Journal of Biomolecular NMR, **26:** 375–376, 2003.
KLUWER/ESCOM

Letter to the Editor: Sequence-specific assignment of the B-Myb DNA-binding domain (B-MybR2R3) bound to a 16 base-pair DNA target site corresponding to a regulatory site from the *tom-1* gene

Gareth Jones[a], Mark Howard[b], Pauline McIntosh[c], Richard A. Williamson[b] and Mark D. Carr[a,*]
[a]*Department of Biochemistry, University of Leicester, Adrian Building, University Road, Leicester LE1 7RH, U.K.;* [b]*Department of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, U.K.;* [c]*National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.*

Received 28 January 2003; Accepted 11 April 2003

Key words: B-Myb, B-MybR2R3, DNA-binding domain, *tom-1*

Biological context

The highly conserved vertebrate proteins c-Myb, A-Myb and B-Myb form an important family of transcriptional activators, which play key roles in regulating the proliferation and differentiation of progenitor cells. In recent years B-Myb has been shown to be a general, cell cycle regulated transcription factor, which plays a pivotal role in the cell cycle by activating the expression of genes required for the transition of cells from G1 to S-phase (Saville and Watson, 1998; Horstmann et al., 2000).

The Myb proteins are all large, single chain polypeptides (mouse c-Myb 636 residues, A-Myb 751 residues and B-Myb 704 residues), which contain a number of distinct functional regions, including a well defined N-terminal DNA-binding region composed of 3 imperfect repeats of 51 or 52 residues (R1, R2 and R3, Saville and Watson, 1998 and references therein). The sequence of the DNA-binding region is highly conserved across the vertebrate Myb family (about 85% similarity) and for both B-Myb and c-Myb the isolated R2R3 region has been shown to represent the minimal sequence-specific DNA-binding region, which binds as a monomer to a core 6 base pair site with the consensus sequence PyAACG/TG ($K_d \sim 1$–3 nM, Carr et al., 1996; McIntosh et al., 1998). To date, a number of bona fide Myb-regulated genes have been identified, including *mim-1* and *tom-1*, which contain essential Myb target sites in their promoter regions (Saville and Watson, 1998).

Previously, we have determined the solution structure of B-MybR2R3 (110 residues) and carried out preliminary studies of its interaction with a 16 base pair DNA fragment corresponding to the Myb binding site in the tom-1-A promoter of the *tom-1* gene (Carr et al., 1996; McIntosh et al., 1998). In this communication we report the determination of comprehensive sequence-specific resonance assignments for B-MybR2R3 bound to the tom-1-A 16 mer. This work will form the basis for detailed structural comparisons of B-MybR2R3 bound to several variants of the tom-1-A site and allow comparisons with both free B-MybR2R3 and c-MybR2R3 bound to a related target site.

Methods and experiments

The ^{15}N and $^{15}N/^{13}C$ labelled B-MybR2R3 were prepared from a pGEX-3X-based *E. coli* expression vector as described previously (Carr et al., 1996). The *E. coli* HB101 strain used for expression is auxotrophic for both proline and leucine and so produced samples of B-MybR2R3 in which all of the residues apart from proline and leucine were uniformly ^{15}N or $^{15}N/^{13}C$ labelled.

The 16 base pair DNA fragment (TCCTTAACGGACTGAG) corresponding to the Myb binding site was prepared as described previously using deoxyoligonucleotides supplied by Oswell (McIntosh et al., 1998). The ^{15}N and $^{15}N/^{13}C$ B-MybR2R3-DNA complexes were produced by mixing equal volumes of the two binding partners as 250 μM solutions in a 25 mM potassium phosphate, 100 mM potassium chloride and 50 μM EDTA buffer at pH 6.0. The NMR experiments were carried out on 0.35 ml samples of 1.0 mM ^{15}N and 1.6 mm $^{15}N/^{13}C$ labelled B-MybR2R3 bound to the tom-1-A 16-mer in either a 90% H_2O/10% D_2O or 100% D_2O buffer.

*To whom correspondence should be addressed.

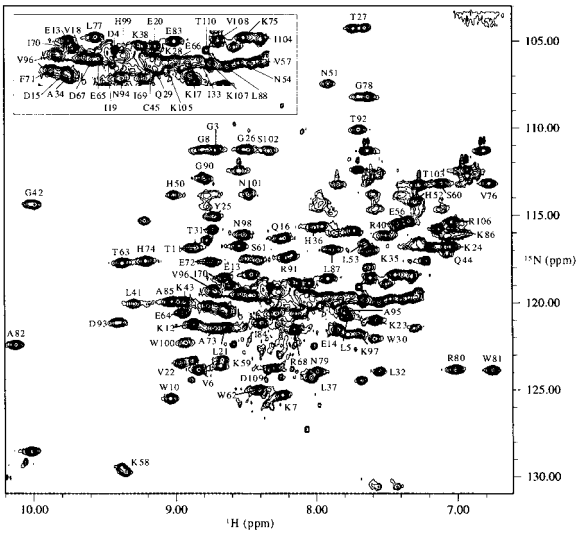

Figure 1. An $^{15}N/^{1}H$ HSQC spectrum recorded from a 1.0 mm sample of ^{15}N labelled B-MybR2R3 bound to a 16 base pair DNA fragment corresponding to the Myb binding site in the tom-1-A promoter. The assignments of the signals from backbone amide groups of the protein are indicated with the somewhat overlapped central region of the spectrum shown as an expanded area at the top left hand side of the figure.

