J. Mol. Biol. (1995) **254**, 692–703

JMB

Solution Structure of the HU Protein from *Bacillus stearothermophilus*

Hans Vis[1], Matteo Mariani[1,2], Constantin E. Vorgias[3], Keith S. Wilson[3] Robert Kaptein[1] and Rolf Boelens[1]*

[1]*Bijvoet Center for Biomolecular Research Utrecht University Padualaan 8, 3584 CH Utrecht, The Netherlands*

[2]*Istituto di Biofisica Facolta' di Medicina Universita' di Genova Via Giotto 2, Genova Sestri Ponente, Italy*

[3]*European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestrasse 85 D-22603, Hamburg, Germany*

**Corresponding author*

The histone-like protein HU from *Bacillus stearothermophilus* is a dimer with a molecular mass of 19.5 kDa that is capable of bending DNA. An X-ray structure has been determined, but no structure could be established for a large part of the supposed DNA-binding β-arms. Using distance and dihedral constraints derived from triple-resonance NMR data of a $^{13}C/^{15}N$ doubly-labelled HU protein 49 distance geometry structures were calculated, which were refined by means of restrained Molecular Dynamics. From this set a total of 25 refined structures were selected having low constraint energy and few constraint violations. The ensemble of 25 structures display a root-mean-square co-ordinate deviation of 0.36 Å with respect to the average structure, calculated over the backbone heavy atoms of residues 2 to 54 and 75 to 90 (and residues 2' to 54' and 75' to 90' of the second monomer). The structure of the core is very similar to that observed in the X-ray structure, with a pairwise r.m.s.d. of 1.06 Å. The structure of the β-hairpin arm contains a double flip-over at the prolines in the two strands of the β-arm. Strong ^{15}N-NH heteronuclear nuclear Overhauser effects indicate that the β-arm and especially the tip is flexible. This explains the disorder observed in the solution and X-ray structures of the β-arm, in respect of the core of the protein. Overlayed onto itself the β-arm is better defined, with an r.m.s.d. of 1.0 Å calculated over the backbone heavy atoms of residues 54 to 59 and 69 to 74. The tip of the arm adopts a well-defined 4:6 β-hairpin conformation similar to the iron co-ordinating β-arms of rubredoxin.

Keywords: DNA-binding protein; structure refinement; restrained molecular dynamics; triple-resonance NMR; β-turn

Introduction

The bacterial protein HU is an abundant DNA-binding protein that is thought to play an important role in the structure of the bacterial nucleoid and binds to DNA in a sequence-independent manner (Drlica & Rouviere-Yaniv, 1987; Pettijohn, 1988; Schmid, 1990). HU is also required in the assembly of higher order nucleoprotein structures. These complexes mediate a variety of DNA metabolic events (Echols, 1986, 1990; Surette & Chaconas, 1992; Lavoie & Chaconas, 1993), such as replication, transcription, site-specific recombination and transposition, during which the structure of the DNA helix is altered by, for example, melting, wrapping, bending or looping (Hochschild, 1990; Nash, 1990; Harrington, 1992; Matthews, 1992). It was shown that HU promotes DNA flexibility (Hodges-Garcia *et al*., 1989; Paull *et al*., 1993) and that it has, like other members of its family such as the integration host factor (IHF; Goodman *et al*., 1990) and TF1 from bacteriophage SPO1 (Geiduschek *et al*., 1990), the ability to bend DNA.

HU is a homotypic dimer in solution. In some organisms, for instance in *Escherichia coli*, HU consists of two different monomers, but in many organisms HU has identical monomers. The HU protein from *Bacillus stearothermophilus* (HUBst) consists of two identical monomers of 90 amino acid residues with a total molecular mass of 19.5 kDa. A three-dimensional structure has been determined

Abbreviations used: NOE, nuclear Overhauser effect; 3D, three-dimensional; HUBst, HU protein from *Bacillus stearothermophilus*; CT-HSQC, constant-time heternuclear single quantum coherence; NOESY, NOE spectroscopy; TOCSY, total correlated spectroscopy.

0022–2836/95/490692–12 $12.00/0

by X-ray crystallography (Tanaka *et al*., 1984), which has been refined to a resolution of 2.1 Å (White *et al*., 1989). A pair of extended β-hairpin arms protrude from the hydrophobic core of the dimer into the solvent. These β-arms were proposed as the DNA-binding region of HUBst, which would bind *via* the minor groove of DNA and invoke dramatic bending of the DNA helix (Tanaka *et al*., 1984). However, since eight residues in the tip of the β-hairpin were not visible in the X-ray structure (White *et al*., 1989), no structural evidence was found to confirm this binding model at that time. On the other hand, a considerable amount of biochemical evidence supports the view that these arms are involved in DNA binding (Gualerzi *et al*., 1986; Drlica & Rouviere-Yaniv, 1987; Lammi *et al*., 1984; Härd *et al*., 1989).

An NMR study of the HUBst dimer has been started with the aim of resolving the structure and dynamics of the DNA-binding region both free in solution and bound to DNA. The ^{1}H, ^{13}C and ^{15}N nuclear magnetic resonance assignments and the analysis of the secondary structure of HUBst in solution have been described recently (Vis *et al*., 1994). It was found that the region of residues 55 to 74 is basically an anti-parallel β-hairpin arm, but that there are irregularities at Pro63 and at Pro72, which had previously been unnoticed (White *et al*., 1989). Similarly, NMR studies have shown that the secondary structure of the core of TF1 from bacteriophage SPO1 shows high similarity to that of HUBst (Jia *et al*., 1994). In the present paper the determination of the three-dimensional solution structure of HUBst from NMR data is presented as derived from distance geometry and restrained molecular dynamics calculations, and the structure of the DNA-binding arms is analysed in more detail. Parts of the DNA-binding arms, including the tips of the arms, are more structured than other regions and these parts are characterized by relatively stable hydrogen bonds. It will be pointed out that the structure of the β-hairpin turn belongs to the 4:6 class (Sibanda *et al*., 1989), similar to conformations found in rubredoxin and the human immunoglobulin fragment Fab New.

Results and Discussion

Distance constraints

Most of the NOEs were assigned from the 3D time-shared NOESY-(^{13}C,^{15}N)-HSQC spectrum and from the 3D aromatic ^{13}C-edited NOESY spectrum. Several NOEs, especially those involving the aromatic phenylalanine protons, were assigned from 2D NOE spectra recorded at 750 MHz. The first series of distance geometry calculations on HU were performed using short- and medium-range distance constraints for the secondary structure elements and using distance constraints derived from other unambiguously assigned NOE cross-peaks. The structure determination of a symmetric dimeric protein presents a problem for NMR, since intra- and intermonomer NOEs cannot easily be discerned. We used a protocol for assigning such intra- and intermonomer NOEs, initially using the X-ray structure as a starting model, as developed on the dimeric Arc repressor (Breg *et al*., 1990; Bonvin *et al*., 1994). A few (eight) constraints were assigned as being intermonomer, since they would correspond to intramonomer proton–proton distances of more than 16 Å, while the intermonomer distances would be less than 4.5 Å on the basis of the X-ray structure. These intermonomer constraints appeared to be sufficient to calculate initial structures with a comparable folding topology to that of the X-ray structure. Further intermonomer NOEs were all based upon iteratively improved NMR structures. Structures with low energies, a low number of distance violations and good stereochemical properties were used in further exploratory runs to find as many unambiguous distance constraints as possible from multiple assigned NOEs. Inconsistencies in the data were checked simultaneously and unreasonable assignments were eliminated.

The final structures were calculated on the basis of 2700 experimental NMR constraints for the dimer. These comprised 1245 approximate interproton and 25 hydrogen-bond distance constraints, 41 approximate ϕ and 39 approximate χ_1 torsional angle constraints per monomer. Interproton distance constraints per monomer could be subdivided into 173 intraresidue, 420 sequential, 316 interresidue medium-range and 336 interresidue long-range ($|i - j| > 4$) constraints. In total, 166 long-range interproton distance constraints between the two monomers were collected for the dimer.

