Journal of Biomolecular NMR **29:** 549–550, 2004.

Letter to the Editor: Resonance assignments for the 18 kDa protein CC1736 from *Caulobacter crescentus*

Yang Shen[a], Hanudatta S. Atreya[a], Rong Xiao[b], Thomas B. Acton[b], Ritu Shastry[b], LiChung Ma[b], Gaetano T. Montelione[b,c] & Thomas Szyperski[a,*]

[a]*Department of Chemistry and Structural Biology, The State University of New York, Buffalo, NY 14260, U.S.A. and Northeast Structural Genomics Consortium;* [b]*Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ 08854, U.S.A. and Northeast Structural Genomics Consortium;* [c]*Department of Biochemistry, Robert Wood Johnson Medical School, UMDNJ, Piscataway, NJ 08854, U.S.A.*

Received 11 March 2004; Accepted 15 April 2004

Key words: alpha-beta protein, Northeast Structural Genomics Consortium, reduced-dimensionality NMR

Biological context

Structural genomics aims at making three-dimensional (3D) structural information readily available for each protein domain family in nature. In the Northeast Structural Genomics Consortium (NESGC; www.nesg.org), our efforts to solve three-dimensional protein structures focus on protein families encoded in eukaryotic genomes and their homologues in bacterial and archaeabacterial 'reagent' genomes (Wunderlich et al., 2004). Among these reagent organisms is *Caulobacter crescentus*, which is a proteobacterium known to differentiate and divide asymmetrically in each cell cycle (Laub et al., 2000). Owing to the small size of the genome containing only about 4000 genes, this organism has been selected as a model system to study cellular differentiation and cell cycle regulation (Laub et al., 2000). The recent sequencing of the genome of *C. crescentus* has fostered novel insights into its comparably complex cell cycle (Nierman et al., 2001). Here, we report nearly complete ^{1}H, ^{13}C and ^{15}N resonance assignments of a 148-residue protein of unknown function from *C. crescentus* that is encoded in gene CC1736 (NCBI ID: 155892; Swiss-Prot ID: Q9A7I7; NESGC target ID: CcR19; NESGC Rost-cluster ID: 17538).

Methods and experiments

Uniformly ^{13}C,^{15}N-enriched CC1736 was expressed and purified following standard protocols used in our structural genomics consortium. Briefly, the full-length CC1736 gene from *Caulobacter crescentus* was cloned into a pET21d (Novagen) derivative, yielding the plasmid pCcR19-21.1. The resulting open reading frame contains eight non-native residues at the C-terminus (LEHHHHHH) of the protein. *E. coli* BL21 (DE3) pMGK cells, a rare codon enhanced strain, were transformed with pCcR19-21.1, and cultured in MJ9 minimal medium (Jansson et al., 1996) containing $(^{15}NH_4)_2SO_4$ and *U*-^{13}C-glucose as the sole nitrogen and carbon sources. Initial growth was carried out at 37 °C until the optical density of the culture medium ($O.D_{600}$) reached 0.7 units. The incubation temperature was then decreased to 17 °C and protein expression was induced by the addition of IPTG (isopropyl-β-D-thiogalactopyranoside) at a final concentration of 1 mM. Following overnight incubation at 17 °C, cells were harvested by centrifugation and lysed by sonication. *U*-^{13}C, ^{15}N CC1736 was purified in a two step protocol consisting of Ni-NTA affinity column (Qiagen) and gel filtration column (HiLoad 26/60 Sephadex 75, Amersham Pharmacia Biotech) chromatography. The final yield of the protein (>97% homogeneity by SDS-PAGE; 17.7 kDa by MALDI-TOF mass spectrometry) was ∼10 mg/L.

U-^{13}C,^{15}N CC1736 samples for NMR spectroscopy were prepared at a concentration of 1.1 mM in 95% H_2O/5% D_2O solution (20 mM MES, 100 mM NaCl, 5 mM $CaCl_2$, 0.02% NaN3, 10 mM DTT, pH 6.5) and placed in 5-mm Shigemi tubes. NMR spectra were collected at 25 °C on a Varian INOVA

