Journal of Biomolecular NMR, **25:** 167–168, 2003.
KLUWER/ESCOM

Letter to the Editor: NMR assignment of the conserved hypothetical protein TM1290 of *Thermotoga maritima*

Touraj Etezady-Esfarjani[a,b], Wolfgang Peti[b] & Kurt Wüthrich[a,b,*]
[a]*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Zürich, CH-8093 Zürich, Switzerland;* [b]*The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, U.S.A.*

Received 29 October 2002; Accepted 6 November 2002

Key words: hypothetical proteins, NMR assignment, structural genomics, *Thermotoga maritima*

Biological context

Thermotoga maritima is a hyperthermophilic bacterium that was originally isolated in geothermally heated marine sediments in Italy. The genome of this organism has been completely sequenced and consists of 1877 coding regions, of which 863 (46%) do not yet have functional assignments (Nelson et al., 1999). Comparative studies revealed that 24% of the *T. maritima* genes are more similar to archaeal genes than to other bacterial genes, indicating that considerable lateral gene transfer occurred between *T. maritima* and archaea (Nelson et al., 1999). In fact, *T. maritima* is believed to be one of the oldest and most slowly evolving eubacterial lineages (Achenbach-Richter et al., 1987).

In association with The Joint Center for Structural Genomics (JCSG), our laboratory has started investigations of the proteome of *T. maritima* (Lesley et al., 2002). We use a NMR-based high-throughput screening method (to be described elsewhere) for targeting proteins of interest, and TM1290, a 115-residue conserved hypothetical protein, is the first protein that was chosen for a complete structure determination.

Methods and experiments

The plasmid pET-25b(+) encoding the TM1290 gene was transformed into *E. coli* strain BL21-CodonPlus (DE3)-IRL (Stratagene). Expression of uniformly $^{13}C/^{15}N$-labeled TM1290 was carried out by growing the cells in M9 minimal medium containing $^{15}NH_4Cl$ (1 g l^{-1}) and [$^{13}C_6$]-D-glucose (4 g l^{-1}) as the sole nitrogen and carbon sources, respectively. Cell cultures were grown at 37 °C with vigorous shaking to an OD_{600} of 0.6 before expression of the protein was induced with IPTG (final concentration 1 mM). The cells were grown for an additional 4 h before harvesting. The cell pellet was resuspended in extraction buffer (50 mM Tris-HCl at pH 8.0, 1 mM EDTA, 0.1 mg ml^{-1} DNase I) containing protease inhibitors (Complete, Roche), lysed by sonication, and the cell debris were removed by centrifugation. In the first purification step, ammonium sulphate (40% w/v) was gradually added to the soluble fraction at 4 °C for 1 h, which was then centrifuged to remove the precipitate. The soluble fraction was first desalted by a HiPrep 26/10 Desalting column (Pharmacia), then loaded onto a Q-Sepharose column (Pharmacia) equilibrated with Buffer A (50 mM Tris-HCl at pH 7.2, 1 mM EDTA), and then eluted with a 0–500 mM NaCl gradient. Fractions containing TM1290 were pooled and concentrated to a final volume of 2 ml and heated at 75 °C for 15 min, whereupon the precipitated protein was removed by centrifugation. The homogeneity of the purified protein was evaluated by SDS-PAGE electrophoresis, and the solution was concentrated to a final volume of 500 µl.

In addition to the unlabeled protein, two NMR samples were used, containing, respectively, 2.3 mM uniformly ^{15}N-labeled or 3.8 mM uniformly $^{13}C/^{15}N$-labeled TM1290 in 600 µl 95% H_2O/5% D_2O containing 20 mM sodium phosphate at pH 6.0 and 2 mM NaN_3. NMR measurements were performed at 313 K on Bruker Avance600 and Avance900 spectrometers using TXI-HCN-xyz gradient probes. Proton chemical shifts are referenced to internal 3-(trimethyl-silyl)-1-

