Protein Science (1994), *3*:1806–1821. Cambridge University Press. Printed in the USA.

Solution structure of the DNA-binding domain of the heat shock transcription factor determined by multidimensional heteronuclear magnetic resonance spectroscopy

FRED F. DAMBERGER,[1] JEFFREY G. PELTON,[2] CELIA J. HARRISON,[3] HILLARY C.M. NELSON,[3] AND DAVID E. WEMMER[4]

[1] Biophysics Graduate Group, University of California, Berkeley, California 94720
[2] Structural Biology Division of Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, California 94720
[3] Department of Molecular and Cell Biology, University of California, Berkeley, California 94720
[4] Department of Chemistry, University of California, Berkeley, California 94720

(RECEIVED June 24, 1994; ACCEPTED July 26, 1994)

Abstract

The solution structure of the 92-residue DNA-binding domain of the heat shock transcription factor from *Kluyveromyces lactis* has been determined using multidimensional NMR methods. Three-dimensional (3D) triple resonance, ^{1}H-^{13}C-^{13}C-^{1}H total correlation spectroscopy, and ^{15}N-separated total correlation spectroscopy-heteronuclear multiple quantum correlation experiments were used along with various 2D spectra to make nearly complete assignments for the backbone and side-chain ^{1}H, ^{15}N, and ^{13}C resonances. Five-hundred eighty-three NOE constraints identified in 3D ^{13}C- and ^{15}N-separated NOE spectroscopy (NOESY)-heteronuclear multiple quantum correlation spectra and a 4-dimensional $^{13}C/^{13}C$-edited NOESY spectrum, along with 35 ϕ, 9 χ_1, and 30 hydrogen bond constraints, were used to calculate 30 structures by a hybrid distance geometry/simulated annealing protocol, of which 24 were used for structural comparison. The calculations revealed that a 3-helix bundle packs against a small 4-stranded antiparallel β-sheet. The backbone RMS deviation (RMSD) for the family of structures was 1.03 ± 0.19 Å with respect to the average structure. The topology is analogous to that of the C-terminal domain of the catabolite gene activator protein and appears to be in the helix-turn-helix family of DNA-binding proteins. The overall fold determined by the NMR data is consistent with recent crystallographic work on this domain (Harrison CJ, Bohm AA, Nelson HCM, 1994, *Science 263*:224) as evidenced by RMSD between backbone atoms in the NMR and X-ray structures of 1.77 ± 0.20 Å. Several differences were identified some of which may be due to protein–protein interactions in the crystal.

Keywords: comparison to crystal structure; DNA-binding; heat shock factor; helix-turn-helix; NMR; protein structure

Reprint requests to: David E. Wemmer, Department of Chemistry, University of California, Berkeley, California 94720; e-mail: dewemmer@lbl.gov.

Abbreviations: CAP, catabolite gene activator protein; CBCA(CO)NH, C_β to C_α to carbonyl to amide N/HN correlation experiment; CBCANH, C_β to C_α to amide N/HN correlation experiment; DG, distance geometry; DQF-COSY, double-quantum filtered correlation spectroscopy; HCCH-TOCSY, ^{1}H-^{13}C-^{13}C-^{1}H total correlation spectroscopy; ct, constant time; HNF-3_γ, hepatocyte nuclear factor/*forkhead*; HSE, heat shock element; HSF, heat shock transcription factor; HTH, helix-turn-helix; HMQC, heteronuclear multiple quantum correlation; HSQC, heteronuclear single quantum correlation; NOESY, NOE spectroscopy; rf, radio frequency; RMSD, RMS deviation; ⟨RMSD⟩, RMS deviation with respect to the average coordinates; SA, simulated annealing; TOCSY, total correlation spectroscopy; TPPI, time proportional phase incrementation; TSP, 3-(2,2,3,3,$^{2}H_4$) trimethylsilyl propionate; 2D, 2-dimensional; 3D, 3-dimensional; 4D, 4-dimensional.

When eukaryotic cells are exposed to elevated temperatures or environmental stress, a new set of proteins is produced in the "heat shock response." The production of these proteins is regulated at the transcriptional level by the heat shock transcription factor (for recent reviews see Craig et al., 1993; Lis & Wu, 1993; Morimoto, 1993). In yeast, HSF forms a trimer in solution and also binds DNA in this form (Sorger & Nelson, 1989). HSF is constitutively bound to DNA and acts as a transcription factor even under nonstressed conditions (Jakobsen & Pelham, 1988; Sorger & Pelham, 1988; Wiederrecht et al., 1988; Park & Craig, 1989; Gross et al., 1990; Chen & Pederson, 1993; Gallo et al., 1993). Heat shock conditions induce a higher level of transcriptional activation mediated by HSF through mechanisms that are not yet clear (Nieto-Sotelo et al., 1990; Sorger, 1990;

Jakobsen & Pelham, 1991; Gallo et al., 1993). The DNA-binding activity directs HSF to heat shock elements, defined as a series of inverted repeats of the 5-base pair sequence nGAAn (Amin et al., 1988; Xiao & Lis, 1988; Perisic et al., 1989; Boorstein & Craig, 1990; Xiao et al., 1991). Most heat shock promoters contain several HSEs of varying length and location.

The DNA-binding domain of HSF was mapped genetically to a sequence of roughly 130 amino acids (Wiederrecht et al., 1988) that was subsequently reduced to an 89-amino acid sequence in yeast. This fragment was shown to be capable of binding HSE sites sequence specifically (Flick et al., 1994). The sequence in this region is highly conserved across all HSFs but does not show obvious sequence homology to other known DNA-binding proteins. The high degree of sequence conservation suggests that this is an independently folded domain whose structure will have wide-ranging implications for the DNA recognition of all HSFs. The structure is of particular interest for 2 reasons. First, HSF binds to DNA as a homotrimer, which is unusual among sequence-specific DNA-binding proteins. Second, the sequence cannot be classified into any of the known DNA-binding motifs based on secondary structure prediction programs. In the present work, we describe the structure of this domain as defined by heteronuclear multidimensional NMR studies and identify the DNA-binding motif and likely DNA-contacting residues. In addition, we compare the NMR structure of this DNA-binding protein to the recently determined crystal structure of the same domain (Harrison et al., 1994) and those of other DNA-binding proteins. The eventual goal of this NMR study is to characterize interactions of HSF with DNA and the dynamics of the protein in the free and bound state.

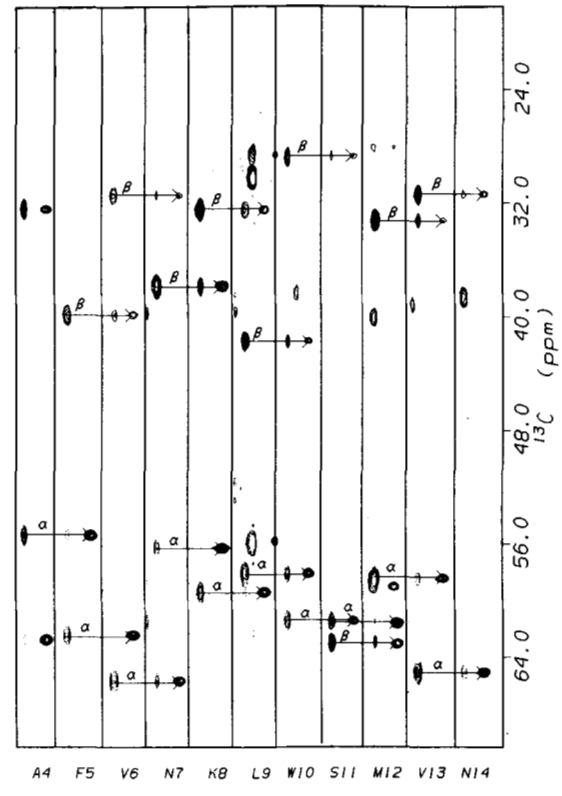

Fig. 1. Sequential strips for residues Ala-4 to Asn-14 extracted from 3D triple-resonance CBCANH (left side of strip) and CBCA(CO)NH (right side of strip) spectra that were used to make sequential assignments of HSF. Residues are denoted by 1-letter amino acid code and residue number. Intraresidue C_α and C_β correlations (CBCANH) are labeled with Greek letters. Arrows link intraresidue C_α and C_β correlations to the interresidue C_α and C_β signals observed in both CBCANH and CBCA(CO)NH strips for the following residue. Cross peaks rendered with dashed lines are of opposite phase to those drawn with solid lines.

Results

^1H, ^{15}N, and ^{13}C assignments

Using the methods described below, complete sequential assignments were obtained for 85 of 89 backbone N-H signals in the ^1H-^{15}N HSQC spectrum. This includes all N-H signals except those of Ala-1, Gly-28, Asn-80, and Asp-81. Aside from Ala-1, tentative assignments could be made for these residues although not confirmed due to overlap (Asn-80, Asp-81) or lack of NOEs (Gly-28). In addition, all the ϵN-H pairs of the 6 Arg and 1 ϵN-H pairs of the 4 Trp residues were assigned based on characteristic Arg-$^{15}N_\epsilon$ (~85 ppm) and Trp-$NH_{\epsilon 1}$ (~10 ppm) shifts. Twelve pairs of signals expected from the 9 Asn and 3 Gln NH2 groups were also identified, although not assigned because of severe overlap of their $^1H_\gamma$ and $^1H_\beta$ resonances. The completeness of the backbone N-H assignments (95%) was achieved despite a significant number of degeneracies observed: 10 backbone amides within 0.3 ppm (^{15}N) and 0.02 ppm (^1H) of another resonance.

The 3D triple-resonance CBCANH (Grzesiek & Bax, 1992a) and CBCA(CO)NH (Grzesiek & Bax, 1992b) experiments were used to sequentially assign the backbone ^{15}N, NH, C_α, and side-chain C_β signals. This method has recently been successfully applied to *Acanthamoeba* profilin-I (Archer et al., 1993), human profilin (Metzler et al., 1993), and human macrophage inflammatory protein-1β (Lodi et al., 1994). As can be seen in Figure 1, the CBCANH experiment correlates both the C_β and C_α frequencies of a residue with its own ^{15}N and NH frequencies via the 1-bond $^1J_{C_\alpha N}$ coupling (~12 Hz) and with the ^{15}N and ^1H of the following residue via the 2-bond $^2J_{C_\alpha N}$ coupling (~7 Hz). In contrast, the CBCA(CO)NH experiment correlates the C_β and C_α frequencies of a residue to only the ^{15}N and ^1H of the following residue. Thus, through comparison of the 2 spectra, intra- and interresidue correlations can be identified and used to define which residues are adjacent in the sequence. For example, for Trp-10 the interresidue C_α and C_β correlations identified from the CBCA(CO)NH spectrum (C_α 58.0 ppm, C_β 41.7 ppm) correspond to the intraresidue C_α and C_β correlations for Leu-9 from the CBCANH spectrum, indicating that these residues are adjacent in the sequence. Using this approach, we were able to correlate C_α, C_β, ^{15}N, and NH signals for 65 of the 89 residues identified in the triple-resonance spectra and link fragments of sequential residues from 2 to 11 amino acids in length.

Spin system identification

Through analysis of an ^{15}N-separated TOCSY-HMQC spectrum, we were able to correlate ^{15}N, HN, H_α, and H_β signals for approximately 70% of the residues. These data, combined with the intraresidue C_α and C_β chemical shifts for each residue identified in the triple-resonance data, provided a convenient

starting point for assignment of the side-chain ^{1}H and ^{13}C signals through analysis of a 3D HCCH-TOCSY spectrum (Bax et al., 1990b; Clore et al., 1990) as described previously (Ikura et al., 1991b; Pelton et al., 1991). Representative data are shown in Figure 2. In Figure 2A, $^{13}C_{\alpha}$-H_{α} ($^{13}C_{\alpha}$ 55.9 ppm) to side-chain ^{1}H correlations for several residues including Ala-4, Gln-69, Met-76, Leu-77, and Glu-89 are shown. Once the ^{1}H spin systems were identified using the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ signals, the remaining side-chain ^{13}C spin systems were assigned by searching for the characteristic ^{1}H chemical-shift pattern of a residue in the appropriate ^{13}C chemical-shift range (plane) of the 3D spectrum. For example, in Figure 2B, the ^{1}H correlations obtained from several side-chain Val $^{13}C_{\gamma}$-^{1}H pairs are shown. The $^{13}C_{\gamma}$-$^{1}H_{\alpha}$ signals in this plane were matched to $^{13}C_{\alpha}$-$^{1}H_{\alpha}$/$^{1}H_{\gamma}$ signals in the corresponding C_{α} plane, thus assigning the valine $^{13}C_{\gamma}$-$^{1}H_{\gamma}$ signals. The spin system assignments were used to register many of the fragments of sequential residues identified in the CBCANH and CBCA(CO)NH data with the sequence of HSF. In addition, analysis of the HCCH-TOCSY spectrum allowed us to identify C_{β} and C_{α} frequencies in cases where these signals were weak or absent in the CBCANH/CBCA(CO)NH data, allowing for the extension of the fragments. Moreover, by making use of type assignments and characteristic ^{13}C chemical shifts (Horvath & Lilley, 1978), complete side-chain ^{1}H and ^{13}C assignments were made without use of a 3D HCCH-COSY spectrum (Bax et al., 1990a). The CBCANH and CBCA(CO)NH data again supplied useful complementary information by eliminating ambiguities in distinguishing H_{β}-C_{β} and H_{γ}-C_{γ} pairs in the HCCH-TOCSY spectrum because C_{β} chemical shifts were already determined from the triple resonance experiments.

