

# Solution Structure of the First HMG Box Domain in Human Upstream Binding Factor<sup>†,‡</sup>

Yingqi Xu, Wulin Yang, Jihui Wu,\* and Yunyu Shi\*

Laboratory of Structural Biology, School of Life Science, University of Science and Technology of China, Hefei, Anhui 230026, Peoples Republic of China

Received November 26, 2001

**ABSTRACT:** Human upstream binding factor is a nucleolar transcription factor involved in transcription by RNA polymerase I. It contains six HMG box domains; the HMG box is a minor groove DNA-binding domain that has been found in hundreds of proteins with different functions. Among the six HMG box domains in hUBF, the first one can bind to the ribosomal promoter specifically by itself and is essential for the whole protein's DNA binding specificity. Here we report the three-dimensional structure of this first HMG box free in solution determined by multidimensional NMR using <sup>13</sup>C,<sup>15</sup>N-labeled protein. Like the previously determined HMG box structures, hUBF HMG box 1 adopts a twisted L-shape consisting of three α-helices: helix 1 (17–30) and helix 2 (38–51) pack onto each other to form the short arm, while helix 3 (57–76) is associated with an extended strand N-terminal to helix 1 and forms the long arm. A cluster of conserved residues, in particular the aromatic residues F21, Y49, and Y60, is important to maintain the fold. The short arm is rigid due to extensive hydrophobic interaction between helix 1 and helix 2, while the long arm is less rigid.

Efficient transcription by human RNA polymerase I requires at least two transcription factors: upstream binding factor hUBF<sup>1</sup> and selectivity factor hSL1 (1). hSL1 is a complex of TATA binding protein (TBP) and three TBP-associated factors, namely, TAF<sub>48</sub>, TAF<sub>63</sub>, and TAF<sub>110</sub> (2). hSL1 decides promoter selectivity, but it cannot bind to ribosomal promoter effectively on its own. By contrast, hUBF alone is able to complex with the ribosomal promoter, as demonstrated by DNase I footprint experiments (1); the footprint is enhanced and extended with the addition of hSL1, indicating cooperative binding to DNA between the two factors. Transcription activation by hUBF necessitates the formation of a triple complex among hSL1, hUBF, and the ribosomal promoter, and it is proposed that the major role of hUBF is to recruit hSL1 to the ribosomal promoter (2).

Analysis of the hUBF primary sequence has revealed six tandem arranged domains with sequence similarity to HMG proteins 1 and 2, which were coined the HMG box (1). These HMG boxes in hUBF were involved in promoter binding as well as interaction with hSL1. Among them, the first HMG box, HMG box 1, played a dominant role in determining

DNA-binding specificity, while the other HMG boxes and the amino terminus enhanced the interaction with DNA (3).

Many other proteins have been found to contain one or more HMG box domains. Structural studies on HMG boxes revealed a common twisted L-shape fold formed by intersection of three α-helices (4–13, 37). The conserved amino acids defining the HMG box were involved in either maintaining this structure or interacting with target DNA. Figure 1 compares the sequence of some HMG boxes with hUBF HMG box 1.

HMG box proteins were generally divided into two subclasses according to their DNA binding preferences (14). The first subclass included sequence-specific transcription factors, such as mammalian testis-determining factor SRY (9) and lymphoid enhancer binding factor LEF-1 (10). The second subclass was typified by HMG1 and HMG2, which bind to DNA with little or no specificity; instead, they recognize DNA structural features, for example, four-way junction and *cis*-platinum adducts. Recent studies emphasized the common architectural role for both subclass proteins of modulating DNA bending and facilitating the assembly of higher order nucleoprotein complexes (38).

hUBF binds to the ribosomal promoter with only a relaxed specificity (1–3): no discernible recognition site could be defined. It may represent another subclass of HMG box proteins lying between the upper two subclasses. This was supported by a homology study of 121 HMG boxes, which classified three (namely, the HMG boxes 1, 3, and 6) of the six HMG boxes in hUBF into a middle subgroup (15). As the first step in RNA pol I transcription, hUBF must bind to the ribosomal promoter specifically; how this specificity is encoded in either the promoter or the multiple DNA-binding domains of hUBF remains unclear. Another interesting

<sup>†</sup> This work is supported by the Chinese National Fundamental Research Project, Grant G1999075605, the Chinese National Natural Science Foundation, Grant 39990600, and the National High Technology Research and Development Program of China and by a grant from the Chinese Academy of Science.

<sup>‡</sup> The atomic coordinate file of hUBF HMG box 1 has been deposited in the Brookhaven Protein Data Bank (code 1K99).