*To whom correspondence should be addressed. E-mail: szypersk@nsm.buffalo.edu

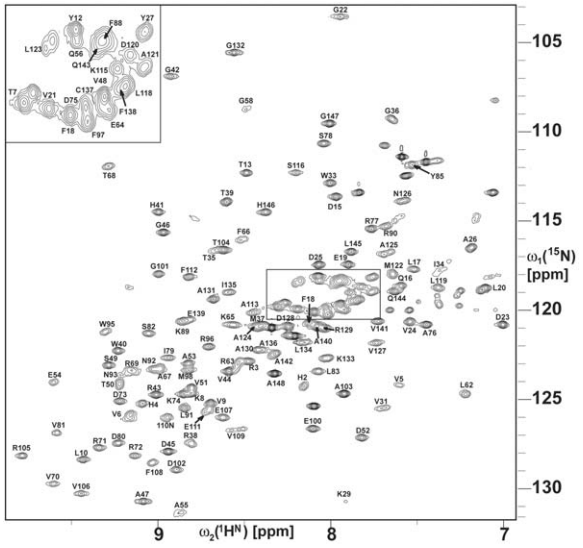

Figure 1. 2D [^{1}H,^{15}N]-HSQC spectrum of CC1736 from *C. crescentus*. Peaks are labeled with their respective sequential resonance assignments using the one-letter code of amino acids and the amino acid sequence number. For clarity, the assignments of the crowded central region of the spectrum are shown in an insert at the upper left corner.

spectrometer operating at 750 MHz ^{1}H resonance frequency, processed with the program PROSA (Güntert et al., 1992) and analyzed using the program XEASY (Bartels et al., 1995). Sequence specific backbone (H^N, H^α, N, C′, C^α) and H^β/C^β resonance assignments were achieved by using the reduced dimensionality (RD) experiments 3D $\underline{H}^{\alpha/\beta}\underline{C}^{\alpha/\beta}$(CO)NHN and HNN<u>CAHA</u> (Szyperski et al., 1998; 2002) in conjunction with conventional 3D HNNCACB and 3D HNN(CO)CA (Cavanagh et al., 1996). Side chain ^{1}H and ^{13}C resonance assignments were obtained using RD 3D <u>HC</u>CH-COSY (Szyperski et al., 2002). 3D ^{13}C- and ^{15}N-resolved [^{1}H-^{1}H] NOESY spectra (Cavanagh et al., 1996) were recorded to support (i) the sequential resonance assignment and (ii) the identification of regular secondary structure elements by observation of ^{1}H–^{1}H nuclear Overhauser enhancements (NOEs).

Extent of assignments and data deposition

Nearly complete sequential resonance assignment were obtained for CC1736, that is, 95.9% of the backbone shifts, excluding the N-terminal NH_3^+, the Pro ^{15}N and the ^{13}C′ shifts of residues preceding the Pro residues, and $^{13}C^\beta$ resonances, and 94.7% of the side chain chemical shifts, excluding the Lys NH_3+, the Arg NH_2, the OH, the side chain ^{13}C′ and the aromatic quaternary ^{13}C shifts, were obtained. The assigned 2D [^{15}N, ^{1}H] HSQC spectrum is shown in Figure 1. Chemical shifts are deposited in the BioMagResBank (accession number 6120). Based on these chemical shifts, and supported by observation of sequential and medium-range NOEs, the regular secondary structure elements of CC1736 were identified. Seven β-strands (residues 3–10, 34–44, 47–56, 62–71, 76–82, 89–98, 101–111) and two α-helices (residues 15–21, 116–144) were found, demonstrating that CC1736 belongs to the class of α/β proteins. The ongoing 3D structure determination will provide the first 3D structural information for a protein domain family of currently unknown function. Moreover, the CC1736 structure will be among the first to be made available for *C. crescentus*.

Acknowledgements

This work was supported by the Protein Structure Initiative of the *National Institutes of Health* (P50 GM62413) and the *National Science Foundation* (MCB 00075773 to T.S.).

References

Bartels, C., Xia, T., Güntert, P., Billeter, M. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy*, Wiley, New York, NY.

Güntert, P., Dötsch, V., Wider, G. and Wüthrich K. (1992). *J. Biomol. NMR*, **2**, 619–629.

Jansson, M., Li, Y.-C., Jendeberg, L., Anderson, S., Montelione, G. T. and Nilsson, B. (1996) *J. Biomol. NMR*, **7**, 131–141.

Laub, M.T., McAdams, H., Fledblyum, T., Fraser, C. and Shapiro, L. (2000) *Science*, **290**, 2144–2148.

Nierman, W.C. et al. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 4136–4141.

Szyperski, T., Banecki, B. Braun, D. and Glaser, R. (1998) *J. Biomol. NMR*, **11**, 387–405.

Szyperski, T., Yeh, D.C., Sukumaran, D.K., Moseley, H.N. and Montelione, G.T. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 8009–8014.

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

Wunderlich, Z., Acton, T.B., Liu, J., Kornhaber, G., Everett, J., Carter, P., Lan, N., Echols, N., Gerstein, M., Rost, B. and Montelione, G.T. (2004) *Protein Sci.*, in press.