In addition to the HCCH-TOCSY spectrum, 2D ^{1}H DQF-COSY and TOCSY spectra and a 3D ^{15}N-separated TOCSY-HMQC spectrum were used to assign some spin systems using standard methods (Wüthrich, 1986). For example, the Arg spin systems were identified in the 3D ^{15}N-separated TOCSY-HMQC experiment by matching backbone NH and side-chain NH_{ϵ} cross-peak patterns. Fifteen other residues were identified by looking for spin system patterns in 2D DQF-COSY and TOCSY spectra and correlating them with the 3D ^{15}N-separated TOCSY-HMQC spectrum. Many of the remaining residues were in the large class of AMX spin systems which, excluding the distinctive serines, comprise 27 of the 89 spin systems observable in the ^{15}N-separated 3D experiments (there are 3 prolines in the sequence). Finally, the 4 Trp, 2 Tyr, and 4 of the 7 Phe aromatic spin systems were assigned through analysis of 2D ct-HSQC and 2D ct-HCCH-TOCSY spectra (Ikura et al., 1991a) that were optimized for aromatic spin systems. Subsequently, sequential assignments were obtained through identification of intraresidue NOEs between H_{δ} and both H_{α} and H_{β} and between H_{δ} and HN in 3D ^{13}C-separated NOESY-HMQC and 3D ^{15}N-separated NOESY-HMQC spectra, respectively. In total, using the methods described, 95% of the backbone and 80% of the side-chain ^{1}H and ^{13}C signals have been assigned. The assignments are summarized in Table 1.

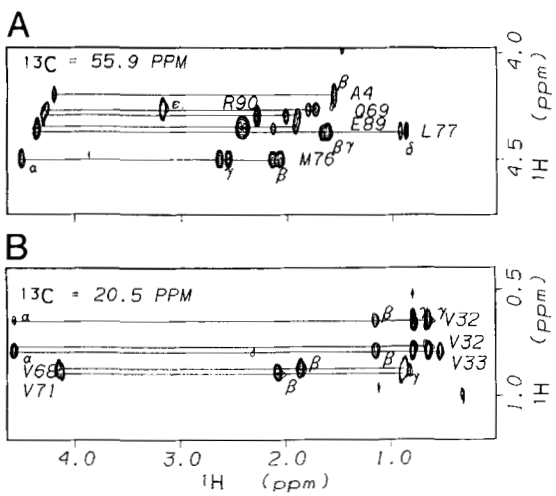

Fig. 2. Expansion of representative planes from the 3D HCCH-TOCSY spectrum **(A)** showing TOCSY correlations from labeled $^{13}C_{\alpha}$-$^{1}H_{\alpha}$ pairs (^{13}C = 55.9 ppm) and **(B)** valine ^{13}C-^{1}H pairs (^{13}C = 20.5 ppm).

Secondary structure

Upon completion of the resonance assignments, we undertook analysis of 2D ^{1}H NOESY and 3D ^{15}N-separated and ^{13}C-separated NOESY-HMQC spectra and a 4D $^{13}C/^{13}C$-edited NOESY spectrum to define the secondary structure of HSF. Characteristic patterns of sequential, medium-range, and long-range NOEs observed in these spectra can be used to identify elements of regular secondary structure (Wüthrich, 1986). A portion of a 3D ^{15}N-separated NOESY-HMQC spectrum of HSF is shown in Figure 3, and a complete analysis of the sequential and medium-range NOEs is presented in Figure 4. In Figure 3, the presence of strong NOEs between adjacent NH protons (d_{NN} NOEs) for residues Phe-56 to Tyr-63 (H3) indicate that this segment adopts a helical structure. The identification of $d_{\alpha N(i,i+3)}$ and $d_{\alpha\beta(i,i+3)}$ NOEs, slowly exchanging amide protons, small $^{3}J_{NH_{\alpha}}$ coupling constants, and upfield-shifted $^{13}C_{\alpha}$ signals (Spera & Bax, 1991; Wishart et al., 1991) (Fig. 4) confirm that these residues form part of an α-helix. The segments Ala-4–Asn-14 (H1) and Arg-36–Phe-48 (H2) also show a series of d_{NN} NOEs indicative of helical secondary structure. For segment H1 the expected patterns of $d_{\alpha N(i,i+3)}$ and $d_{\alpha\beta(i,i+3)}$ NOEs and upfield-shifted $^{13}C_{\alpha}$ signals confirm that this segment is α-helical.

For segment H2, the first residues (Glu-37–Gln-41) show all the evidence for α-helical conformation exhibited by segments H1 and H3. However the segment surrounding Pro-45 (Gln-41–Phe-48) is characterized by several weak intensity $d_{\beta N}$ NOEs and far fewer $d_{\alpha N(i,i+3)}$ and $d_{\alpha\beta(i,i+3)}$ NOEs (Fig. 4). Furthermore, unlike segments H1 and H3, where $^{13}C_{\alpha}$ shifts are consistently upfield, several residues centered around Glu-42 show small or no upfield shifts and, in addition, the $^{3}J_{NH_{\alpha}}$ coupling constant of Val-43 is large ($\phi = -120 \pm 60°$), indicating a non-helical backbone structure. These data suggest that the central portion of helix H2 contains a distortion centered between residues Gln-41 and Val-43.

Between segments H1 and B1 is a short segment (Lys-16–Lys-20) that shows medium d_{NN} and $d_{\beta N}$ NOEs, a number of $d_{\alpha N(i,i+2)}$ and $d_{\alpha N(i,i+3)}$ NOEs, a single $d_{\alpha\beta(i,i+3)}$ NOE, and a number of upfield-shifted $^{13}C_{\alpha}$ signals indicative of a helical/turn structure.

In addition to the 3 helices, 4 segments characterized by strong $d_{\alpha N}$ and weak d_{NN} NOEs were observed, indicating that these

Table 1. *Assignments at pH 3.4 and 25.0 °C for HSF*[a]

Residue	Amide	C^{α}	C^{β}	Other
A1				
R2	120.9 (8.68)	54.6 (4.66)	30.1 (1.65,1.59)	H^{γ} (1.68), H^{δ} (2.94), N^{ϵ} 83.2 (7.07)
P3		63.3 (4.50)	32.5 (1.63, 2.57)	C^{γ} 28.2 (2.17, 2.23), C^{δ} 50.9 (3.46, 4.05)
A4	126.6 (8.79)	55.6 (4.16)	18.0 (1.51)	
F5	116.6 (9.09)	62.6 (4.01)	39.9 (2.26, 3.36)	H^{δ} (6.92)
V6	115.4 (6.75)	65.4 (3.25)	31.6 (1.89)	$C^{\gamma 1}$ 21.6 (0.90), $C^{\gamma 2}$ 23.0 (0.80)
N7	117.5 (7.73)	56.2 (4.47)	37.8 (2.74, 2.80)	
K8	120.8 (7.97)	59.6 (3.94)	32.3 (1.34, 1.78)	C^{γ} 25.5 (1.25 1.50), C^{δ} 28.9 (1.46), C^{ϵ} 43.0 (3.21)
L9	120.5 (8.27)	58.0 (3.61)	41.7 (1.77, 0.75)	C^{γ} 26.5 (0.89), C^{δ} 22.9 (0.35), C^{δ} 25.1 (0.09)
W10	117.7 (8.50)	61.4 (4.03)	28.8 (3.30, 3.33)	$C^{\delta 1}$ 126.2 (7.36), $C^{\epsilon 3}$ 120.7 (7.68), $C^{\zeta 3}$ 121.9 (7.11) $C^{\eta 2}$ 124.5 (7.23), $C^{\zeta 2}$ 114.5 (7.53), $N^{\epsilon 1}$ 127.7 (10.23)
S11	110.6 (7.83)	61.0 (4.10[t])	63.1 (4.00[t])	
M12	119.3 (7.92)	58.8 (3.38)	33.2 (1.01, 1.83)	C^{γ} 30.0 (1.55[t], 1.66[t])
V13	115.7 (8.03)	65.0 (3.27)	31.3 (1.08)	$C^{\gamma 1}$ 21.1 (−0.41), $C^{\gamma 2}$ 22.2 (0.13)
N14	114.2 (7.84)	53.5 (4.70)	38.5 (2.79, 2.85)	
D15	118.6 (6.97)	53.1[t](4.70)	38.6 (2.98, 3.14)	
K16	129.1 (9.16)	59.3 (4.10)	31.9 (1.94 deg[b])	C^{γ} 24.8 (1.60, 1.63), C^{δ} 28.9 (1.78), C^{ϵ} 42.0 (3.09)
S17	114.1 (8.34)	61.2 (4.33)	62.7 (3.83, 3.95)	
N18	117.0 (8.30)	53.1 (5.15)	39.2 (3.15[t], 3.02)	
E19	120.6 (7.42)	59.7 (4.57)	28.3 (2.17, 2.25)	C^{γ} 33.0 (2.60, 2.64)
K20	116.5 (8.48)	57.9 (3.91)	31.5 (1.00 deg)	C^{γ} 24.8 (1.01), C^{δ} 29.0 (1.46), C^{ϵ} 41.7 (2.86)
F21	114.4 (8.16)	56.4 (4.97)	42.3 (3.09, 3.23)	C^{δ} 131.3 (7.11)
I22	117.8 (8.16)	60.0 (4.98)	37.7 (1.99)	$C^{\gamma 2}$ 14.1 (0.88), $C^{\delta 2}$ 14.0 (0.30)
H23	116.6 (8.33)	53.8 (4.87)	30.4 (3.47, 3.65)	
W24	120.8 (8.93)	56.5 (5.32)	31.3 (3.16, 3.34)	$C^{\delta 1}$ 128.7 (7.39), $C^{\epsilon 3}$ 120.9 (7.70), $C^{\zeta 3}$ 121.7 (6.83) $C^{\eta 2}$ 123.2 (6.18), $C^{\zeta 2}$ 113.9 (6.42), $N^{\epsilon 1}$ 128.6 (9.91)
S25	119.4 (9.05)	58.2 (4.87)	63.3 (3.94, 4.10)	
T26	116.4 (8.55)	64.5 (4.19)	69.1	C^{γ} 22.1 (1.27)
S27	117.1 (8.58)	(3.85[t])		
G28	101.9[t](6.70[t])	45.6 (3.43, 4.07)		
E29	116.7 (8.06)	56.4 (4.50)	29.1 (1.91, 2.27)	C^{γ} 32.6 (2.36, 2.40)
S30	112.0 (8.24)	57.4 (5.83)	65.3 (3.94, 4.01)	
I31	122.3 (9.19)	60.9 (4.32)	41.8 (1.29)	H^{γ} (0.20[t]), $C^{\gamma 2}$ 17.5 (0.34), $C^{\delta 2}$ 12.6 (−0.38)
V32	125.3 (8.92)	60.9 (4.52)	34.2 (1.10)	$C^{\gamma 1}$ 20.5 (0.60), $C^{\gamma 2}$ 20.8 (0.74)
V33	127.0 (9.08)	58.7 (4.57)	33.5 (2.27)	$C^{\gamma 1}$ 21.1 (0.76), $C^{\gamma 2}$ 20.2 (0.49)
P34		65.3 (4.13)	(2.05[t], 2.50[t])	C^{γ} 27.9 (1.80[t], 2.18[t]), C^{δ} 50.8 (3.73, 3.49)
N35	111.35 (8.24)	52.4 (4.80)	40.3 (2.74, 2.76)	
R36	131.3 (9.56)	60.7 (3.51)	29.6 (1.41, 1.70)	C^{γ} 24.9 (0.82, −0.01), C^{δ} 43.2 (2.32, 2.44), N^{ϵ} 83.4 (6.97)
E37	118.4 (8.50)	59.8 (3.96)	27.2 (2.09, 2.21)	C^{γ} 32.7 (2.41)
R38	116.8 (8.32)	58.8 (4.08)	30.0 (1.83, 1.92)	C^{δ} 42.8 (3.21, 3.26), N^{ϵ} 83.4 (7.32)
F39	117.8 (7.82)	61.6 (4.00)	39.8 (3.20, 3.23)	C^{δ} 131.2 (6.96), C^{ϵ} 130.3 (6.11), C^{ζ} 128.0 (6.10)
V40	121.8 (8.48)	66.1 (3.20)	31.3 (2.13)	$C^{\gamma 1}$ 21.6 (0.96), $C^{\gamma 2}$ 24.7 (1.11)
Q41	114.7 (7.83)	58.1 (3.99)	29.3 (1.95, 2.03)	C^{γ} 33.5 (2.31, 2.53)
E42	110.5 (7.79)	56.5 (4.29)	29.6 (1.12, 1.59)	C^{γ} 32.6 (2.11, 2.16)
V43	116.2 (7.40)	64.1 (4.05)	33.2 (1.65)	$C^{\gamma 1}$ 23.0 (0.87), $C^{\gamma 2}$ 23.5[t] (0.84)
L44	118.9 (6.79)	60.1 (3.73)	39.6 (1.65, 0.56)	C^{γ} 26.3 (1.35), C^{δ} 25.4 (0.56), C^{δ} 24.6 (0.48)
P45		64.9 (4.45)		C^{δ} 50.9 (2.82, 3.83)
K46	113.9 (7.36)	57.6 (3.80)	31.9 (1.73, 1.31)	C^{γ} 25.5 (1.08, 1.32), C^{δ} 28.9 (1.63), C^{ϵ} 41.7 (2.92)
Y47	114.3 (7.26)	59.1 (4.02)	42.4 (2.22, 2.15)	C^{δ} 133.8 (6.96), C^{ϵ} 117.4 (6.68)
F48	117.6 (8.06)	55.8 (4.79)	40.3 (2.61, 3.12)	C^{δ} 132.8 (7.35), C^{ϵ} 130.6 (7.20), C^{ζ} 129.3 (6.85)
K49	122.0 (8.45)	57.2 (4.07)	32.6 (1.48, 1.68)	C^{γ} 24.8 (1.19, 1.26), C^{δ} 28.9 (1.58), C^{ϵ} 41.7 (2.87)
H50	116.6 (8.89)	55.0 (4.62)	27.5 (3.22, 3.39)	
S51	114.0 (8.02)	57.3 (4.65)	65.0 (3.65 deg)	
N52	116.6 (8.31)	51.8 (4.69)	39.2 (3.01, 3.16)	
F53	121.5 (9.41)	62.9 (4.03)	39.0 (2.85, 3.05)	C^{δ} 132.6 (6.92), C^{ϵ} 130.2 (6.76), C^{ζ} 128.8 (6.42)
A54	119.2 (8.49)	55.1 (3.95)	20.3 (1.43)	
S55	114.3 (8.33)	61.5 (4.08)	62.7 (3.90, 4.00)	
F56	124.2 (7.81)	60.7 (4.05)	38.7 (2.90, 3.02)	C^{δ} 131.5 (6.62), C^{ϵ} 130.6 (6.60), C^{ζ} 128.3 (6.47)
V57	118.5 (8.23)	66.7 (2.85)	31.1 (1.84)	$C^{\gamma 1}$ 21.6 (0.96), $C^{\gamma 2}$ 23.9 (0.27)
R58	120.0 (7.85)	59.7 (3.94)	29.8 (1.84 deg)	H^{γ} (1.48, 1.73), C^{δ} 43.2 (3.21 deg), N^{ϵ} 83.8 (7.25)
Q59	119.1 (7.67)	59.2 (3.66)	28.0 (0.60, 1.34)	C^{γ} 34.5 (1.83, 2.07)
L60	118.9 (7.93)	58.8 (3.68)	39.9 (1.06, 1.53)	C^{γ} 25.8 (0.50), C^{δ} 23.2 (−0.12), C^{δ} 24.8 (−0.92)

