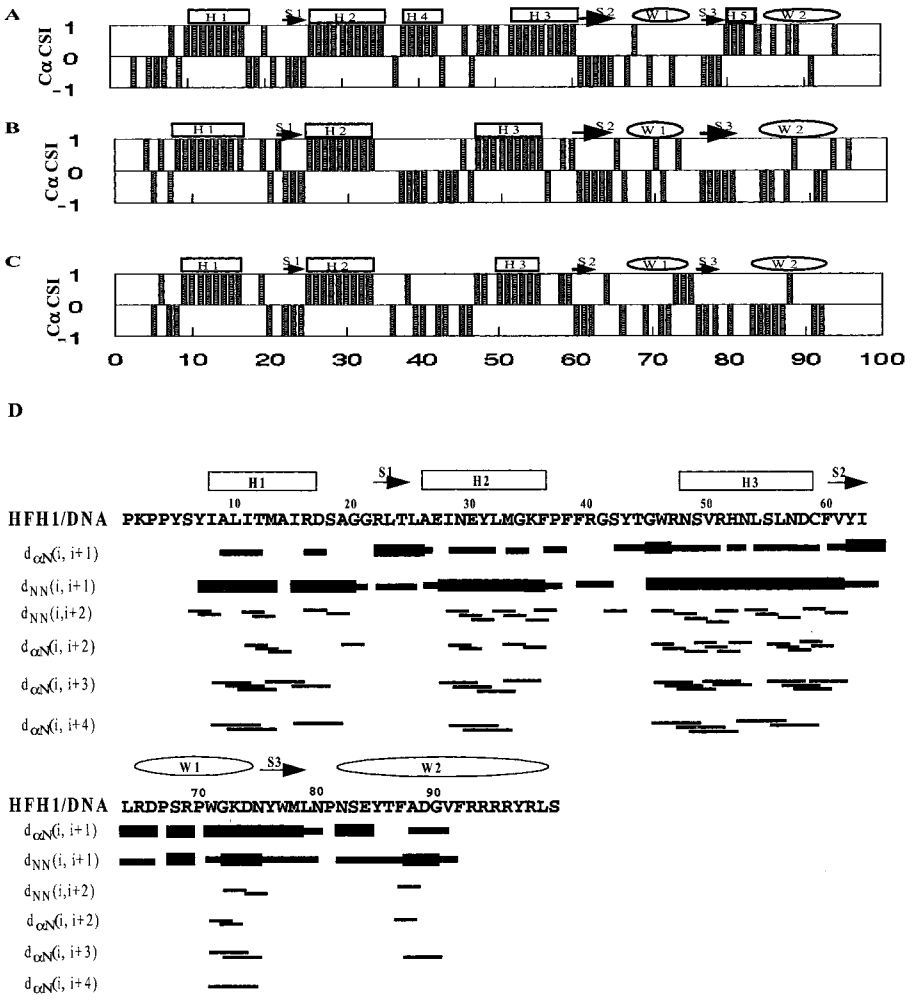

FIGURE 4: $^{13}C^{\alpha}$ secondary shift indices [$\Delta\delta(^{13}C^{\alpha})(PPM_{obs} - PPM_{random\ coil})$] plotted as a function of amino acid residues for the (A) Genesis/HFH-2#12−DNA complex, (B) HFH-1/HFH-2#12−DNA complex, and (C) HFH-1/HFH-1#25−DNA complex. (D) NOE pattern of the HFH-1/HFH-1#25−DNA complex. The positions of the predicted secondary structural elements are shown on top of the corresponding plot.

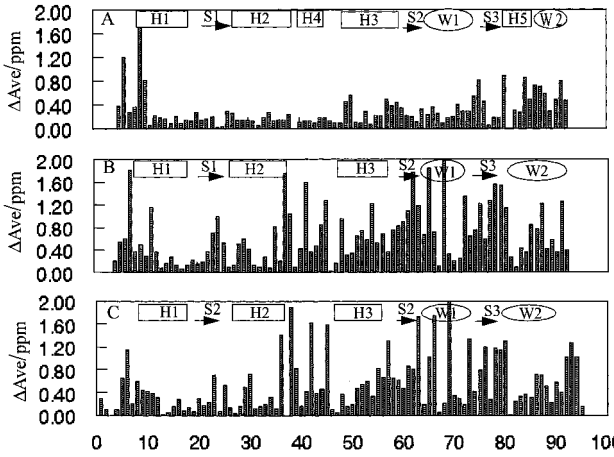

FIGURE 5: Backbone average chemical shift changes (see text for definition) caused by DNA interactions plotted as a function of amino acid residues. (A) $\Delta\delta$Av ($PPM_{Genesis/HFH-2\#12} - PPM_{DNA-free\ Genesis}$); (B) $\Delta\delta$Av ($PPM_{HFH-1/HFH-2\#12} - PPM_{DNA-free\ HFH-1}$); (C) $\Delta\delta$Av ($PPM_{HFH-1/HFH-1\#25} - PPM_{DNA-free\ HFH-1}$).

