NMR Structure of the DNA-Binding Domain of the Cell Cycle Protein Mbp1 from *Saccharomyces cerevisiae*†

Margie Nair,‡ Pauline B. McIntosh,‡ Thomas A. Frenkiel,§ Geoff Kelly,§ Ian A. Taylor,|| Stephen J. Smerdon,|| and Andrew N. Lane*,‡,⊥

Division of Molecular Structure, MRC Biomedical Research Centre, and Division of Protein Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom, and James Graham Brown Cancer Center, University of Louisville, Louisville, Kentucky 40202

Received August 7, 2002; Revised Manuscript Received December 2, 2002

ABSTRACT: The three-dimensional solution structure of the DNA-binding domain of Mlu-1 box binding protein (Mbp1) has been determined by multidimensional NMR spectroscopy. Mbp1 is a cell cycle transcription factor from *Saccharomyces cerevisiae* and consists of an N-terminal DNA-binding domain, a series of ankyrin repeats, and a heterodimerization domain at the C-terminus. A set of conformers comprising 19 refined structures was calculated via a molecular dynamics simulated annealing protocol using distance, dihedral angle, and residual dipolar coupling restraints derived from either double or triple resonance NMR experiments. The solution structure consists of a six-stranded β-sheet segment folded against two pairs of α-helices in the topology of the winged helix−turn−helix family of proteins and is in agreement with the X-ray structures. In addition, the solution structure shows that the C-terminal tail region of this domain folds back and makes specific interactions with the N-terminal β-strand of the protein. This C-terminal region is essential for full DNA-binding activity but appears in the X-ray structure to be disordered. The fold-back structure extends the region of positive electrostatic potential, and this may enhance the nonspecific contribution to binding by favorable electrostatic interactions with the DNA backbone.

Transcriptional regulation in eukaryotes is a complex process involving interactions between arrays of transcription factors that are often modular in architecture. These domains may be required for detection of extranuclear signals, recruitment, sequence-specific DNA binding, or binding of coregulatory proteins. Interplay between these protein assemblies is essential for the execution of a precise program of gene expression that occurs in response to many different signals. Entry into the eukaryotic cell cycle is regulated at a point in late G1-phase where factors such as cell size, nutrient availability, and the presence of cytokines are monitored. If conditions are favorable, cells become committed to cell division and will not arrest until mitosis has taken place. Progression into S-phase is regulated through a complex interplay between cyclin-dependent kinases and their inhibitors, G1/S-specific transcription factors and other regulatory molecules. In mammalian cells, mutations within many of these proteins result in uncontrolled cell proliferation and tumor development.

In budding yeast, the decision to enter S-phase is taken at a point called "start". This is the committed step in cell cycle progression and probably corresponds to the "restriction" checkpoint of the mammalian cell cycle. DNA replication is initiated, cells duplicate their spindle-pole bodies, and bud formation begins. Passage through start and entry into S-phase require the integration of a plethora of intra- and extracellular signals and their communication to two transcriptional complexes known as MBF[1] and SBF. Swi6 (92 kDa) is a common subunit of both complexes and is required for transcriptional activation but is not capable of DNA binding. This activity is supplied by two related proteins, Swi4 (120 kDa) and Mbp1 (90 kDa) (*1*, *2*). Heteromers of Swi6/Swi4 (SBF) bind to upstream sequences in the cyclin promoters known as SCB's which exist in multiple copies and have the consensus sequence 5′-CACGAAA-3′ (*2*). Similarly, genes required for DNA replication are regulated by binding of Swi6/Mbp1 heteromers (MBF) to sequences called MCB's with the consensus 5′-ACGCGTNA-3′ (*3*, *4*).

Mbp1 is a 93.9 kDa protein that comprises three domains which include an N-terminal DNA-binding domain followed by a series of ankyrin repeats and a heterodimerization domain at the C-terminus (*5*). Our construct of Mbp1 comprises a 15.4 kDa segment (residues 2−124) of the

† This work was supported by the U.K. Medical Research Council.

* To whom correspondence should be addressed at the James Graham Brown Cancer Center. E-mail: anlane01@gwise.louisville.edu. Fax: (502) 852-5679.

‡ Division of Molecular Structure, National Institute for Medical Research.

§ MRC Biomedical Research Centre, National Institute for Medical Research.

|| Division of Protein Structure, National Institute for Medical Research.

⊥ James Graham Brown Cancer Center, University of Louisville.

[1] Abbreviations: HTH, helix−turn−helix; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; Mbp1, Mlu-1 box binding protein; MBF, MCB binding factor; SBF, SCB binding factor; MBC, Mlu-1 cell cycle box; SCB, Swi4/6 cell cycle box; RDC, residual dipolar coupling.

