Journal of Biomolecular NMR, **25:** 83–84, 2003.
KLUWER/ESCOM

Letter to the Editor: Resonance assignment and topology of a clostridial repetitive oligopeptide (CROP) region of toxin A from *Clostridium difficile*

Rappas Mathieu[a], Jenson Lim[b], Peter Simpson[a], Sunil Prasannan[a], Neil Fairweather[b] & Stephen Matthews[a,*]

[a]*Centre for Structural Biology and* [b]*Centre for Molecular Microbiology and Infection, Department of Biological Sciences, Imperial College of Science, Technology and Medicine, Exhibition Road, London SW7 2AY, U.K.*

Received 19 September 2002; Accepted 9 October 2002

Key words: Clostridium difficile, CROPs, NMR, toxin A

Biological Context

Clostridium difficile is a recently recognised bacterial pathogen, causing antibiotic associated diarrhoea and pseudomembraneous colitis in elderly patients. *C. difficile* produces two protein toxins, termed TcdA and TcdB (toxins A and B), each possessing glycosyltransferase activity that modifies small GTP-binding proteins within sensitive cells, reviewed in von Eichel-Streiber et al. (1996). TcdA and TcdB are both cytotoxic to cultured cells, and in addition TcdA has enterotoxic activity. Although the precise mechanism whereby these toxins cause symptoms of disease is not clear, they are absolutely required for the pathogenesis of *C. difficile*. These proteins are related both functionally and by amino acid sequence to the TcsL, TcsH and Tcnα toxins produced by *C. sordellii* and *C. novyi*. This group of toxins are distinct in many ways. They possess glycosyltransferase activity, an activity not found in other bacterial toxins, are unusually large (over 270–300 kDa) and contain multiple clostridial repetitive oligopeptides (CROPs) within their C-termini (von Eichel-Streiber et al., 1996).

CROPs account for 30% of the toxin sequence and contain a conserved stretch of three aromatic amino acids (YYF) (Wren, 1991; von Eichel-Streiber et al., 1992). In toxin A from *C. difficile* strain VPI10463 there are 7 class I repeats (30 a.a. in length) and 31 class II repeats (20–21 a.a in length) (Dove et al., 1990). Evidence exists indicating that TcdA binds to cellular receptors through the C-terminal CROP sequences. For example, cell binding and the cytopathic effects of toxin are inhibited by monoclonal antibodies that recognise the CROPs (Frey and Wilkins, 1992). The cellular receptor is believed to be a glycoprotein that contains modifications similar to the human Lewis antigens X, Y and Z (Tucker and Wilkins, 1991). Furthermore, a 14 CROP repeat fragment derived from the C-terminus of the protein (TAC14) has also been shown to retain receptor binding activity (Ward et al., 1997). We have embarked of a structural analysis of the C-terminal receptor binding fragment from *Clostridium difficile* toxin A. We present comprehensive NMR assignments for the five C-terminal CROPs, termed TAC5.

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

Methods

A 116-residue C-terminal fragment of toxin A (residues 2591-2706) that contains 5 CROPs was subcloned by PCR from TAC14 (14 CROP repeat fragment) into pET28a at NcoI (utilising the ATG as its start codon) and XhoI sites. The histidine tagged protein was over-expressed in *Escherichia coli* strain BL21 grown on minimal media containing 0.7 g/l ^{15}N-ammonium chloride and 2 g/l ^{13}C-glucose and was purified on TALON beads. For NMR analysis the purified TAC5 protein was concentrated to ca. 0.5 mM in 300 μl of 10 mM, pH 5.5 sodium acetate buffer. NMR spectra were acquired at 313K on a four channel Bruker DRX500 equipped with a triple resonance, gradient cryoprobe. The sequence specific ^{1}HN, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and $^{13}C'$ assignments were achieved using HNCACB, CBCA(CO)NH, HN(CA)CO, HNCO, HNCA and HN(CO)CA experiments (for review see Bax, 1994). All the experiments use gradients for