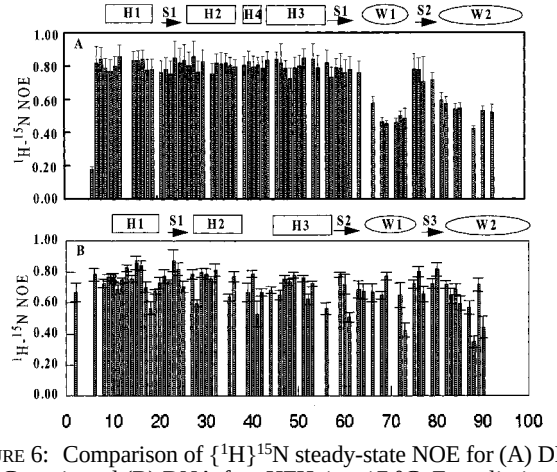

FIGURE 6: Comparison of $\{^1H\}^{15}N$ steady-state NOE for (A) DNA-free Genesis and (B) DNA-free HFH-1 at 17 °C. Error limits were determined as described previously (*23*).

modify the overall folding of HNF-3/fkh proteins (*7*−*11*). One question, then, is how the folding differences in this sequence influence the binding specificity of HNF-3/fkh family members. A nonspecific DNA contact was made from the Arg residue immediately preceding the highly conserved Trp residue in HNF-3γ. This Arg resides at the N terminus of H3 of HNF-3γ. Neither Genesis nor HFH-1 contains Arg at the corresponding position, and the residue at the corresponding position of HFH-1 and Genesis resides on the outside of H3. The sequence between H2 and H3 in HNF-3/fkh proteins is highly divergent. Thus, a combination of

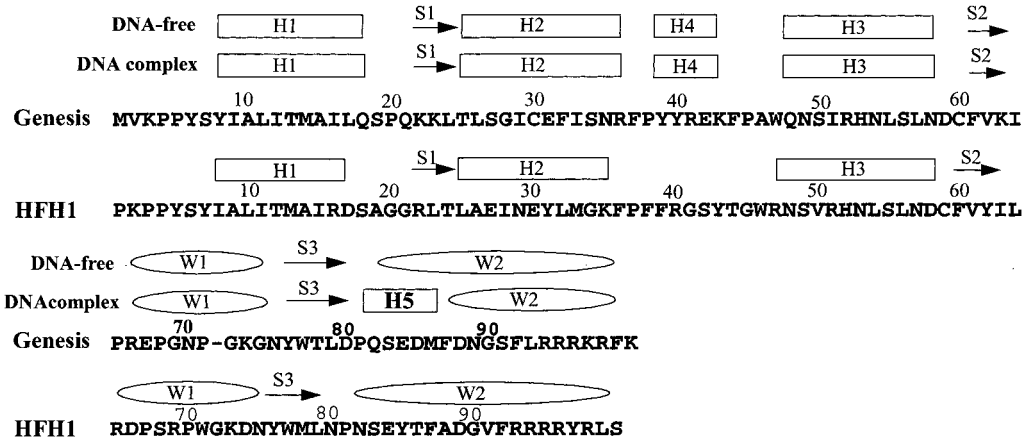

FIGURE 7: Amino acid sequences and the corresponding secondary structures of Genesis and HFH-1 and their DNA complexes. The detailed supporting data are plotted in Figures 3 and 4.

divergent amino acid residues and folding differences in the sequence influences the DNA binding specificity and affinity of HNF-3/fkh proteins. This conclusion was strongly supported by thermodynamic studies of other DNA binding proteins and an "induced-fit" model proposed previously (*38*−*40*).

A large number of HNF-3γ/fkh proteins have been identified in divergent organisms and tissues. In many cases, a tissue can express more than one HNF-3γ/fkh homologue, and each homologue controls the expression of a distinct set of genes in the tissue. Genesis and HFH-1 are both mammalian HNF-3 proteins. Because Genesis and HFH-1 are highly homologous (Figure 7) and have shared DNA binding sequences, there is an important question as to how other components of cellular control machinery can distinguish between the two HNF-3/fkh proteins. One possibility is that the structural differences of the two proteins, when bound to DNA, can be recognized by other proteins. Our results indicate that the two obvious structural differences between Genesis and HFH-1 on the HFH-2#12 binding site occur in the sequence between H2 and H3 and in the sequence at the N terminus of wing 2 (Figure 7). In Genesis, both of these sequences are folded as helices, while in HFH-1, both sequences exist as extended structures on the basis of the NOE patterns and chemical shift indices. These structural differences can potentially function as recognition markers for other proteins to distinguish between HFH-1 and Genesis complexes.