10.1021/bi0205247 CCC: $25.00
Published on Web 01/17/2003

N-terminal DNA-binding domain. We have used multidimensional NMR spectroscopy to determine the solution structure of Mbp1, and we have compared this structure with previously reported crystal structures on this domain. The crystal structures of Mbp1 (PDB ID codes 1BM8 and 1MB1) (*6*, *7*) revealed that the molecule is similar in fold to the winged helix−loop−helix family of proteins, comprising a six-stranded antiparallel β-barrel packed against a bundle of four α-helices. Two β-strands (β5/β6) and a pair of α-helices (αA/αB) together constitute a structural motif characteristic of the "winged helix−turn−helix" (wHTH) subfamily of HTH proteins. The C-terminal residues that are essential for full DNA-binding activity (*5*) are disordered and not visible in the electron density map. Backbone dynamics determined by ^{15}N NMR (*8*) have also indicated that the C-terminal region of the molecule is generally highly mobile relative to the core of the molecule. However, these relaxation studies also show that part of the tail region comprising about eight residues is less mobile, suggesting that part of the tail region may show defined structure. Analysis of NOE interactions between the tail region and the N-terminal β strand suggests that part of the C-terminal tail region may fold onto the core of the molecule. These dynamics studies also indicated slow conformational exchange processes in the core of the molecule adjacent to the putative recognition helix. Such slower motions have been detected in other winged helix−turn−helix proteins and may be involved in interactions with the DNA minor groove (*9*−*11*). The C-terminal extension, which is remote in sequence from the wing region, however, seems to be a property of a subclass of proteins exemplified by Mbp1 and homologues.

We report here the solution structure of residues 2−124 of the DNA-binding domain of yeast transcription factor Mbp1 as determined by double and triple resonance NMR spectroscopy and molecular dynamics simulations. NMR studies indicate strong similarity between the solution and crystal structures for the core of the molecule and confirm the presence of a partially structured C-terminal tail region involving a few residues that make contact with the core of the molecule. The solution structure provides a model for understanding the extended DNA-binding surface and rationalizing the available biochemical data on the protein−DNA interaction.

MATERIALS AND METHODS

Protein Expression and NMR Sample Preparation. Samples of ^{15}N-labeled, ^{15}N,^{13}C doubly labeled, and unlabeled Mbp1 were expressed using *Escherichia coli* (*12*). All NMR samples were prepared using 40 mM sodium phosphate and 100 mM NaCl in a 90% H_2O/10% D_2O solution at pH 7.6.

NMR Data Collection, Methods, and Analysis. The backbone and side-chain assignments of Mbp1 have been reported (*13*) and included chemical shift data deposition in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession number 4254. Proline residues were assigned by analysis of the HACAN experiment that correlates the intraresidue Hα, Cα chemical shift with the ^{15}N chemical shift assignment of the following residue (*14*). Carbonyl carbons were assigned using HNCO (*15*) and correlation with the known amide proton and hydrogen shifts (*13*).

NMR data were recorded at 288 K using Varian Unity, UnityPlus, and Inova spectrometers operating at 500 and 600 MHz. Data were processed using NMRPipe software (*16*). NOE volumes were determined from 3D {^{1}H,^{15}N} and {^{1}H,^{13}C} NOESY−HSQC spectra (*17*) recorded at mixing times of 100 and 200 ms, respectively. NOE peak volumes were measured in XEASY (*18*) using the standard integration protocol. Peak distances were calibrated in DYANA1.5 (*19*). Hydrogen bond restraints were determined using the software program HBPlus (*20*), and fast exchanging amides were confirmed using the CLEANEX-HSQC experiment recorded with a series of mixing times (*21*). ϕ angle restraints were determined from HNHA spectra where the intensity ratio of the diagonal HN peak and the H_α cross-peak was measured using the formula $S_{cross}/S_{diag} = -\tan^2 2\pi J_{HH}\zeta$, where 2ζ are the dephasing and rephasing delays (*22*). Additional dihedral angle restraints were calculated with the program TALOS (*23*).

Residual dipolar couplings were measured using the *n*-alkylpoly(ethylene glycol)/*n*-alkyl alcohol alignment method, which has been shown to form dilute liquid crystalline phases in aqueous solution appropriate for protein studies (*24*). Deuterium quadrupolar splittings of the solvent were observed in ^{2}H NMR spectra for a polyoxyethylene 5 octyl ether ($C_{18}H_{36}O_6$) and octanol solution in 90% H_2O/10% D_2O. Two-dimensional *J*-modulated HSQC experiments (*25*) were recorded on the ^{15}N-labeled partially aligned sample at 288 K to measure residual NH dipolar couplings. The axial and rhombic components of the alignment tensor were estimated by analysis of the distribution of dipolar couplings using the maximum likelihood method (*26*) and the powder pattern approximation method (*27*).