\* To whom correspondence should be addressed. J.W.: fax, +86-551-3603754; e-mail, wujihui@ustc.edu.cn. Y.S.: fax, +86-551-3603754; e-mail, yyshi@ustc.edu.cn.

<sup>1</sup> Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; CSI, chemical shift index; RMSD, root mean square deviation; hUBF, human upstream binding factor; HMG, high mobility group.

	2	10	20	30	40	50	60	70	80																																																																	
HMG1 A	GK	GD	PK	PK	RG	KMS	TA	FT	VQ	T	C	EE	KK	K	H	P	D	A	S	V	N	F	S	E	F	S	K	C	E	R	K	T	M	S	A	K	K	G	K	F	E	D	M	A	K	A	D	K	A	R	E	R	E	M	K	T	I	P	P	K	G	E	83											
HMG1 B	F	K	D	P	N	A	P	K	R	P	P	S	A	F	T	L	C	S	E	Y	R	P	K	I	K	E	G	H	P	--	G	L	S	I	G	D	V	A	K	K	I	G	E	M	N	N	T	A	A	D	D	K	Q	Y	E	K	K	A	A	K	L	E	K	D	I	A	A	R	A	K	77			
HMG-D	S	D	K	P	K	R	L	S	A	M	L	W	L	S	A	N	E	S	K	R	E	N	P	--	G	I	K	V	T	E	A	K	R	G	E	L	W	R	A	M	K	--	D	K	S	E	W	E	A	K	A	K	A	D	D	R	A	V	E	F	E	A	N	G	73									
NHP6A	T	R	K	K	D	P	N	A	P	K	R	A	L	S	A	M	L	W	L	S	A	N	E	S	K	R	E	N	P	--	D	I	T	G	Q	V	G	K	L	G	E	K	K	A	L	T	P	E	E	K	Q	Y	E	K	A	Q	A	D	K	K	R	E	S	E	K	E	L	N	A	T	L	A	82	
hUBF-box1	K	K	L	K	K	H	P	D	F	P	K	K	P	L	P	Y	T	R	F	F	M	E	K	R	A	K	A	L	H	P	--	E	M	S	L	D	L	T	K	L	E	K	K	I	Q	D	F	Q	R	E	K	Q	E	F	E	R	N	L	A	R	F	R	E	D	H	P	D	L	I	Q	N	A	K	89
mSox-5	G	S	P	H	I	K	R	P	M	A	F	M	V	A	K	D	E	R	R	K	L	Q	A	F	P	--	D	M	H	S	N	I	S	K	I	G	S	R	K	A	M	N	L	E	K	Q	Y	E	E	A	R	L	S	K	O	L	E	K	P	D	K	Y	K	P	R	P	K	R	T	81				
hSRY	V	Q	D	R	V	K	R	P	M	A	F	I	V	S	R	D	Q	R	K	M	A	L	E	N	P	--	R	M	R	S	E	I	S	K	O	I	G	Y	Q	N	K	M	L	T	E	A	K	W	P	I	T	Q	E	A	Q	K	L	O	A	M	R	E	K	P	N	I	K	Y	R	P	76			
LEF-1	Q	E	P	K	R	P	H	I	K	P	L	A	F	M	L	M	K	E	M	R	A	N	V	A	E	C	--	L	K	S	A	A	I	N	I	T	G	R	R	H	A	L	S	R	E	E	A	K	Y	E	L	A	R	K	E	R	Q	L	M	Q	L	P	G	S	A	R	D	N	Y	G	K	K	84	

FIGURE 1: Amino acid sequence and alignment of rat HMG1A, rat HMG1B, *Drosophila melanogaster* HMG-D, *Saccharomyces cerevisiae* NHP6A, hUBF HMG box 1, mouse Sox-5, human SRY, and mouse LEF-1. The conserved positions are highlighted in gray. The numbering at the top refers to hUBF HMG box 1 in this paper.

feature of hUBF is that it binds to G-C-rich sequences (1), whereas other HMG box proteins usually prefer A-T-rich sequences, which may have larger bendability to adopt bending induced by HMG box binding.

A straightforward strategy is to obtain the structure of hUBF and its complex with (part of) the ribosomal promoter to make clear the details of their interaction and understand the mechanism underlying the relaxed specificity. Here we report the solution structure of the first HMG box of hUBF as the first step in this direction.

## MATERIALS AND METHODS

**Expression, Purification, and Labeling of hUBF HMG Box 1.** Details for subcloning, expression, and purification of the hUBF HMG box 1 domain were described elsewhere (16). The recombinant protein contains residues 103–192 of hUBF (1), plus an N-terminal Met and a C-terminal His tag (-LEHHHHHH). For preparation of isotopically labeled proteins, cultures were grown on an SV40 medium containing 2.5 g/L [ $^{13}\text{C}_6$ ]glucose and/or 0.5 g/L  $^{15}\text{NH}_4\text{Cl}$  as the sole carbon and nitrogen sources, respectively. The yield was  $\sim 20$  mg/L.