(*continued*)

Table 1. *Continued*

Residue	Amide	C^{α}	C^{β}	Other
N61	116.2 (8.06)	56.2 (4.46)	39.0 (2.87, 2.95)	
M62	120.1 (7.79)	58.3 (4.00)	31.6 (1.77, 1.94)	C^{γ} 31.3 (2.17, 2.24)
Y63	115.4 (7.43)	58.2 (4.67)	38.0 (2.69, 3.76)	C^{δ} 133.9 (7.25), C^{ϵ} 117.6 (6.71)
G64	105.5 (7.59)	45.8 (3.62, 4.17)		
W65	121.4 (8.39)	57.0 (4.84)	30.4 (2.66, 2.98)	$C^{\delta 1}$ 126.1 (7.02), $C^{\epsilon 3}$ 119.5 (7.37), $C^{\zeta 3}$ 121.9 (7.15) $C^{\eta 2}$ 124.5[t] (7.31[t]), $C^{\zeta 2}$ 114.5 (7.70), $N^{\epsilon 1}$ 128.8 (10.52)
H66	118.3 (8.91)	53.5 (5.12)	31.5 (3.26, 3.36)	
K67	125.6 (9.36)	56.9 (4.49)	33.2 (1.72, 1.88)	C^{ϵ} 41.7 (2.12, 2.40)
V68	126.3 (8.48)	62.0 (4.15)	33.1 (1.82)	C^{γ} 21.1 (0.87), H^{γ} (0.87)
Q69	125.4 (8.60)	55.5 (4.28)	29.4 (1.91, 2.00)	C^{γ} 33.8 (1.95, 2.24)
D70	122.0 (8.45)	52.8 (4.65[t])	38.6 (2.70[t], 2.80[t])	
V71	120.1 (8.14)	62.4 (4.08)	32.4 (2.01)	C^{γ} 21.2 (0.85), C^{γ} 20.2 (0.85)
K72	123.9 (8.34)	56.6 (4.29)	32.6 (1.71, 1.80)	C^{γ} 25.5 (1.40, 1.45), C^{δ} 28.9 (1.68), C^{ϵ} 41.7 (2.96)
S73	115.5 (8.19)	58.3 (4.43)	63.8 (3.84, 3.88)	
G74	110.6 (8.45)	45.6 (3.92, 4.01)		
S75	115.2 (8.20)	58.2 (4.43)	63.8 (3.85, 3.93)	
M76	121.6 (8.43)	55.8 (4.48)	31.9 (2.05, 2.11)	C^{γ} 31.9 (2.52, 2.61)
L77	121.7 (8.18)	55.0 (4.32)	42.0 (1.58, 1.63)	C^{γ} 27.1 (1.59), C^{δ} 23.2 (0.85), C^{δ} 25.1 (0.89)
S78	115.0 (8.14)	58.3 (4.41)	63.7 (3.86, 3.93)	
N79	120.2 (8.42)	53.2 (4.70)	38.5 (2.84[t])	
N80	118.6[t] (8.42[t])			
D81	118.9[t] (8.37[t])			
S82	116.4 (8.20)	59.0 (4.55)	64.4 (4.00 deg)	
R83	122.5 (7.93)	55.2 (4.85)	31.0 (1.72, 1.80)	H^{γ} (1.35, 1.50), C^{δ} 43.6 (3.12 deg), N^{ϵ} 84.4 (7.33)
W84	127.3 (9.00)	56.6 (4.52)	31.5 (2.93, 3.00)	$C^{\delta 1}$ 125.6 (7.20), $C^{\epsilon 3}$ 120.3 (7.50), $C^{\zeta 3}$ 122.0 (7.33) $C^{\eta 2}$ 124.6 (7.47), $C^{\zeta 2}$ 114.8 (7.67), $N^{\epsilon 1}$ 129.6 (10.08)
E85	121.9 (8.51)	53.5 (5.33)	31.0 (1.80, 1.89)	C^{γ} 33.4 (2.19, 2.36)
F86	120.7 (9.29)	55.7 (5.52)	43.1 (2.80, 3.01)	H^{δ} (6.92)
E87	121.0 (9.38)	54.4 (5.33)	32.2 (1.99, 2.02)	C^{γ} 32.6 (2.43, 2.46)
N88	123.9 (8.56)	52.1 (4.37)	37.6 (0.96[t], 1.90[t])	
E89	123.0 (8.50)	55.9 (4.35)	28.5 (1.91, 2.11)	C^{γ} 32.6 (2.40, 2.42)
R90	120.4 (8.19)	56.0 (4.22)	30.1 (1.71, 1.77)	C^{δ} 43.4 (3.18), N^{ϵ} 84.2 (7.18)
H91	119.6 (8.29)	54.8 (4.66)	28.8 (3.22, 3.25)	
A92	128.3 (8.33)	52.5 (4.24)	19.1 (1.34)	

[a] Chemical shifts for ^{13}C and ^{1}H (in parentheses) are given in ppm.
[b] Both signals have the same chemical shift.
[t] Tentative assignment.

residues adopt an extended conformation. As shown in Figure 5, these segments form a 4-stranded antiparallel β-sheet, as evidenced by the long-range $d_{\alpha N}$, d_{NN}, and $d_{\alpha\alpha}$ NOEs observed in 3D ^{15}N- and ^{13}C-separated NOESY-HMQC and a 4D $^{13}C/^{13}C$-edited NOESY spectrum. Another segment consisting of residues Gln-69–Ser-83 (L1) also exhibited medium intensity $d_{\alpha N}$ NOEs. However, no medium-range or long-range NOEs were identified in the NOESY spectra. Furthermore, no protected amides were found, and the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ shifts for these residues show relatively small shifts (Fig. 4). Together, these data indicate that this region does not conform to a regular pattern of secondary structure.

Stereoassignments and χ_1 angles

Stereoassignments and χ_1 angles for 7 of the 9 valines, both isoleucines, and the single threonine were determined on the basis of $^{3}J_{NC_{\gamma}}$ and $^{3}J_{C'C_{\gamma}}$ coupling constants (Table 2) (Grzesiek et al., 1993; Vuister et al., 1993). For the valine residues, the relatively large $^{3}J_{NC_{\gamma}1}$ and $^{3}J_{C'C_{\gamma}2}$ coupling constants and the comparatively small $^{3}J_{NC_{\gamma}2}$ and $^{3}J_{C'C_{\gamma}1}$ coupling constants indicate that each adopts the *trans* configuration. The $^{3}J_{NC_{\gamma}2}$ coupling for Val-13 is slightly larger than for the other valines, which may reflect some conformational averaging or skewing of χ_1 toward the *gauche*-+ conformation. By comparison, in peptides $^{3}J_{C'C_{\gamma}}$ for valine typically falls in the range of 2–2.5 Hz, which is significantly different than the measured values of $^{3}J_{C'C_{\gamma}2}$ (3.4 Hz) and $^{3}J_{C'C_{\gamma}1}$ (<1 Hz) for Val-13. These data, together with a value of 1.7 Hz for $^{3}J_{NC_{\gamma}1}$, show that although there may be some averaging or skewing toward *gauch*-+, Val-13 exists predominantly as the *trans* rotamer. Similar arguments were used to determine stereospecific assignments and χ_1 angles for Ile-22, Ile-31, and Thr-26. Where possible (including Val-13) intraresidue $d_{N\gamma m}$ and $d_{\alpha\gamma m}$ NOEs were used to confirm the χ_1 determinations (Zuiderweg et al., 1985).

Tertiary structure

To study the tertiary structure of HSF, distance and dihedral constraints determined from the NMR data were used in a hy-

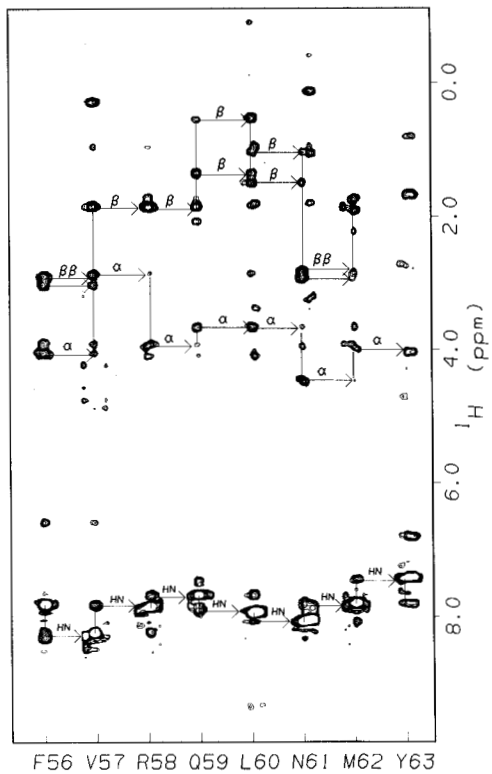

Fig. 3. Sequential strips for residues Phe-56 to Tyr-63 extracted from a 3D ^{15}N-separated NOESY-HMQC spectrum of HSF. Residues are indicated by 1-letter amino acid code and residue number. Intraresidue $d_{\alpha N}$ and $d_{\beta N}$ NOEs are labeled with Greek letters. Arrows indicate NOEs to sequential amide protons. The helical secondary structure of this region is indicated by the strong sequential d_{NN}, $d_{\beta N}$, and weak sequential $d_{\alpha N}$ NOEs.

brid DG/SA protocol (Nilges et al., 1988; Brünger, 1992). The final structures were calculated on the basis of 175 structurally significant intraresidue, 211 sequential, 93 medium-range, and 104 long-range distance constraints, 35 ϕ and 9 χ_1 dihedral restraints, and 30 hydrogen bond constraints (15 hydrogen bonds). The distribution of NOE constraints is shown in Figure 8A. The H bond assignments accounted for 15 of the 34 most slowly exchanging amides. Four H bonds were assigned in the β-sheet and 4 and 7 H bonds, respectively, were assigned in helices H1 and H3. Because of the distortion identified in helix H2, which may cause nonregular H bonding patterns, no H bonds were assigned for this helix. Of the 30 structures calculated, 24 had acceptably low energies and no large distance or torsion angle violations (see Materials and methods). The best fit superposition of the 24 structures is shown in Figure 6. The RMSD relative to the mean ⟨RMSD⟩ for the backbone positions of the protein core (residues 4–14, 31–68, and 84–86) was 1.03 ± 0.19 Å (Fig. 6A). This value should be compared to ⟨RMSD⟩s of 0.4–0.5 Å for highly refined structures (Powers et al., 1993; Lodi et al., 1994). The present resolution of these structures is sufficient, however, to define the backbone coordinates of HSF (Fig. 7) and will improve with the addition of further long-range NOEs, stereoassignment of the β-protons, and additional χ_1 dihedral constraints. The protein topology is as follows: helix H1; a helical/turn region (residues Lys-16–Lys-20); β-strand B1, a short Ω-loop (residues Ser-25–Ser-30); β-strand B2; helix H2; helix H3; β-strand B3; extended loop L1, and β-strand B4. As can be seen in Figure 7, the β-sheet forms a platform against which the C-termini of helices H1 and H3 and the N-terminus of helix H2 rest. The 3 amphipathic helices together form a 3-helix bundle with the hydrophobic faces packing against one another and against the hydrophobic face of the antiparallel β-sheet. The β-strand B3 is significantly rotated away from the plane formed

Table 2. *$^3J_{NC}$ and $^3J_{CC}$ coupling constants,[a] stereo assignments,[b] and χ_1 angles for valine, isoleucine, and threonine residues of HSF*

Residue	Chemical shift (ppm)	$^3J_{NC_\gamma}$ (Hz)	$^3J_{C'C_\gamma}$ (Hz)	Residue	Chemical shift (ppm)	$^3J_{NC_\gamma}$ (Hz)	$^3J_{C'C_\gamma}$ (Hz)	χ_1 NMR (±20°)	χ_1 X-ray (deg)
V6γ1	21.6 (0.90)	>1.6[c]	<1.0	V6γ2	23.0 (0.80)	<1.0	~3[c]	180	180
V13γ1	21.1 (−0.41)	1.8 ± 0.1	<1.0	V13γ2	22.2 (0.13)	1.2 ± 0.1	3.4 ± 0.3	~180[d]	74
V32γ1	20.5 (0.60)	1.7 ± 0.1	~1.0	V32γ2	20.8 (0.74)	<1.0	~3[c]	180	−176
V33γ1	21.1 (0.76)	~1.8[c]	<1.0	V33γ2	20.2 (0.49)	<0.5	3.8 ± 0.4	180	177
V40γ1	21.6 (0.96)	nd[e]	<1.0	V40γ2	24.7 (1.11)	<1.0	~3.7[c]	180	165
V43γ1	23.0 (0.87)	~1.6[c]	<1.0	V43γ2	23.5 (0.84)	<1.0	~3[c]	180	177
V57γ1	21.6 (0.96)	~1.8[c]	<1.0	V57γ2	23.9 (0.27)	nd	3.5± 0.3	180	174
I22γ_m	14.1 (0.88)	~1.0[f]	<1.0					180	−179
I31γ_m	17.5 (0.34)	2.2 ± 0.1	<1.0					−60	−64
T26γ_m	22.1 (1.27)	1.6 ± 0.1	<1.0					−60	58

[a] Coupling constants were determined from spin-echo difference ct-HSQC spectra of HSF (Grzesiek et al., 1993; Vuister et al., 1993). The precision in the coupling constants was determined by calculating RMS differences in intensities of leucine and isoleucine C_δ methyl groups. The coupling constants have been corrected for a 5–7% ($^3J_{CC}$) and 7–10% ($^3J_{NC}$) underestimate due to systematic errors inherent in the method by multiplication of the measured values by 1.06 ($^3J_{CC}$) and 1.08 ($^3J_{NC}$). Systematic errors increase linearly with the size of the coupling constant.