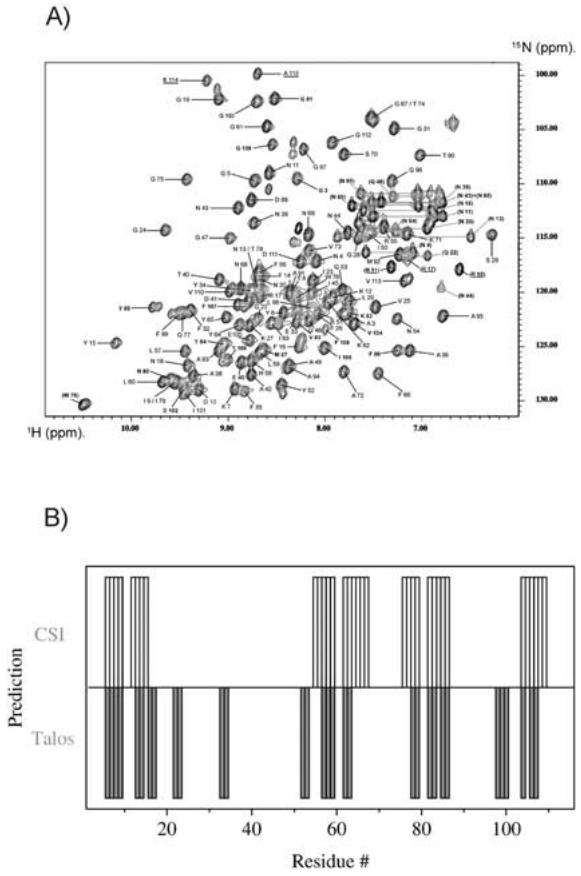

Figure 1. (A) Assigned ^{1}H-^{15}N HSQC spectrum of TAC5. (B) CSI consensus plot (Wishart and Sykes, 1994) of TACS generated using $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{13}C'$ chemical shifts. β-sheet secondary structure elements are exclusively observed and therefore shown. Also, residues involved in β-sheet as predicted from backbone torsion angle prediction package TALOS (Cornilescu et al., 1999) are indicated as bars.

coherence selection, together with sensitivity enhancement. H_{α} and H_{β} assignments were obtained using HBHA(CBCACO)NH (Bax, 1994). 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 and topology

Sequence-specific assignments were made primarily *via* the HNCACB/CBCA(CO)NH pair, using the carbonyl experiments to resolve any ambiguities arising from degeneracy in the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ frequencies. Using this strategy, backbone ^{1}HN, ^{15}N, H_{α}, H_{β}, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and $^{13}C'$ assignments could be made for the over 98% of residues TAC5, with the only omissions being G109 and M87. The side chain assignments were achieved using HCCH-total correlation (TOCSY) spectroscopy (Bax, 1990) and was assessed to be essential complete with the exception of overlapping resonances in regions containing the aromatic, asparagine and glutamine residues. The chemical shift data were used to identify secondary structure elements (Figure 1), which indicate exclusively β structure and the presence of periodicity, i.e., two strands per class II repeat. Furthermore, despite little sequence homology this β structure is reminiscent of the β-solenoid fold adopted for choline binding repeats from the pneumococcal virulence factor LytA (Fernandez-Tornero, 2001). A table of the assignments is available as supplementary material and has been deposited in the BioMagResBank in Madison, WI (accession code 5524).

Acknowledgements

The authors are indebted for the financial support of BBSRC and The Wellcome Trust. A list of NMR assignments is also available from S.M. (s.j.matthews@ic.ac.uk).

References

Bax, A., Clore, G.M. and Gronenborn, A.M.J. (1990) *J. Magn. Reson.*, **88**, 425–431.
Cornilescu, G., Delaglio, F. and Bax, A. (1999) *J. Biomol. NMR*, **13**, 289–302.
Dove, C.H., Wang, S.Z., Price, S.B., Phelps, C.J., Lyerly, D.M., Wilkins, T.D. and Johnson, J.L. (1990) *Infect. Imm.*, **58**, 480–488.
Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G. and Romero, A. (2001) *Nat. Struct. Biol.*, **8**, 1020–1024.
Frey, S.M. and Wilkins, T.D. (1992) *Infect. Immun.*, **60**, 2488–2492.
Tucker, K.D. and Wilkins, T.D. (1991) *Infect. Immun.*, **59**, 73–78.
von Eichel-Streiber, C., Boquet, P., Sauerborn, M. and Thelestam, M. (1996) *Trends Microbiol.*, **4**, 375–382.
von Eichel-Streiber, C., Laufenberg-Feldmann, R., Sartingen, S., Schulze, J. and Sauerborn, M. (1992) *Mol. Gen. Genet.*, **233**, 260–268.
Ward, S.J., Douce, G., Dougan, G. and Wren, B.W. (1997) *Rev. Med. Microbiol.*, **8**, S34–S36.
Wishart, D.S. and Sykes, B.D. (1994) *J. Biomol. NMR*, **4**, 171–180.
Wren, B.W. (1991) *Mol. Microbiol.*, **5**, 797–803.