Journal of Biomolecular NMR (2005) 32: 83–88
DOI 10.1007/s10858-005-3662-3

NMR structure note

Solution structure of the HMG-box domain in the SSRP1 subunit of FACT

Nobuyuki Kasai[a], Yasuo Tsunaka[a], Izuru Ohki[a], Susumu Hirose[b], Kosuke Morikawa[a] & Shin-ichi Tate[a,*]

[a]*Department of Structural Biology, Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka, 565-0874, Japan;* [b]*Department of Developmental Genetics, National Institute of Genetics, Mishima, Shizuoka-ken, 411-8540, Japan*

Received 28 January 2005; Accepted 04 March 2005

Key words: FACT, HMG, NMR structure

Abbreviations: FACT – facilitates chromatin transcription; SSRP1 – structure-specific recognition protein-1; HMG – high mobility group; RDC – residual dipolar coupling; RMSD – root mean square deviation; FACT-HMG – the HMG-box domain of FACT; dFACT-HMG – the HMG-box domain of FACT from *Drosophila melanogaster*.

Biological context

Nucleosomes dynamically alter their conformations in response to transcription, replication, repair and recombination processes. In recent years, studies have revealed that facilitates chromatin transcription (FACT) might behave as a histone chaperone (Belotserkovskaya et al., 2004). FACT dissociates the histone H2A/H2B dimers from the nucleosomes and thereby facilitates RNA polymerase II transcription (Belotserkovskaya et al., 2004). Biochemical and genetic studies have shown that FACT associates with many elongation related factors (Belotserkovskaya et al., 2004). *Drosophila* FACT binds to a nucleosome and promotes GAGA factor-directed chromatin remodeling (Shimojima et al., 2003), highlighting its important roles in epigenetic regulation for homeotic genes. The FACT proteins, which are highly conserved in all eukaryotes, form heterodimers consisting of two subunits, structure-specific recognition protein-1 (SSRP1) and SPT16 (also known as CDC68). The smaller SSRP1 subunit contains a High Mobility Group (HMG)-box domain. The FACT HMG-box domain (FACT-HMG) is categorized in the non-sequence-specific HMG-box protein group, which includes HMGB1, HMGB2 and NHP6A. The intact FACT heterodimers also interact with nucleosomes and DNA in a sequence independent manner (Belotserkovskaya et al., 2004).

The solution structures of free yeast homolog NHP6A and the NHP6A-DNA complex, which shared the FACT function, were already reported (Allain et al., 1999). However, the structure of the HMG-box domain belonging to SSRP1 has not been determined yet. Here, we report the solution structure of the HMG-box domain (dFACT-HMG) in the SSRP1 subunit (residues 555–624) of FACT from *Drosophila melanogaster*.

Methods and results

We used the same *D. melanogaster* FACT cDNA as that in the previous study (Shimojima et al., 2003). The HMG-box domain cDNA, encoding residues 555–624 of the SSRP1 subunit of FACT (dFACT-HMG domain), was cloned into the pET28a vector between the *Nde*I and *Xho*I sites.

*To whom correspondence should be addressed. E-mail: tate@beri.or.jp

The dFACT-HMG was overexpressed in *Escherichia coli* strain BL21 (DE3), grown on minimal media containing $^{15}NH_4Cl$ (0.5 g/l) as the sole nitrogen source, with or without [U−^{13}C]-glucose (1 g/l) as the sole carbon source, to produce the uniformly ^{15}N- or ^{15}N/^{13}C-labeled protein. Protein expression was induced at an $OD_{600} = 0.6$ by the addition of IPTG to a final concentration of 1 mM. After Ni-NTA column (Qiagen) purification, the sample was treated with thrombin protease for 12 h (4 °C) to cleave the His_6-tag fragment. Finally, dFACT-HMG was purified on a Hi-Trap SP-Sepharose cation exchange column (Amersham) by elution with a concentration gradient of 0–1 M sodium chloride. The Gly, Ser, His and Met residues, which were derived from the vector sequence, remained at the N-terminus. To prepare the NMR sample, the purified dFACT-HMG was dialyzed against 2 mM sodium acetate buffer (pH 5.2) and was concentrated with Centriprep-3 and Centricon-3 filters (Millipore). The sample for the NMR experiment comprised 2.2 mM of the dFACT-HMG protein in a buffer containing 90% H_2O/10% D_2O or 100% D_2O.