[b] Chemical shifts of ^{13}C and ^1H (in parentheses) are referenced to TSP. Stereoassignments were not determined for Val-68 or Val-71 due to overlap.

[c] Estimate due to partial overlap

[d] Compared to the other valines, the larger than normal value for V13γ2 $^3J_{NC}$ suggests some degree of rotamer averaging. This measurement was not included as a constraint in structure calculations.

[e] Not determined due to overlap.

[f] Estimate. Could not be integrated due to partial overlap with a t_1 streak.

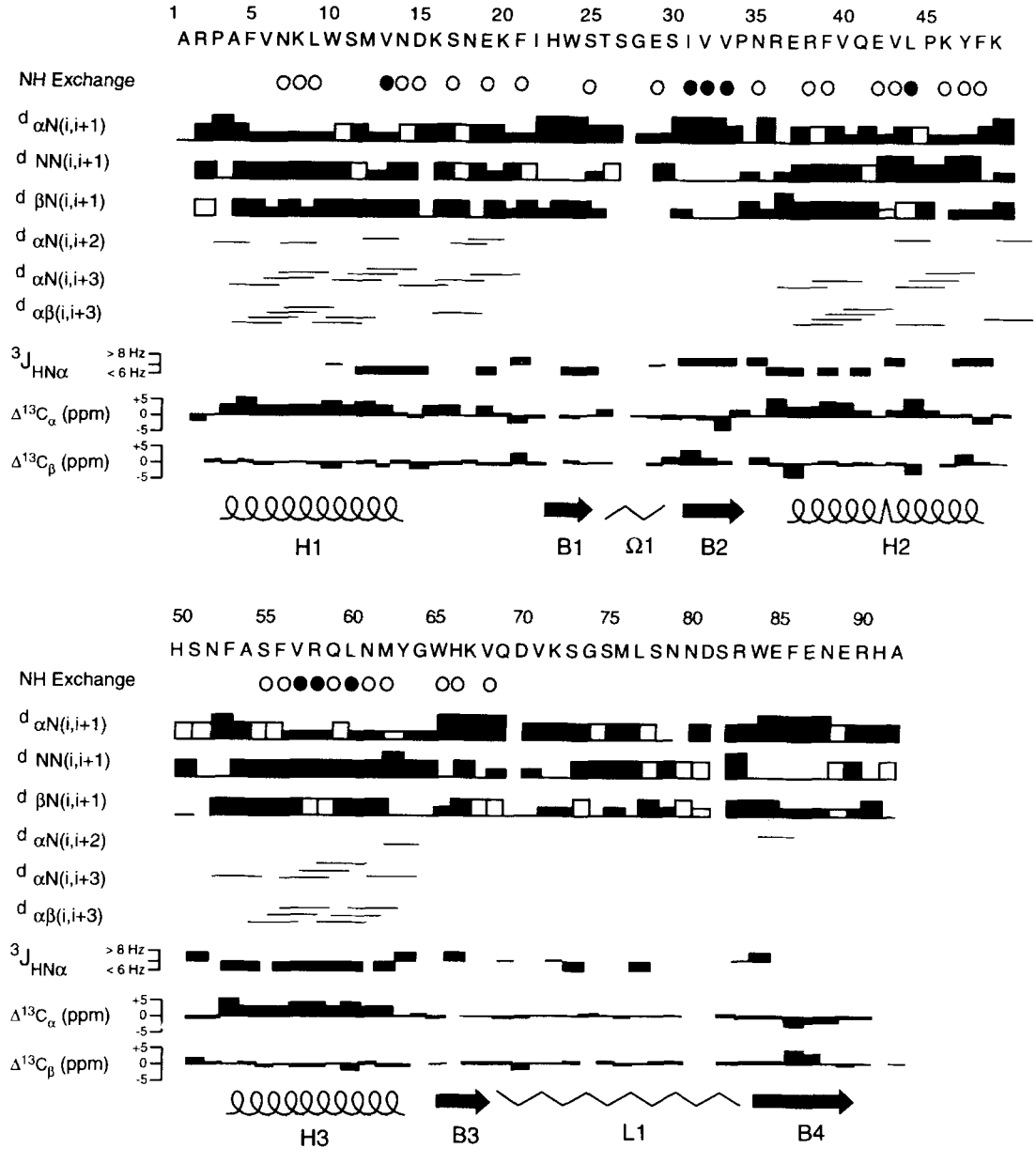

Fig. 4. Summary of sequential and intermediate-range NOE connectivities, amide hydrogen exchange rates, $^3J_{NH_\alpha}$ coupling constants, and deviations of $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts from random coil values. The NOE correlations were determined from 2D NOESY, 3D ^{15}N- and ^{13}C-separated NOESY-HMQC spectra, and a 4D $^{13}C/^{13}C$ edited NOESY spectrum, all recorded at 25 °C and pH 3.4. Sequential NOE intensities are denoted by bar height as strong, medium, or weak. Open boxes indicate upper bounds on sequential NOEs in cases where overlap prevented accurate intensity determination. Lines indicate that no sequential NOE could be detected. For proline residues, NOE intensities to their H_δ protons were substituted for NOE intensities to HN protons. For medium-range NOE connectivities, lines denote that an NOE was observed correlating the relevant protons. Open and filled circles indicate residues with amide HN resonances that are not fully exchanged 370 min and several weeks after dissolution in D_2O (25 °C, pH 2.75), respectively. Secondary structure deduced from the NOESY data are indicated below the NOE data: helix by coils, β-strands by arrows, and unstructured regions by jagged lines.

by β-strands B1, B2, and B4. In addition, the loop L1 extends away from the outside face of the β-sheet.

Superpositions for the 3-helix bundle (⟨RMSD⟩ of 0.81 ± 0.15 Å) and the 4-stranded antiparallel β-sheet (⟨RMSD⟩ of 0.99 ± 0.27 Å) are shown in Figure 6B and C, respectively. The distortion in H2, identified previously, is clearly evident near Gln-41. It can also be seen that H2 and H3 are connected by a short segment such that the axis of the C-terminal portion of H2 lies almost perpendicular to the axis of H3. It is clear from Figure 6B and C that the 3-helix bundle and the β-sheet are well defined and that the Ω-loop connecting strands B1 and B2 and the extended loop L1 connecting B3 and B4 are less well defined. This is confirmed by examining the ⟨RMSD⟩ of backbone coordinates and the ϕ angular order parameters (S(ϕ))

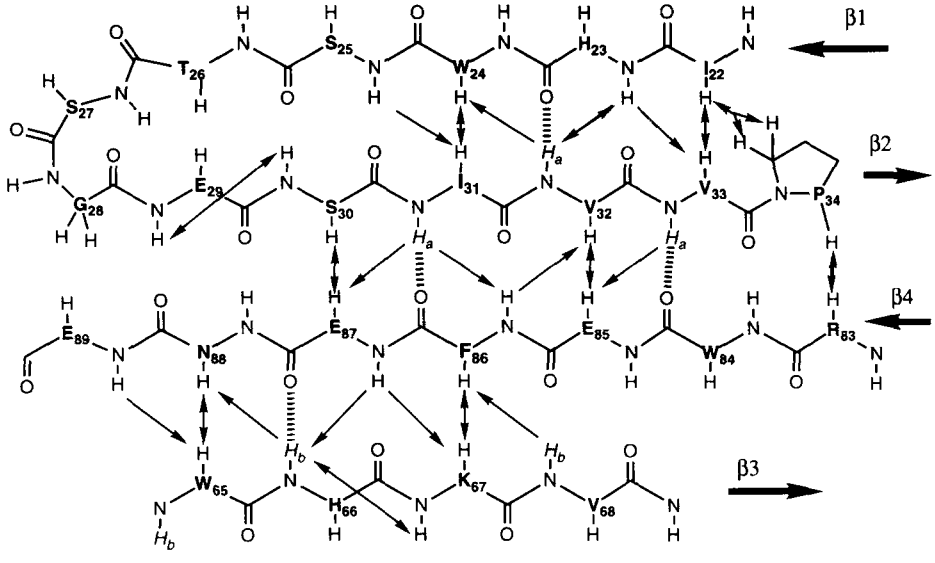

Fig. 5. Schematic diagram of the β-sheet topology of the DNA-binding domain of HSF. Arrows denote long-range d_{NN}, $d_{\alpha N}$, and $d_{\alpha\alpha}$ NOEs. Dashed lines represent hydrogen bonds derived from the hydrogen exchange and NOE data. Thick arrows indicate the direction and location of the 4 strands of antiparallel β-sheet. H_a and H_b indicate amides that were still present after dissolution in D_2O for 2 weeks and 370 min, respectively.

(Hyberts et al., 1992) as a function of sequence (Fig. 8B,C). The N- and C-termini, the Ωloop, and the extended loop (L1) all show ⟨RMSD⟩s greater than 2 Å and S(φ) values less than 0.6. These regions were excluded from the overall superposition. Segments that showed moderate ⟨RMSD⟩s and small S(φ) values that were included in the superposition were the helical/turn region between H1 and B1 (Lys-16–Lys-20) (⟨RMSD⟩ of 1.12 ± 0.19 Å and average S(φ) of 0.85 ± 0.17), the turn between H2 and H3 (Lys-49–Phe-53) (⟨RMSD⟩ of 1.05 ± 0.27 Å and S(φ) of 0.73 ± 0.27), and the distorted β-strand B3 (Trp-65–Val-68), which shows the largest ⟨RMSD⟩s (1.88 ± 0.60 Å) and smallest S(φ) (0.63 ± 0.24). Indeed, when this strand is excluded from the superposition of the β-sheet backbone coordinates, the ⟨RMSD⟩ is reduced from 0.99 ± 0.27 Å to 0.54 ± 0.11 Å. This region shows a greater range of conformations in the family of structures (Fig. 6C) because there are fewer long-range NOE distance constraints between strand B3 and other regions of the structure (Fig. 8A) as compared to the other β-strands.

Discussion

Using a combination of 3D triple-resonance CBCANH and CBCA(CO)NH (Grzesiek & Bax, 1992a, 1992b), double-resonance HCCH-TOCSY (Bax et al., 1990b; Clore et al., 1990), 3D ^{15}N TOCSY-HMQC (Driscoll et al., 1990), and various 2D experiments, we have assigned almost all of the ^1H, ^{15}N, and ^{13}C signals of the 92-residue DNA-binding domain of HSF from *Kluyveromyces lactis* (Table 1). The CBCANH, CBCA(CO)NH, and HCCH-TOCSY experiments rely primarily on strong 1- and 2-bond coupling constants, which allow for assignment of the NMR signals without reference to the secondary structure of the protein as in more traditional methods (Wüthrich, 1986). In addition, for HSF, this combination of experiments provided for the relatively efficient assignment of the NMR signals with a minimal use of spectrometer time (all of the above-mentioned 3D experiments were collected in approximately 9 days).

Analysis of 2D NOESY, 3D ^{15}N- and ^{13}C-separated NOESY-HMQC spectra, a 4D $^{13}C/^{13}C$-edited NOESY spectrum, and D_2O exchange spectra, yielded a total of 583 NOE distance constraints and 15 H bonds (30 H bond constraints) that were used to characterize the secondary and tertiary structure of HSF. Analysis of HMQC-J (Kay & Bax, 1990) and sp¹¬-echo difference ct-HSQC spectra (Grzesiek et al., 1993; Vuister et al., 1993) also provided important φ and χ_1 torsion angle restraints as well as stereospecific assignments for many of the C_γ methyl groups. Together, the distance and dihedral constraints were sufficient to define the conformation of HSF using a hybrid DG/SA protocol (Brünger, 1992).

Unlike other DNA-binding proteins characterized to date which typically bind as monomers or dimers, HSF binds to DNA as a homotrimer. Specifically, a C-terminal triple-stranded coiled coil (Sorger & Nelson, 1989; Peteranderl & Nelson, 1992) directs 3 N-terminal DNA-binding units consisting of approximately 90 residues (Flick et al., 1994) to a series of 5 base pair-inverted repeats of the form nGAAn (Amin et al., 1988; Xiao & Lis, 1988). Most of the residues that are conserved among all known HSFs (Harrison et al., 1994) are located in helices H2 and H3 (Kinemage 1) and include all of the surface side chains of helix H3. The high degree of conservation of the residues in these helices combined with their arrangement in the structure (Figs. 6B, 7) lead us to conclude that these helices constitute a DNA-binding region and that HSF is a member of the HTH family of DNA-binding proteins. Similar suggestions have been made based on the location of secondary structure elements in *Drosophila melanogaster* HSF (Vuister et al., 1994) and a crystal structure of *K. lactis* HSF (Harrison et al., 1994). Moreover, our NMR data show that the secondary structure and topology (α1-β1-β2-α2-α3-β3-β4) of HSF and CAP are almost identical. Except for an additional β-strand (B1), this structure also resembles that of the HTH eukaryotic transcription factor hepatocyte nuclear factor/*forkhead* (HNF-3_γ) (Clark et al., 1993) as suggested previously (Harrison et al., 1994; Vuister et al., 1994).

