Journal of Biomolecular NMR, **17:** 175–176, 2000.
KLUWER/ESCOM

Letter to the Editor: Sequence-specific ^{1}H, ^{13}C, and ^{15}N assignments of the MAR-binding domain of chicken MeCP2/ARBP

Eike Brunner[a], Joachim Weitzel[b], Björn Heitmann[a], Till Maurer[a], Wolf H. Strätling[b] & Hans Robert Kalbitzer[a]
[a]*Universität Regensburg, Institut für Biophysik und physikalische Biochemie, D-93040 Regensburg, Germany;* [b]*Institut für Medizinische Biochemie und Molekularbiologie, Universitäts-Krankenhaus Eppendorf, D-20246 Hamburg, Germany*

Received 7 December 1999; Accepted 31 March 2000

Key words: chicken methyl-CpG-binding protein 2 (cMeCP2/ARBP), NMR assignments

Biological context

Methyl-CpG-binding protein 2 (MeCP2) is an abundant chromosomal protein, which specifically recognizes symmetrically methylated CpG dinucleotides (Lewis et al., 1992). In addition, it was independently identified through its ability to preferentially bind nuclear matrix attached regions (MARs), the putative bases of chromatin loop domains, and named attachment region binding protein (ARBP; von Kries et al., 1991). In humans, mutations of the *MECP2* gene cause Rett syndrome, a progressive neurodevelopmental disorder (Amir et al., 1999). A highly conserved region of 125 amino acid residues in the N-terminal half of the protein is necessary for binding to MARs while a smaller subregion of 85 residues within this region (methyl binding domain, MBD) is sufficient for recognition of methylated CpGs (Nan et al., 1993; Weitzel et al., 1997). An NMR study of the latter domain from rat MeCP2 (rMeCP2 MBD) was published recently (Wakefield et al., 1999). MeCP2 furthermore contains a transcriptional-repression domain that interacts with a large corepressor complex containing the corepressor mSin3A and histone deacetylases HDAC1 and 2 (Nan et al., 1998). It is thought that long-term silencing of methylated sequences operates through the generation of a deacetylated closed chromatin structure.

Methods and results

The DNA sequence corresponding to amino acids 72 to 196, which comprises the MAR-binding domain of chicken MeCP2/ARBP (cMeCP2 ARBD) and to which a polyhistidine tag (MGHHHHHH-) is added at the N-terminus, was PCR-amplified using as template a partial cDNA (Weitzel et al., 1997) and primers of sequence: 5′-GATATACCATGGGCCATCACCATCACCATCACGCCCCGGCCGTG-3′ and 5′-CCGGATCCATTAGCCGCTGCCTTTG-3′. NcoI and BamHI cut amplified DNA was ligated into NcoI and BamHI cut pET-16b vector DNA to give pET-cARBP-Ex4.2. *E. coli* BL21(DE3)pLysS transformed with pET-cARBP-Ex4.2 were grown to an OD of 0.6 at 600 nm and induced with 1 mM IPTG for 3 h. To prepare uniformly ^{15}N- or ^{13}C-^{15}N-labeled proteins, cells were grown in isotopically labeled Bio-Express media (Cambridge Isotope Laboratories). Pelleted cells were resuspended in PBS (154 mM NaCl, 10 mM sodium phosphate buffer, pH 7.5) containing 5 mM 2-mercaptoethanol, lysed by shock freezing in liquid nitrogen and, after addition of 250 mM NaCl, by incubation with lysozyme (50 μg/ml), DNase I (10 μg/ml), and RNase A (10 μg/ml) for 30 min at 37 °C. Debris was spun down at 14×10^3 *g*. The supernatant was added to Ni^{2+}-agarose beads and rolled for 20 min. Beads were washed once with PBS containing 10 mM imidazole and twice with PBS containing 50 mM imidazole. cMeCP2 ARBD was then eluted with PBS containing 0.5 mM EDTA and 500 mM imidazole. The eluate was loaded on a Mono S HR5/5 FPLC column which was developed with a linear gradient from 150 mM to 600 mM NaCl in PBS. All fractions were monitored by SDS/PAGE on 13% acrylamide gels. cMeCP2 ARBD was eluted at ∼ 550 mM NaCl concentration. Fractions containing the ARBD at a purity