Structure Calculation Protocols. Structures were initially calculated using DYANA 1.5 (*19*) with a total of 894 distance restraints (which included 270 medium-range and 205 long-range restraints), 106 ϕ angles (which included 74 TALOS-derived restraints and 32 restraints determined from HNHA spectra), 100 ψ angle restraints derived from TALOS, and 72 hydrogen bond restraints (Table 1). A set of 100 structures with low target function and low restraint violation was selected for refinement in X-PLOR (*28*), incorporating 87 residual dipolar coupling (RDC) restraints. The upper limit distance restraints in DYANA were converted to X-PLOR upper limit restraints. The error range used for the RDC restraints was ±0.2. The refinement protocol in X-PLOR comprised an initial stage of 2000 steps of conjugate gradient minimization of the structures. This was followed by 3000 steps of Verlet dynamics starting at 2000 K and slow cooling to 500 K in steps of 250 K with NOE and dihedral force constants set to 200 kcal mol^{-1} Å^{-1} and 150 kcal mol^{-1} deg^{-1}, respectively. RDC force constants were increased from 0.001 to 0.1 kcal mol^{-1} Hz^{-1}. The final step comprised 3000 steps of Verlet dynamics at 100 K followed by 2000 steps of conjugate gradient minimization with NOE and dihedral force constants of 30 kcal mol^{-1}Å^{-1} and 175 kcal mol^{-1}deg^{-1}, respectively, and the RDC force constant set to 1.1 kcal mol^{-1} Hz^{-1}. A family of conformers comprising 19 structures was selected on the basis of low overall energy and least NOE, dihedral, and RDC violations after final minimization. Selected structures were analyzed using the software programs MOLMOL (*29*) and INSIGHT2000 (MSI). Electrostatic potential maps were calculated using the program GRASP v-1.2.5 (*30*). PROCHECK-NMR (*31*) was used to analyze the quality of the structures. The 13 C-terminal

Table 1: Restraint and Structural Statistics for Mbp1 Structures

restraints	
NOE distance restraints	894
hydrogen bonds	72
residual dipolar coupling restraints	87
dihedral angle restraints	106 ϕ
	100 ψ
DYANA structural statistics	
no. of conformers selected	100
range of target function	2.30−4.38
mean target function and SD	3.47 ± 0.47
refinement statistics for selected 19 conformers	
NOE violations >0.5 Å	0
dihedral angles >5°	0
residual dipolar coupling violation >1.0 Hz	0
mean RMSD from experimental restraints	
distance (Å)	0.015
dihedral angle (deg)	0.71
residual dipolar coupling (Hz)	0.21
mean RMSD from idealized geometry	
bond (Å)	0.0012
angle (deg)	0.49
RMSD to mean structure (Å)	
backbone core (4−101)	0.84 ± 0.15
all heavy atoms (4−101)	1.91 ± 0.23
backbone dihedral angles (Ramachandran plot)	
most favored region (%)	73.5
allowed region	23.2
generously allowed	1.8
disallowed	1.5

residues comprising the histidine affinity tag were not included in the calculations. Coordinates and restraints for a set of 19 structures have been deposited in the Protein Data Bank (http://www.pdb.org/) under accession number 1L3G.

RESULTS AND DISCUSSION

Since the original report (*13*), we have assigned the C′ resonances, and we have completed the assignment of the proline residues. The proline residues in the core of the molecule were unambiguously identified in HNCA (*32*) and CBCA(CO)NH (*33*) experiments. However, in the tail region, the sequence PPPAP is problematic owing to the absence of H_N for the sequential proline residues. We resolved the ambiguity using the HACAN experiment (*14*), which also confirmed the other proline residues.

Secondary Structure. Figure 1a shows a summary of the sequential and medium-range NOEs involving the H_N, Hα, and Hβ protons and the dihedral angle restraints. The restraint data clearly indicate the presence of secondary structure elements, and these are confirmed by values of the Cα chemical shift (*34*) (data not shown). The conserved core of the molecule comprises residues 1−101 as defined by the crystal structures. Figure 1b shows a comparison of the secondary structure elements determined in solution from these NMR data with those determined from the X-ray structures. The remaining C-terminal residues (102−124) are classified as the disordered "tail" of the molecule which comprises residues involved in conformational exchange or high-frequency motions of substantial amplitude according to ^{15}N relaxation data. Analysis of the ϕ/ψ angles for 19 refined structures of Mbp1 indicates that most of the backbone torsion angles (98.5%) lie within the allowed regions of the Ramachandran plot. Most of the residues in the disallowed regions (1.5%) of the Ramachandran plot are positioned in poorly defined segments across the backbone (*31*) (Table 1).

Tertiary Structure. The three-dimensional conformation of Mbp1 was calculated using a combination of NMR data as summarized in Table 1 and as described in Materials and Methods.The backbone atoms of residues I4−D101 from 19 refined structures were superimposed using MOLMOL (*29*). The best fit to the mean coordinate position was determined to be 0.84 ± 0.15 Å. The superimposed structures are shown in Figure 2; poorly defined portions of the molecule correspond to a lack of NOE data (Figure 3).