**NMR Sample Preparation and NMR Spectroscopy.** Both the purified  $^{15}\text{N}$ -labeled and  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled proteins were dissolved to a final concentration of 3 mM in 500  $\mu\text{L}$  of water containing 10 mM  $\text{NaPO}_4$ , pH 5.5, and 10%  $\text{D}_2\text{O}$ . All NMR experiments were performed on a Bruker DMX500 spectrometer with self-shielded z-axis gradients.

The following spectra were recorded to obtain backbone and side chain resonance assignments: 2D  $^1\text{H}$ – $^{15}\text{N}$  HSQC (17), 3D triple-resonance spectra HNCO (18), HNCA (19), HN(CO)CA (19), HN(CA)CO (20), CBCA(CO)NH (21), CBCANH (22), H(CA)NH (23), C(CO)NH (24), and H(C-CO)NH (24), 2D  $^1\text{H}$ – $^{13}\text{C}$  HSQC (17), 3D HCCH-TOCSY (25) and HCCH-COSY (26), and 2D CB(CC)HD-COSY (aromatics) (27). 3D  $^{15}\text{N}$ -separated or  $^{13}\text{C}$ -separated NOESY (28) were acquired with a mixing time of 130 ms. Most of the experiments were performed at 300 K on the doubly labeled sample, except that some HSQC and NOESY spectra were recorded at temperatures varying from 290 to 310 K on the  $^{15}\text{N}$ -labeled protein to search the best sample condition.

The  $^{15}\text{N}$ -labeled sample was then lyophilized and dissolved in 99.96%  $\text{D}_2\text{O}$ , which was followed immediately with HSQC experiments to monitor the disappearance of NH signals at 290 K. After all of the peaks had vanished, 2D homonuclear TOCSY and NOESY were recorded on this sample at 300 K, which exhibited proton resonance from aromatic rings exclusively in the region beyond 6.0 ppm.

All spectra were processed using the NMRPipe (29) package and analyzed with the PIPP program; both software were run on the Linux system. Linear prediction (30) was used to improve spectral resolution in the indirect dimensions where constant-time acquisition was used, for example, the  $^{15}\text{N}$  evolution dimension in all the above triple-resonance experiments.

**Experimental Restraints and NMR Structure Calculation.** NOE peaks from three spectra, i.e., 3D  $^{15}\text{N}$ -separated NOESY, 3D  $^{13}\text{C}$ -separated NOESY in water, and 2D NOESY in  $\text{D}_2\text{O}$ , were used to delineate the interproton distance restraints. According to their intensities these peaks were grouped into four classes: strong, 1.8–3.0 Å; medium, 1.8–4.0 Å; weak, 1.8–5.0 Å; and very weak, 1.8–6.0 Å. Considering that the spin diffusion effect could be serious for aliphatic and aromatic protons, a more conservative distance estimation was used for the latter two NOESY; therefore, most medium-range and long-range NOEs from these two spectra were put into the weak or very weak groups, while most intraresidue and sequential NOEs were not used. The 1.8 Å lower limits were imposed only implicitly by the van der Waals repulsion force. For methyl protons, nonstereospecifically assigned methylene protons, and aromatic ring protons,  $r^{-6}$  summation averages (31) were applied.

The chemical shift index (CSI) (32) was calculated for four types of nuclear:  $\text{C}\alpha$ ,  $\text{C}\beta$ ,  $\text{C}'$ , and  $\text{H}\alpha$ . The derived secondary structures based on the consensus CSI were converted into restraints on  $\phi$  and  $\psi$  angles; for  $\alpha$ -helix residues,  $\phi$  was limited in  $-60 \pm 40^\circ$  while  $\psi$  was in  $-50 \pm 50^\circ$ . For the identified slow-exchange amide protons located in regular secondary structures, hydrogen bond restraints were added.

Structures were calculated using the program CNS v1.0 (33), employing a simulated annealing protocol for torsion angle dynamics (34). Simple impulsion nonbonded interactions were used during structure calculation. Structural figures were produced with MOLMOL (39).