All NMR experiments were carried out at 25 °C on Bruker DMX600 and DMX750 spectrometers. The sequence specific NMR backbone assignment of the dFACT-HMG domain was gained from the 3D HNCA, HN(CO)CA, HNCACB and CBCA(CO)NH spectra. The ^{1}H–^{15}N HSQC spectrum of the dFACT-HMG domain is shown in Figure 1. The amino acid residue numbers 6–74 correspond to residues 556–624 of the SSRP1 subunit of *Drosophila* FACT. All of the backbone signals were assigned, except for that of amide proton of Gly1 and Cα of Trp52. The intensities of the crosspeaks of some residues (Trp15, Trp52, Asp54 and Ala55) were very small in the ^{1}H–^{15}N HSQC spectrum. The labels of these four residues are outlined in Figure 1. The 3D HBHA(CO)NH, C(CO)NH, H(CCO)NH, HCCH-TOCSY, HCCH-COSY, CCH-TOCSY, CCH-COSY and ^{15}N-edited TOCSY spectra were used to assign sidechain resonances. Aromatic sidechain assignments were achieved from 2D TOCSY, NOESY and DQF-COSY. The sidechain of Trp52 showed line broadening signals. The Leu and Val stereospecific methyl assignments were obtained from the ^{13}C constant-time HSQC spectra of a 10% ^{13}C-labeled sample (Neri et al., 1989). The chemical shift assignments were deposited in the BioMagResBank (accession code BMRB-6469). All data were processed with the NMRPipe program (Delaglio et al., 1995) and were analyzed with the PIPP program (Garrett et al., 1991) and an in-house, semi-automatic backbone resonance assignment program, JASS (Tate et al. unpublished), to accelerate the backbone assignment process.

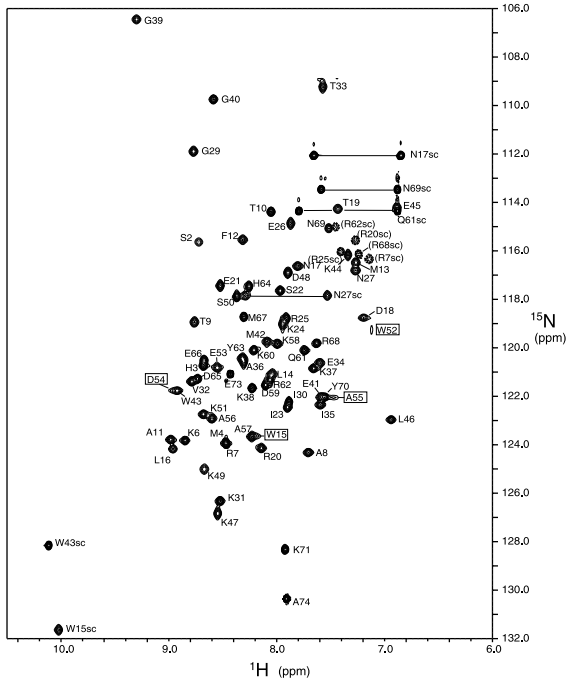

Figure 1. The ^{1}H–^{15}N HSQC spectrum of the dFACT-HMG domain. The assignments of the backbone amide groups are labeled. The residue numbers 6–74 correspond to residues 556–624 of the SSRP1 subunit of *Drosophila* FACT. The sc labels indicate sidechain peaks from tryptophans, asparagines or glutamine residues. The labels surrounded by circular brackets mean negative peaks resulting form aliasing. The labels surrounded by squares are indicative of the crosspeaks with relatively small intensities.