Analysis of the final structures

From 49 calculated structures, 17 structures were rejected because of the presence of one or two D-amino acids. From the remaining 32 structures, a total of 25 were selected, having a low distance and dihedral constraint energy, and a low number of constraint violations. The set of 25 structures fit the experimental NMR data quite well, since on average no distance constraint violations are larger than 0.55 Å. NMR observable interproton distances (<6.0 Å) were calculated from the final structures and compared with the upper-bound constraint values of the observed NOEs. These calculated distances (''expected'' distances) were arranged into distance shells. For each shell the list of experimental NOE distances (''observed'' distances) were matched with the expected distances and grouped according to their upper-bound value (see Table 1). The majority of NOE upper-bound distances correspond with the calculated interproton distances in the same distance range. For longer distances, the percentage of observed NOEs decreases as the interproton distances in the structures increase. The non-observed interproton distances in the shells 2.0 to 2.8 Å and 2.8 to 3.5 Å involve mainly distances involving surface residues with long side-chains (Arg, Lys and Glu) that have

Table 1. Comparison of expected and observed numbers of interproton distances of HUBst

		Observed (Å)					
Shell (Å)	Expected[a]	2.0–2.8	2.8–3.5	3.5–4.0	4.0–5.0	5.0–5.5	Total %
2.0–2.8	617	173	50	76	59	0	358 (58)
2.8–3.5	962	151	93	145	115	5	509 (53)
3.5–4.0	1000	36	64	188	168	5	461 (46)
4.0–5.0	3219	22	49	214	497	46	828 (26)
5.0–5.5	2164	5	4	38	201	30	278 (13)
5.5–6.0	2599	2	3	19	71	49	144 (5)

The expected number of distances per distance shell was calculated and averaged over the final 25 structures. For each shell the list of observed experimental NOE distances was matched with the expected distances and categorized according to their applied upper-bound value.

[a] Methylene protons of the side-chains, except for all β-protons, were taken as a single entry in the calculations. Lys-H^{ζ} Arg-H^{η}, -OH and -COOH protons were not used. Only distances between protons separated by more than three covalent bonds were included. Distances between protons within one aromatic ring were discarded.

a low chemical shift dispersion. Some of the long expected interproton distances correspond with strong NOEs, most likely due to spin diffusion. None of the applied distance constraints derived from these NOEs was significantly violated after adding appropriate pseudo-atom corrections. The opposite case, in which short expected distances correspond with weak NOEs, may be due to fast relaxation of ^{13}C or ^{15}N magnetization in the heteronuclear 3D NOE experiments.

Most of the side-chain χ_1 dihedral angles satisfy the experimental values obtained from *J*-coupling data. The χ_1 dihedral angle constraints are violated for Leu6, Val52 and Asp87 in all structures, and for Glu70 in more than 5% of the structures, respectively. For Lys83 the χ_1 dihedral angle adopts a staggered local conformation. For Lys19, Glu34, Arg61 and Val76 both *trans* and *gauche*(–) conformers of χ_1 are found, suggesting that these side-chains are flexible. None of the ϕ dihedral angle constraints was violated.

The stereochemical quality of the ensemble of 25 structures was checked with the program PROCHECK (Morris *et al*., 1992; Laskowski *et al*., 1993). All 25 HU structures show reasonable to good stereochemical qualities. The structural statistics for HUBst are summarized in Table 2. Most of the structures have dihedral angles close to the ideal values, corresponding to classes 1 and 2, for ϕ,ψ angle distribution and χ_1 angle standard deviation, respectively. In seven cases the classification is 1,3, in two cases 2,2 and in one case 2,3. All 25 structures have the classification 2 for the hydrogen bond energy standard deviation. The

Table 2. Structural statistics for HUBst solution structure

Number of distance constraint violations $\geqslant 0.35$ Å	9
Maximum distance constraint violation (Å)	0.53
r.m.s. deviation from average structure (Å)[a]	
Backbone (N, C^{α}, CO)	0.36 ± 0.01
All heavy atoms	0.91 ± 0.01
Deviations from idealized covalent geometry[b]	
Bonds (Å)	0.0191 ± 0.0001
Angles (deg.)	1.375 ± 0.011
Average number of bad non-covalent contacts[b]	0
% of residues[b,c] with ϕ/ψ in:	
Most favoured regions	74.0
Additional allowed regions	18.8

	Values in HUBst[b]	Ideal values[d]
$\chi 1$ *g*(–)	71.3 ± 17.8	64.1 ± 15.7
$\chi 1$ *t*	191.9 ± 20.0	183.6 ± 16.8
$\chi 1$ *g*(+)	–61.0 ± 16.9	–66.7 ± 15.0
$\chi 2$	178.8 ± 11.5	177.4 ± 18.5
ϕ Helix	–68.7 ± 17.0	–65.3 ± 11.9
ψ Helix	–38.3 ± 18.7	–39.4 ± 11.3
ω	177.5 ± 7.3	180.0 ± 5.8
Chirality C^{α}	32.8 ± 3.9	33.9 ± 3.5

[a] Residues 2 to 53 and 75 to 90 (and 2' to 53' and 75' to 90' of the second monomer).

[b] Calculated from the final set of structures with PROCHECK; *g*(–), *gauche*(–); *t*, *trans*.

[c] Residues excluding glycine and proline.

[d] Ideal values from Morris *et al*. (1992; PROCHECK).

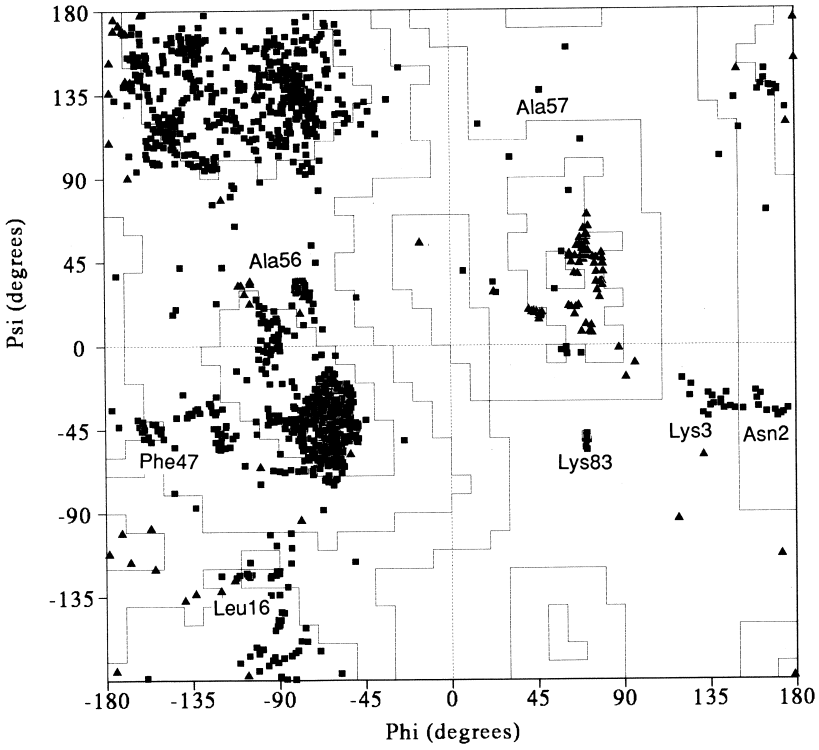

Figure 1. Ramachandran plot of the ϕ,ψ dihedral angles of the 25 structures of HU. Glycine residues are indicated as triangles. Non-glycine residues found in the unfavourable area are indicated in the Figure. The plot was generated using the programme PROCHECK (Laskowski *et al*., 1993).