## RESULTS AND DISCUSSION

**Chemical Shift Assignment.** A set of six experiments was used to sequentially assign the backbone resonances. These experiments included HNCA and HN(CO)CA, CBCANH and CBCA(CO)NH, and HNCO and HN(CA)CO. According to  $\text{C}\alpha$ ,  $\text{C}\beta$ , and  $\text{C}'$  chemical shifts, unambiguous backbone assignment was obtained for all residues except the C-terminal His tag. The assignment was checked with H(CA)-NH, C(CO)NH and H(CCO)NH TOCSY were used to obtain side chain  $^{13}\text{C}$  and  $^1\text{H}$  chemical shifts, respectively. The better spectral dispersion of amide  $^{15}\text{N}$  chemical shifts helped to

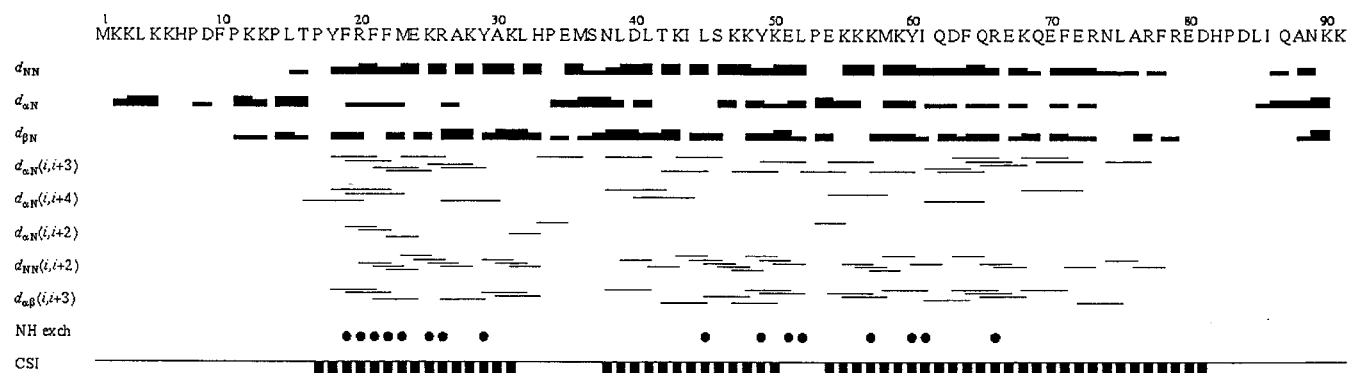


FIGURE 2: Summary of sequential and medium-range NOE connectivities observed in the HMG box 1 of hUBF. The data are derived from  $^{15}\text{N}$ -separated NOESY and  $^{13}\text{C}$ -separated NOESY; the mixing time is 130 ms for both. The three rows of solid bars below the sequence represent the observed sequential  $d_{\text{NN}}$ ,  $d_{\alpha\text{N}}$ , and  $d_{\beta\text{N}}$  NOE connectivities; the thickness of the bars indicates the intensities of the corresponding cross-peaks in NOESY. Lines below the sequential connectivities represent the  $d_{\alpha\text{N}}(i,i+3)$ ,  $d_{\alpha\text{N}}(i,i+4)$ ,  $d_{\alpha\text{N}}(i,i+2)$ ,  $d_{\text{NN}}(i,i+2)$ , and  $d_{\alpha\beta}(i,i+3)$  NOE connectivities. Filled circles represent the location of slowly exchanging amide protons. The short bars at the bottom represent consensus CSI predictions from  $\text{C}\alpha$ ,  $\text{C}\beta$ ,  $\text{C}'$ , and  $\text{H}\alpha$  chemical shifts; bars below the line mean  $-1$  index, while those above the line mean  $+1$  index.

resolve most of the aliphatic side chain resonances, which could be of problem since there are 20 lysine residues in the protein. HCCH-TOCSY and HCCH-COSY were then used to resolve the ambiguities inside side chains; some side chain resonances missed in  $\text{C}(\text{CO})\text{NH}$  or  $\text{H}(\text{CCO})\text{NH}$  were also found and assigned. The water peak did not interfere with these two spectra seriously since most  $\text{H}\alpha$  chemical shifts were below 4.50 ppm. Some aromatic ring proton chemical shifts were assigned by correlating  $\text{H}\delta$  to  $\text{C}\beta$  in  $\text{CB}(\text{CC})\text{HD}$ -COSY, as well as 2D TOCSY in  $\text{D}_2\text{O}$ ; the others were assigned through NOEs between ring protons and the intraresidue  $\text{H}\beta$  and  $\text{HN}$ . In total, more than 90% of the side chain resonances were assigned, which were listed in Table S1 in Supporting Information.

The HMG box protein NHP6A was reported to be unfolded under 310 K (11). We recorded  $^1\text{H}$ - $^{15}\text{N}$  HSQC at different temperatures from 290 to 310 K, which were all similar, suggesting that the global structure of hUBF HMG box 1 was sustained at 310 K.

