Journal of Biomolecular NMR **29:** 453–454, 2004.

Letter to the Editor: ^{1}H-, ^{13}C- and ^{15}N-NMR assignment of the conserved hypothetical protein TM0487 from *Thermotoga maritima*

Marcius S. Almeida*, Wolfgang Peti & Kurt Wüthrich
The Scripps Research Institute, Department of Molecular Biology and Joint Center for Structural Genomics, 10550 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

Received 7 January 2004; Accepted 16 January 2004

Key words: conserved hypothetical protein, NMR chemical shift assignment, structural proteomics, *Thermotoga maritima*

Biological context

As a part of the Joint Center for Structural Genomics, our laboratory investigates proteins from the hyperthermophilic bacterium *Thermotoga maritima* (Lesley et al., 2002). TM0487, a 103-residue conserved hypothetical protein, could readily be crystallized but diffracted poorly. Nevertheless, as seen by 1D ^{1}H NMR screening, TM0487 is a well folded protein suitable for 3D structure determination using NMR spectroscopy. With regard to the biological relevancy of a structure determination of TM0487, the lowest fold and function assignment score (Rychlewski et al., 2000) with the splicing factor U2Af subunit (PDBid: 1jmt; 15% sequence similarity) is only −6.41, making TM0487 an interesting NMR structural proteomics target.

Methods and experiments

The plasmid pET-25b(+) containing the TM0487 cDNA was transformed into *E. coli* strain BL21-CodonPlus (DE3)-IRL (Stratagene). Expression of uniformly ^{15}N-labeled and ^{13}C/^{15}N-labeled TM0487 was carried out by growing cells in M9 minimal medium containing either 1 g/L of $^{15}NH_4Cl$ and 4 g of D-glucose, or 1 g/L of $^{15}NH_4Cl$ and 4 g/L of [$^{13}C_6$]-D-glucose as the sole nitrogen and carbon sources. Cell cultures were grown at 37 °C and induced with 1 mM IPTG for 4 hours. The cells were harvested by centrifugation, resuspended in extraction buffer (50 mM Tris-HCl at pH 8.0, 1 mM EDTA, 0.1 mg/ml DNase I, 0.1% Triton X-100, Complete Tabs. (Roche)) and lysed by sonication. After centrifugation, the fraction containing the soluble proteins was loaded onto a Hitrap Q FF column (Pharmacia) equilibrated with 50 mM Tris-HCl at pH 7.0. The bound material was eluted with a linear gradient of NaCl (0–412.5 mM) at 20.6 mM per min. Fractions containing the protein, as checked by SDS polyacrylamide gel electrophoresis, were pooled and heated at 85 °C for 15 min with constant shaking (750 rpm), whereupon the precipitated protein was removed by centrifugation. The resulting solution was concentrated by ultrafiltration, incubated for 1 hour at 37 °C with 30 mM DTT, and loaded onto a Superdex 75 column (Pharmacia) equilibrated with 20 mM sodium phosphate buffer (pH 6.0). Fractions containing the protein, as checked by SDS polyacrylamide gel electrophoresis, were pooled and further analyzed by MALDI-TOF mass spectrometry. This sample was concentrated by ultrafiltration, and supplemented with 5 mM [$^{2}H_{10}$]-DL-DTT, 10% (v/v) D_2O and 2 mM NaN_3. The final concentration of TM0487 was between 2.3 to 4.4 mM for the different samples. NMR measurements were performed at 40 °C on Bruker Avance600 and Avance900 spectrometers, using TXI-HCN z- or xyz-gradient probes. Proton chemical shifts are referenced to internal 3-(trimethyl-silyl)-1-propanesulfonic acid, sodium salt (DSS). Using the absolute frequency ratios, the ^{13}C and ^{15}N chemical shifts were referenced indirectly to DSS. 2D [^{1}H,^{15}N]-HSQC, 3D HNCACB, 3D CBCA (CO)NH and 3D HNCO spectra (Bax and Grzesiek, 1993) were used to obtain sequence-specific assignments for the polypeptide backbone. The sequence-specific assignment was completed with the assign-

*To whom correspondence should be addressed. E-mail: marcius@scripps.edu

ment of P86, which is followed by P87, through $d_{\alpha\beta}$ and $d_{\beta\delta}$ sequential NOEs identified in the 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum. ^{1}H and ^{13}C assignments of all side chains were obtained using 3D HBHA(CO)NH, 3D ^{15}N-resolved [^{1}H,^{1}H]-TOCSY and 3D HC(C)H-TOCSY spectra. The spin systems of the aromatic rings were identified using the 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum. The resonances of the side chain εNH groups of the arginines, and the NH_2 groups of the asparagines and glutamines were identified in the 2D [^{15}N,^{1}H]-HSQC spectrum, and sequence-specific assignments were then obtained using the 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY spectrum. The εCH_3 groups of the methionines were initially identified in the 2D [^{13}C,^{1}H]-HSQC spectrum, and sequence-specifically assigned using the 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum. The NMR spectra were processed with PROSA (Güntert et al., 1992) or XWINNMR3.5 (Bruker, Billerica, U.S.A.), and analyzed with XEASY (Bartels et al., 1995).