Ramachandran plot of all 25 structures, presented in Figure 1, shows that most of the residues have favourable ϕ,ψ dihedral angles. The majority of ϕ,ψ combinations found in the unfavourable area correspond to glycine residues. Exceptions in some structures, as indicated in Figure 1, are the N-terminal residues Asn2 and Lys3, Ala56 and Ala57 in the β-hairpin arm opposite to Pro72, and Lys83. None of the ω dihedral angles shows deviations from ideal values larger than 8°. Figure 2 shows the variation in mutual backbone C^{α}, CO and N distances averaged over the set of 25

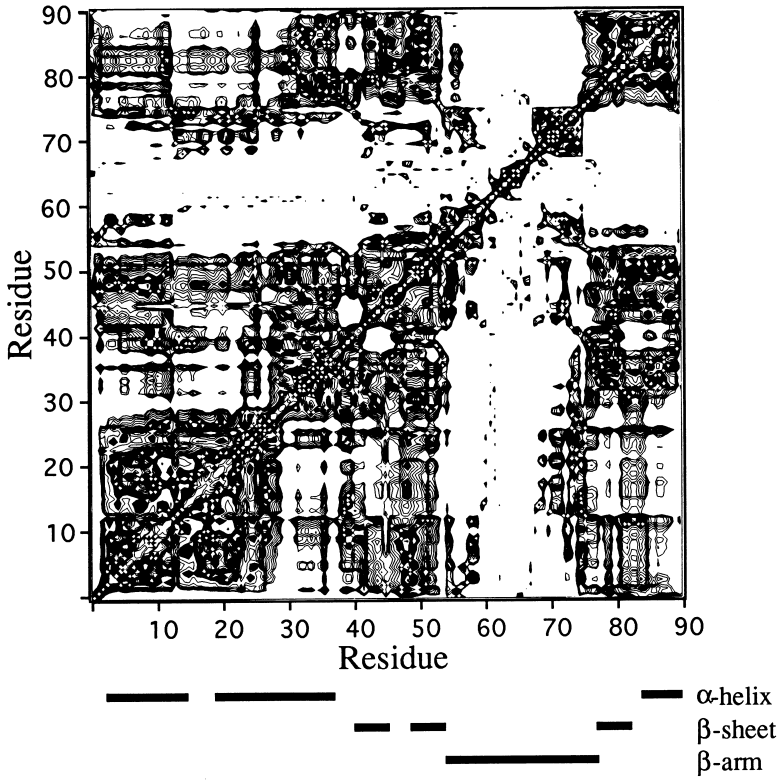

Figure 2. Backbone N, C^{α} and CO distance variation map of the first monomer of HU averaged over the best 25 structures. The dark areas in the plot indicate well-defined regions of the protein, showing small distance variations. The distance variation of the second monomer shows a similar result. Distances between the well-defined regions of the monomers also displayed small variations, indicating that the structure of one monomer relative to the other is well defined.

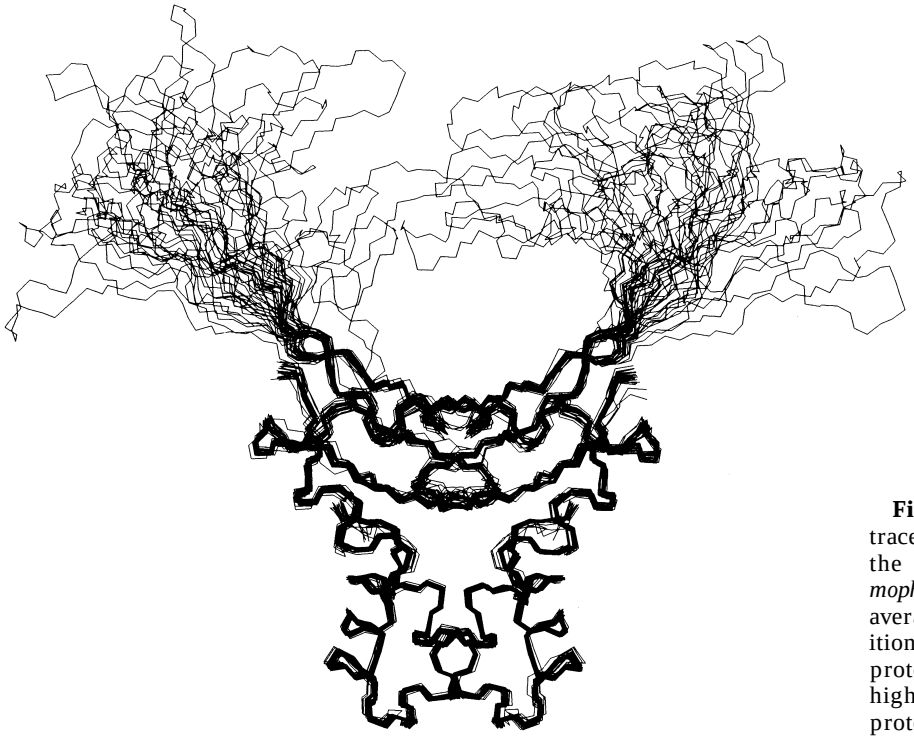

Figure 3. Polypeptide backbone trace of the 25 best structures of the HU protein from *B. stearothermophilus* after superposition on the average backbone heavy-atom positions of the well-defined part of the protein. Clearly, the β-arms show a high variability with respect to the protein core.

structures. The well-defined regions of HU, which show only little variation in mutual backbone atom distances, contain residues 2 to 53 and 75 to 90 (and 2' to 53' and 75' to 90' of the second monomer). Furthermore, the variation in backbone atom distances within the β-arm is low between atoms of residues 54 to 59 and of residues 69 to 74, indicating that part of its structure is well defined. Figure 3 shows the backbone trace of these 25 structures after superposition on the backbone heavy atoms of the well-defined part of the dimer, for which many long-range distance constraints were found. The final r.m.s.d. (from the average) is 0.36 Å on the backbone atoms and 0.91 Å on all heavy atoms of the well-defined residues. For residues 54 to 74 and 54' to 74' in the two β-hairpin arms (residues 53 to 76 and 53' to 76') no long-range constraints were determined to the residues of the well-defined part. Therefore, the β-arms show a higher conformational variability with respect to the well-defined part of the protein. As discussed below, this structural variability is associated with a high mobility for the arm region.

For the 25 structures the r.m.s.d. (from the average) after superposition of the backbone of the well-defined residues of the first and the second monomer are 0.33 Å and 0.34 Å, respectively. To analyse the symmetry of the dimer, co-ordinates between two monomers were interchanged within structures for which the pairwise r.m.s.d. with respect to the structure most like the average structure was lower after such an interchange of co-ordinates. For the new 25 structures the r.m.s.d. values (from the average) after superposition of the backbone of the well-defined residues of the first and the second monomer are 0.32 and 0.33, respectively. The small difference in these values from the values for the original co-ordinates indicates that the dimer is symmetric and, since the β-hairpin arms have uncorrelated structures, only averaged results of the two monomers will be presented.