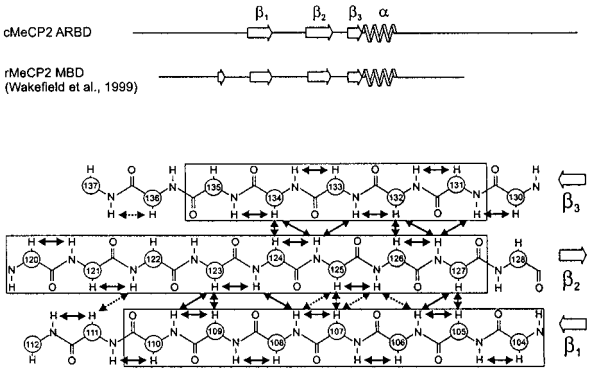

Figure 1. Secondary structure of cMeCP2 ARBD compared with rMeCP2 MBD and β-sheet topology. NOE contacts are marked by arrows. Dotted arrows indicate NOE contacts which are not completely unambiguous due to signal overlap or/and low signal-to-noise ratio. The β-strands predicted by CSI are indicated by boxes.

of 99% were concentrated and adjusted to PBS containing 0.1 mM NaN_3 using Centriprep 10 (Amicon). The identity and integrity of the ARBD was confirmed by mass spectrometry and Southwestern blot analysis.

NMR experiments were carried out at 298 K on Bruker DRX-800 and DRX-600 spectrometers (chemical shift referencing: ^{1}H to internal DSS, ^{15}N and ^{13}C indirectly following the IUPAC recommendations, Markley et al., 1998). Sequence-specific resonance assignments were carried out on the basis of HNCA, CBCA(CO)NH, ^{1}H-^{15}N-HSQC, ^{1}H-^{15}N-NOESY HSQC, ^{1}H-^{15}N-TOCSY HSQC, HNCO, and HCCH-TOCSY experiments. Chemical shift indices (CSI) were calculated using the computer program csi v. 1.1 (Wishart and Sykes, 1994; courtesy of Sykes et al., http://www.pence.ualberta.ca/ftp).

Extent of assignments and data deposition

The NMR data of cMeCP2 ARBD were recorded at 2 mM concentration in 154 mM NaCl, 10 mM sodium phosphate buffer, and at a pH of 6.8 where the protein remains soluble for at least three months. We assigned the chemical shifts of 99% of the backbone $^1H^N$, $^{15}N^H$, $^{13}C^\alpha$, 81% of the $^{13}C^\beta$, 83% of the $^{13}C'$, and 86% of the $^1H^\alpha$atoms (deposited in the BioMagResBank, http://www.bmrb.wisc.edu, under accession number 4467). From the CSI we predict the secondary structure shown in Figure 1. The N-terminal sequence (residues 72 to 103) does not exhibit canonic secondary structure elements. Three β-strands (β_1: residues 104–110 (GWTRKLK), β_2: residues 120–127 (KYDVYLIN), and β_3: residues 131–135 (KAFRS)), immediately followed by an α-helix (residues 136–145 (KVELIAYFEK)), are detected in the central region of cMeCP2 ARBD. The C-terminal last 51 residues again form a region not exhibiting canonic secondary structure elements, although this region is important for binding to MARs. The topology of the three-stranded antiparallel β-sheet was determined from NOE contacts (Figure 1).

It is interesting to compare our results for cMeCP2 ARBD (125 amino acid residues) with the data reported by Wakefield et al. (1999) for rMeCP2 MBD (88 amino acid residues), which is identical in sequence with the chicken protein. In contrast to the four-stranded β-sheet structure of rMeCP2 MBD, we observe only three β-strands for cMeCP2 ARBD. It should be noted that this observation is not only based on CSI calculations but also on the analysis of the NOESY spectra. Furthermore, the length of each strand differs from the respective one in the reported rMeCP2 MBD structure, while both studies predict a three-turn α-helix at identical corresponding positions.

Acknowledgements

We thank S. Giehler for excellent technical assistance. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged (grants SFB 545-B2 and Str 145/13-1 to W.H.S.; Heisenberg fellowship and grant Br 1278/8-1 to E.B.).

References

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. and Zoghbi, H.Y. (1999) *Nat. Genet.*, **23**, 185–188.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992) *Cell*, **69**, 905–914.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *Pure Appl. Chem.*, **70**, 117–142.

Nan, X., Meehan, R.R. and Bird, A. (1993) *Nucleic Acids Res.*, **21**, 4886–4892.

Nan, X., Ng, H.-H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenmann, R.N. and Bird, A. (1998) *Nature*, **393**, 386–389.

von Kries, J.P., Buhrmester, H. and Strätling, W.H. (1991) *Cell*, **64**, 123–135.

Wakefield, R.I.D., Smith, B.O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A.P. and Barlow, P.N. (1999) *J. Mol. Biol.*, **291**, 1055–1065.

Weitzel, J.M., Buhrmester, H. and Strätling, W.H. (1997) *Mol. Cell. Biol.*, **17**, 5656–5666.

Wishart, D.S. and Sykes, B.D. (1994) *J. Biomol. NMR*, **4**, 171–180.