Winged helix−turn−helix motif DNA-binding proteins typically contain a pair of α-helices and one or two β-hairpin wing segments that display flexibility in terms of DNA recognition and binding (*35*, *36*). The three-dimensional solution structure of Mbp1 establishes the presence of these secondary structure elements between residues 6 and 101 (Figure 2). The structure consists of an N-terminal double antiparallel β-stranded segment (sheet β1−β4) followed by a short α-helix, helix A (A36−N44), which is interrupted by a single residue turn (L40). Helix B (A48−E57), a relatively longer α-helical segment, is separated from helix A by a three-residue loop segment (F45−K47). Helix A folds against strands β2 and β3 as illustrated by a number of NOE interactions between these segments of the protein. The A/B helical pair corresponds to the HTH motif. The classical HTH motif consists of two helices joined by a tight turn, and their axes make an angle of approximately 120°. The second helix is termed the recognition helix and binds to the major groove of the DNA (*36*). Helix B of Mbp1 has been identified as the putative binding surface for DNA by homology comparison as well as sequence and chemical shift analysis (*5*), electrostatic potential mapping methods (*6*, *7*), and cross-saturation experiments (*37*). Lack of NOE data results in poorly defined structure in the latter part of helix A and the adjoining region of helix B for the solution structure. The low restraint density correlates with the ^{15}N relaxation data which showed the presence of slow conformational exchange processes involving the residues in the loop region neighboring helix B and extending into the region of β-sheet strand 5 (*8*).

Analysis of the NMR data indicates that helix D may extend from about residue 93 to residue 100. The coupling constant data derived from NMR experiments suggest helical torsion angles for residues Y93−Q95 in the solution structure, and the NOEs and chemical shifts indicate further helical nature from residues 98 to 100 (Figure 1a) with a bend-inducing proline residue (P98). This is consistent with one of the X-ray structures (1BM8).

In the solution structure of the Mbp1 DNA-binding domain, the antiparallel β-stranded winged segment (sheet β5 and β6) folds onto the core of the molecule and lies at nearly right angles to the four-stranded β segment. This six-stranded β-barrel is restrained by side-chain and backbone interactions between residues of the four-strand β segment and the β-hairpin. These NOE interactions include backbone side-chain interactions between S11 and F70. The β-hairpin segment consists of residues E64−K65 and T76−P79 and forms part of the putative DNA-binding surface together with the helix−turn−helix bundle. Measurement of chemical shift changes by NMR upon DNA binding (*5*) indicates that significant changes occur for a cluster of residues, R50−Q74, that corresponds to helix B and strands β5 and β6.