The arrangement of the elements of secondary structure elements is very similar to that of the X-ray structure (White *et al*., 1989) and to the previously described secondary structure analysis (Vis *et al*., 1994). Three α-helices are found from residues 4 to 14, 18 to 38 and 84 to 89. A three-stranded anti-parallel β-sheet comprises residues 41 to 45, 48 to 52 and 77 to 81, while the second and the third strands are extended up to residues 62 and 67 and are arranged as a β-hairpin arm. A regular anti-parallel β-sheet is observed from residues 53 to 56 and 73 to 76. Then, at Ala57 and at Pro72 the strands flip inside out and continue as a regular anti-parallel β-sheet up to Asn62 and Glu67. The loop region of the β-arm, involving Asn62 to Glu67, will be discussed in more detail below. Figure 4 shows two ribbon views of HU in which the flip-over of the β-hairpin arms is illustrated. The structures show typical characteristics of solution structures with more disorder in the side-chains on the surface and a well-defined hydrophobic core formed by residues of both monomers. The interior of the dimer is formed by a number of aromatic rings and aliphatic side-chains, originating from Leu6, Val10, Leu16, Val25, Val28, Phe29, Ile32, Leu36, Val42, Leu44, Phe47, Phe50, Phe79, Leu85, Val89. The r.m.s.d. between the average NMR structure and the X-ray structure

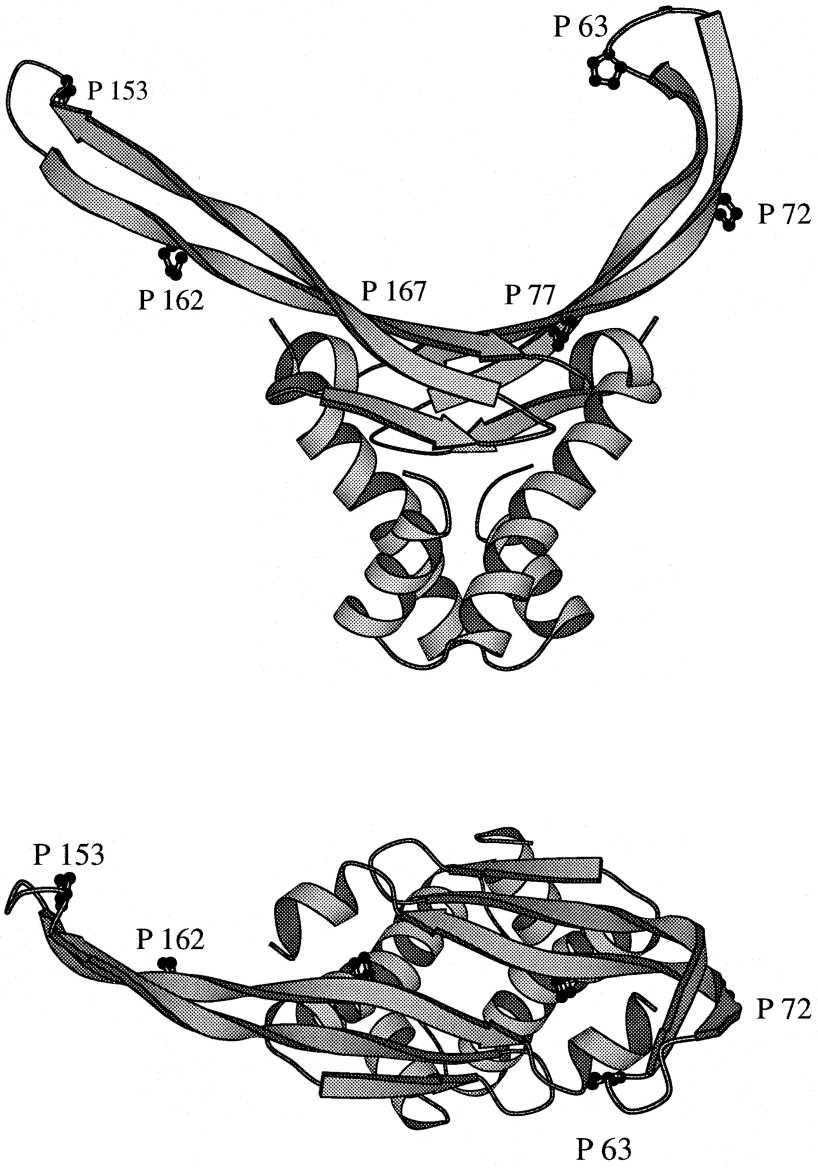

Figure 4. Ribbon views of HU in two different orthogonal orientations. (This pictures were created with the MOLSCRIPT programme (Kraulis, 1991).

(White *et al*., 1989), superimposed on the backbone heavy atoms of the well-defined residues of the dimer, is 1.06 Å.

Analysis of the β-hairpin arms

The supposed DNA-binding β-arms were searched for hydrogen bonds, using as criteria a maximum distance of 2.5 Å and minimum angles between donor and acceptor of 135° and of 125°. Results of this search are presented in Table 3. Since fast exchange has been observed for most amide protons in the β-arms (Vis *et al*., 1994), no hydrogen-bond distance constraints were applied in this part of the protein during structure refinement. Hydrogen bonds are indicated if found in 75, 50 and 25% of the structures, respectively. The hydrogen-bond pattern for the β-arms of HU is consistent with the long-range backbone-to-backbone NOEs observed in the β-arms between the second and the third β-strand, which had indicated a double flip-over of the two strands at Ala57 and Pro72 and at Pro63 and Gly66 (Vis *et al*., 1994). The slow and intermediate exchange of the amide protons of Gln64 and Gly66, respectively, is well accounted for by weak hydrogen bonds to CO and O^{δ} of Asn62, respectively. No hydrogen bond was found, however, involving the amide proton of Glu67, for which intermediate exchange was also observed. A type I β-turn with a common 4:4 conformation (Sibanda *et al*., 1989), as proposed previously by us (Vis *et al*., 1994) for β-hairpin residues Pro63 to Gly66, would contain appropriate β-sheet hydrogen bonds both between Asn62-NH and Glu67-CO, and between Glu67-NH and Asn62-CO. In the solution structure of HU, however, the latter hydrogen bond is absent, which indicates that the turn in the tip of the β-hairpin arm has a 4:6 conformation (see Sibanda *et al*., 1989).

The deviation from the mean of the C^{α} atom positions of residues in the β-arm, after superposition on all backbone atoms and on the backbone

Table 3. Hydrogen bonds observed in the DNA-binding β-arm of the final NMR structures of HUBst

Amide	Acceptor		135°	12.5°
Ala53	Val76	C=O	±	+
Arg55	Ser74	C=O	+ +	+ +
Arg58	Ile71	C=O		+
Gly60	Met69	C=O	+ +	+ +
Asn62	Glu67	C=O	+	+
Glu66	Asn62	C=O	±	+
Met69	Gly60	C=O	+	+
Ile71	Arg58	C=O	+	+ +
Val76	Ala53	C=O	+ +	+ +
Gln64	Asn62	O^{δ}	±	±

The hydrogen bonds were calculated from the final set of 25 structures. A maximum distance of 2.5 Å and a minimum angle as indicated were used as criteria to indicate a hydrogen bond. + +, + and ± indicate hydrogen bonds found in at least 75, 50 and 25% of the structures, respectively. For all these hydrogen bonds fast exchange has been observed (Vis *et al*., 1994).

atoms of the well-defined residues of the dimer, is shown in Figure 5(a), which emphasizes that the structure of the arm relative to the core of the dimer is poorly defined. The variability of residues in the β-arm is also reflected in the smaller ϕ and ψ dihedral angle order parameters, shown in Figure 5(b) and (c), respectively. Small ϕ and ψ angular order parameters are also displayed by residues 42 to 46 in the first β-strand. Distance variations in Figure 2 are mainly observed between C^{α}, CO and N atoms of the β-arm and between C^{α}, CO and N atoms of the β-arm and the core of the protein. However, the variation of backbone atom distances within the β-arm is low between atoms of residues 54 to 59 and of residues 69 to 74, indicating that the first part of the β-arm has a reasonably well-defined structure by itself. In addition, a low variation of backbone atom distances for residues 63 to 66 in the β-turn is observed. A view of the backbone of 50 β-arm structures of both monomers, superimposed on the backbone of residues 54 to 59 and residues 69 to 74 and displaying an r.m.s.d of 1.0 Å, is presented in Figure 6. The structure of the β-turn relative to the first part of the β-hairpin arm seems ill-defined. At Asn62 and Glu67 the variation of mutual C^{α}, CO and N atom distances with the neighbouring residues increases (Figure 2) and the ϕ and ψ dihedral angle order parameters are low (Figure 5), indicating that the largest distortions in the structure of the β-hairpin are located at these residues.