Distance restraints were obtained from 2D NOESY, 3D ^{15}N-edited NOESY and ^{13}C-edited NOESY spectra with a 100 ms mixing time. The inter-proton distance restraints were classified into three categories by the NOE peak intensities, corresponding to 1.8–3.0 (strong), 1.8–4.0 (medium) and 1.8–5.0 Å (weak), respectively. The amide protons that slowly exchanged with the bulk water were detected with a series of ^{1}H–^{15}N HSQC spectra. The distance restraints for the hydrogen bonds were applied for slowly exchanging amides, i.e. 2.8–3.4 Å for N–O and 1.8–2.4 Å for H–O pairs. The backbone torsion angles were estimated from the scalar $^3JH_{NH\alpha}$ coupling constants

Table 1. Structural statistics for the dFACT-HMG domain

Restraints for structural calculation	
NOE distance restraints	1530
Unambiguous	1511
intra-residue	484
sequential (|i−j| = 1)	352
Medium-range (|i−j| < = 4)	431
long-range (|i−j| > 4)	244
Ambiguous	19
Dihedral angle restraints	60
Hydrogen bonds	17
RDC restraints	54
RMSD from ideal stereochemistry	
Bonds (Å)	0.0054 ± 0.0002
Angles (°)	0.698 ± 0.014
Impropers (°)	0.605 ± 0.032
PROCHECK-NMR (Laskowski et al., 1996)	
most favored (%)	80.4
additionally allowed (%)	16.3
generously allowed (%)	1.8
disallowed (%)	1.5
Coordinate Precision	
RMSD from the average structure(Å)	
5–70 backbone atoms	0.71
5–70 heavy atoms	1.63
11–25, 32–44, 50–68 backbone atoms	0.42
11–25, 32–44, 50–68 heavy atoms	1.32
Comparison between the lowest energy dFACT-HMG structure and the other HMG-box domain structures	
dFACT-HMG (11–25, 32–44, 54–68) vs. NHP6A (27–40, 47–60, 71–85)	1.91 Å
dFACT-HMG (11–68) vs. HMG-D (11–68)	1.86 Å
dFACT-HMG (11–28, 29–47, 48–68) vs. HMGB1 A domain (14–31, 34–52, 55–75)	2.56 Å
dFACT-HMG (11–47,48–68) vs. HMGB1 B domain (13–49,52–72)	2.30 Å

derived from the *J*-modulation HSQC spectrum (Kuboniwa et al., 1994) and the TALOS analysis (Cornilescu et al., 1999). The $1D_{NH}$ RDC spectrum was measured in a 5% solution of $C_{12}E_5$/n-hexanol (Ruckert and Otting, 2000), using the IPAP (in-phase/anti-phase)-HSQC spectra (Ottiger et al., 1998). At the first stage of the structural determination, we used the DYANA program (Guntert et al., 1997), and then the structures were calculated using a standard simulated annealing protocol with CNS version 1.1 (Brunger et al., 1998). The Da (−15 Hz) and rhombicity (0.6) values were obtained for dFACT-HMG, through a grid search by minimizing the total energy of the calculated structures.

The solution structure of the dFACT-HMG domain was determined on the basis of the distance, dihedral angle and RDC restraints. The NOE restraints of the Trp52 sidechain were treated as ambiguous NOE restraints. The N- (residues 1–4) and C-termini (residues 71–74) were disordered, as judged from the heteronuclear ^{15}N-{^{1}H} NOE data (values < 0.4) (data not shown). A summary of the structural statistics for the dFACT-HMG domain is shown in Table 1. The 30 lowest energy structures from 100 initial structures showed no distance violations larger than 0.5 Å, and no angle violations larger than 5°. The Quality factor for the RDC was 12.1%. The residues of these 30 structures show good covalent geometries with well defined secondary structures. A backbone view of the 30 final structures of the dFACT-HMG domain is displayed in Figure 2a. The structures were overlaid