Unusual structural features of HSF

Although the HTH is a familiar DNA-binding motif, 2 novel features were found in HSF that add to the considerable vari-

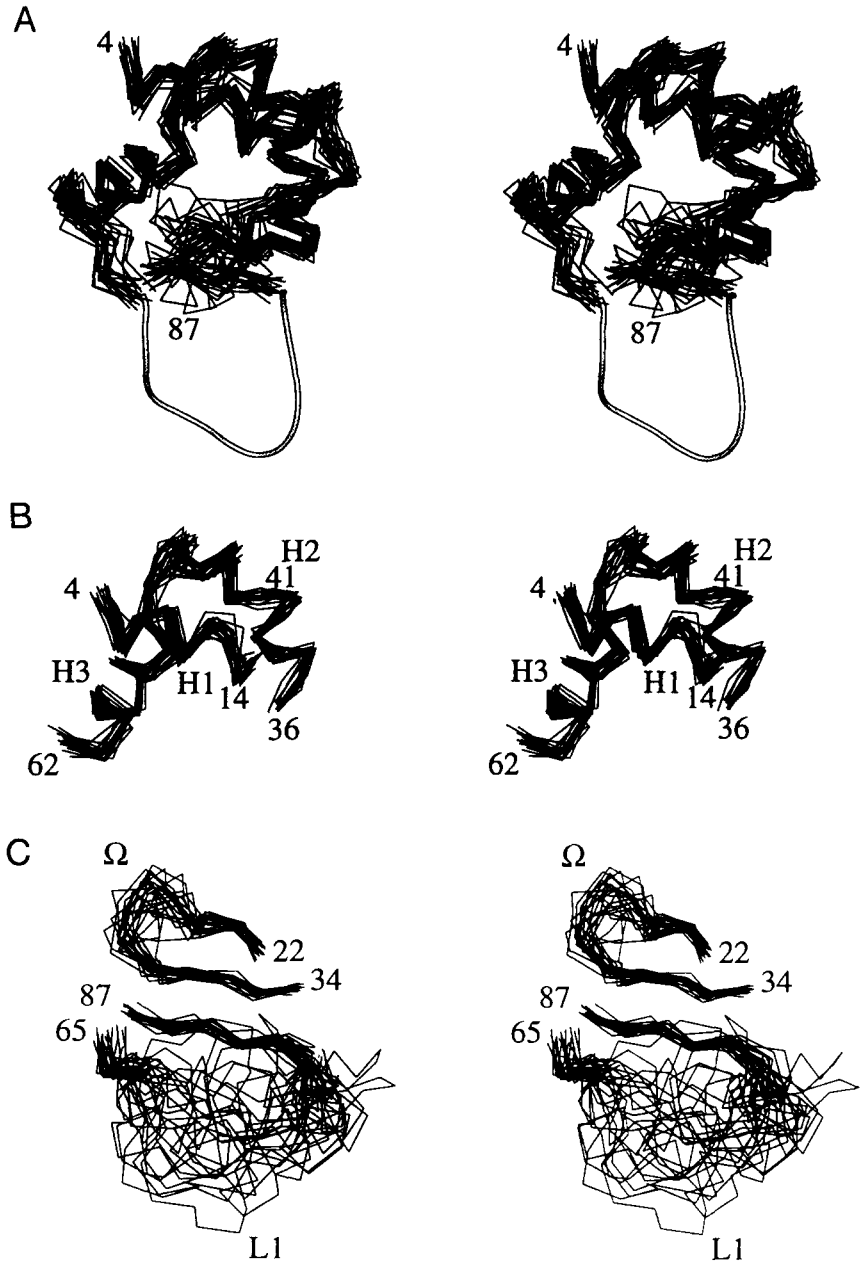

Fig. 6. A: Stereo view of the superposed C_α coordinates for 24 accepted structures of HSF. Shown are the C_α coordinates for residues 4–68 and 83–87, along with a tube indicating the position of the C_α atoms of the poorly defined loop (L1) for 1 of the 24 structures. For the alignment, only backbone coordinates of residues in well-defined segments of structure (4–25, 31–68, 84–86) were used. RMSD for these backbone coordinates is 1.03 ± 0.19 Å. Ala-4 and Glu-87 are labeled to indicate the termini of the chain trace. **B:** Stereo view of the superimposed C_α coordinates of the 3-helix bundle for the 24 accepted structures. For the alignment, the backbone coordinates of the residues in the 3 helices and the turn between helices H2 and H3 (residues 4–14, 36–62) were used. RMSD of these backbone atoms relative to those of the mean structure was 0.81 ± 0.15 Å. The helices are labeled as defined in Figure 4. The N- and C-termini and Gln-41 are labeled by residue number. The central bulge is visible in helix H2 near Gln-41. **C:** Stereo view of the superimposed C_α coordinates of the β-sheet and connecting loops for the 24 accepted structures. Structures were aligned with the mean structure using backbone coordinates for the 4 strands (residues 22–24, 31–34, 65–68, and 84–86) (RMSD for these coordinates is 0.99 ± 0.27 Å). The Ω-loop connecting β-strands B1 and B2 is clearly somewhat ill-defined and the extended loop connecting β-strands B3 and B4 is highly undefined. The N- and C-terminal amino acids of the 4 termini are labeled.

ations of the HTH observed in other proteins (Harrison & Aggarwal, 1990; Steitz, 1990; Brennan, 1993; Feng et al., 1994). One such feature of HSF is the irregular structure in the central portion of helix H2 as evidenced by $^3J_{NH_\alpha}$ coupling constants, $^{13}C_\alpha$ chemical shifts, and NOE data (Fig. 4), and the presence of a conserved proline (Pro-45) toward the C-terminus of H2. The structure calculations show a discontinuity in helix H2 between residues Glu-42 and Pro-45 (Fig. 7). Typically, the presence of a proline residue in the central portion of a helix produces a slight bend but does not significantly distort the ϕ/ψ angles of residues on either side of the bend (Barlow & Thornton, 1988). The distortion observed in HSF appears to be atypical with distortions on either side of the proline, although a detailed study will require further refinement of the data. Another unusual feature is the ill-defined loop (L1) between residues Gln-69 and Lys-83 that connects β-strands 3 and 4 (Fig. 7). This region, lacking in CAP, is reminiscent of loop L1 that connects β-strands 2 and 3 in HNF-3_γ (Clark et al., 1993), which makes important contacts to DNA in the HNF-3_γ-DNA complex.

An interesting feature of HSF is that helix H3 terminates with a Shellman-like CAP sequence (Shellman, 1980; Aurora et al., 1994) (C3[Leu-60]-C2-C1-Ccap-C′-C″-C‴[Trp-65]) except that the C1 position is occupied by an apolar residue (Met-62). This termination sequence is characterized by an apolar side-chain contact between C3 and C″ (confirmed by side-chain NOEs between Leu-60 and Leu-65), 2 H bonds from C1 to C′ and C2 to C″ (confirmed by slowed amide proton exchange of Gly-64 [fully exchanged in just under 370 min] and Trp-65), and positive ϕ angle for the Gly (average ϕ in NMR family of +88°, S(ϕ) of 0.8).

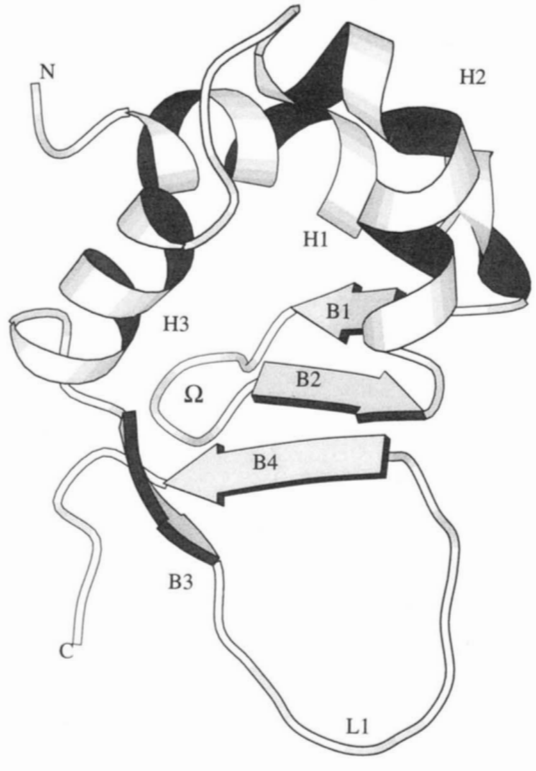

Fig. 7. Ribbon diagram of HSF showing the relative orientations of the 3 helices and the 4-stranded antiparallel β-sheet. The structure with the the lowest target function energy at the end of the last X-PLOR refinement is represented. The helices and sheets and 2 loops are labeled as defined in Figure 4. The N- and C-termini are also labeled. The figure was created with MOLSCRIPT software (Kraulis, 1991).

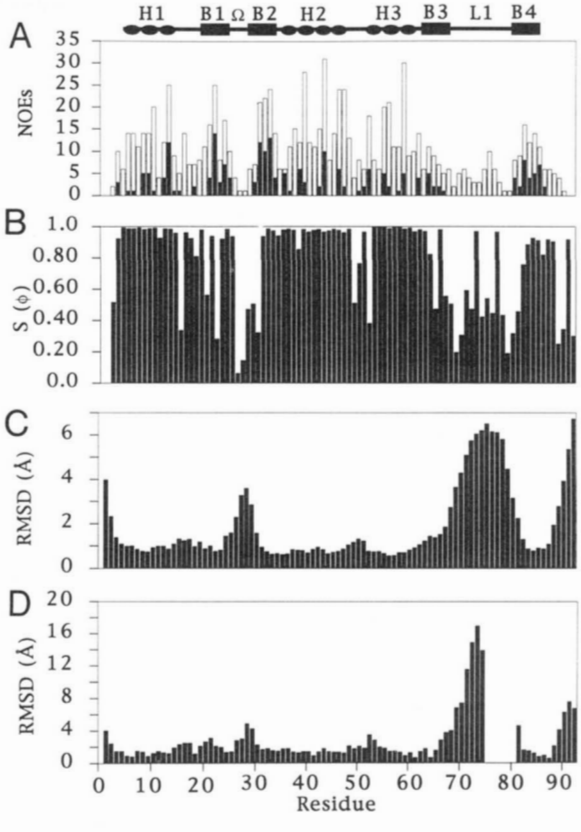

Fig. 8. A: Distribution of NOE constraints for HSF. The height of each bar represents the total number of NOEs used for that residue in the structure calculations. Shaded bars represent long-range NOEs. For each constraint, the originating and destination residues were each counted once. **B:** The angular order parameter S^{angle} for ϕ dihedral angles as a function of residue for the 24 accepted structures. An S value of 0.9 corresponds to a standard deviation in the ϕ angle of $\pm 24°$ (Hyberts et al., 1992). **C:** RMS differences of backbone (N, CA, C, O) coordinates for the family of 24 accepted structures relative to the average structure. The superposition included residue Ala-4–Ser-25, Ile-31–Val-68, and Trp-84–Phe-86. The Ω-loop (Ser-25–Ser-30) and the L1-loop (Val-68–Trp-84) show large deviations and are poorly defined by the NMR constraints. The β-strand B3 (residues Trp-65–Val-68) also shows larger deviations than other regions of secondary structure. **D:** RMS differences of backbone coordinates for the family of 24 structures relative to the 1.8-Å X-ray structure. Superposition as for Figure 8C.

Hydrophobic packing

Side chains whose packing is crucial to preserving the overall fold of HSF should correspond to conserved hydrophobic residues in the sequence. The interior face of the β-sheet is made up of hydrophobic residues, most of which are conserved, whereas the exterior face is hydrophilic and poorly conserved. The orientation of H3 with respect to the β-sheet is defined by the contacts between their hydrophobic surfaces. This arrangement must be crucial to the functioning of the domain judging from the high degree of sequence conservation at the interface. The N-terminus of helix H1, whose axis is at right angles to that of H3, makes numerous contacts with β-strand B1 and to a lesser extent with β-strand B2. Helix H1 also makes many contacts to the other 2 helices. The orientation of helix H2 is determined primarily by side-chain contacts with H3 across the turn region of the HTH and a number of contacts to H1 and B4.

Proposed residues that interact with DNA

Considering the DNA contacts seen in other HTH proteins complexed with DNA, it is very likely that the majority of DNA contacts in the HSF–DNA complex will be made by the exposed residues of helix H3 (the recognition helix) such as Ala-54, Ser-55, Arg-58, Gln-59, Asn-61, and Met-62. The positions of several conserved positively charged residues in the solution structure suggest that the side chains of Lys-8, Lys-46, and Lys-49 located above the N-terminus of H3 and Lys-67 below the C-terminus of H3 are important to the functioning of the domain. Because they are highly conserved, positively charged, and located so close to the recognition helix (H3), it is likely that these side chains make important contacts to the phosphate backbone.

Although the evolutionary link between bacteria and yeast is very distant, it is interesting to note that the putative HTH of the *Escherichia coli* alternative σ factor for heat shock (σ-32) shows 40% sequence identity (65% sequence similarity) with the HTH region of HSF (based on BESTFIT alignment [GCG version 7.3; Devereux et al., 1984] of residues 253–272 of *E. coli* σ-32 with residues 36–62 of *K. lactis* HSF). Comparing HSF sequences reported recently for several organisms (Harrison et al., 1994) using the same alignment gives an average of 32% sequence identity/50% sequence similarity. Moreover, the −35

promoter consensus of σ-32, TNtCNcCTTGAA (Cowing et al., 1985), contains the sequence GAA, and genetic evidence derived from σ-70 (a homolog of σ-32) implicate Arg-268 and Ala-264 in the putative recognition helix of σ-32 as making contacts with the G and the last A, respectively, of the GAA sequence (Gardella et al., 1989; Siegele et al., 1989; Dombroski et al., 1992). The 2 homologous residues in HSF, Arg-58 and Ala-54 are on the exposed face of the recognition helix H3 and may make sequence-specific contacts to the GAA sequences within the HSEs in a fashion analogous to that proposed for σ-32. Confirmation of this hypothesis must await further genetic and biophysical characterization of the HSF-DNA complex.