Figure 5(d) shows the backbone ^{15}N-NH heteronuclear NOE, observed in the ^{15}N-NH heteronuclear NOE difference experiment, as a function of the residue number. The intensity of ^{15}N-NH NOEs indicates high mobility of the NH vectors (Peng & Wagner, 1992). Strong ^{15}N-NH backbone NOEs are observed for many residues of the β-arm. Furthermore, the side-chain NH_2 groups of Asn43, Asn62 and Gln64 show high mobility, observed as strong heteronuclear NOEs. The heteronuclear NOEs of amides of residues 62 to 69, which comprise the tip of the β-hairpin arms, are even stronger than of

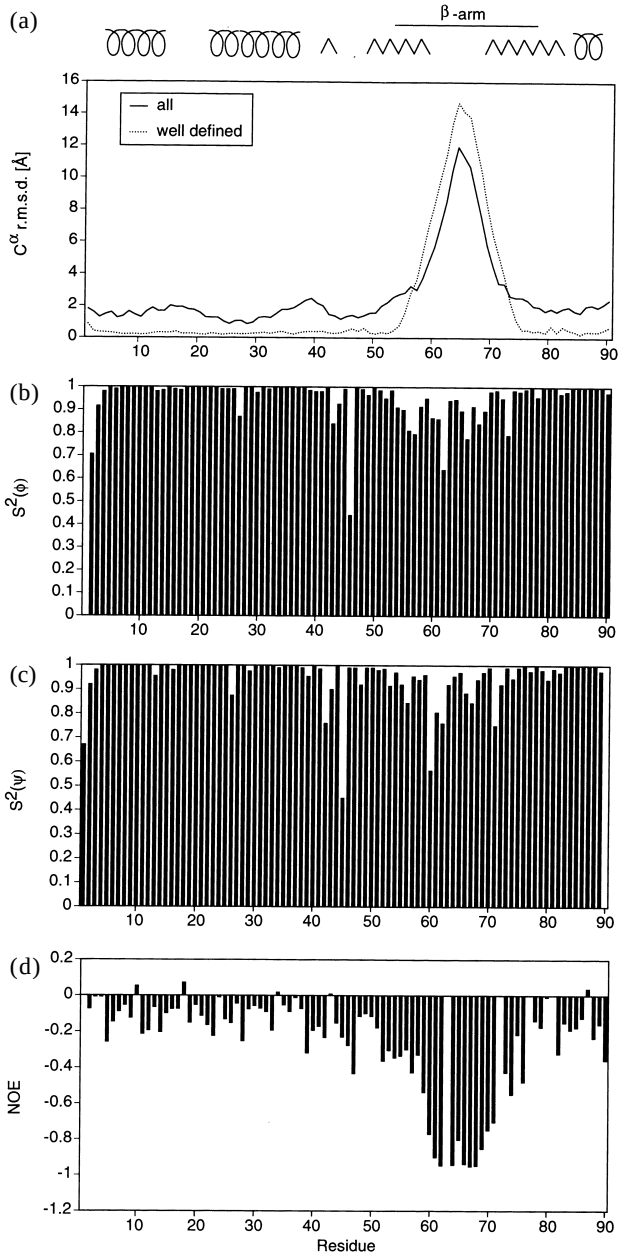

Figure 5. r.m.s.d (from the average) on C^{α} atoms (a); the structures were superimposed on C^{α}, CO and N atoms of all and of only well-defined residues of the dimer, backbone ϕ (b) and ψ (c) angular order parameters S^2 and ^{15}N-NH heteronuclear NOE (d) as function of the residue sequence of HU. The average values for the two monomers are shown.

residues in the rest of the arms, indicating the highest flexibility in this region. The results from the heteronuclear NOE experiment indicate that the β-arms of HU are highly flexible relative to the core of the dimer and the tip of the arms show even higher flexibility. This is consistent with the structure determination using NOE and *J*-coupling data, shown above, which showed significant disorder for these regions. Other residues showing high mobility are Gly39, Phe47 and Gly82, which

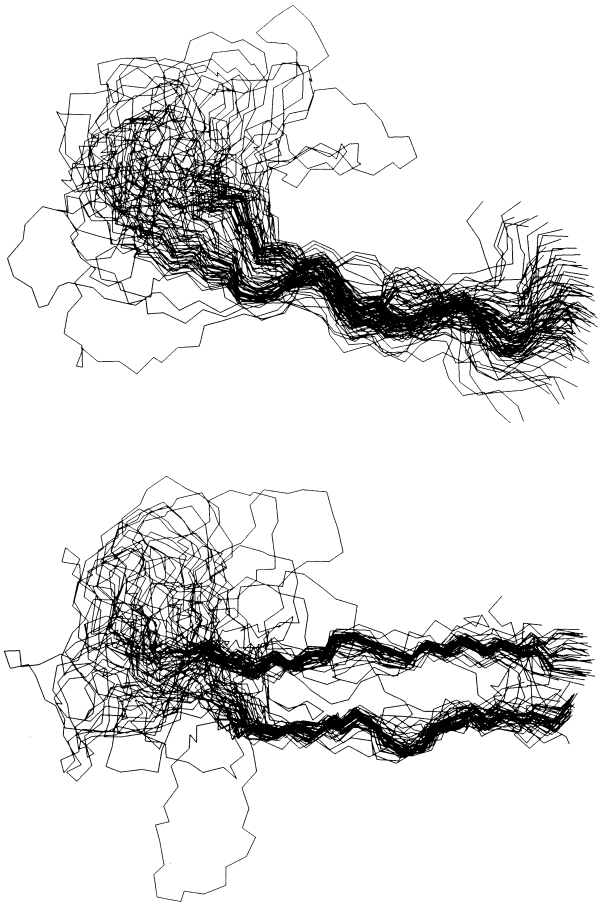

Figure 6. View of the structures of the β-hairpin arm of HU (residues to 53 to 76). The structures of both arms were superimposed on the backbone heavy-atoms of residues to 54 to 59 and 69 to 74, in the first part of the β-arm.

are present in loop regions connecting secondary structure elements, and Lys90, the C terminus.

The structure of the β-hairpin turn is presented in Figure 7(a), which shows a superposition on the N, C^{α} and CO atoms from residues 62 to 67 of 40 structures. The r.m.s.d. (from the average) is 0.35 Å. From an analysis of X-ray protein structures a classification of β-hairpin turn conformations has been developed (Sibanda *et al*., 1989). The set of structures used by Sibanda *et al*. (1989) revealed five β-hairpin turns with a 4:6 structure. We have compared the structure of the β-turn of HU with those 4:6 β-hairpin conformations by calculating the averaged pairwise r.m.s.d. of these structures relative to HU after superposition of the corresponding six residues of the β-turn on the N, C^{α} and CO atoms of residues 62 to 67 of HU (Asn-Pro-Gln-Thr-Gly-Glu). For two β-turns this r.m.s.d. is smaller than 1.0 Å, in particular, for residues 72 to 77 (Asn-Thr-Ser-Lys-Asn-Gln) in the heavy chain of human immunoglobulin fragment Fab New (Saul *et al*., 1978; Saul & Poljak, 1992) and for residues 6 to 11 (Cys-Thr-Val-Cys-Gly-Tyr) in the first iron co-ordinating arm of the iron-sulphur protein rubredoxin from *Desulfovibrio vulgaris* (Adman *et al*., 1977, 1991) the r.m.s.d. values are 0.84 Å and 0.50 Å, respectively. In addition, a high sequence homology between the first and the second iron co-ordinating arms of rubredoxin suggests also that this second β-hairpin, comprising the residues 39 to 44 (Cys-Pro-Val-Cys-Gly-Ala), adopts a 4:6 conformation. The averaged pairwise r.m.s.d. for this β-turn with respect to HU is 0.49 Å. The backbone conformations of the two iron co-ordinating β-turns of rubredoxin and of Fab New, overlayed on the β-turn of HU with the lowest pairwise r.m.s.d. from the average 4:6 β-turn conformation, are shown in Figure 7(b) and (c), respectively. The conformations of the two iron co-ordinating β-hairpins of rubredoxin are very similar to the conformations of the DNA-binding β-hairpins of HU, whereas the structure of the 4:6 β-hairpin of Fab New differs only slightly. Interestingly, the same residues at position 2 (Pro63) and position 5 (Gly66) in the β-hairpin arm of HU are also found in the second β-hairpin arm of rubredoxin, while the first β-hairpin arm of rubredoxin has an identical glycine only.