Extent of the assignments and data deposition

Sequence-specific assignments were obtained for all spin systems of TM0487 (Figure 1). Unassigned backbone resonances include the ^{15}N chemical shifts of all the prolines and of M2, ^{13}CO of the residues T52, C55, D85, T89 and S94, which precede the five prolines, and of the chain-terminal residues P1 and V103, the N-terminal amino protons, and H^N of M2. Assignments were obtained for all the non-labile hydrogen atoms. Only the εCH_3 chemical shifts of 4 out of the 6 methionines could not be unambiguously assigned. Furthermore, the chemical shifts of all side chain labile protons of asparagine, glutamine and tryptophan, and the Hε resonances of arginine were assigned. The ^{1}H, ^{13}C and ^{15}N chemical shifts have been deposited in the BioMagResBank (http://www.bmrm.wisc.edu) under the BMRB accession number 5976.

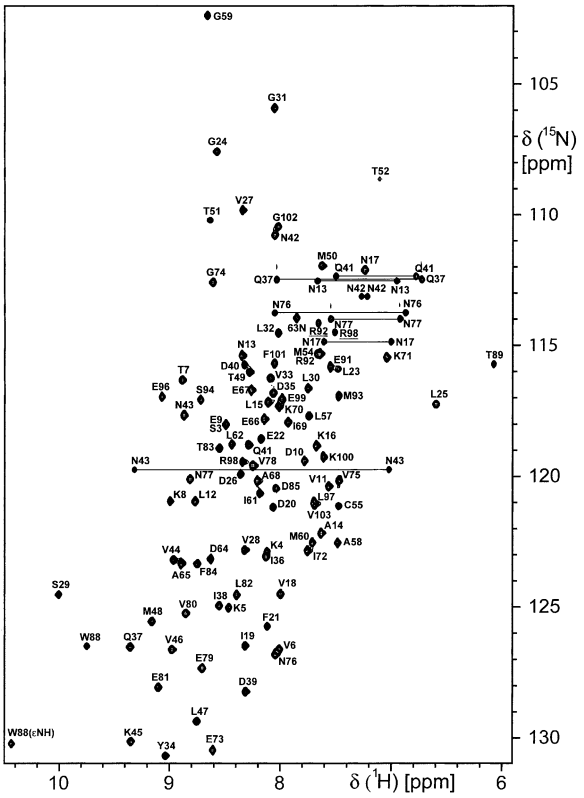

Figure 1. 2D [^{1}H,^{15}N]-HSQC spectrum of uniformly ^{15}N-labelled TM0487 (2.9 mM), collected at a ^{1}H frequency of 600 MHz. All peaks are annotated with the one-letter amino acid symbol and the sequence position. The two underlined labels correspond to folded side chain εNH cross peaks of R. The cross peaks connected by horizontal lines correspond to side chain NH_2 groups of Q and N.

Acknowledgments

We thank Drs S.A. Lesley and H.E. Klock from the JCSG consortium funded by NIGMS GM062411 for providing us with the genomic materials. M.S.A. is supported by the Pew Latin American Fellows Program in the Biological Sciences; W.P. is supported by an E. Schrödinger Fellowship (J2145); K.W. is the Cecil H. and Ida M. Green Visiting Professor of Structural Biology at TSRI.

References

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

Bax, A. and Grzesiek, S. (1993) *Accounts Chem. Res.*, **26**, 131–138.

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

Lesley, S.A., Kuhn, P., Godzik, A., Deacon, A.M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H.E., McMullan, D., Shin, T., Vincent, J., Robb, A., Brinen, L.S., Miller, M.D., McPhillips, T.M., Miller, M.A., Scheibe, D., Canaves, J.M., Guda, C., Jaroszewski, L., Selby, T.L., Elsliger, M.A., Wooley, J., Taylor, S.S., Hodgson, K.O., Wilson, I.A., Schultz, P.G. and Stevens, R.C. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 11664–11669.

Rychlewski, L., Jaroszewski, L., Li, W. and Godzik, A. (2000) *Protein Sci.*, **9**, 232–241.