Comparison to other structural studies of HSF

After the signal assignments for HSF had been completed and the secondary structure had been defined, a 1.8-Å X-ray-derived model of the protein was solved and made available (Harrison et al., 1994). We note that, through the use of the heteronuclear experiments described herein, it was possible to make assignments and determine the overall fold of the protein without reference to the crystal structure. The topology of the X-ray model is similar to the model derived from the NMR data. Both structures contain a 4-stranded antiparallel β-sheet and a 3-helix bundle. An irregular segment defined as an α-helical bulge was identified in helix H2 and is consistent with the NMR results. The loop (L1) with ill-defined structure connecting β-strands B3 and B4 (Fig. 6C) is consistent with the results of the X-ray model, which shows larger *B*-factors in this region than in the rest of the structure, and absence of electron density for residues 76–79 (Harrison et al., 1994).

Recently, the assignments and secondary structure of HSF from *D. melanogaster* were reported (Vuister et al., 1994). The secondary structure from *Drosophila* contains 3 helices similar to our NMR and X-ray (Harrison et al., 1994) results on HSF from *K. lactis*. A significant difference, however, is the identification of a 3-stranded rather than 4-stranded antiparallel β-sheet. The segment corresponding to B3 in *K. lactis* HSF is reported to be in an extended conformation in *D. melanogaster*. A more detailed comparison must await further analysis of the NMR data for both proteins.

RMSD comparison to X-ray model

At the current stage of refinement, it is appropriate to compare the overall fold of the family of NMR structures with the X-ray structure and to compare in more detail the distance constraints used in the X-PLOR calculations with those predicted from the crystal structure. Superposition of the backbone coordinates for all residues except the Ω-loop (26–30), the loop L1 (69–83), and the first 3 and last 6 residues onto the X-ray model reveals good overall agreement with an RMSD of 1.77 ± 0.19 Å (Fig. 8D; Kinemage 2). Comparison of the β-sheet coordinates reveals similarly good agreement (RMSD of 1.65 ± 0.24 Å). If β-strand B3 is excluded (residues 65–68), the agreement is significantly better (RMSD of 0.85 ± 0.14 Å), indicating that the orientation of B3 deviates from the X-ray model. As stated previously, this β-strand is not precisely defined due to the small number of cross-strand constraints connecting it to the rest of the structure (Fig. 8A). The 3-helix bundle also shows good agreement with the crystal structure with an RMSD of 1.50 ± 0.18 Å for backbone coordinates. The similarity of the NMR and X-ray results is significant given that the solvent content of the crystals used in the X-ray analysis was unusually low (~30%).

Backbone NOE comparison to X-ray model

To identify possible differences between the solution and crystal structures, we compared the backbone distances derived from NOEs and ϕ angles derived from $^3J_{HNH_\alpha}$ coupling constants with the corresponding distances and angles in the X-ray model. The comparison revealed that the constraints derived from the majority of backbone NOEs and all of the measured ϕ dihedral constraints were in agreement with the X-ray model. A few differences in measured and predicted backbone NOE intensities were noted, which could be grouped into 2 classes: those where the observed NOE was weaker than expected compared to the crystal structure, and those where the NOE was stronger than expected. In the first class we found 5 examples; $d_{\alpha N}(5,6)$, $d_{\alpha N}(9,10)$, $d_{\alpha N}(40,41)$, $d_{\alpha N}(42,43)$, and $d_{NN}(15,16)$. For the 4 $d_{\alpha N}$ NOEs, a weak NOE was observed although a medium intensity NOE was expected (3.4–3.5 Å in the X-ray model), whereas for $d_{NN}(15,16)$ no NOE was observed, although a weak NOE was expected (4.3 Å in the X-ray model). There are many reasons why an NOE might be weaker, including exchange broadening or H_α saturation. In the present cases, the differences could not be explained by these reasons because other NOEs to these amide protons were observed and because these α-protons all resonate far from H_2O (4.77 ppm). Hence, the data suggest true structural differences. In the second class, 2 differences were observed, namely $d_{\beta N}(13,14)$ and $d_{NN}(66,67)$. In both cases, the observed NOE is of medium intensity but was predicted to be weak based on the X-ray coordinates (4.1 Å, and 4.3 Å). These disagreements between the NMR data and X-ray results may reflect differences in the 2 structures, although spin diffusion cannot be ruled out. Taken together, the differences observed for the backbone are relatively small, although it is interesting to note that the locations of the "perturbed" sequential NOEs correspond roughly to locations of protein–protein contacts in the crystal. Asn-7 is involved in such contacts, as are Asn-14, Gln-41, Glu-42, and His-66. We have also observed several side-chain NOEs with altered intensities involving residues in these regions of HSF. These will be analyzed in detail at a later stage of refinement.

χ_1 Angle comparison to X-ray model

Of the 10 measured χ_1 angles (Table 2), 2 show significant differences with those observed in the X-ray structure. The NMR data for Val-13 and Thr-26 indicate χ_1 values of 180° (with evidence for some rotamer averaging near 180°) and −60°, respectively, whereas in the X-ray structure, the measured angles were 74° (Val-13) and 58° (Thr-26). Val-13 is located at the C-terminus of H1. The methyl groups point toward the interior of the protein and show NOEs to residues in β-strands B1 and B2. We have also noted several differences between side-chain NMR distance constraints that involve Val-13 and residues close to it. One previously mentioned is the sequential side-chain–backbone NOE $d_{\beta N}(13,14)$, which the X-ray structure predicts to be weak but shows a medium-intensity NOE. This result is consistent with the change in the side-chain conformation of Val-13 in solution, placing Val-13 H_β closer to Asn-14 H_N than pre-

dicted from the χ_1 angle observed in the X-ray structure. The structural difference in this region is probably related to differences in side-chain packing rather than global backbone conformation, as evidenced by the close agreement of the backbone coordinates for the family of NMR structures with the crystal structure. Further refinement is needed to characterize these differences in detail. Thr-26 is located in the Ω-loop (Ser-25–Ser-30), which is on the surface of the structure. This region shows no long-range NOEs to other parts of the structure except for the 2 terminal residues of the Ω-loop (Ser-25 and Ser-30), which show NOEs to the neighboring β-strands. Consequently, it appears less well defined in the NMR family of structures (Fig. 6B). In the crystal structure on the other hand, the Ω-loop is involved in extensive protein–protein interactions and has *B*-factors that are similar to the rest of the structure. It is possible that Thr-26 prefers 1 rotamer state in the solution structure due to local packing of its γ methyl group with other atoms in the Ω-loop or hydrogen bond interactions with the side-chain hydroxyl group and that this interaction is perturbed in the crystal by protein–protein contacts.

Summary

In conclusion, the backbone geometry of the X-ray structure is very similar to the solution structure (Kinemage 2), although subtle differences appear to exist, some of which probably relate to protein–protein contacts in the crystal. These differences are located at the end of helix H1 ($d_{\alpha N}(5,6)$, $d_{\alpha N}(9,10)$, $d_{\beta N}(13,14)$, and $d_{NN}(15,16)$ NOEs, and Val-13 χ_1), in the Ω-loop (Thr-26 χ_1), at the distortion of helix H2 ($d_{\alpha N}(40,41)$ and $d_{\alpha N}(42,43)$), and in the β-strand B3 $d_{NN}(66,67)$. To make a more detailed comparison and characterize the nature of these changes, particularly with respect to side-chain packing and distortions in the DNA-binding surface if they exist, a high-resolution solution structure will be needed that includes stereoassignments of the β-protons and a greater number of distance and χ_1 constraints.

The NMR signal assignments and the solution structure of HSF presented here provide the basis for the design of new biochemical and biophysical experiments necessary for an understanding of HSF–DNA and HSF–HSF interactions and the effects of crystal packing on the protein structure. For example, the heteronuclear ^{15}N and ^{13}C assignments can be used to probe the dynamics of this DNA-binding domain and should also aid studies of larger HSF–DNA complexes through the use of 3D and 4D NMR experiments. The solution structure and, in particular, the identification of an HTH motif should also help direct mutagenesis studies aimed at the identification of residues important for complex formation. Moreover, further refinement of the solution structure should allow for the identification of any structural changes that occur in HSF upon DNA binding.

Materials and methods

Sample preparation

The HSF DNA-binding domain from *K. lactis* (residues 194–282) was expressed in *E. coli* using a T7 expression system (Studier et al., 1990). The amino acid sequence encoded by the plasmid construct includes an N-terminal Met and the C-terminal sequence Arg-His-Ala. Samples labeled with ^{15}N and $^{13}C/^{15}N$ were prepared by growing bacteria on minimal media containing $^{15}NH_4Cl$ or $^{15}NH_4Cl/^{13}C_6$-glucose as the sole nitrogen and carbon sources, respectively. The protein overexpression and purification protocol have been described elsewhere (Harrison, 1994). Approximately 20 mg of purified protein was obtained from each liter of cell growth. Purity was determined to be greater than 99% for all samples as determined by electrospray ionization mass spectrometry. The molecular weights obtained indicated the absence of the N-terminal Met as well as complete incorporation of the isotopic labels (C.J. Harrison & D. King, unpubl. results).

NMR samples in H_2O were prepared by dissolving the lyophilized protein in 450 μL H_2O and 50 μL D_2O. NMR samples in D_2O were prepared by dissolving protein in 500 μL of 99.96% D_2O, relyophilizing the protein, and redissolving it again in 99.96% D_2O. All samples were buffered to pH 3.4 using 10 mM potassium phosphate buffer. Typical protein concentrations were 2 to 4 mM. Samples were stored in NMR tubes at 4 °C.

NMR spectroscopy

NMR experiments were performed on a Bruker AMX 600 spectrometer equipped with a 3-channel interface, external 50-W linear amplifier (ENI, Rochester, New York) for ^{13}C frequencies, triple-resonance (^1H, ^{15}N, ^{13}C) carbon-optimized probe, and UXNMR software version 930301. All experiments were conducted at 25 °C. Presaturation was accomplished using rf field strengths of 35 Hz for samples in D_2O and 140 Hz for samples in H_2O with the carrier placed on the water resonance (4.77 ppm). All NMR data were processed using the FELIX program (version 2.0) distributed by Biosym Technologies (San Diego, California). All linear prediction was implemented using FELIX (version 2.0). DG and SA calculations were performed using the X-PLOR program version 3.0 (Brünger, 1992). All calculations were performed on SGI workstations (Silicon Graphics, Mountain View, California).

^1H-^1H 2D NMR

Two-dimensional ^1H-^1H spectra were collected with 512 t_1 and 1,024 complex t_2 points using time proportional phase incrementation (Marion & Wüthrich, 1983) for quadrature detection in t_1 and with solvent presaturation. Spectral widths were 14.12 ppm in both dimensions. NOESY (Macura et al., 1981), DQF-COSY (Rance et al., 1983), and clean-TOCSY experiments (Griesinger et al., 1988) were performed in both H_2O and D_2O. Two-dimensional NOESY spectra were recorded with a mixing time of 100 ms, during which the solvent was irradiated. Two-dimensional clean-TOCSY spectra (Griesinger et al., 1988) were acquired with a 100-ms MLEV17 (Bax & Davis, 1985) mixing sequence. The ratio of clean-TOCSY delay Δ, to the spinlock 90° pulse length τ_{90}, was 2.5, with an average rf power for the spinlock of 7.5 kHz. All ^1H-^1H 2D spectra were 0 filled to 1K by 1K points. In some cases, linear prediction was applied in the second dimension as implemented in FELIX version 2.0. Data were apodized in both dimensions using shifted sine-bell window functions with phase shifts ranging from 45° to 90° for different data sets.

Heteronuclear 2D NMR

Quadrature detection for all heteronuclear 2D NMR experiments was accomplished using States-TPPI (Marion et al., 1989b). Unless otherwise stated, in all heteronuclear 2D experiments, WALTZ16 modulation (Shaka et al., 1983) was used for ^{15}N decoupling and GARP (Shaka et al., 1985) was used for ^{13}C decoupling. Two-dimensional ^{1}H-^{15}N HSQC spectra (Bodenhausen & Ruben, 1980; Bax et al., 1990c) and 2D ^{1}H-^{15}N HMQC spectra (Mueller, 1979; Bax et al., 1990c) were acquired on uniformly ^{15}N- or $^{13}C/^{15}N$-labeled HSF in 90% H_2O/10% D_2O. The water signal was suppressed using presaturation in the HMQC experiments and both presaturation and a 1.8-ms spinlock pulse (Messerle et al., 1989) were used in HSQC experiments. Typically the data were collected with spectral widths of 29.78 ppm in t_1 (512 complex points) and 14.12 ppm in t_2 (1,024 complex points).

Amide exchange rates were probed by lyophilizing the protein from H_2O and redissolving the protein in D_2O (pH 2.75, uncorrected for ^{2}H isotope effects) and acquiring a series of 2D ^{1}H-^{15}N HMQC spectra at 13.1, 37.7, 76.8, 156, and 370 min after addition of D_2O. Very slowly exchanging amide signals were observed in spectra obtained from a sample dissolved in D_2O for several weeks. Spectra were 0 filled to 1K by 1K points and apodized using 60° skewed sine-bell window functions.

Two-dimensional ^{1}H-^{13}C ct-HSQC spectra of the aliphatic region of HSF uniformly labeled with $^{13}C/^{15}N$ were recorded as described (Vuister & Bax, 1992). Solvent suppression was achieved using presaturation and a 1-ms spinlock pulse (Messerle et al., 1989). Carbon decoupling during detection was accomplished using GARP (Shaka et al., 1985) modulation of a 4.2-kHz rf field. A constant time delay of 26.6 ms was used that resulted in oppositely phased signals for carbons attached to an even or odd number of ^{13}C neighbors. The data were collected with spectral widths of 33.13 ppm (^{13}C) in t_1 (128 complex points) and 14.12 ppm (^{1}H) in t_2 (512 complex points). Spectra were processed as described (Vuister & Bax, 1992) except that linear prediction as implemented in FELIX (version 2.0) was used.