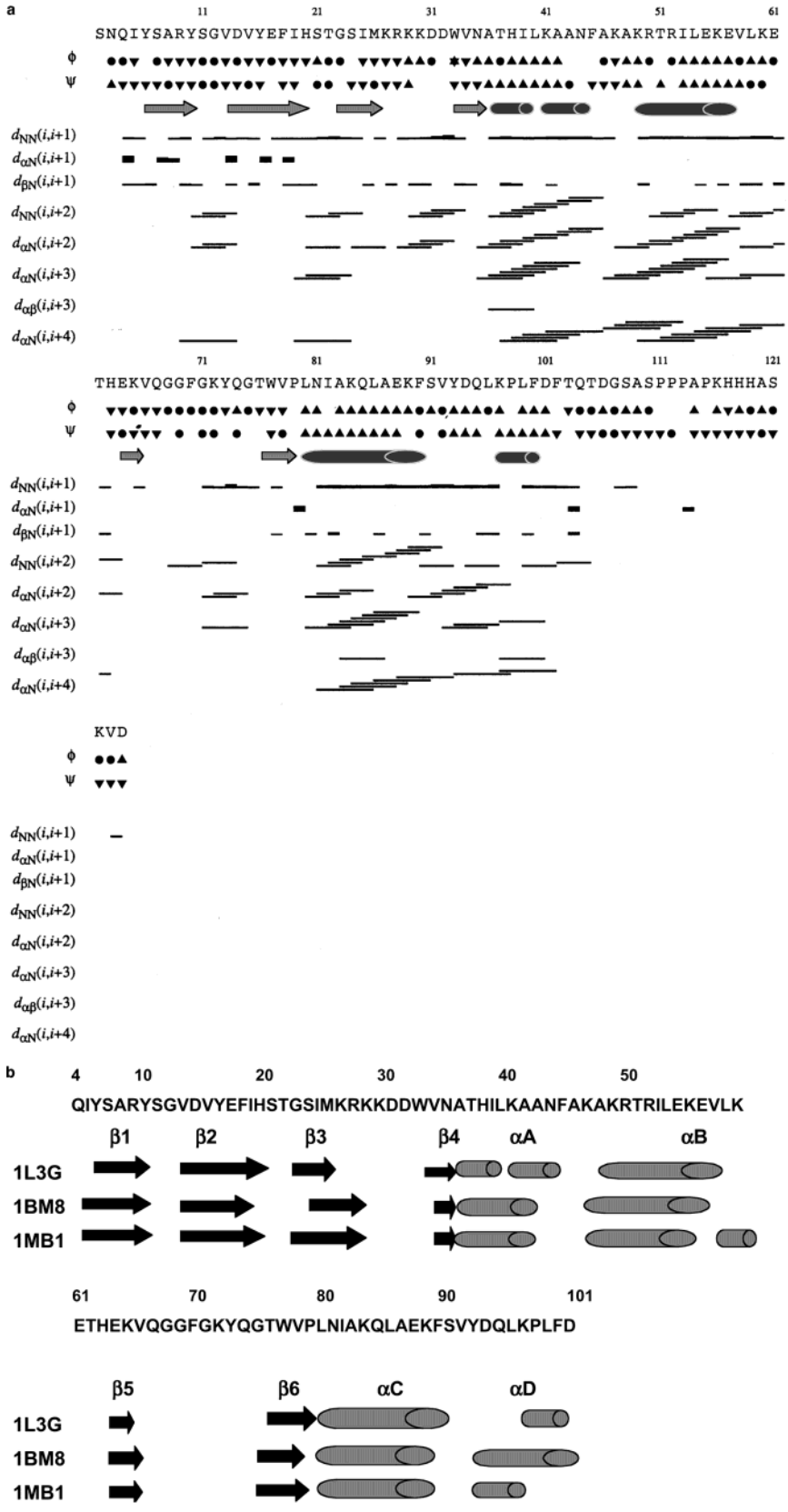

FIGURE 1: Secondary structures of Mbp1 in solution compared with the crystal state. (a) Summary of the upper distance limit sequential and medium-range NOEs for the H_N, Hα, and Hβ protons and the backbone dihedral angles (ϕ and ψ) generated using DYANA 1.5 (*19*). Horizontal lines indicate NOE interaction between two residues. The symbol ▼ indicates torsion angle values compatible with β-strands, and the symbol ▲ indicates torsion angle values observed in the α-helix. The symbol * refers to conformations of both α and β secondary types. The symbol ● refers to torsion angle restraint that excludes values of regular secondary structure elements. Arrows and cylinders represent β-sheet and α-helical secondary structure elements, respectively. (b) Comparison of the secondary structure elements derived from NMR (1L3G, this study) with the X-ray structures (1BM8 and 1MB1) using MOLMOL *(29)*. This comparison extends only to residue 101 because of the absence of electron density beyond this residue. Arrows and cylinders represent β-sheet and α-helical secondary structure elements, respectively.

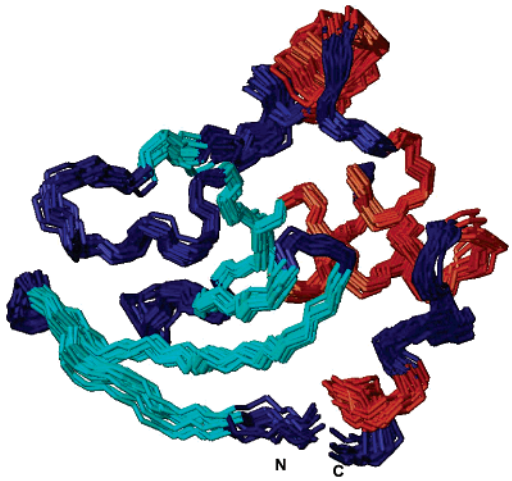

FIGURE 2: Solution structure of the core of Mbp1. An ensemble of 19 refined structures of Mbp1 (4−101) created in MOLMOL (*29*) indicates regions of secondary structure elements, β-sheet (light blue), α-helix (orange), and turns plus other (dark blue).

Conformation of the Tail Region (Residues 102−124). NOE contacts for the tail region (residues 102−124) comprise mainly intraresidue and sequential contacts. However, 12 long-range distance restraints between the core of the molecule and the C-terminal region were identified in the NOE spectra. These restraints involve three residues in the N-terminal β-strand, namely, Y6, S7, and A8, and residues P113 and A114 in the tail region. The HN atom of residue A114 shows interactions with the side-chain atoms Hα and Hβ of Y6 and the HN atom of residue S7. NOE interactions between HN of S7 and the side-chain atoms of P113 were also observed. The HN atom of A8 interacts with the side-chain Hγ of P113. Four NOE interactions involving F102 HN and the side-chain atoms Hγ and Hδ of I5 were also observed. This finding is in agreement with ^{15}N relaxation rate constants for residues A109−K116 which were similar to those of core residues, thus suggesting similar motional correlation times (*8*). Figure 4 shows the structure of the average conformer of Mbp1 including a portion of the fold-back tail region (residues F102−H118) of the molecule.

DSSP (discrimination of protein secondary structure class) analysis (*38*) indicates a bend conformation for residues P113−P115. This segment is folded back onto the core of the molecule and interacts with residues of the N-terminal β-sheet region. As a result of contacts observed in NOESY spectra between residues in the triple Pro segment and β-strand 1, this C-terminal region forms a loop around residue S110 which allows the region of the triple Pro segment to lie nearly parallel to a segment of β-strand 1. Classification of secondary structure elements for the tail region from atomic coordinates using the program STRIDE (*39*) indicates that a type IV turn is present for residues T105−S108 for the majority of conformers. NOE contacts for this segment are few and are primarily sequential NOEs involving G107, S108, and A109. Although relaxation and