For HUBst most NOE cross-peaks were obtained from a 3D time-shared NOESY- (^{13}C,^{15}N)-HSQC spectrum and from a 3D NOESY-(^{13}C)-HSQC spectrum edited on aromatic carbon resonances. Furthermore, qualitative *J*-coupling data were translated into ϕ and χ_1 dihedral angle constraints. A high number of NOEs (on average about 14 per residue) were identified, which allowed us to obtain a high-precision structure of HUBst with reasonable to good stereochemical properties. A double flip-over of the two strands of the β-arms, indicated by long-range backbone-to-backbone NOEs, has been found using NOE data. The curved shape of the β-hairpin arms resulting from this backbone geometry may improve the capability of binding to DNA. The DNA-binding β-hairpin arms of HU showed high mobility relative to the core. However, within the β-hairpin arms some of the residues show significantly higher structural disorder than others, but the residues in the β-turn are mutually relatively well-defined. Structural analysis of the tip of the β-hairpin arms of HU has indicated that the β-turn adopts a 4:6 conformation, similar to the conformations found in rubredoxin and Fab New. The β-arms are highly mobile, as can be deduced from strong ^{15}N-NH heteronuclear NOEs. The β-hairpin turn unit appears to move as a rigid body. The segmental mobility in the arms allows for conformations that support a previously suggested model of DNA binding (Tanaka *et al*., 1984). In this model the DNA-binding arms can move as rigid arms, creating sufficient room for accepting DNA. The tips of the arms are highly flexible, and once the DNA has moved inwards, the arms close and the tip of the arms wrap around the DNA.

Materials and Methods

Sample preparation

Cloning of the HUBst gene, overproduction and purification of the protein and the DNA-binding assay

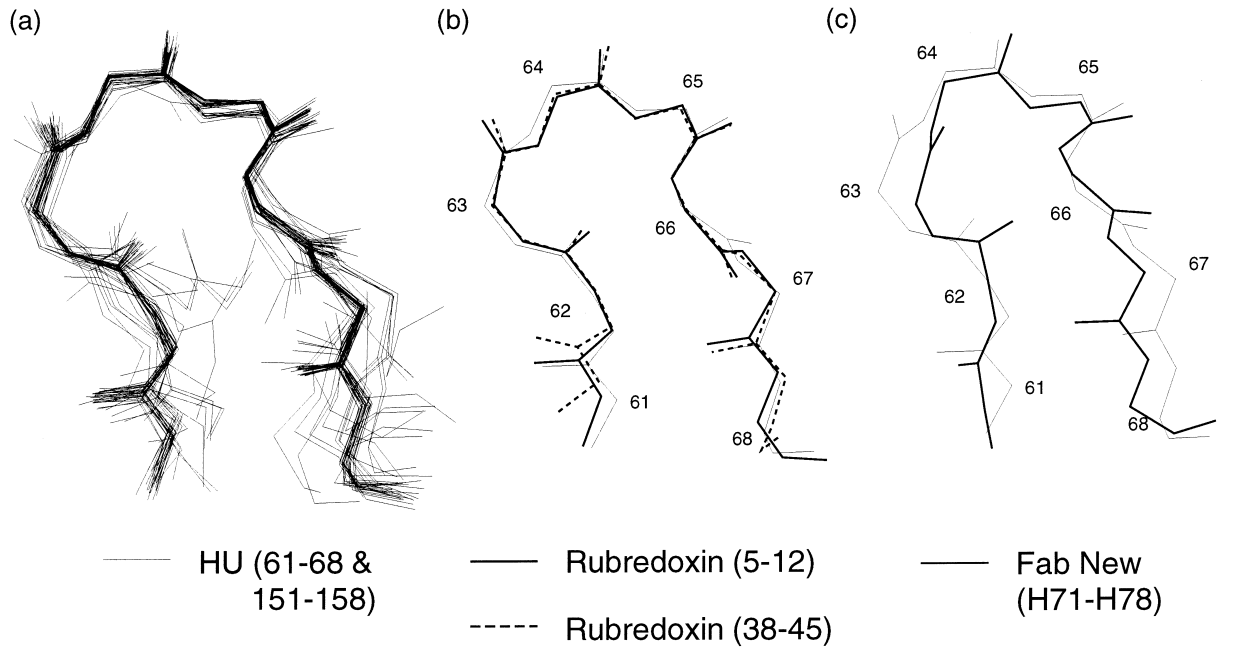

Figure 7. (a) Conformation of the β-hairpin turn of HU. (a) Forty structures of the HU β-hairpin turn (residues 61 to 68) were superimposed on the N, C^{α} and CO atoms of residues 62 to 67. (b) Comparison of the β-hairpin of HU with the first and second iron co-ordinating β-hairpin arms of the iron-sulphur protein rubredoxin from *D. vulgaris*. (c) Comparison with a 4:6 β-hairpin in the heavy chain of human immunoglobulin fragment Fab New. The conformations of rubredoxin and of Fab New were overlayed on residues 62 to 67 of the β-turn of HU. For the HU β-hairpin turn in (b) and (c) the conformation closest to the average β-turn conformation was used (thin line).

have been described by Padas *et al*. (1992). $^{13}C/^{15}N$-labelled protein was prepared using $^{15}NH_4Cl$ and [^{13}C]glucose in a minimal medium. The protein was purified employing affinity chromatography on heparin-Sepharose and Mono-S FPLC. The purity was over 95% as judged by SDS-PAGE and silver staining. The recombinant HU was as active as the protein purified from *B. stearothermophilus* with respect to its DNA-binding properties.

NMR spectroscopy

NMR experiments were performed with one 2.2 mM protein sample in 95% $^1H_2O/5\%$ ^2H_2O at 311 K, containing 50 mM KP_i buffer at pH 4.6 and 200 mM KCl. Heteronuclear NMR spectra were recorded on a Bruker AMX/2 600, which was equipped with three rf channels, a triple-resonance HCN probe with a shielded gradient coil, a 10 amp gradient amplifier (Bruker BGU) and a 300 W linear amplifier for the ^{13}C channel (Bruker BLTX300). A 2D NOE buildup series was recorded at 750 MHz on a Varian Unity Plus 750 MHz spectrometer using non-labelled protein dissolved in ^2H_2O at 311 K, containing 50 mM KP_i buffer at pH 4.6 and 200 mM KCl.

2D NOE (Jeener *et al*., 1982; States *et al*., 1982), 3D time-shared NOESY-(^{13}C,^{15}N)-HSQC (Pascal *et al*., 1994; Vis *et al*., 1994), 3D time-shared TOCSY-(^{13}C,^{15}N)-HSQC (Vis *et al*., 1994), 3D NOESY-(^{13}C)-HSQC for aromatic nuclei, $\{^{15}N\}$ spin-echo difference CT-HSQC (Vuister *et al*., 1993) and $\{^{13}CO\}$ spin-echo difference CT-HSQC (Grzesiek & Bax, 1993) spectra were recorded in a phase-sensitive mode using States-TPPI (Marion & Wüthrich, 1989). The 2D heteronuclear NOE difference experiment was recorded essentially as described by Dayie & Wagner (1994). Phase-sensitive detection in 3D HNHA (Vuister & Bax, 1993) and 3D HNHB (Archer *et al*., 1991) spectra was achieved according to the TPPI method (Marion & Wüthrich, 1983). A summary of all recorded NMR spectra with concomitant NMR parameters is presented in Table 4. Folding was used to increase the resolution of the ^{13}C direction of the 3D TOCSY-HSQC and 3D NOESY-HSQC experiments. The NOE mixing times were 150 ms. 2D NOE spectra were recorded in ^2H_2O at 750 MHz with mixing times of 75, 100 and 150 ms, respectively. All spectra were processed using our NMR programme package TRITON and analysed using the NMR analysis programme ALISON on a Silicon Graphics Indy workstation.