Two-dimensional ^{1}H-^{13}C ct-HSQC and ct-HCCH-TOCSY spectra (Ikura et al., 1991a) optimized for detection of aromatic resonances were collected with the ^{13}C carrier set to 125 ppm and INEPT transfer delays set to 1.3 ms. The spectra were collected with spectral widths of 26.29 ppm (^{13}C) in t_1 (56 complex points) and 11.73 ppm (^{1}H) in t_2 (512 complex points). The constant time delay was tuned to ^{13}C-^{13}C couplings of 70 Hz (14.3 ms) and no carbonyl pulses were applied. The ^{13}C mixing time in the ct-HCCH-TOCSY experiment was accomplished using a 14.34 ms DIPSI-2 sequence (Shaka et al., 1988) with a 7.6-kHz spinlock field. Both data sets were apodized with a 15° skewed sine-bell in t_2 and a 35° skewed sine-bell in t_1. Linear prediction as implemented in FELIX (version 2.0) was applied in the t_1 time domain. The final data sets were 1,024 × 1,024 real points in size.

$^{3}J_{NH_\alpha}$ coupling constants were measured semiquantitatively (either less than 6 Hz or greater than 8 Hz) from an ^{15}N HMQC-J spectrum (Kay & Bax, 1990) in which 650 complex points (t_{1max} = 214 ms) were collected. Exponential line broadening (7 Hz) and exponential line narrowing (−2 Hz) were applied in the ^{1}H and ^{15}N dimensions, respectively. After 0 filling, the final digital resolution was 4 Hz (^{1}H) and 2 Hz (^{15}N). These data were used to estimate ϕ torsion angles for 36 residues.

$^{3}J_{NC_\gamma}$ and $^{3}J_{C'C_\gamma}$ coupling constants for Val, Ile, and Thr residues were determined from 2D ^{15}N and ^{13}C spin-echo difference ct-HSQC experiments as described (Grzesiek et al., 1993; Vuister et al., 1993) using a constant time period of 28.6 ms. These data were used to determine χ_1 torsion angle constraints for 7 of 9 Val, both Ile, and the single Thr residue, and to stereospecifically assign the $^{13}C_\gamma$ signals for 7 of the 9 Val residues.

^{15}N-Separated 3D NMR

Three-dimensional ^{15}N-separated TOCSY-HMQC (Driscoll et al., 1990) and 3D ^{15}N-separated NOESY-HMQC experiments (Kay et al., 1989; Marion et al., 1989c) were acquired on uniformly ^{15}N-labeled samples of HSF in 90% H_2O/10% D_2O. Quadrature detection in the indirectly detected dimensions was accomplished using TPPI (Marion & Wüthrich, 1983). In both experiments, GARP modulation (Shaka et al., 1985) was used to decouple ^{15}N. Solvent suppression was implemented using presaturation. The TOCSY-HMQC experiment was performed using a 100-ms clean TOCSY MLEV17 mixing sequence (Bax & Davis, 1985; Griesinger et al., 1988). The ratio of the clean TOCSY delay, Δ, to the spinlock 90° pulse length, τ_{90}, was equal to 2.5 with an average spin-lock field strength of 6.7 kHz. The 3D-^{15}N-separated NOESY-HMQC experiment was performed using a 100-ms mixing time. Both the TOCSY and NOESY experiments were acquired with spectral widths of 14.12 ppm, 60.0 ppm, and 14.12 ppm in F1 (^{1}H), F2 (^{15}N), and F3 (^{1}H), respectively, and with 200 real points in t_1, 62 real points in t_2, and 1,024 real points in t_3. A total of 16 scans were signal averaged for each real t_1, t_2 pair. Shifted sine-bell window functions were applied in all 3 dimensions and linear prediction (FELIX program, version 2.0) and 0 filling in t_1 and t_2 resulted in a final matrix for both experiments of 512 (^{1}H) × 512 (^{1}H) × 64 (^{15}N) points.

Triple resonance NMR

CBCANH (Grzesiek & Bax, 1992a) and CBCA(CO)NH (Grzesiek & Bax, 1992b) triple resonance experiments were performed as described on uniformly $^{13}C/^{15}N$-labeled samples in 90% H_2O/10% D_2O. Presaturation and a 1.8-ms spinlock pulse (Messerle et al., 1989) were used to suppress the solvent in the CBCANH experiment, whereas only the 1.8-ms spinlock pulse was retained for solvent suppression in the CBCA(CO)NH experiment. Both the CBCANH and CBCA(CO)NH spectra were acquired with spectral widths of 56.0, 24.9, and 14.12 ppm for F1 (^{13}C; 52 complex points), F2 (^{15}N; 32 complex points), and F3 (^{1}H; 512 complex points), respectively, and with 32 scans per complex t_1, t_2 pair. The carriers were placed at 43.0 ppm, 117.9 ppm, and 4.77 ppm for the ^{13}C, ^{15}N, and ^{1}H channels, respectively. ^{15}N decoupling during acquisition was accomplished using WALTZ16 (Shaka et al., 1983) modulation of a 1.6-kHz rf field. ^{1}H decoupling during the ^{13}C evolution period was achieved using DIPSI-2 modulation (Shaka et al., 1988) of a 5.4-kHz rf field. Time domain convolution (Marion et al., 1989a) was used to eliminate the residual solvent signal in both experiments and a 50° skewed sine-bell window function was applied in the t_3 dimension. Linear prediction (FELIX version 2.0) was applied in t_1 and t_2 followed by multiplication by unshifted cosine-

squared window functions. Both matrices were 512 (^{1}H) × 256 (^{13}C) × 64 (^{15}N) real points.

^{13}C-separated 3D experiments

3D-HCCH-TOCSY (Bax et al., 1990b; Clore et al., 1990) and 3D-^{13}C-separated NOESY-HMQC spectra (Ikura et al., 1990) were recorded using uniformly ^{13}C/^{15}N-labeled HSF samples in 99.96% D_2O. Quadrature detection was achieved using the States-TPPI method (Marion et al., 1989b). In both experiments the ^{13}C carrier was set to 43 ppm and ^{13}C decoupling during acquisition was accomplished using GARP modulation (Shaka et al., 1985) of a 4.1-kHz field. Solvent suppression was achieved in the HCCH-TOCSY experiment using presaturation and a 1.8-ms spinlock pulse. In the ^{13}C-separated NOESY-HMQC spectrum, 0.5-ms and 1.8-ms spinlock pulses were used for solvent suppression (Messerle et al., 1989). Mixing in the ^{13}C HCCH-TOCSY experiment was accomplished using a DIPSI-2 sequence (Shaka et al., 1988) with a mixing time of 22.6 ms and a spinlock field of 5.4 kHz. The mixing time in the ^{13}C-separated NOESY-HMQC experiment was 100 ms. Both experiments were collected with 128 complex points in t_1 (^{1}H), 32 complex points in t_2 (^{13}C), and 256 complex points in t_3 (^{1}H), and with 32 scans per complex t_1, t_2 pair. Spectral widths for the HCCH-TOCSY spectrum were 8.01, 23.8, and 8.01 ppm in F1 (^{1}H), F2 (^{13}C), and F3 (^{1}H), respectively, and for the ^{13}C-separated NOESY-HMQC spectrum they were 10.93, 23.8, and 13.44 ppm in F1 (^{1}H), F2 (^{13}C), and F3 (^{1}H), respectively. The HCCH-TOCSY data were multiplied by 50°-shifted sine-bell window functions in all 3 dimensions and linear prediction (FELIX version 2.0) was applied in t_1, resulting in a final matrix size of 512 (^{1}H) × 128 (^{1}H) × 64 (^{13}C) points. The ^{13}C-separated NOESY data were 0 filled and then apodized with a 90° skewed sine-bell squared window function in t_3. Data in the t_1 and t_2 dimensions were multiplied by a skewed sine-bell window function and linear prediction was applied in t_1 to arrive at a final matrix size of 256 (^{1}H) × 128 (^{1}H) × 64 (^{13}C) real points.

4D ^{13}C/^{13}C-edited NOESY spectrum

A 4D ^{13}C/^{13}C-edited NOESY spectrum was recorded as described (Clore et al., 1991) on a ^{13}C/^{15}N-labeled sample of HSF dissolved in 99.96% D_2O with a mixing time of 100 ms. The acquired data matrix consisted of 8 (t_1, ^{13}C) × 64 (t_2, ^{1}H) × 8 (t_3, ^{13}C) × 256 (t_4, ^{1}H) complex points with spectral widths of 23.8 ppm in both ^{13}C dimensions and 9.15 ppm in both ^{1}H dimensions. The ^{13}C carrier was centered at 43 ppm during labeling of carbon magnetization and the ^{1}H carrier was positioned at 3.5 ppm except for solvent presaturation, which was accomplished at 4.77 ppm. Data were processed as described (Clore et al., 1991) except that linear prediction as implemented in FELIX (version 2.0) was used to extend each ^{13}C dimension from 8 to 12 points followed by 0 filling to 32 complex points. The final data matrix consisted of 32 (F1) × 128 (F2) × 32 (F3) × 256 (F4) real points.

Chemical-shift referencing

Chemical shifts were directly referenced to 3-(2,2,3,3-$^2H_4$) trimethylsilyl propionate (^{1}H) and indirectly referenced to liquid ammonia (^{15}N) (Live et al., 1984) and TSP (^{13}C) (Bax & Subramanian, 1986).

Structure calculations

NOEs were classified qualitatively as strong (1.8–2.7 Å), medium (1.8–3.5 Å), and weak (1.8–5.0 Å) (Williamson et al., 1985; Clore et al., 1986). A correction was added to the upper limit for constraints involving methyl protons and nonstereospecifically assigned protons (Wüthrich et al., 1983; Clore et al., 1986). ϕ Torsion angles were constrained to $-60° \pm 30°$ for $^3J_{NH_\alpha}$ values that were less than 6 Hz, and to $-120° \pm 40°$ for $^3J_{HN_\alpha}$ values that were greater than 8 Hz (Pardi et al., 1984). χ_1 Angles were constrained to within ±20° of the measured values (Kraulis et al., 1989). Hydrogen bonds were identified on the basis of NOEs identified in regular secondary structures and low amide proton exchange rates (still present after 370 min of exchange). Constraints were applied between N and O atoms (2.8–3.3 Å) and between H_N and O atoms (1.8–2.3 Å) (Clore et al., 1990).

Structures were calculated using a hybrid DG/SA method (Nilges et al., 1988) as contained in the X-PLOR program 3.0 (Brünger, 1992). Random starting structures were calculated using the X-PLOR routine *generate-template*. A subset of atoms from this extended conformation was used for embedding using the distance and torsion angle constraint matrix with the *dg_subembed* routine. The hybrid DG/SA routine *dgsa* was applied to the full structure as described (Brünger, 1992). Final refinement was applied using the routine *refine*.

The initial set of structures calculated on the basis of approximately 300 NOE and 24 hydrogen bond constraints (12 H-bonds) was used to resolve ambiguities in NOE assignments, allowing new NOE constraints to be added in further rounds of refinement. Thirty final structures were calculated and the 24 best were chosen for structural comparison. The chosen 24 structures had no distance violations greater than 0.5 Å and no torsion angle violations greater than 5°. The average total target function energy for the family of structures at the end of X-PLOR refinement was 150.9 ± 13.4 kcal/mol. The average NOE and dihedral violation energies were 25.9 ± 5.4 kcal/mol and 0.96 ± 1.6 kcal/mol, respectively, using force constants of 50 kcal/mol Å^2 and 200 kcal/mol rad^2. The ⟨RMSD⟩s between the distance constraints and calculated structures for the family of 24 structures was 0.028 ± 0.003 Å and the ⟨RMSD⟩ between the dihedral constraints and calculated structures was 0.35 ± 0.13°. The coordinates for these 24 structures have been deposited with the Brookhaven Protein Data Bank (file 2HSF). Superpositions and analysis of the structures was accomplished using X-PLOR and Insight II software.

Acknowledgments

This work was supported through NIH grant GM 44086 and an award from the PEW Scholars Program in Biomedical Sciences (to H.C.M.N.) and by the Director, Office of Energy Research, Office of Biological & Environmental Research, General Life Science Division of the US Department of Energy under Contract No. DE-AC03-76F00098 and through instrumentation grants from the US Department of Energy, DE-FG05-86ER75281, and the National Science Foundation, DMB 86-09305 and BBS 87-20134 (to D.E.W.).

We thank Eva Grotkopp for help with protein expression; Dr. David King for performing electrospray-ionization mass spectrometry and a

critical reading of the manuscript, Drs. Sharon Archer and Stephan Grzesiek for advice on CBCANH and CBCA(CO)NH experiments, and Biosym for β-test versions of the program FELIX.

Supplementary material on Diskette Appendix

A complete list of the restraints used to calculate the structures appears on the Diskette Appendix (file Dambergr.NMR, SUPLEMNT directory). Two kinemages (file Dambergr.kin, KINEMAGE directory) are also provided. Kinemage 1 shows the C_α coordinates of HSF with the conserved regions highlighted, along with the ill-defined regions, and HTH. The location of Gln-41 is also indicated near the distortion in helix H2. Kinemage 2 shows a superposition of the C_α coordinates of one of the NMR structures with the 1.8-Å X-ray structure minimizing the RMSD between the 2 structures for the backbone coordinates in regions of regular secondary structure.

References

Amin J, Ananthan J, Voellmy R. 1988. Key features of heat shock regulatory elements. *J Mol Biol 8*:3761-3769.

Archer SJ, Vinson VK, Pollard TD, Torchia DA. 1993. Secondary structure and topology of *Acanthamoeba* profilin I as determined by heteronuclear nuclear magnetic resonance spectroscopy. *Biochemistry 32*: 6680-6687.

Aurora R, Srinivasan R, Rose GD. 1994. Rules for α-helix termination by glycine. *Science 264*:1126-1130.

Barlow DJ, Thornton JM. 1988. Helix geometry in proteins. *J Mol Biol 201*:601-619.

Bax A, Clore GM, Driscoll PC, Gronenborn AM, Ikura M, Kay LE. 1990a. Practical aspects of proton carbon carbon proton 3-dimensional correlation spectroscopy of ^{13}C-labeled proteins. *J Magn Reson 87*:620-627.