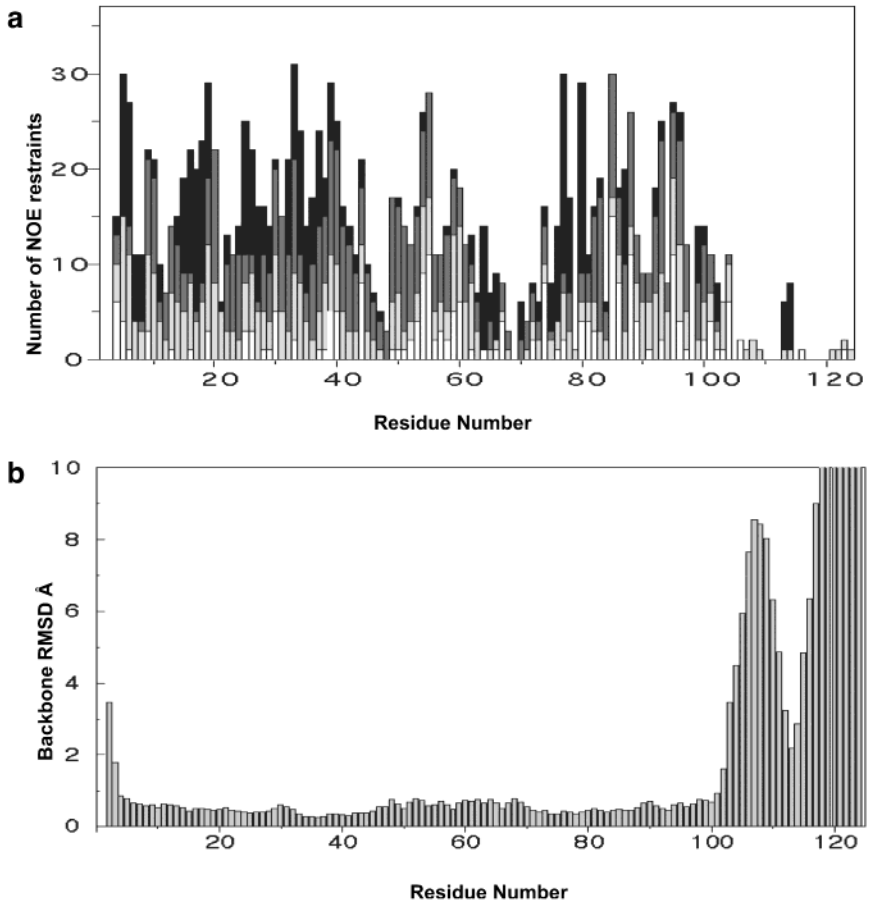

FIGURE 3: Restraint data for Mbp1. (a) Distance restraint data per residue [intraresidual (white), short range (vertically hatched), medium range (horizontally hatched), and long range (black)] generated using DYANA 1.5 (*19*). (b) Mean RMSD across backbone (Å) for 19 conformers of Mbp1 (2−124). As residues beyond 118 in the free protein are completely disordered, the *y* axis is truncated in this plot. The maximum RMSD is 26 Å.

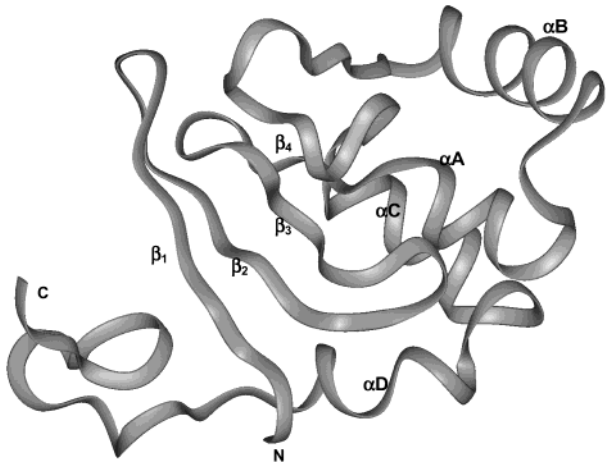

FIGURE 4: Solution structure of Mbp1 including the tail. Structure of the average conformer of Mbp1 showing a portion of the tail region (residues 102−118) folding onto the core region of the molecule.

chemical shift indices indicate a lack of structure beyond about residue 118 in the free protein, substantial changes in chemical shifts and relaxation rate constants occur for residues A120, S121, and K122 upon DNA binding. This suggests that these residues become more ordered on DNA binding and therefore should be considered as part of the DNA-binding site.

Comparison with the Crystal Structures. Superimposition of the crystal structures and the average NMR structure indicates low backbone RMSD values for regions encompassing secondary structure elements (Figure 1b). In the average NMR structure, regions identified as having secondary structure elements show RMSD deviations of <1 Å as compared with either crystal structure. In the NMR structure, there is a correlation between the lack of NOE restraints and high RMSD value as noted for the winged β-sheet segment (T62−K72), the C-terminus of helix B, and the loop between helix C and helix D (Figure 5).

In the crystal structure 1MB1 (Figure 5), helix B is segmented into an α-helix (helix B) and a 3_{10}-helix (helix B′) by a short turn comprising residues K56 and E57. In the solution structure, helix C (L80−F90) and helix D (K97−F100) are separated by a relatively longer loop segment compared to either crystal structures. Crystal structures 1BM8 and 1MB1 show helices C and D in an antiparallel arrangement separated by a three-residue loop segment. In the crystal structures, a proline residue produces a bend between β-strand 5 and helix C. Both 1BM8 and the solution structure incorporate a bend-inducing proline residue as part of helix D. Helix D is restrained by interactions with residues F18 and S21 of β-strand 2. Comparison of the helical pairs for the solution and crystal structures indicates that secondary structure elements are not equally defined (Figure 5).