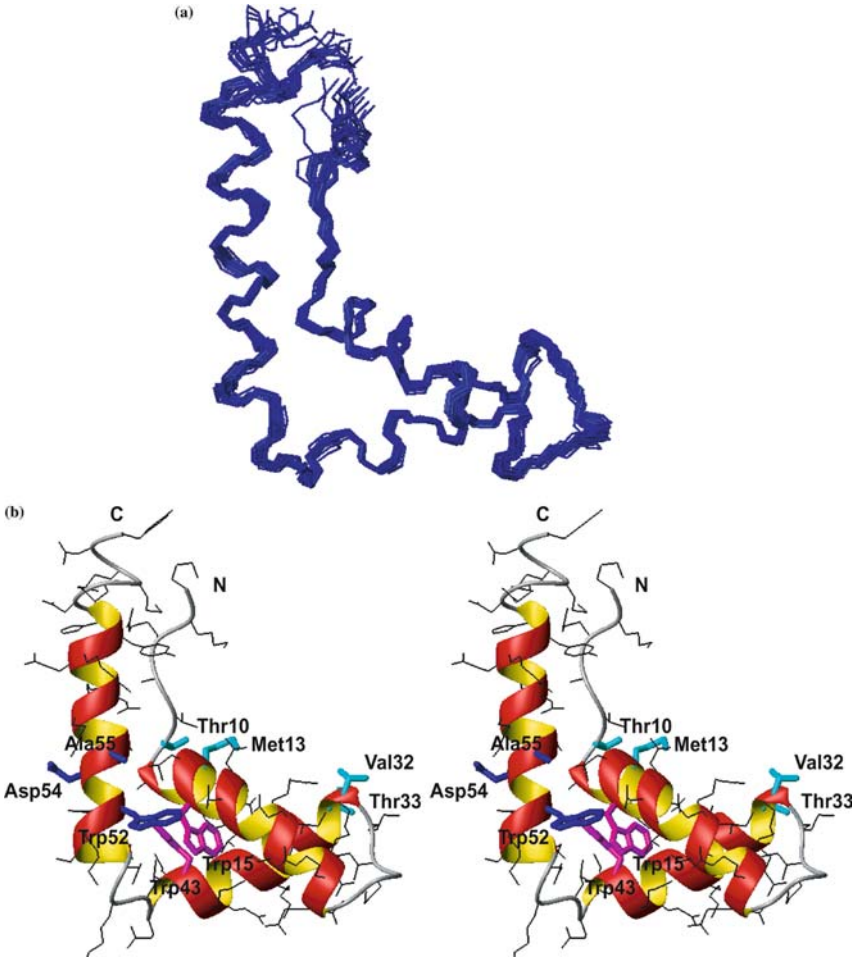

Figure 2. Solution structure of dFACT-HMG. The N-terminal (residues 1–4) and C-terminal (residues 71–74) disordered tails are removed for clarity. a, Ensemble of the final set of 30 lowest energy structures for the dFACT-HMG domain gained by best-fit superpositions of the Cα, C, O and N backbone atoms of the helices. b, Stereoview of a ribbon diagram of the lowest energy structure. The three residues that were broadened in the ^{1}H–^{15}N HSQC spectrum are colored in blue. The two tryptophans of the central major hydrophobic core are colored in magenta. The four DNA binding residues are colored in cyan. These images were prepared using MOLMOL (Koradi et al., 1996).

by a best-fit superposition of the backbone heavy atoms of residues Ala11–Arg25, Val32–Lys44 and Ser50–Arg68, which were located in the secondary structural regions. The atom coordinates have been deposited in the protein data bank (PDB ID:1WXL). The structure consists of three α-helices, composed of Ala11–Arg25, Val32–Lys44 and Ser50–Arg68. These helices are connected by two loops, formed by residues Glu26–Lys31 between helices 1 and 2, and by residues Glu45–Lys49 between helices 2 and 3. They jointly form the typical L-shaped fold of the HMG-box domains.