**Secondary Structures.** Analysis of the chemical shifts predicted the existence of three helical regions located at 17–31, 38–50, and 54–81, which coincided with the three amino acid segments separated by Pro17, Pro34, Pro53, and Pro83. From the characteristic sequential and medium-range NOE links (35) (Figure 2), a similar prediction of secondary structures was made, although some subtle difference existed. In particular, the second  $\alpha$ -helix seemed to extend until L52, while toward the end of the third  $\alpha$ -helix NOE peaks became very weak, indicating a less rigid region. The prediction is consistent with secondary structure studies by circular dichroism and FTIR, which indicated a helix content of 54.3% and 59.9%, respectively (16).

The exchange rates of amide protons with solvents were measured. Amide protons still present after 16 h of exchange were considered slow-exchange amide protons, which may be engaged in hydrogen bonds. We identified 16 slow-exchange amide protons, which were all in the predicted helical regions (Figure 2).

**Structure Calculations.** The final set of constraints contained 152 intraresidue, 273 sequential, 379 medium-range, and 228 long-range distance constraints (Figure 3), together with 100 backbone dihedral angle constraints and 11

hydrogen bond constraints. Since the C-terminal His tag was not assigned and thus had no NMR restraint, it was excluded from structure calculation. From 60 calculated structures, the 20 structures with minimal NOE violations and lowest total energy were selected. All of these 20 conformations were in good agreement with the experimental data, with no distance violations larger than 0.3 Å and no angle violations of more than 5°. The covalent geometry was well respected (Table 1). A Ramachandran plot was produced by PROCHECK-NMR (36), which exhibited 73.2% of the residues in the most favored regions and 20.5% and 4.9%, respectively, in additional and generously allowed regions. Some terminal residues had disallowed backbone conformation (1.3%), which should be due to lack of restraints.

The structure was well-defined for the first two helices and the N-terminal part of helix 3, with an atomic root mean square deviation about the average structure of 0.55 Å for the backbone atoms and 1.53 Å for all heavy atoms, for residues 13–68. Figure 4A shows the backbone superimposition of the final 20 structures. Toward the end of helix 3, the structure was less rigid, as evident from Figure 3, which shows the average RMS deviation to mean structure plotted against sequence. In addition, superimposition of residues 61–78 gave an RMSD of 1.03 Å for backbone atoms and 2.32 Å for heavy atoms, indicating a locally ordered structure with some variability.

**Structure Description.** Figure 4B shows the ribbon representation for the lowest energy structure of hUBF box 1. As expected from the secondary structure analysis, the structure contains three  $\alpha$ -helices, which are located at 17–30, 38–51, and 57–76. They formed an L-shape fold together with an extended stretch of amino acids preceding helix 1 (residues 11–16). The global structure is similar to other HMG box structures available.

The angle between helix 2 and helix 3 is about 80°, while helix 1 intersects with helix 2 and helix 3 with an angle of 40° and 60°, respectively. The angle between helix 1 and 2 is somewhat smaller than in other structures (13). Like other HMG structures, helix 1 is bent by about 40° in the middle, in this case at Lys25. A type I turn (8, 35) centered at Pro34-Glu35 is observed following helix 1.



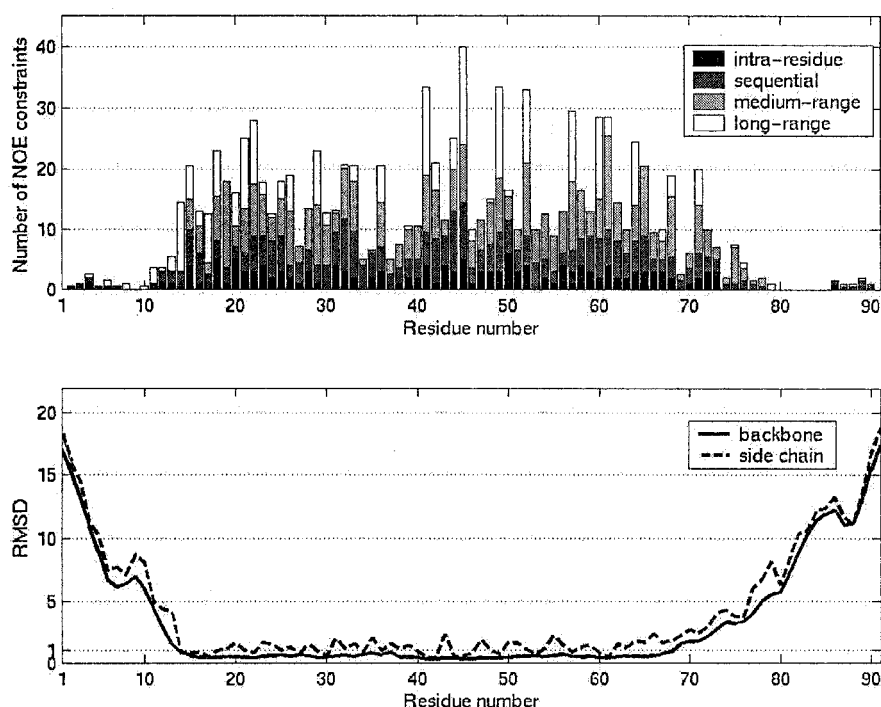


FIGURE 3: Plot of the number of NOE constraints per residue used in the calculation of the hUBF box 1 structure (upper) and the average RMS deviation from the mean structure (lower) versus the sequence of hUBF box 1.