Interhelical angles for the solution structure and the two crystal structures were calculated in MOLMOL (*29*) using cylinder axes (data not shown). Overall, there is good agreement for these values for all three structures although some differences for interhelical angle values involving helix D were noted. This may be related to the differences in the regions of defined secondary structure elements as shown in Figures 1 and 5.

In the crystal structure 1BM8 (Figure 5) the loop of the β-hairpin shows flexibility as suggested by weak electron density and high thermal factors for these core residues (*6*). It was suggested that the region may undergo a conforma-

FIGURE 5: Comparison of the solution structure with the X-ray structures of Mbp1. Ribbon diagrams were created in MOLMOL (*29*) showing secondary structure elements for the core region of the Mbp1 (a), 1BM8 (b), and 1MB1 (c) average NMR conformer.

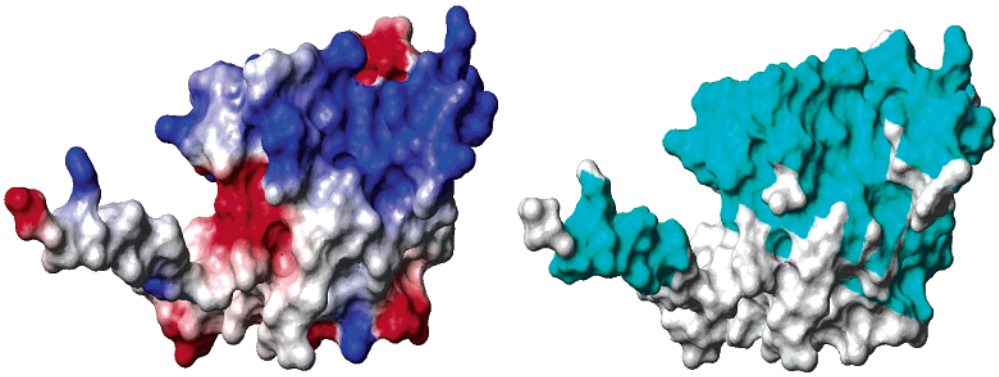

FIGURE 6: Electrostatic potential and DNA-binding surface map of NMR structure. The structure of residues 2−124 closest to the average was selected for calculating the electrostatic potential. The orientation of both surface plots is the same as that shown in Figure 4. The electrostatic potential map on the left indicates acidic (red) and basic (blue) sites, respectively. The regions of positive electrostatic potential comprise helix A and helix B and the β-hairpin segment. The DNA-binding surface defined by chemical shift mapping and cross-saturation experiments is shown on the right (in cyan).

tional change upon DNA binding. Few NMR restraints are available for the loop region of the β-hairpin segment. The large RMSD values across the backbone for this region are in agreement with the lack of NMR data (Figure 3).

In the crystal structure 1MB1, residue M26 imposed a high degree of ordering on the crystal structure as it makes a large number of van der Waals contacts with apolar and aromatic residues. The significance of the M26 residue is also evident in the solution structure where M26 shows NOE interactions with β-strand 2 (Y16, E17, F18) and β-strand 4 (S35) and with residue 73 of the winged β-hairpin.

The precision with which the NMR structure could be defined was in some regions limited by the high RMSD for some segments of the molecules within the ensemble. This is due to the low number of observable NOE interactions in these regions. The low restraint density is an indication of dynamical processes, either fast processes in the tail region or slower conformational exchange in the wing region. This in turn limits a complete comparison between NMR and crystal structures, and it is thus difficult to assess whether real differences between these structures do exist. However, the existence of these dynamical processes, which are related to functional properties, does imply differences with the single conformation observed in the crystal state.

The basic surface of the electrostatic potential map of the solution structure for residues S2−D124 (Figure 6) confirms that residues of helix B, the putative recognition helix, and surface residues of helix A together with the β-hairpin segment confer positive electrostatic potential on the molecule. This reflects the number of conserved lysine and arginine residues. The fold-back of the tail region places residues H117 and K116 in a position to make interactions with the phosphodiester backbone of DNA as these residues confer positive electrostatic potential on the molecule. The mutation K116A causes a 10-fold decrease in affinity for DNA but not necessarily a corresponding change in specificity (*6*). Furthermore, K122 presumably also makes contact on folding in the presence of the DNA, as described above, consistent with the decreased DNA-binding affinity of the K122A variant of Mbp1. The location of the remainder of the Mbp1 protein cannot be defined from the present data as the truncated protein is disordered beyond residue ca. 120. The core protease-resistant fragment in the presence of DNA ends at about residue 124; there is a long, unstructured stretch of residues prior to the start of the ankyrin repeats (*5*). This suggests that the DNA-binding domain and the ankyrin repeats are linked by a long flexible tether.