The backbone RMSD values between the lowest energy dFACT-HMG structure and other HMG-box domain structures are listed in Table 1. In terms of the RMSD values around 2 Å, the dFACT-HMG is somewhat similar to all of the other HMG-box domains. The surfaces opposite from the DNA binding surface in dFACT-HMG and NHP6A exhibit more negative charges than the others (Figure 3). The most remarkable difference is represented in the left column of Figure 3. In contrast, the corresponding surface of HMG-D, with its high sequence identity and conformational similarity to

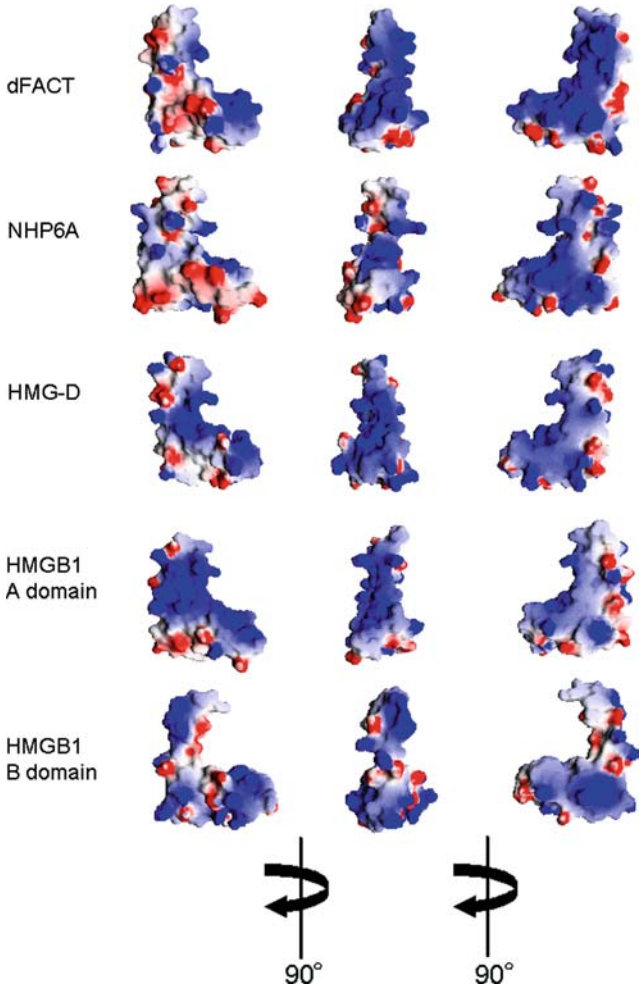

Figure 3. Comparison of the electrostatic potential surfaces among the HMG-box domains. These images represent the molecular surfaces of dFACT-HMG (this paper), NHP6A (Allain et al., 1999), HMG-D (Jones et al., 1994), and the HMGB1 A (Hardman et al., 1995) and B (Weir et al., 1993) domains with their electrostatic potentials. The molecular orientation of dFACT-HMG in the left column image is almost the same as that in Figure 2b. The central and right column images show the DNA binding interfaces. Calculations were performed with GRASP (Nicholls et al., 1991). Red and blue denote regions of negative and positive electrostatic potential, respectively (red < -5 kBT, blue $> +5$ kBT, where kB is the Boltzmann constant and T is the temperature in Kelvin).

dFACT-HMG, is covered by more positive charges.