Table 1: NMR Structure Statistics for HMG Box 1 of hUBF

(a) NMR Constraints	
distance constraints	1032
intraresidue	152
sequential ( $ i - j  = 1$ )	273
medium range ( $ i - j  < 5$ )	379
long range ( $ i - j  = 5$ )	228
dihedral angle restraints	100
hydrogen bonds	11
(b) Statistics for 20 SA Structures	
NOE violations	
number $> 0.1$ Å	$4.8 \pm 1.4$
maximum violations (Å)	$0.16 \pm 0.03$
RMSD from idealized covalent geometry	
bonds (Å)	$0.0019 \pm 0.0002$
angles (deg)	$0.39 \pm 0.01$
impropers (deg)	$0.21 \pm 0.01$
Lennard-Jones potential energy (kcal mol <sup>-1</sup> )	$-206 \pm 20$
(c) Coordinate Precision	
pairwise RMSD for backbone atoms (13–68) (Å)	$0.80 \pm 0.13$
pairwise RMSD for heavy atoms (13–68) (Å)	$1.73 \pm 0.18$
RMSD to the mean	bb/heavy
11–75 (Å)	0.98/1.88
17–31, 38–51, 57–75 (Å)	0.78/1.67
13–68 (Å)	0.55/1.53
61–78 (Å)	1.03/2.32
(d) Comparison between the Lowest Energy of Box 1 with Other HMG Box Proteins	
box 1 (14–65) vs HMG1B (10–61) (Å)	1.99 (4)
box 1 (14–52, 55–65) vs HMG-D (8–57) (Å)	2.24 (6)
box 1 (14–65) vs HMG1A (11–35, 38–64) (Å)	2.01 (7)
box 1 (14–65) vs LEF-1 (6–57) (Å)	2.27 (10)
box 1 (14–65) vs NHP6A (24–75) (Å)	1.85 (11)
box 1 (14–65) vs mSox-5 (6–57) (Å)	2.46 (13)
box 1 (14–65) vs hSRY (8–59) (Å)	1.49 (37)
box 1 (14–65) vs hSRY <sup>M91</sup> (8–59) (Å)	1.47 (37)

The structure is stabilized by extensive hydrophobic interactions centered at the turn. The conserved aromatic residues F21, Y49, and Y60 pack onto each other, forming the core at the turn. This core is extended into both arms

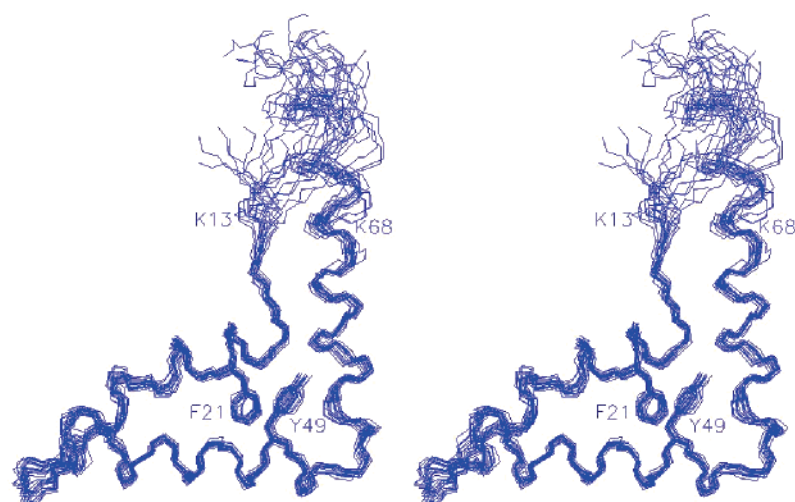
and the angle apex, as illustrated in Figure 4C. Namely, the packing of helix 1 onto helix 2 involves Y18, F22, K25, Y29, M36, L41, I44, and L45 in addition; the turn between helix 2 and helix 3 is stabilized by L52 and K57; helix 3 is aligned with the strand N-terminal to helix 1 by packing of F64, E67, K68, and F71 to P14 and F71, L75, and A76 to P11.