The additional positive electrostatic potential of the ordered tail region adjacent to the "classical" winged HTH domain implies an extended DNA-binding surface. We presume that this extension from the tail provides primarily nonspecific electrostatic interactions. We note that, in order for such an extended binding surface to interact with DNA, either the DNA has to bend or the protein has to undergo substantial changes in conformation on DNA binding (or both). The NOE and chemical shift profiles of the DNA-bound state of the protein argue against large-scale alteration in the shape of the protein.

These structural data along with previously reported biochemical and dynamical data together demonstrate that Mbp1 is a member of the winged helix−turn−helix class of DNA-binding proteins. The addition of a tail extension increases the overall surface area of positive electrostatic potential, which is presumably responsible for much of the nonspecific part of the binding affinity. This further suggests that the electrostatic contribution to the free energy of binding should be fairly large.

ACKNOWLEDGMENT

NMR spectra were recorded at the MRC Biomedical NMR Centre in London.

REFERENCES

1. Koch, C., Moll, T., Neuberg, M., Ahorn, H., and Nasmyth, K. (1993) *Science 261*, 1551−1557.
2. Primig, M., Sockanathan, S., Auer, H., and Nasmyth, K. (1992) *Nature 358*, 593−597.
3. Lowndes, N. F., Johnson, A. L., Breeden, L., and Johnston, L. H. (1992) *Nature 357*, 505−508.
4. McIntosh, E. M., Atkinson, T., Storms, R. K., and Smith, M. (1991) *Mol. Cell. Biol. 11*, 329−337.
5. Taylor, I. A., McIntosh, P. B., Pala, P., Treiber, M. K., Howell, S., Lane, A. N., and Smerdon, S. J. (2000) *Biochemistry 39*, 3943−3954.
6. Xu, R. M., Koch, C., Liu, Y., Horton, J. R., Knapp, D., Nasmyth, K., and Cheng, X. (1997) *Structure 5*, 349−358.
7. Taylor, I. A., Treiber, M. K., Olivi, L., and Smerdon, S. J. (1997) *J. Mol. Biol. 272*, 1−8.
8. McIntosh, P. B., Taylor, I. A., Frenkiel, T. A., Smerdon, S. J., and Lane, A. N. (2000) *J. Biomol. NMR 16*, 183−196.
9. Furui, J., Uegaki, K., Yamazaki, T., Shirakawa, M., Swindells, M. B., Harada, H., Taniguchi, T., and Kyogoku, Y. (1998) *Structure 15*, 491−500.
10. Jin, C., Marsden, I., Chen, X., and Liao, X. (1998) *Biochemistry 37*, 6179−6187.
11. Jin, C., and Liao, X. (1999) *J. Mol. Biol. 292*, 641−651.
12. Taylor, I. A, and Smerdon, S. J. (1997) *Proteins 27*, 325−327.
13. McIntosh, P. B., Taylor, I. A., Smerdon, S. J., Frenkiel, T. A., and Lane, A. N. (1999) *J. Biomol. NMR 13*, 397−398.
14. Kanelis, V., Donaldson, L., Muhandiram, D. R., Rotin, D., Forman-Kay, J. D., and Kay, L. E. (2000) *J. Biomol. NMR 16*, 253−259.
15. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
16. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
17. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.
18. Bartels, C. H., Xia, T.-H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR 5*, 1−10.
19. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
20. McDonald, I., Naylor, D., Jones, D., and Thornton, J. (1993) *Hydrogen Bond Calculator* v3.15.
21. Hwang, T. L., van Zijl, P. C., and Mori, S. (1998) *J. Biomol. NMR 11*, 221−226.
22. Vuister, G. W., Delaglio, F., and Bax A. (1993) *J. Biomol. NMR 3*, 67−80.
23. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR 13*, 289−302.
24. Ruckert, M., and Otting, G. (2000) *J. Am. Chem. Soc. 122*, 7793−7797.
25. Tjandra, N., Grzesiek, S., and Bax, A. (1996) *J. Am. Chem. Soc. 118*, 6264−6272.
26. Warren, J. J., and Moore, P. B. (2001) *J. Magn. Reson. 149*, 271−275.
27. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson. 133*, 216−221.
28. Brunger, A. T. (1988) in *X-PLOR Version 3.1. A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
29. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
30. Nicholls, A., Bharadwaj, R., and Honig, B. (1993) *Biophys. J. 64*, a166−a166.

31. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
32. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
33. Grzesiek, S., and Bax, A. (1992*) J. Magn. Reson. 99*, 201−207.
34. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
35. Gajiwala, K. S., and Burley, S. K. (2000) *Curr. Opin. Struct. Biol. 10*, 110−116.
36. Huffman, J. L., and Brennan, R. G. (2002) *Curr. Opin. Struct. Biol. 12*, 98−106.
37. Lane, A. N., Kelly, G., Ramos, A., and Frenkiel, T. A. (2001) *J. Biomol. NMR 21*, 127−139.
38. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
39. Frishman, D., and Argos, P. (1995) *Proteins 23*, 566−579.

BI0205247