Stereospecific assignments and torsional angle restraints

Backbone HN-H^{α} vicinal coupling constants on HU were estimated from the ratio of cross-and diagonal peaks in a 3D HNHA spectrum (Vuister & Bax, 1993) and from a 3D time-shared TOCSY-(^{13}C,^{15}N)-HSQC spectrum (Vis *et al*., 1994; Constantine *et al*., 1994) recorded with an isotropic mixing time of 9.5 ms. The torsion angle ϕ was restrained to values of $-120° (\pm 60)°$ when its $^3J_{H^NH^{\alpha}}$ value was larger than 9 Hz and, for residues that were previously found in an α-helix, to values of $-60° (\pm 60)°$ when its $^3J_{H^NH^{\alpha}}$ value was significantly smaller than 9 Hz. Heteronuclear $^3J_{NH^{\beta}}$ coupling constants were estimated from a 3D HNHB spectrum and $^3J_{H^{\alpha}H^{\beta}}$, $^3J_{H^{\beta 1}H^{\gamma}}$ and $^3J_{H^{\beta 2}H^{\gamma}}$ coupling constants were estimated from the TOCSY-(^{13}C,^{15}N)-HSQC spectrum. The heteronuclear $^3J_{NC^{\gamma}}$ and $^3J_{COC^{\gamma}}$ coupling constants for threonine, valine and isoleucine residues were measured using $\{^{15}N\}$ and $\{^{13}CO\}$ spin-echo difference CT-HSQC spectra, respectively (Vuister *et al*., 1993; Grzesiek & Bax, 1993). All coupling constants were estimated as small or large, corresponding with a *gauche* or *trans* orientation of the

Table 4. Summary of NMR parameters used in the NMR experiments recorded for HUBst

Experiment	Number of points[a] $F_1 \times F_2 \times F_3$	Spectral widths (p.p.m.)[b] F_1	F_2	F_3	Recording time (h)
NOESY-(^{13}C,^{15}N)-HSQC	90(C) × 64(C) × 384(C)	11.11(H)	33.13(C) 32.88(N)	11.11	77
NOESY-(^{13}C)-HSQC (aromatics)	22(C) × 64(C) × 512(C)	11.11(H)	8.32(C)	11.11	55
TOCSY-(^{13}C,^{15}N)-HSQC	100(C) × 48(C) × 512(C)	11.11(H)	33.13(C) 32.88(N)	11.11	85
2D NOE	400(C) × 1024(C)	12.31(H)	12.31(H)		12
3D HNHA	64(R) × 160(R) × 512(C)	32.88(N)	11.11(H)	11.11	36
3D HNHB	64(R) × 160(R) × 512(C)	32.88(N)	11.11(H)	11.11	36
$\{^{15}N\}$ spin-echo difference CT-HSQC	256(C) × 1024(C)	33.13(C)	11.11		10
$\{^{13}CO\}$ spin-echo difference CT-HSQC	256(C) × 1024(C)	33.13(C)	11.11		10
^{15}N-^{1}H NOE	128(C) × 512(C)	32.88(N)	11.11		12

[a] Data were recorded in the indirectly detected dimensions, either Real (R) or Complex (C).
[b] The type of nucleus for the indirectly detected dimensions is indicated: (H) for proton, (C) for carbon and (N) for nitrogen.

two nuclei. Stereospecific assignments of β-protons of AMX spin systems and γ-methyl groups of valine residues were obtained from $^3J_{H^\alpha H^\beta}$, $^3J_{NH^\beta}$, $^3J_{NC^\gamma}$ and $^3J_{COC^\gamma}$ coupling constants and intraresidue NOEs involving H^β (H^γ for valines), H^α and HN protons. For structure calculations, the χ_1 angles were restrained to values of 60°, 180° or –60° (±60)°. Dihedral angle constraints were only imposed in the refinement protocol after distance geometry. Stereospecific assignments of δ-methyl groups of leucine residues were obtained from $^3J_{H^{\beta1}H^\gamma}$ and $^3J_{H^{\beta2}H^\gamma}$ coupling constants and intraresidue NOEs involving H^δ and the other protons. Stereospecific assignments were not based on any model structure.

Distance constraints

All NOEs were classified as either strong, strong-medium, medium, weak or very weak, corresponding to upper limit interproton distance constraints of 2.8 Å, 3.5 Å; 4.0 Å, 5.0 Å and 5.5 Å, respectively. Appropriate pseudo-atom corrections were added to all upper-bounds involving methylene, methyl or aromatic groups, according to the method of Wüthrich (1986), except for methyl groups, for which a pseudo-atom correction of 0.3 Å was used (Koning *et al*., 1990). All lower bounds were set to 1.8 Å. Preliminary structures were searched for hydrogen bonds using as criteria a maximum distance of 2.4 Å and a minimum angle between donor and acceptor of 135° (Bonvin *et al*., 1994). For those hydrogen bonds that were in agreement with the experimental exchange data (Vis *et al*., 1994) distance restraints between the corresponding carbonyl oxygen and the amide proton were added using upperbounds of 2.3 Å.

Structure calculations

Protein structures were generated with distance geometry (Havel, 1991) using the DGII programme (BIOSYM, San Diego). The distance geometry procedure consisted of three steps: triangulation, embedding in three dimensions and an optimization step for both distances and chiralities. The structures were refined using the programme Discover (BIOSYM, San Diego). The refinement protocol started with a restrained energy minimization step (50 iterations using steepest descents minimization followed by 100 iterations using conjugate gradients minimization), followed by restrained Molecular Dynamics simulations *in vacuo* for 20 picoseconds with a timestep of one femtosecond, using a Consistence Valence ForceField (CVFF). Initial velocities were randomly taken from a Maxwellian distribution at 800 K. The temperature was lowered 100 K every picosecond until a temperature of 311 K was established. Finally, another energy minimization step was performed (100 steepest descents iterations followed by 500 conjugate gradients iterations). All distance geometry and restrained Molecular Dynamics calculations were performed on a Silicon Graphics Challenge computer. Structures were examined on Silicon Graphics IRIS workstations using the programme InsightII (BIOSYM, San Diego). The order parameters for the ϕ and ψ dihedral angles were calculated as described by Hyberts *et al*. (1992). Atomic co-ordinates of the structures and experimental NMR constraints have been deposited with the Brookhaven Protein Data Bank: entries 1HUE and R1HUEMR, respectively.

Acknowledgements

This work was supported in part by the Netherlands Foundation for Chemical Research (SON) with financial support from the Netherlands Organization for Scientific Research (NWO).

References

Adman, E. T., Sieker, L. C. & Jensen, L. H. (1977). A structural model of rubredoxin from *Desulfovibrio vulgaris* at 2 Å resolution. *J. Mol. Biol.* **112**, 113–120.

Adman, E. T., Sieker, L. C. & Jensen, L. H. (1991). Structure of rubredoxin from *Desulfovibrio vulgaris* at 1.5 Å resolution. *J. Mol. Biol.* **217**, 337–352.

Archer, S. J., Ikura, M., Torchia, D. A. & Bax, A. (1991). An alternative 3D NMR technique for correlating backbone ^{15}N with side chain H^β resonances in larger proteins. *J. Magn. Reson.* **95**, 636–641.

Bonvin, A. M. J. J., Vis, H., Breg, J. N., Burgering, M. J. M., Boelens, R. & Kaptein, R. (1994). Nuclear magnetic resonance solution structure of the Arc repressor

using relaxation matrix calculations. *J. Mol. Biol.* **236**, 328–341.

Breg, J. N., van Opheusden, J. H. J., Burgering, M. J. M., Boelens, R. & Kaptein, R. (1990). Structure of Arc repressor in solution; evidence for a family of β-sheet DNA binding proteins. *Nature*, **346**, 586–589.