Bax A, Clore GM, Gronenborn AM. 1990b. ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new 3-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J Magn Reson 88*:425-431.

Bax A, Davis G. 1985. MLEV-17-based two dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson 65*:355-360.

Bax A, Ikura M, Kay LE, Torchia DA, Tschudin R. 1990c. Comparison of different modes of 2-dimensional reverse-correlation NMR for the study of proteins. *J Magn Reson 86*:304-318.

Bax A, Subramanian S. 1986. Sensitivity enhanced two-dimensional heteronuclear shift correlation NMR spectroscopy. *J Magn Reson 67*:565-569.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett 69*:185-189.

Boorstein WR, Craig EA. 1990. Structure and regulation of the SSA4 HSP70 gene of *Saccharomyces cerevisiae*. *J Biol Chem 265*:18912-18921.

Brennan RG. 1993. The winged helix motif: Another helix-turn-helix takeoff. *Cell 74*:773-776.

Brünger AT. 1992. *X-PLOR manual version 3.0: A system for crystallography and NMR*. New Haven, Connecticut: Yale University Press.

Chen J, Pederson DS. 1993. A distal heat shock element promotes the rapid response to heat shock of the HSP26 gene in the yeast *Saccharomyces cerevisiae*. *J Biol Chem 268*:7442-7448.

Clark KL, Halay ED, Lai E, Burley SK. 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature 364*:412-420.

Clore GM, Bax A, Driscoll PC, Wingfield PT, Gronenborn AM. 1990. Assignment of the side-chain ^{1}H and ^{13}C resonances of interleukin-1β using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy. *Biochemistry 29*:8172-8184.

Clore GM, Kay LE, Bax A, Gronenborn AM. 1991. Four-dimensional ^{13}C/^{13}C-edited nuclear Overhauser enhancement spectroscopy of a protein in solution: Application to interleukin 1β. *Biochemistry 30*:12-18.

Clore GM, Nilges M, Sukumaran DK, Brünger AT, Karplus M, Gronenborn AM. 1986. The three-dimensional structure of α1-purothionin in solution: Combined use of nuclear magnetic resonance, distance geometry, and restrained molecular dynamics. *EMBO J 5*:2729-2735.

Cowing DW, Bardwell JC, Craig EA, Woolford C, Hendrix RW, Gross CA. 1985. Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc Natl Acad Sci USA 82*:2679-2683.

Craig EA, Gambill BD, Nelson RJ. 1993. Heat shock proteins: Molecular chaperones of protein biogenesis. *Microbiol Rev 57*:402-414.

Devereux J, Haeberli, P, Smithies O. 1984. A comprehensive set of sequence analysis programs for VAX. *Nucl Acids Res 12*:387-395.

Dombroski AJ, Walter WA, Record MTJ, Siegele DA, Gross CA. 1992. Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell 70*:501-512.

Driscoll PC, Clore GM, Marion D, Wingfield PT, Gronenborn AM. 1990. Complete resonance assignment for the polypeptide backbone of interleukin 1β using three-dimensional heteronuclear NMR spectroscopy. *Biochemistry 29*:3542-3556.

Feng JA, Johnson RC, Dickerson RE. 1994. Hin recombinase bound to DNA: The origin of specificity in major and minor groove interactions. *Science 263*:348-355.

Flick K, Gonzalez LJ, Harrison CJ, Nelson HCM. 1994. Yeast heat shock factor contains a flexible linker between the DNA-binding and trimerization domains; implications for DNA-binding by trimeric proteins. *J Biol Chem 269*:12475-12481.

Gallo GJ, Prentice H, Kingston RE. 1993. Heat shock factor is required for growth at normal temperatures in the fission yeast *Schizosaccharomyces pombe*. *Mol Cell Biol 13*:749-761.

Gardella T, Moyle H, Susskind MM. 1989. A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J Mol Biol 206*:579-590.

Griesinger C, Otting G, Wüthrich K, Ernst RR. 1988. Clean TOCSY for ^{1}H spins system identification in macromolecules. *J Magn Reson 110*: 7870-7820.

Gross DS, English KE, Collins KW, Lee SW. 1990. Genomic footprinting of the yeast HSP82 promoter reveals marked distortion of the DNA helix and constitutive occupancy of heat shock and TATA elements. *J Mol Biol 216*:611-631.

Grzesiek S, Bax A. 1992a. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J Magn Reson 99*:201-207.

Grzesiek S, Bax A. 1992b. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc 114*:6291-6293.

Grzesiek S, Vuister GW, Bax A. 1993. A simple and sensitive experiment for measurement of $J_{CC'}$ couplings between backbone carbonyl and methyl carbons in isotopically enriched proteins. *J Biomol NMR 3*:487-493.

Harrison CJ. 1994. The structure of the DNA-binding domain from *Kluyveromyces lactis* heat shock transcription factor [thesis]. Berkeley, California: University of California.

Harrison CJ, Bohm AA, Nelson HCM. 1994. Crystal structure of the DNA binding domain of the heat shock transcription factor. *Science 263*: 224-227.

Harrison SC, Aggarwal AK. 1990. DNA recognition by proteins with the helix-turn-helix motif. *Annu Rev Biochem 59*:933-969.

Horvath OW, Lilley DM. 1978. ^{13}C NMR of peptides. *Prog NMR Spectrosc 12*:1-40.

Hyberts SG, Goldberg MS, Havel TF, Wagner G. 1992. The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci 1*:736-751.

Ikura M, Kay LE, Bax A. 1991a. Improved three dimensional ^{1}H-^{13}C-^{1}H spectroscopy of a ^{13}C-labeled protein using constant-time evolution. *J Biomol NMR 1*:299-304.

Ikura M, Kay LE, Tschudin R, Bax A. 1990. 3-dimensional NOESY-HMQC spectroscopy of a ^{13}C-labeled protein. *J Magn Reson 86*:204-209.

Ikura M, Spera S, Barbato G, Kay LE, Krinks M, Bax A. 1991b. Secondary structure and side-chain ^{1}H and ^{13}C resonance assignments of calmodulin in solution by heteronuclear multidimensional NMR spectroscopy. *Biochemistry 30*:9216-9228.

Jakobsen BK, Pelham HR. 1988. Constitutive binding of yeast heat shock factor to DNA in vivo. *Mol Cell Biol 8*:5040-5042.

Jakobsen BK, Pelham HR. 1991. A conserved heptapeptide restrains the activity of the yeast heat shock transcription factor. *EMBO J 10*:369-375.

Kay LE, Bax A. 1990. New methods for the measurement of NH-C-α-H coupling constants in ^{15}N-labeled proteins. *J Magn Reson 86*:110-126.

Kay LE, Marion D, Bax A. 1989. Practical aspects of 3D heteronuclear NMR of proteins. *J Magn Reson 84*:72-84.

Kraulis J, Clore GM, Nilges M, Jones TA, Pettersson G, Knowles J, Gronenborn AM. 1989. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry 28*:7241-7257.

Lis J, Wu C. 1993. Protein traffic on the heat shock promoter: Parking, stalling, and trucking along. *Cell 74*:1-4.

Live DH, Davis DG, Agosta WC, Cowburn D. 1984. Long-range hydrogen bond mediated effects in peptides: ^{15}N NMR study of gramacidin S in water and organic solvents. *J Am Chem Soc 106*:1934-1941.

Lodi PJ, Garrett DS, Kuszewski J, Tsang ML, Weatherbee JA, Leonard WJ, Gronenborn AM, Clore GM. 1994. High-resolution solution structure of the β chemokine hMIP-1 β by multidimensional NMR. *Science 263*:1762–1767.

Macura S, Huang Y, Suter D, Ernst RR. 1981. Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J Magn Reson 43*:259–281.

Marion D, Ikura M, Bax A. 1989a. Improved solvent suppression in one-dimensional and 2-dimensional NMR spectra by convolution of time-domain data. *J Magn Reson 84*:425–430.

Marion D, Ikura M, Tschudin R, Bax A. 1989b. Rapid recording of 2D NMR spectra without phase cycling—Application to the study of hydrogen exchange in proteins. *J Magn Reson 85*:393–399.

Marion D, Kay LE, Sparks SW, Torchia DA, Bax A. 1989c. Three dimensional heteronuclear NMR of ^{15}N labeled proteins. *J Am Chem Soc 111*:1515–1517.

Marion D, Wüthrich K. 1983. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H-^1H spin-spin coupling constants in proteins. *Biochem Biophys Res Commun 113*: 967–974.

Messerle BA, Wider G, Otting G, Weber C, Wüthrich K. 1989. Solvent suppression using a spin lock in 2D NMR and 3D NMR spectroscopy with H_2O solutions. *J Magn Reson 85*:608–613.

Metzler WJ, Constantine KL, Friedrichs MS, Bell AJ, Ernst EG, Lavoie TB, Mueller L. 1993. Characterization of the three-dimensional solution structure of human profilin: ^1H, ^{13}C, and ^{15}N NMR assignments and global folding pattern. *Biochemistry 32*:13818–13829.

Morimoto RI. 1993. Cells in stress: Transcriptional activation of heat shock genes. *Science 259*:1409–1410.

Mueller L. 1979. Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J Am Chem Soc 101*:4481–4484.

Nieto-Sotelo J, Wiederrecht G, Okuda A, Parker CS. 1990. The yeast heat shock transcription factor contains a transcriptional activation domain whose activity is repressed under nonshock conditions. *Cell 62*:807–817.

Nilges M, Clore GM, Gronenborn AM. 1988. Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett 229*:317–324.

Pardi A, Billeter M, Wüthrich K. 1984. Calibration of the angular dependence of the amide proton-C alpha proton coupling constants, $^3J_{HN_\alpha}$, in a globular protein. Use of $^3J_{HN_\alpha}$ for identification of helical secondary structure. *J Mol Biol 180*:741–751.

Park HO, Craig EA. 1989. Positive and negative regulation of basal expression of a yeast HSP70 gene. *Mol Cell Biol 9*:2025–2033.

Pelton JG, Torchia DA, Meadow ND, Wong CY, Roseman S. 1991. ^1H, ^{15}N, and ^{13}C NMR signal assignments of IIIGlc, a signal-transducing protein of *Escherichia coli*, using three-dimensional triple-resonance techniques. *Biochemistry 30*:10043–10057.

Perisic O, Xiao H, Lis JT. 1989. Stable binding of *Drosophila* heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. *Cell 59*:797–806.

Peteranderl R, Nelson HCM. 1992. Trimerization of the heat shock transcription factor by a triple-stranded alpha-helical coiled-coil. *Biochemistry 31*:12272–12276.

Powers R, Garrett DS, March CJ, Frieden EA, Gronenborn AM, Clore GM. 1993. The high-resolution, three-dimensional solution structure of human interleukin-4 determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry 32*:6744–6762.

Rance M, Sorensen OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K. 1983. Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering. *Biochem Biophys Res Commun 117*:479–485.

Shaka AJ, Barker P, Freeman R. 1985. Computer-optimized decoupling scheme for wideband applications and low level operation. *J Magn Reson 64*:547–552.

Shaka AJ, Keeler J, Freeman R. 1983. Evaluation of a new broadband decoupling sequence: Waltz16. *J Magn Reson 53*:313–340.

Shaka AJ, Lee CJ, Pines A. 1988. Iterative schemes for bilinear operators: Application to spin decoupling. *J Magn Reson 77*:274–293.

Shellman C. 1980. The α_L conformation at the ends of helices. In: Jaenicke R, ed. *Protein folding: Proceedings of the 28th Conference of the German Biochemical Society*. New York: Elsevier/North-Holland. pp 53–61.

Siegele DA, Hu JC, Walter WA, Gross CA. 1989. Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. *J Mol Biol 206*:591–603.

Sorger PK. 1990. Yeast heat shock factor contains separable transient and sustained response transcriptional activators. *Cell 62*:793–805.

Sorger PK, Nelson HCM. 1989. Trimerization of a yeast transcriptional activator via a coiled-coil motif. *Cell 59*:807–813.

Sorger PK, Pelham HR. 1988. Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. *Cell 54*:855–864.

Spera S, Bax A. 1991. Empirical correlation between protein backbone conformation and C_α and C_β ^{13}C nuclear magnetic resonance chemical shifts. *J Am Chem Soc 113*:5490–5492.

Steitz TA. 1990. Structural studies of protein-nucleic acid interaction: The sources of sequence-specific binding. *Q Rev Biophys 23*:205–280.

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol 185*:60–89.

Vuister GW, Bax A. 1992. Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Magn Reson 98*:428–435.

Vuister GW, Kim SJ, Wu C, Bax A. 1994. NMR evidence for similarities between the DNA-binding regions of *Drosophila melanogaster* heat shock factor and the helix-turn-helix and HNF-3/forkhead families of transcription factors. *Biochemistry 33*:10–16.

Vuister GW, Wang AC, Bax A. 1993. Measurement of 3-bond nitrogen carbon-J couplings in proteins uniformly enriched in ^{15}N and ^{13}C. *J Am Chem Soc 115*:5334–5335.

Wiederrecht G, Seto D, Parker CS. 1988. Isolation of the gene encoding the *S. cerevisiae* heat shock transcription factor. *Cell 54*:841–853.

Williamson MP, Havel TF, Wüthrich K. 1985. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^1H nuclear magnetic resonance and distance geometry. *J Mol Biol 182*:295–315.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol 222*:311–333.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: John Wiley & Sons.

Wüthrich K, Billeter M, Braun W. 1983. Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J Mol Biol 169*:949–961.

Xiao H, Lis JT. 1988. Germline transformation used to define key features of heat-shock response elements. *Science 239*:1139–1142.

Xiao H, Perisic O, Lis JT. 1991. Cooperative binding of *Drosophila* heat shock factor to arrays of a conserved 5 bp unit. *Cell 64*:585–593.

Zuiderweg ER, Boelens R, Kaptein R. 1985. Stereospecific assignments of ^1H NMR methyl lines and conformation of valyl residues in the *lac* repressor headpiece. *Biopolymers 24*:601–611.