Both the N-terminal and the C-terminal parts are disordered in our structure. They may be flexible as suggested by their narrow peaks, averaged chemical shifts, and the very few NOEs. Anyway, the orientation of the tails in Figure 4B is somewhat reminiscent of the basic tails in LEF-1 (10), NHP6A (11), and SRY (37), which bend toward the concave face of the molecule and interact with DNA in the major groove. In fact, the N-terminus of hUBF HMG box 1 is basic rich (K2, K3, K5, and K6), although its role in DNA binding has not been studied.

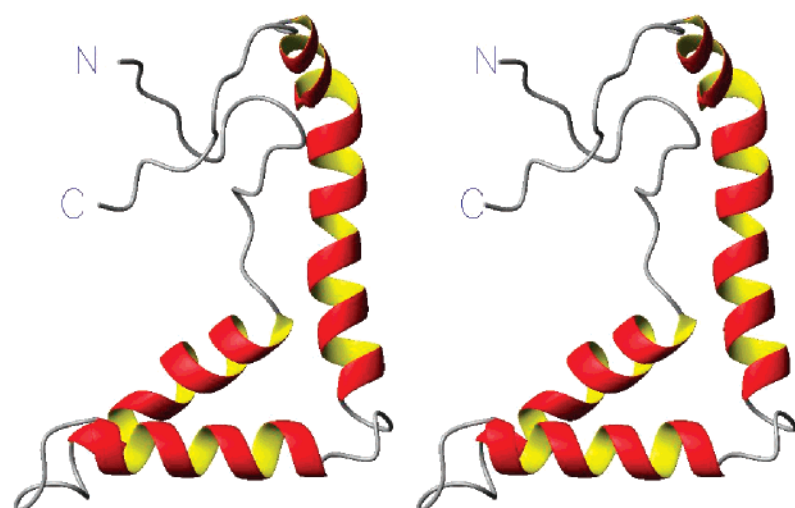
**Comparison with Other HMG Box Proteins.** The backbone RMS deviations of the lowest energy hUBF box1 structure to other available HMG box structures are listed in Table 1d. The most similar structure is a recently reported structure of hSRY HMG box complexed with a 14 bp DNA fragment, with an RMSD of 1.49 Å (1.47 Å for the mutant hSRY complex); to the other structures the RMSD is around 2 Å. For HMG1A, the gap was chosen at S37–N38 in box 1 to minimize the RMSD; this is a little striking, however, since in most structures N38 had been involved in the second helix. It seems that our structure is equally similar to HMG1B and to HMG1A.

One of the most significant differences of hUBF HMG box 1 to other HMG box proteins is that it has a tyrosine rather than a tryptophan at position 49. Its smaller ring seems to possess more freedom than its counterpart in other structures, as evidenced by the different orientation of the ring in backbone superimposition (Figure 4A).

A.



B.



C.

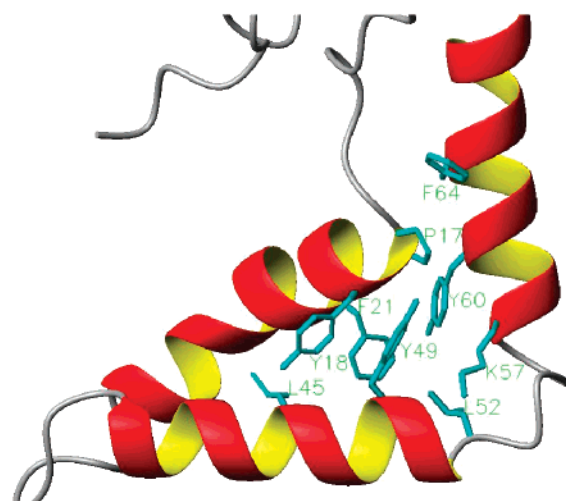


FIGURE 4: (A) Stereoview of the backbone superimposition of the 20 structures for hUBF HMG box 1; the N-terminal (1–10) and C-terminal (81–91) disordered tails are removed for clarity. Side chains of F21 and Y49 are also shown. (B) Stereoview of the ribbon representation of the lowest energy structure. (C) Packing of side chains in the hydrophobic core located at the turn of “L”, which is important for maintaining the fold.

The long arm in our structure is less rigid than the short arm. This was also noticed in earlier studies of HMG1B (4), Sox4 (8), and Sox5 (13). The reason must be the much less hydrophobic interaction in the long arm.

*Interaction with Nucleic Acid.* Previous study has demonstrated that the recombinant hUBF HMG box 1 can bind to the four-way DNA junction and a synthetic 15 bp linear DNA, with the dissociation constants being  $3.0 \times 10^{-8}$  and

$2.6 \times 10^{-7}$  M, respectively (16). hUBF box 1 bears several conserved residues that had been revealed to be important in interacting with the DNA minor groove in other HMG box proteins (9–12, 37). However, whether these residues play similar roles in this case has to wait for a structure of box 1 complexed with a DNA to be decided. With the structure of free hUBF HMG box 1, we are now at a good start to do the work.