Constantine, K. L., Friedrichs, M. S. & Mueller, L. (1994). Simple approaches for estimating vicinal ^{1}H-^{1}H coupling constants and for obtaining stereospecific resonance assignments in leucine side chains. *J. Magn. Reson. B*, **104**, 62–68.

Dayie, K. T. & Wagner, G. (1994). Relaxation-rate measurements for ^{15}N-^{1}H groups with pulsed-field gradients and preservation of coherence pathways. *J. Magn. Reson. A*, **111**, 121–126.

Drlica, K. & Rouviere-Yaniv, J. (1987). Histonelike proteins of bacteria. *Microbiol. Rev.* **51**, 301–319.

Echols, H. (1986). Multiple DNA-protein interactions governing high-precision DNA interactions. *Science*, **233**, 1050–1056.

Echols, H. (1990). Nucleoprotein structures initiating DNA replication, transcription, and site-specific recombination. *J. Biol. Chem.* **265**, 14697–14700.

Geiduschek, E. P., Schneider, G. J. & Sayre, M. H. (1990). TF1, A bacteriophage-specific DNA-binding and DNA-bending protein. *J. Struct. Biol.* **104**, 84–90.

Goodman, S., Yang, C., Nash, H. Sarai, A. & Jernigan, R. (1990). Bending of DNA by IHF protein. In *Structure & Methods*, vol. 2, *DNA Protein Complexes & Proteins* (Sarma, R. H. & Sarma, M. H., eds), pp. 51–62, Adenine Press, New York.

Grzesiek, S. & Bax, A. (1993). Amino acid type determination in the sequential assignment procedure of uniformly ^{13}C/^{15}N-enriched proteins. *J. Biomol. NMR*, **3**, 185–204.

Gualerzi, C. O., Losso, M. A., Lammi, M., Friedrich, K., Pawlik, R. T., Canonaco, M. A., Gianfranceschi, G., Pingoud, A. & Pon, C. L. (1986). Proteins from the prokaryotic nucleoid. Structural and functional characterization of the Escherichia coli DNA-binding proteins NS (HU) and H-NS. In *Bacterial Chromatin* (Gualerzi, C. O. & Pon, C. L., eds), pp. 101–134, Springer-Verlag, Heidelberg.

Härd, T., Sayre, M. H., Geiduschek, E. P. & Kearns, D. R. (1989). A type II DNA-binding protein genetically engineered for fluorescence spectroscopy: The ''arm'' of Transcription Factor I binds in the DNA grooves. *Biochemistry*, **28**, 2813–2819.

Harrington, R. E. (1992). DNA curving and bending in protein-DNA recognition. *Mol. Microbiol.* **6**, 2549–2555.

Havel, T. F. (1991). An evaluation of computational strategies for use in the determination of protein structures from distance constraints obtained by nuclear magnetic resonance. *Prog. Biophys. Mol. Biol.* **56**, 43–78.

Hochschild, A. (1990). Protein-protein interaction and DNA loop formation. In *DNA Topology and its Biological Effects*, (Cozzarelli, N. R. & Wang, J. C., eds), pp. 107–138, Cold Spring Harbor Press, Cold Spring Harbor, New York.

Hodges-Garcia, Y., Hagerman, P. J. & Pettijohn, D. E. (1989). DNA ring closure mediated by protein HU. *J. Biol. Chem.* **264**, 14621–14623.

Hyberts, S. G., Goldberg, M. S., Havel, T. F. & 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.

Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. (1982). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553.

Jia, X., Reisman, J. M., Hsu, V., Geiduschek, E. P., Parello, J. & Kearns, D. R. (1994). Proton and nitrogen NMR sequence-specific assignments and secondary structure determination of the *Bacillus subtilis* SPO1-encoded transcription factor 1. *Biochemistry*, **33**, 8842–8852.

Koning, T. G. M., Boelens, R. & Kaptein, R. (1990). Calculation of the nuclear Overhauser effect and the determination of proton-proton distances in the presence of internal motions. *J. Magn. Reson.* **90**, 111–123.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 945–949.

Lammi, M., Paci, M. & Gualerzi, C. B. (1984). Proteins from the prokaryotic nucleoid. The interaction between protein NS and DNA involves the oligomeric form of the protein and at least one Arg residue. *FEBS Letters*, **170**, 99–104.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Lavoie, B. D. & Chaconas, G. (1993). Site-specific HU binding in the Mu transposome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. *Genes Dev.* **7**, 2510–2519.

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

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

Matthews, K. S. (1992). DNA looping. *Microbiol. Rev.* **56**, 123–136.

Morris, A. L., McArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Protein: Struct. Funct. Genet.* **12**, 345–364.

Nash, H. A. (1990). Bending and supercoiling of DNA at the attachment site of bactereophage lambda. *Trends Biochem. Sci.* **15**, 222–227.

Padas, P. M., Wilson, K. S. & Vorgias, C. E. (1992). The DNA-binding protein HU from mesophilic and thermophilic Bacilli: gene cloning, overproduction and purification. *Gene*, **17**, 39–44.

Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D. & Kay, L. E. (1994). Simultaneous acquisition of ^{15}N- and ^{13}C-edited NOE spectra of proteins dissolved in H_2O. *J. Magn. Reson. B*, **103**, 197–201.

Paull, T. T., Haykinson, M. J. & Johnson, R. C. (1993). The non-specific DNA-binding and bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes Dev.* **7**, 1521–1534.

Peng, J. W. & Wagner, G. (1992). Mapping of spectral density functions using heteronuclear NMR relaxation measurements. *J. Magn. Reson.* **98**, 308–332.

Pettijohn, D. E. (1988). Histone-like proteins and bacterial chromosome structures. *J. Biol. Chem.* **263**, 12793–12796.

Saul, F. A. & Poljak, R. J. (1992). Crystal structure of human immunoglobulin fragment Fab New refined

at 2.0 Å resolution. *Protein: Struct. Funct. Genet.* **14**, 363–371.

Saul, F. A., Amzel, L. M. & Poljak, R. J. (1978). Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin New at 2.0 Å resolution. *J. Biol. Chem.* **253**, 585–597.

Schmid, M. B. (1990). More than just ''histone-like'' proteins. *Cell*, **63**, 451–453.

Sibanda, B. L., Blundell, T. L. & Thornton, J. M. (1989). Conformation of β-hairpins in protein structures: systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J. Mol. Biol.* **206**, 759–777.

States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982). A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.* **48**, 286–297.

Surette, M. G. & Chaconas, G. (1992). The Mu transpositional enhancer can function in trans: requirement of the enhancer for synapsis but not strand cleavage. *Cell*, **8**, 1101–1108.

Tanaka, I., Appelt, K., Dijk, J., White, S. W. & Wilson, K. S. (1984). 3-Å resolution structure of a protein with histone-like properties in prokaryotes. *Nature*, **310**, 376–381.

Vis, H., Boelens, R., Mariani, M., Stroop, R., Vorgias, C. E., Wilson, K. S. & Kaptein, R. (1994). ^{1}H, ^{13}C, and ^{15}N resonance assignments and secondary structure analysis of the HU protein from *Bacillus stearothermophilus* using two- and three-dimensional double- and triple-resonance heteronuclear magnetic resonance spectroscopy. *Biochemistry*, **33**, 14858–14870.

Vuister, G. W. & Bax, A. (1993). Quantitative J correlation: a new approach for measuring homonuclear three-bond $J(H^{N}H^{\alpha})$ coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc.* **115**, 7772–7777.

Vuister, G. W., Wang, A. C. & Bax, A. (1993). Measurement of three-bond nitrogen-narbon J couplings in proteins uniformly enriched in ^{15}N and ^{13}C. *J. Am. Chem. Soc.* **115**, 5334–5335.

White, S. W., Appelt, K., Wilson, K. S. & Tanaka, I. (1989). A structural motif that bends DNA. *Proteins*, **5**, 281–288.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Edited by P. E. Wright

(Received 25 May 1995; accepted 22 September 1995)