Journal of Biomolecular NMR **28:** 91–92, 2004.
KLUWER/ESCOM

Letter to the Editor: Backbone ^{1}H, ^{15}N and ^{13}C assignments for the 21 kDa *Caenorhabditis elegans* homologue of 'brain-specific' protein

Daniel Monleón[a], YiWen Chiang[a], James M. Aramini[a], G.V.T. Swapna[a], Daphne Macapagal[a], Kristin C. Gunsalus[a], Seho Kim[b], Thomas Szyperski[b] & Gaetano T. Montelione[a,c,*]

[a]*Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry and Northeast Structural Genomics Consortium, Rutgers University, Piscataway, NJ 08854, U.S.A.;* [b]*Department of Chemistry and Northeast Structural Genomics Consortium, University at Buffalo, The State University of New York, 816 Natural Science Complex, Buffalo, NY 14260, U.S.A.;* [c]*Department of Biochemistry, Robert Wood Johnson Medical School, Piscataway, NJ 08854, U.S.A.*

Received 4 June 2003; Accepted 9 June 2003

Key words: AutoAssign, reduced-dimensionality triple resonance NMR, residual dipolar coupling for assignments, structural proteomics

Biological context

The Northeast Structural Genomics Consortium (NESG) is a pilot project designed to evaluate the feasibility and value of structural genomics. The 21 kDa *Caenorhabditis elegans* protein coded by gene CE32E8.3 (TrEMBL protein P91127, referred to here as WR33) is one of several hundred targets identified for structural analysis by the Northeast Structural Genomics Consortium (www.nesg.org). WR33 belongs to a large protein domain family with homologues in several eukaryotic genomes, including those of *Homo sapiens* (TrEMBL proteins Q9Y326, O94811, and Q9Y6H0), *Mus musculis* (TrEMBL protein Q9CRB6), and *Drosophila melanogaster* (TrEMBL protein Q9VV43). The 25 kDa *Bos taurus* (bovine) homologue from this family (Q27957), with 38% sequence identity with WR33 over 175 residues, is characterized as 'brain specific protein P25' (Shiratsuchi et al., 1995) and is expressed in oligodendrocytes and neutrophils of bovine brain tissue. However, none of the members of this strongly conserved protein domain family has a characterized biological function. As a step toward characterizing its biochemical and biophysical functions, we have overexpressed and purified WR33, and determined nearly complete ^{15}N, $^{1}H^{N}$, $^{13}C'$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and $^{1}H^{\alpha}$ resonance assignments. These assignments, together with ^{1}H-^{15}N heteronuclear NOE data, show that ∼1/3 of the residues in WR33 are conformationally disordered. Thus, the sidechain resonance assignments and structure determination for full length WR33 is not going to be pursued. However, the backbone resonance assignments for WR33 described here can be used for designing smaller constructs more suitable for high-resolution structure analysis, and are valuable for future studies of potential protein-protein interactions.

*To whom correspondence should be addressed. E-mail: guy@cabm.rutgers.edu

Methods and experiments

The gene coding for the WR33 protein was subcloned from cDNA clone YK37508 into expression vector pET15b with a hexa-His N-terminal purification tag, generating plasmid pET15b-WR33. *E. coli* strain BL21(DE3) cells transformed with pET15b-WR33 were grown at 37 °C in minimal medium containing $(^{15}NH_4)_2SO_4$ and $^{13}C_6$-glucose as sole nitrogen and carbon sources. Purification was carried out using standard Ni-NTA affinity chromatography and gel filtration methods. Sample purity (>95%) and molecular weight (20.8 kDa) were verified by SDS-PAGE and MALDI-TOF mass spectrometry.

Uniformly ^{13}C, ^{15}N-enriched WR33 was prepared at ∼1.4 mM protein concentration in H_2O solution containing 5% D_2O, 20 mM sodium phosphate, 50 mM sodium chloride and 5 mM DTT at pH 6.50 ± 0.05. A second ^{13}C, ^{15}N-enriched WR33 sample was prepared for residual dipolar coupling measurements at 0.55 mM concentration in the same pH 6.5 buffer containing 4.5% (w/v) DMPC/DHPC/CTAB bicelle (30:10:1) (Losonczi and Prestegard, 1998).