The high-resolution structural determination of a protein−DNA complex by NMR methods requires spectra of high quality. In this study, the spectral quality of HFH-1/HFH-2#12 DNA complex is low, which is likely due to the less stable complex. To study the structural differences between Genesis and HFH-1 on an identical DNA site, the $^{13}C^{\alpha}$ chemical shift indices were used to deduce the secondary structure of HFH-1 in complex with the HFH-2#12 DNA site. The values of the $^{13}C^{\alpha}$ chemical shift indices and average chemical shift changes obtained for the HFH-1/HFH-2#12 complex were compared with those obtained for the HFH-1#25 DNA complex and the DNA-free HFH-1 with confirmed secondary structures. The parameters closely resemble one another, and the comparisons lead to the conclusions that HFH-1 adopts a different local structure from Genesis on an identical DNA sequence and that HFH-1 adopts similar structures on the two different DNA sequences. This strategy allows us to study the structure of a protein−DNA complex whose spectra are of unfavorable quality.

Comparison of the chemical shift changes of Genesis and HFH-1 binding to DNA showed that larger perturbation values were observed for HFH-1, even when the proteins were bound to an identical DNA sequence, HFH-2#12. One possibility is that although HFH-1 and Genesis can both bind to HFH-2#12 the two proteins reside on different positions of the sequence. However, a previous experiment demonstrated that H3 of both Genesis and HFH-1 binds to the center AAAACAA sequence in the HFH-2#12 binding site with similar DNA contact schemes (*41*). Thus, the large chemical shift perturbations observed in the three helices of HFH-1 in the HFH-2#12 complex should not be the result of a different DNA contact scheme. It is also possible that the 3D folding was perturbed slightly in HFH-1 when it contacted the DNA. Genesis is tightly packed and shows a better temperature-dependent stability than does HFH-1. Genesis was found to be stable at the NMR conditions used in our study over the temperature range 0−22 °C, while HFH-1 showed both low- and high-temperature-induced denaturation. Thus, the NMR experiments on DNA-free HFH-1 can be performed only over a narrow temperature range around 17 °C. This observation suggests that HFH-1 may have weaker core packing and that a structural change may be more readily achievable when it interacts with other molecules. This possibility is further supported by comparison of average chemical shift changes of Genesis and HFH-1 in which two HFH-1/DNA complexes show similar chemical shift change patterns, while the HFH-1/HFH-2#12 and Genesis/HFH-2#12 DNA complexes show rather different patterns.

One of the divergent sequences in the HNF-3 homologous proteins is the wing 1 sequence. This sequence was shown to interact with DNA in a Genesis−DNA complex (*9*). The wing 1 sequence of Genesis is highly dynamic in both the DNA-free protein and in a DNA complex, as judged by backbone $\{^1H\}^{15}N$ heteronuclear NOE data (*9*,*23*). Interestingly, the wing 1 sequence of HFH-1 is less dynamic on the basis of heteronuclear NOE data. The data indicate that wing 1 of HFH-1 has smaller amplitudes of internal motions on the picosecond-to-nanosecond time scale than does wing 1 of Genesis. Wing 1 sequences are rather divergent in the HNF-3 family members. One of the obvious differences is that wing 1 of Genesis contains three Gly residues, while

wing 1 of HFH-1 only has one Gly residue. Another difference is the insertion of a Trp residue (W71) in wing 1 of HFH-1. The orientation of the Trp side chain is rather disordered in our solution structures because of the lack of NOE connectivities. However, the side-chain stacking effect is still possible because the hydrophobic side chain of Leu23 and the aromatic ring of Trp71 are in close proximity in several structures. This putative interaction can also reduce the amplitudes of the internal motions in wing 1 of HFH-1. Therefore, the apparent difference in dynamic properties is likely the result of either fewer Gly residues, one more Trp residue, or both in wing 1 of HFH-1, compared to that of Genesis. Thus, our data indicate that the dynamic properties of a less refined sequence are highly related to its amino acid composition, as also indicated in a previous study on an ArgGlyAsn motif flanked by different residues in homologous fibronectin type III domains (*42*). More detailed studies are necessary to evaluate this hypothesis. The differences in dynamic properties of wing 1 of Genesis and HFH-1 may play important roles in DNA binding and protein−protein interactions in cell regulation, and these possibilities also need further study.

ACKNOWLEDGMENT

We thank Dr. Robert Costa for the gift of the HFH genes. We thank Ms. Karen Reddy for reviewing the manuscript and critical comments.