Discussion and conclusions

The intensities of the NH crosspeaks of Trp15, Trp52, Asp54 and Ala55 in the ^{1}H–^{15}N HSQC spectrum were relatively small (Figure 1), indicating that the mobilities of these residues are distinct from those of the other residues. Three of the four amino acid residues were located in the N-terminus of helix 3 (Figure 2b). This result suggests that helix 3 is dynamic in the intermediate time-scale. Thus, it is likely that the FACT HMG-box domain is more flexible than the other general HMG-box domains. Consistent with this notion, NHP6A, which shares the FACT function, is folded and unfolded at 20 °C and 37 °C, respectively (Allain et al., 1999). This unfolding at 37 °C is also easily recovered by the addition of DNA.

The key residues of HMG-D for DNA binding are Ser10, Met13, Val32 and Thr33 (Jones et al., 1994). The corresponding residues of dFACT-HMG are assumed to be Thr10, Met13, Val32 and Thr33 (Figure 2b). Therefore, it is likely that the

DNA binding modes are very similar between dFACT-HMG and HMG-D, although their molecular surfaces and the properties of their hydrophobic cores remarkably differ.

A comparison of the electrostatic potential surfaces among the non-sequence-specific HMGB family shows the common character of both dFACT-HMG and NHP6A (Figure 3), which share the FACT function. It was reported that the maize FACT-HMG domain interacts directly with nucleosomes (Lichota and Grasser, 2001), and the nuclease sensitivity of the NHP6A interaction with nucleosomes has recently been characterized, in order to examine how the nucleosome structure was altered (Rhoades et al., 2004). These strikingly negative charges, which are specific for the FACT-HMG domain, may be ascribed to the interface with other proteins, such as histone proteins or basic regions within intact FACT subunits. These results provide the first structural basis for further studies on the structure-function relationships of FACT. These pieces of structural information are clues toward understanding the entire view of chromatin remodeling by FACT.

Acknowledgements

We thank Naoko Kajimura and Kyoko Matoba for assistance with sample preparation. This work was partly supported by a research grant endorsed by the Japan New Energy and Industrial Technology Development Organization (NEDO).

References

Allain, F.H., Yen, Y.M., Masse, J.E., Schultze, P., Dieckmann, T., Johnson, R.C., and Feigon, J. (1999) *EMBO J.*, **18**, 2563–2579.

Belotserkovskaya, R., Saunders, A., Lis, J.T., and Reinberg, D. (2004) *Biochim. Biophys. Acta*, **1677**, 87–99.

Brunger, 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. D. Biol. Crystallogr.*, **54**, 905–921.

Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR*, **13**, 289–302.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Garrett, D.S., Powers, R., Gronenborn, A.M., and Clore, G.M. (1991) *J. Magn. Reson.*, **95**, 214–220.

Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Hardman, C.H., Broadhurst, R.W., Raine, A.R., Grasser, K.D., Thomas, J.O., and Laue, E.D. (1995) *Biochemistry*, **34**, 16596–16607.

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.

Koradi R., Billeter M., Wuthrich K. (1996) J. Mol. Graph. **14**, 51–55, 29–32.

Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR*, **4**, 871–878.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.

Lichota, J., and Grasser, K.D. (2001) *Biochemistry*, **40**, 7860–7867.

Neri, D., Szyperski, T., Otting, G., Senn, H., and Wuthrich, K. (1989) *Biochemistry*, **28**, 7510–7516.

Nicholls, A., Sharp, K.A., and Honig, B. (1991) *Proteins*, **11**, 281–296.

Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.*, **131**, 373–378.

Rhoades, A.R., Ruone, S., and Formosa, T. (2004) *Mol. Cell. Biol.*, **24**, 3907–3917.

Ruckert, M., and Otting, M. (2000) *J. Am. Chem. Soc.*, **122**, 7793–7797.

Shimojima, T., Okada, M., Nakayama, T., Ueda, H., Okawa, K., Iwamatsu, A., Handa, H., and Hirose, S. (2003) *Genes. Dev.*, **17**, 1605–1616.

Weir, H.M., Kraulis, P.J., Hill, C.S., Raine, A.R., Laue, E.D., and Thomas, J.O. (1993) *EMBO J.*, **12**, 1311–1319.