*To whom correspondence should be addressed. E-mail: kw@mol.biol.ethz.ch

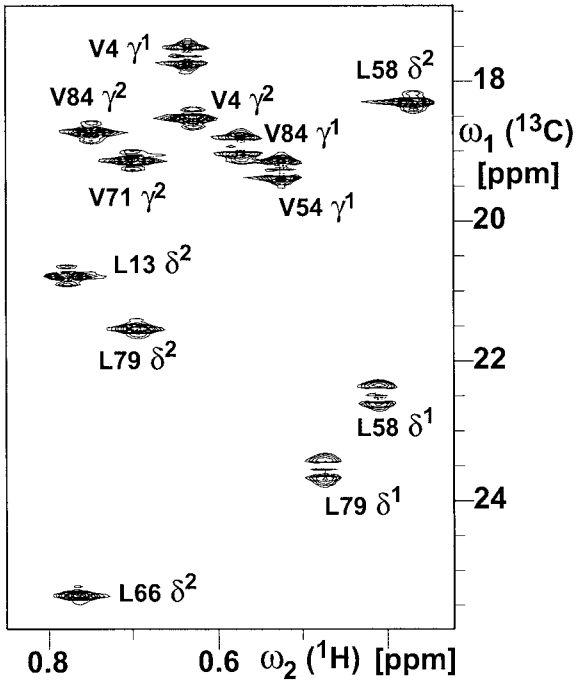

Figure 1. Region [$\omega_1(^{13}C) = 17$–26 ppm, $\omega_2(^1H) = 0.3$–0.8 ppm] of a [^{13}C,^1H]-HSQC spectrum of 10% ^{13}C-labeled TM1290 recorded at 600 MHz and 313 K (2.3 mM TM1290, 600 µl 95% H_2O/5% D_2O containing 20 mM sodium phosphate at pH 6.0 and 2 mM NaN_3). Stereospecific assignments of all the peaks in this region to isopropyl methyls of Val and Leu are indicated.

propanesulfonic acid, sodium salt (DSS). Using the absolute frequency ratios, the ^{13}C and ^{15}N chemical shifts were referenced indirectly to DSS.

2D [^1H,^{15}N]-HSQC, 3D HNCACB, 3D CBCA(CO)NH and 3D HNCO spectra (Bax and Grzesiek, 1993) were used to obtain initial sequence-specific assignments for the polypeptide backbone, and residual gaps and ambiguities were resolved using 3D ^{15}N-resolved [^1H,^1H]-NOESY spectra. Starting with the information derived from the αCH–βCH_n chemical shifts obtained during the backbone assignments, virtually complete ^1H and ^{13}C assignments of all non-aromatic side chain CH_n moieties, including all prolines, were obtained using 2D [^{13}C,^1H]-HSQC, 3D ^{15}N-resolved [^1H,^1H]-TOCSY and 3D H(C)CH-TOCSY (Peti et al., 2000). ^1H spin systems of the aromatic rings of His, Phe and Tyr were identified in a D_2O solution of unlabeled TM1290, using 2D [^1H,^1H]-NOESY and 2D [^1H,^1H]-TOCSY. Sequence-specific assignments of the aromatic side chains were established from NOEs between the βCH_2 group and the aromatic protons (Wüthrich, 1986), using 2D [^1H,^1H]-NOESY and 3D ^{13}C-resolved [^1H,^1H]-NOESY spectra. For the isopropyl methyl groups of Leu and Val, stereospecific assignments were obtained using biosynthetically-directed fractional ^{13}C-labeling (Senn et al., 1989) and 2D [^{13}C,^1H]-HSQC (Figure 1). The NMR spectra were processed with the program PROSA (Güntert et al., 1992), and analysed with the XEASY software package (Bartels et al., 1995).

Extent of assignments and data deposition

The resonance assignments are nearly complete, although there is a loop region of residues 46–52 in which no amide protons resonances could be observed, presumably due to slow dynamics (to be published). The total extent of the assignments for the backbone amide protons and the non-labile protons is 96.9%. Assignments are missing for all protons of residues 48, 49, 50 and 73, the amide protons of residues 1, 46, 47, 51, 52 and 74, the H^α and H^β of His 47, and δCH_2 of Arg 72. Assignments were also obtained for more than 95% of the ^{13}C and ^{15}N atoms. The ^1H, ^{13}C and ^{15}N chemical shifts have been deposited in the BioMagResBank (http://www.bmrm.wisc.edu) under BMRB accession number 5560.

Acknowledgements

We thank Dr. S.A. Lesley and H.E. Klock from the JCSG consortium funded by NIGMS GM062411, for providing us with the genomic materials. Financial support: Schweizerischer Nationalfonds (project 31-66427-01); E. Schrödinger Fellowship to W.P. (J2145); K.W. is the Cecil H. and Ida M. Green Visiting Professor of Structural Biology at TSRI.

References

Achenbach-Richter, L., Gupta, R., Stetter, K.O. and Woese, C.R. (1987) *Syst. Appl. Microbiol.*, **9**, 34–39.

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

Bax, A. and Grzesiek, S. (1993) *Acc. 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. et al. (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 11664–11669.

Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L. et al. (1999) *Nature*, **399**, 323–329.

Peti, W., Griesinger, C. and Bermel, W. (2000) *J. Biomol. NMR*, **18**, 199–205.

Senn, H., Werner, B., Messerle, B.A., Weber, C., Traber, R. and Wüthrich, K. (1989) *FEBS Lett.*, **249**, 113–118.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.