SUPPORTING INFORMATION AVAILABLE

Two tables (S1, S2) of NMR resonance assignments for backbone atoms, H^N, ^{15}N, and ^{13}C of DNA-free Genesis, HFH-1, their DNA complexes, and DNA-free Genesis and HFH-1. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Raumann, B. E., Brown, B. M., and Sauer, R. T. (1994) *Curr. Opin. Struct. Biol. 4*, 36−43.
2. Pabo., C. O., and Sauer R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.
3. Kaufmann, E., and Knöchel, W. (1996) *Mech. Dev. 57*, 3−20.
4. Weigel, D., and Jäckle, H. (1990) *Cell 63*, 455−456.
5. Lai, E., Prezioso, V. P., Smith, E., Litvin, O., Costa, R. H., and Darnell, J. E., Jr. (1990*) Genes Dev. 4*, 1427−1436.
6. Gajiwala, K. S., and Burley, S. K. (2000) *Curr. Opin. Struct. Biol. 10*, 110−116.
7. Clark, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993) *Nature 364*, 412−420.
8. Marsden, I., Jin, C., and Liao, X. (1998) *J. Mol. Biol. 278*, 293−299.
9. Jin, C., Marsden, I., Chen, X., and Liao, X. (1999) *J. Mol. Biol. 289*, 683−690.
10. Van Dongen, M. J, Cederberg, A., Carlsson, P., Enerback, S., Wikstrom, M. (2000) *J. Mol. Biol. 296*, 351−359.
11. Overdier, D. G., Porcell, A., and Costa, R. H. (1994) *Mol. Cell. Biol. 14*, 2755−2766.
12. Pierrou, S., Hellqvist, M., Samuelsson, L., Enerbck, S., and Carlsson, P. (1994) *EMBO J. 13*, 5002−5012.
13. Sutton, J., Costa, R., Klug, M., Field, L., Xu, D., Largaespada, D. A., Fletcher, C. F., Jenkins, N. A., Copeland, N. G., Klemsz, M., and Hromas, R. (1996) *J. Biol. Chem. 271*, 23126−23133.
14. Clevidence, D. E., Overdier, D. G., Tao, W., Qian, X., Pani, L., Lai, E., and Costa, R. H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3948−3952.
15. Fesik, S. W., and Zuiderweg, E. R. (1990) *Q. Rev. Biophys. 23*, 97−131.
16. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson., Ser. B 103*, 203-216.
17. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
18. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 11655−11666.
19. Wittekind, M., and Mueller, L. (1993) J. *Magn. Reson., Ser. B 101*, 201−205
20. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
21. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Formankay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
22. Markley, J. L., Horsley, W. J., and Klein, M. P. (1971) *J. Chem. Phys. 55*, 3604−3605.
23. Jin, C., Marsden, I., Chen, X., and Liao, X. (1998) *Biochemistry 37*, 6179−6187.
24. Delaglio, F., Grzesiek., S, Vuister, G. W., Zhu, G., Pfeifer, J., Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
25. Marsden I., Chen, Y., Jin, C., and Liao, X. (1997) *Biochemistry 36*, 13248−13255.
26. Wishart, D. S., and Skyes, B. D. (1994a*) Methods Enzymol. 239*, 363−392.
27. Wishart, D. S., and Skyes, B. D. (1994b) *J. Biomol. NMR 4*, 171−180.
28. Luginbühl, P., Szyperski, T., and Wüthrich, K. (1995) *J. Magn. Reson., Ser. B 109*, 229−233.
29. Spera, S., and Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490−5492.
30. Zuiderweg, E. R. P., and Fesik, S. W. (1989) *Biochemistry 28*, 2387−2391.
31. Talluri, S., and Wagner, G. (1996) *J. Magn. Reson., Ser. B 112*, 200−205.
32. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
33. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
34. Weigelt, J., Climent, I., Wright, K. D., and Wikström, M. (2001) *Biochemistry 40*, 5861−5869.
35. Günter, P., Mumenthalter, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993*) J. Appl. Crystallogr. 26*, 283−291.
37. Grzesiek, S., Stahl, S. J., Wingfield, P. T., and Bax, A. (1996) *Biochemistry 35*, 10256−10261.
38. Spolar, R., and Record, T. M., Jr. (1994) *Science 263*, 777−784.
39. Craig, M. L., Tsodikov, O. V., McAQuade, K. L., Schlaxc, P. E., Jr., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (1998) *J. Mol. Biol. 283*, 741−756.
40. Jia, X., Lee, L. K., Light, J., Palmer A. G., III, Assa-Munt, N. (1999) *J. Mol. Biol. 289*, 683−689.
41. Bravieri, R., Shiyanova, T., Chen, T., Overdier, D., and Liao, X. *Nucleic Acids Res*. (1997) *25*, 2888−2896.
42. Carr, P. A., Erickson, H. P., and Palmer, A. G., III. (1997) *Structure 5*, 949−959.
43. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI011908K