Journal of Biomolecular NMR, **14:** 383–384, 1999.
KLUWER/ESCOM

Letter to the Editor: Resonance assignment and topology of a 22 kDa C-terminal fragment of the polypyrimidine tract binding protein (PTB) containing two RNA binding domains

Maria R. Conte[a], Tim Grüne[b], Stephen Curry[b,*] & Stephen Matthews[a,*]
[a]*Department of Biochemistry and* [b]*Biophysics Section, Blackett Laboratory, Centre for Structural Biology, Imperial College of Science, Technology and Medicine, Exhibition Road, London SW7 2AY, U.K.*

Received 12 May 1999; Accepted 12 June 1999

Key words: polypyrimidine tract binding protein, RNA binding, RNP, splicing regulation

Biological context

The polypyrimidine tract binding protein (PTB), also known as hnRNP-I, is a negative regulator of pre-mRNA splicing which can also stimulate the internal ribosomal entry site (IRES)-dependent translation initiation of most picornaviruses (reviewed in Valcárcel and Gebauer, 1997). Both of these functions require binding of PTB to RNA sequences, but the mechanisms of action are not well understood. PTB, a homodimer, binds to polypyrimidine tracts within the introns upstream and downstream of regulated exons, repressing their incorporation into mature mRNA. The protein also interacts with IRES elements, which are large RNA segments (∼450 bases) with extensive secondary structure found within the 5′-untranslated region of the picornaviral RNA genome (Valcárcel and Gebauer, 1997). PTB belongs to a family of RNA-binding proteins characterised by possession of at least one ribonucleoprotein consensus sequence (RNP-CS) RNA binding domain. RNP-CS domains are about 90 amino acids long and contain two short but quite strongly conserved motifs. PTB is predicted to have four RNP-CS domains, each of which contains unusual features in the sequences of the conserved motifs; the most important determinants for PTB RNA binding are in the two C-terminal RNP–CS domains (Kaminski et al., 1995; Pérez et al., 1997), whereas homodimer formation is stabilised by the second of the two N-terminal RNP-CS domains (Pérez et al., 1997). We have undertaken NMR-based structural studies of this fragment, in order to provide a clearer understanding of the mode of action of the protein in splicing and translation initiation and of the specificity determinants of its unusual RNP-CS domains (Allain et al., 1997). These studies would also provide information on the relative orientation and dynamics in solution of linked RNP-CS domains.

*To whom correspondence should be addressed. E-mail: s.j.matthews@ic.ac.uk

Methods and results

A 197-residue C-terminal fragment of human PTB (Gil et al., 1991), which contains the third and fourth RNP-CS domains (residues 335–531; designated PTB-C197) was sub-cloned by PCR into expression vector pET-15b (Novagen) using engineered NcoI and NdeI restriction sites. The protein was overexpressed in *E. coli* host strain BL21 (DE3) grown on minimal media containing 0.7 g/l ^{15}N-ammonium chloride and 2 g/l ^{13}C-glucose and purified by cation exchange (Poros HS50; Perseptive Biosystems) and affinity (Hitrap Blue; Pharmacia) chromatography. Following extensive dialysis against 20 mM sodium acetate buffer at pH 5.2, the sample was concentrated to 1 mM in 600 μl. PTB-C197 was adjudged to be strictly monomeric in solution by gel filtration and NMR relaxation.

NMR spectra were acquired at 302 K on a four channel Bruker DRX500 equipped with a z-shielded gradient and triple resonance probe. The sequence-specific ^{1}HN, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and $^{13}C'$ assignments were achieved using HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HN(CA)CO, HNCO experiments (for review see Bax, 1994). All the experiments use gradients for coherence selection, together with sensitivity

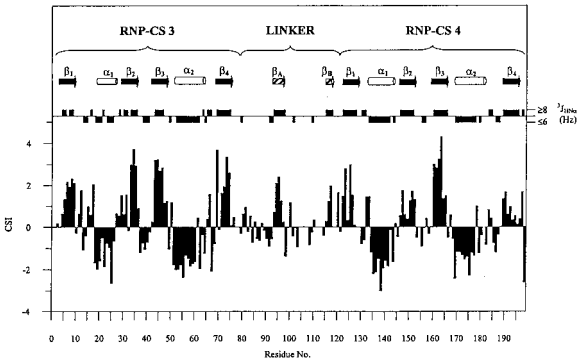

Figure 1. Plot of the chemical shift deviation from random coil and coupling constants. A consensus value for the chemical shift, derived from $[((4 \times \delta_{H\alpha}) + (\delta_{C\beta} - \delta_{C\alpha} - \delta_{C'}))/\text{No. assignments}]$, is used. Coupling constants of 8 Hz or above are indicated by a black bar above the line and values of 6 Hz or below by a black bar below the line. The secondary structure features are also indicated; putative β-strands in addition to the standard βαββαβ motifs are shaded with stripes.