All spectra were acquired at 25 °C unless otherwise indicated. NMR experiments were carried out on Varian INOVA 600 and 750 NMR spectrometers. Data were processed, peak-picked and analyzed with the programs NMRPipe (Delaglio et al., 1995), SPARKY (Goddard and Kneller, Univ. of Calif., San

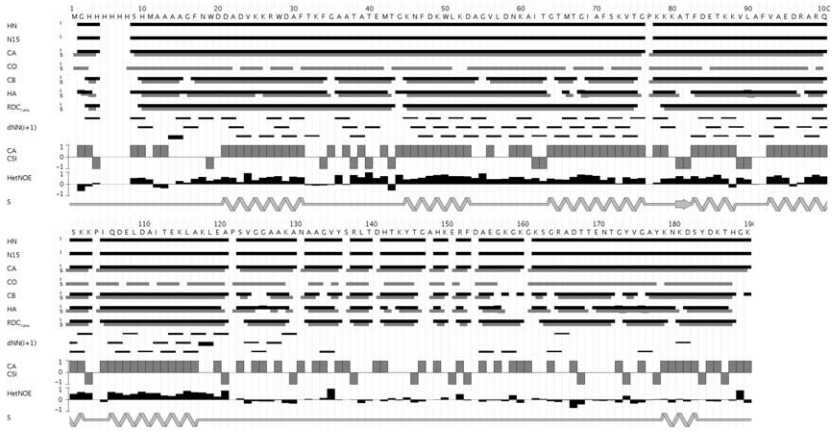

Figure 1. AutoAssign connectivity map showing sequential (s) and intra-residue (i) connectivities used to establish backbone resonance assignments, together with a summary of ^{1}H-^{15}N heteronuclear NOE (HetNOE), H^{N}-H^{N}_{i+1} connectivity ($dNN_{i,i+1}$), C^{α} chemical shift index (CA CSI), and secondary structure (S) data for WR33. All intra-residue (i) and sequential (s) H^{α} connectivities and many C^{α} connectivities were established from reduced dimensionality 3D HNCAHA and 3D HACAcoNH spectra. RDC_{CaHa} denotes sequential connectivities established by comparisons of H^{α}- C^{α} residual dipolar couplings detected on intramolecular and sequential cross peaks of ω_1(^{13}C-^{1}H)-decoupled haCANH spectra (Zweckstetter and Bax, 2001).

Francisco), and AutoAssign (Moseley et al., 2001). Chemical shifts were referenced to internal DSS. The input for AutoAssign included peak lists from 2D ^{1}H-^{15}N HSQC, 3D HNCACB, 3D CBCAcoNH and 3D HNCO spectra recorded essentially as described elsewhere (Montelione et al., 1999). Around 75% of backbone assignments were determined by automatic analysis of manually edited peak lists. The resulting assignments were extended by combined analysis of reduced-dimensionality 3D HACAcoNH and HNCAHA spectra (Szyperski et al., 2002). Additional sequential connectivities were established using ω_1(^{13}C-^{1}H)-decoupled haCANH spectra (Zweckstetter and Bax, 2001) recorded at 32 °C for the partially aligned sample. An annotated [^{15}N-H^{N}]-HSQC spectrum is available as supplementary material.

Extent of assignments and data deposition

The combined use of automated and manual analysis of triple resonance 3D data, along with the 4D information encoded in the reduced dimensionality data and the additional dimension of matching encoded in the residual C^{α}-H^{α} dipolar couplings, provided assignments for ∼92% of assignable backbone atoms (i.e., 175/186 ^{15}N-$^{1}H^{N}$ sites, 185/190 C^{α}, 171/190 C', 164/171 C^{β}, and 169/209 H^{α}) of WR33. ^{15}N and H^{N} assignments for some amides in the segment of residues 125 to 180 could not be determined unambiguously as they exhibit severe degeneracy. A summary of sequential connectivity data used to determine these assignments is presented in Figure 1. These ^{1}H, ^{13}C and ^{15}N chemical shift data have been deposited in BioMagResBank database (accession number 5300). Based on these chemical shifts and additional ^{1}H-^{15}N heteronuclear NOE data, recorded as described by Li and Montelione (1993), residues 20–120 have largely alpha-helical structure, forming a six-helical bundle, while residues 125 to 180 adopt non-regular and/or disordered structures (Figure 1). Interestingly, this largely disordered C-terminal segment is the most strongly conserved region of the protein sequence, suggesting that it may play a role in interactions with other biomolecules.

Acknowledgements

We thank Prof Y. Kohara for providing cDNA clone YK375d8. This work was supported by grants from the National Institutes of Health (P50 GM62413) and the National Science Foundation (MCB 0075773 to T.S.).

References

Delaglio, F. et al. (1995) *J. Biomol. NMR*, **6**, 277–293.
Li, Y.C. and Montelione, G.T. (1993) *J. Magn. Reson. Ser.*, **B101**, 315–319.
Losonczi, J.A. and Prestegard, J.H. (1998) *J. Biomol. NMR*, **12**, 447–451.
Montelione, G.T. et al. (1999) Berliner and N.R. Krishna, Eds., *Modern Techniques in NMR*, pp. 81–130.
Moseley, H.N.B. et al. (2001) *Meth. Enzymol.*, **339**, 91–108.
Shiratsuchi, A. et al. (1995) *Biochim. Biophys. Acta*, **1215**, 66–68.
Szyperski, T. et al. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 8009–8014.
Zweckstetter, M. and Bax, A. (2001) *J. Am. Chem. Soc.*, **123**, 9490–9491.