## ACKNOWLEDGMENT

We thank Dr. F. Delaglio and Prof. A. Bax for providing the software NMRPipe, Dr. Dan Garrett and the Laboratory of Chemical Physics at the NIH for making available the program PIPP, Prof. A. T. Brünger for providing the program CNS, and Dr. R. Koradi and Prof. K. Wüthrich for providing MOLMOL.

## SUPPORTING INFORMATION AVAILABLE

Table S1 listing  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  chemical shifts of hUBF HMG box 1. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Jantzen, H.-M., Admon, A., Bell, S. P., and Tjian, R. (1990) *Nature* 344, 830–836.
- Paule, M. R. (1998) *Transcription of Ribosomal RNA Genes by Eukaryotic RNA Polymerase I*, Springer-Verlag, New York.
- Jantzen, H.-M., Chow, A. M., King, D. S., and Tjian, R. (1992) *Genes Dev.* 6, 1950–1963.
- Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D., and Thomas, J. O. (1993) *EMBO J.* 12, 1311–1319.
- Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., and Norman, D. G. (1993) *Nucleic Acids Res.* 21, 3427–3436.
- Jones, D. N., Searles, M. A., Shaw, G. L., Churchill, M. E., Ner, S. S., Keeler, J., Travers, A. A., and Neuhaus, D. (1994) *Structure* 2, 609–627.
- Hardman, C. H., Broadhurst, R. W., Raine, A. R. C., Grasser, K. D., Thomas, J. O., and Laue, E. D. (1995) *Biochemistry* 34, 16596–16607.
- van Houte, L. P. A., Chuprina, V. P., van der Wetering, M., Boelens, R., Kaptein, R., and Clevers, H. (1995) *J. Biol. Chem.* 270, 30516–30524.
- Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell* 81, 705–714.
- Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature* 376, 791–795.
- Allain, F. H.-T., Yen, Y.-M., Masse, J. E., Schultze, P., Dieckmann, T., Johnson, R. C., and Feigon, J. (1999) *EMBO J.* 18, 2563–2579.
- Murphy, F. V., IV, Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J.* 18, 6610–6618.
- Cary, P. D., Read, C. M., Davis, B., Driscoll, P. C., and Crane-Robinson, C. (2001) *Protein Sci.* 10, 83–98.
- Ner, S. (1992) *Curr. Biol.* 2, 208–210.
- Baxeavanis, A. D., and Landsman, D. (1995) *Nucleic Acids Res.* 23, 1604–1613.
- Yang, W. L., Zhou, D., Zeng, W. Y., and Shi, Y. Y. (2002) *Biochim. Biophys. Acta* (submitted for publication).
- Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., and Tschudin, R. (1990) *J. Magn. Reson.* 86, 304–318.
- Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson., Ser. B* 103, 203–216.
- Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson.* 96, 432–440.
- Clubb, R. T., Thanabal, V., and Wagner, G. (1992) *J. Magn. Reson.* 97, 213–217.
- Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* 114, 6291–6293.
- Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson.* 99, 201–207.
- Kay, L. E., Ikura, M., and Bax, A. (1991) *J. Magn. Reson.* 91, 84–92.
- Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) *J. Biomol. NMR* 3, 225–231.
- Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) *J. Magn. Reson.* 88, 425–431.
- Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., and Kay, L. E. (1990) *J. Magn. Reson.* 87, 620–627.
- Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) *J. Am. Chem. Soc.* 115, 11054–11055.
- Zuiderweg, E. R. P., and Fesik, S. W. (1989) *Biochemistry* 28, 2387–2391.
- Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* 6, 277–293.
- Barkhuijsen, H., de Beer, R., and van Ormondt, D. (1987) *J. Magn. Reson.* 73, 553–557.
- Nilges, M. (1993) *Proteins* 17, 297–309.
- Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* 4, 171–180.
- Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905–921.
- Stein, E. G., Rice, L. M., and Brünger, A. T. (1997) *J. Magn. Reson.* 124, 154–164.
- Wüthrich, K. (1986) in *NMR of proteins and nucleic acids*, John Wiley & Sons, New York.
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* 8, 477–486.
- Murphy, E. C., Zhurkin, V. B., Louis, J. M., Comilescu, G., and Clore, G. M. (2001) *J. Mol. Biol.* 312, 481–499.
- Thomas, J. O., and Travers, A. A. (2001) *Trends Biochem. Sci.* 26, 167–174.
- Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics* 14, 51–55.

BI015977A