enhancement. H_α and H_β assignments were obtained using HBHA(CBCACO)NH (Bax, 1994 and references cited therein), and confirmed by ^{15}N-edited NOESY-HSQC and HNHA (Kuboniwa et al., 1994) experiments. The latter experiment facilitated the estimation of $^3J_{HN\alpha}$ coupling constants.

All triple resonance experiments employed constant-time evolution in the ^{15}N dimension, whereas in the CBCA(CO)NH and HBHA(CBCACO)NH experiments both indirect dimensions were recorded in constant-time mode.

Extent of assignments, topology and data deposition

The NMR assignment of ^{1}HN, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$ nuclei of PTB-C197 is approximately 95% complete. The chemical shift data, in combination with measured $^3J_{HN\alpha}$ coupling constants and NOE patterns from ^{15}N-edited NOESY-HSQC spectra, were used to identify secondary structure elements (Figure 1). These data clearly show the presence of two domains connected by a linker region, and indicate, as predicted from sequence alignments (Ghetti et al., 1992), that they both possess the βαββαβ fold which is characteristic of RNP-CS RNA binding modules. Three of the largest gaps in the assignment occur within the linker (residues 77–79, 104–107 and 111–114). Also, some of the resonances in the loop between the second and third strands in both RNP-CS domains are very weak. Attenuation of these signals could be caused by internal mobility of these regions on a slow timescale. Interestingly, these regions are known to be involved in RNA recognition for other proteins of the same family (Allain et al., 1997 and references cited therein). In addition to the two βαββαβ folds of PTB-C197, further regions of β-sheet type secondary structure are also found within the linker region (Figure 1). This observation is particularly intriguing, as other structural studies on proteins containing multiple RNP domains have so far indicated that, in the absence of bound RNA, the linker region has no unique conformation (Handa et al., 1999 and references cited therein). Analysis of ^{1}H-^{15}N NOE data (not shown) for PTB-C197 indicates that for some of the residues in the linker and in loop regions for the RNP-CS the polypeptide chain is highly flexible. A table of the assignments (including ^{1}HN, $^1H_\alpha$, $^1H_\beta$, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$) is available as supplementary material and has been deposited in the BioMagResBank in Madison, WI, U.S.A. (accession code 4343).

Acknowledgements

The authors are indebted for the financial support of The Wellcome Trust. We are grateful to Professor P.A. Sharp (MIT, U.S.A) for the gift of human PTB cDNA. M.R.C would like to thank Dr. Geoff Kelly for advice and assistance. A list of NMR assignments is also available from S.M. (s.j.matthews@ic.ac.uk) and M.R.C. (s.conte@ic.ac.uk).

References

Allain, F.H.-T., Howe, P.W.A., Neuhaus, D. and Varani, G. (1997) *EMBO J.*, **16**, 5764–5774.
Bax, A. (1994) *Curr. Opin. Struct. Biol.*, **4**, 738–744.
Ghetti, A., Piñol-Roma, S., Michael, W.M., Morandi, C. and Dreyfuss, G. (1992) *Nucleic Acids Res.*, **20**, 3671–3678.
Gil, A., Sharp, P.A., Jamison, S.F. and Garcia Blanco, M.A. (1991) *Genes Dev.*, **5**, 1224–1236.
Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y. and Yokoyama, S., (1999) *Nature*, **398**, 579–585.
Kaminski, A., Hunt, S.L., Patton, J.G. and Jackson, R.J. (1995) *RNA*, **1**, 928–938.
Kuboniwa, H., Grzesiek, S., Delaglio, F. and Bax, A. (1994) *J. Biomol. NMR*, **4**, 871–878.
Pérez, I., McAfee, J.G. and Patton, J.G. (1997) *Biochemistry*, **36**, 11881–11890.
Shoelson, S.E., Pawson, T., Forman-Kay, J.D. and Kay, L.E. (1994) *Biochemistry*, **33**, 5984–6003.
Valcárcel, J. and Gebauer, F. (1997) *Curr. Biol.*, **7**, R705–R708.