All the NMR data were acquired at 20 °C on a 600 MHz Varian Inova spectrometer. The 2D and 3D spectra recorded to obtain sequence-specific resonance assignments for B-MybR2R3 bound to the tom-1-A target site were: $^{15}N/^{1}H$ HSQC, TOCSY-HSQC with a 10.3 kHz MLEV17-based mixing period of 45 ms, NOESY-HSQC using a NOE mixing time of 70 ms; $^{13}C/^{1}H$ HCCH-TOCSY with 7.4 kHz DIPSI-3-based mixing times of 14 and 22 ms; and $^{15}N/^{13}C/^{1}H$ HNCA, HN(CO)CA and CBCA(CO)NH (Grzesiek and Bax, 1992; Bax, 1994 and references therein). The majority of the 3D spectra were acquired over about 85 h but longer acquisition times of about 150 h were used for the TOCSY-HSQC and NOESY-HSQC experiments. Typical acquisitions times in F_1 and F_2 for the 3D spectra were 9–12 ms for ^{15}N, 6–10 ms for ^{13}C and 14–15 ms for ^{1}H and in F_3 were 40–65 ms. $^{15}N/^{1}H$ HSQC spectra were recorded over 2–4 h with acquisition times of 40–60 ms in F_1 and 150 ms in F_2. Decoupling of ^{15}N and ^{13}C from ^{1}H during the acquisition time of experiments was achieved using the GARP1 sequence at fields of about 1.0 and 3.3 kHz, respectively. In the triple resonance experiments, ^{1}H decoupling was carried out with a DIPSI-2 scheme at fields of about 5.3 kHz. The WATERGATE method was used to suppress the water signal when required.

The 3D NMR data were processed using NMRPipe (Delaglio et al., 1995) and analysed using XEASY (Bartels et al., 1995).

Extent of assignments and data deposition

The extensive assignment of resonances for B-MybR2R3 bound to the tom-1-A 16-mer presented a significant challenge due to the combined affects of limited temperature stability ($\leq$ 20 °C), size of the complex (22.6 kDa) and conformational heterogeneity leading to exchange broadening and doubling of resonances from the protein. Despite these difficulties it proved possible to obtain essentially complete sequence-specific backbone resonance assignments (^{15}N, ^{13}C and ^{1}H) for 103 of the 106 non-proline residues in B-MybR2R3 (97.2%), which are indicated for the backbone amide signals in the $^{15}N/^{1}H$ HSQC spectrum shown in Figure 1. The lack of ^{13}C labelling precluded the assignment of signals from the four proline residues in the protein and no assignments were obtained for R46, E47 and R48, which gave rise to no detectable signals in double or triple resonance spectra, presumably due to exchange broadening.

Complete assignments (^{13}C and ^{1}H) were also obtained for the non-exchangeable, aliphatic side chain signals of 76 residues in DNA-bound B-MybR2R3 and partial assignments made for a further 14 residues. In addition to the 4 proline residues, no side chain signals could be assigned for just 9 residues (L21, L32, Q44, R46, E47, R48, W49, L87 and L88), which predominantly fall in a region of R2 that appears to form part of the protein-DNA interface (C45-N51) and is affected by exchange broadening of the NMR signals. To summarise, essentially complete sequence-specific backbone and aliphatic side chain signal assignments were obtained for 83 of the protein's 110 residues (75.5%), complete backbone and partial side chain for a further 14 residues (12.7%) and backbone only for just 6 residues (5.5%). The ^{15}N, ^{13}C and ^{1}H resonance assignments made for B-MybR2R3 bound to the tom-1-A target site have been deposited in the BioMagResBank database (entry number 5517).

References

Bartels, C. et al. (1995) *J. Biomol NMR*, **5**, 1–10.
Bax, A. (1994) *Curr. Opin. Struct. Biol.*, **4**, 738–744.
Carr, M.D. et al. (1996) *Eur. J. Biochem.*, **235**, 721–735.
Delaglio, F. et al. (1995) *J. Biomol NMR*, **6**, 277–293.
Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **96**, 432–440.
Horstmann, S. et al. (2000) *Oncogene*, **19**, 298–306.
McIntosh, P.B. et al. (1998) *Biochemistry*, **37**, 9619–9629.
Saville, M.K. and Watson, R.J. (1998) *Adv. Cancer Res.*, **72**, 109–140.