Journal of Biomolecular NMR **29:** 207–208, 2004.

Letter to the Editor: NMR assignment of the hypothetical rhodanese domain At4g01050 from *Arabidopsis thaliana*

David Pantoja-Uceda[a], Blanca López-Méndez[a], Seizo Koshiba[a], Takanori Kigawa[a], Mikako Shirouzu[a,b], Takaho Terada[a,b], Makoto Inoue[a], Takashi Yabuki[a], Masaaki Aoki[a], Eiko Seki[a], Takayoshi Matsuda[a], Hiroshi Hirota[a], Mayumi Yoshida[a], Akiko Tanaka[a], Takashi Osanai[a], Motoaki Seki[a], Kazuo Shinozaki[a], Shigeyuki Yokoyama[a,b,c] & Peter Güntert[a,*]

[a]*RIKEN Genomic Sciences Center, 1-7-22, Suehiro, Tsurumi, Yokohama 230-0045, Japan;* [b]*RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki, Sayo, Hyogo 679-5148, Japan;* [c]*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan*

Received 10 September 2003; Accepted 18 November 2003

Key words: NMR assignment, structural genomics, *Arabidopsis thaliana*

Biological context

Arabidopsis thaliana (thale cress) is an important model system for identifying plant genes and determining their functions. The complete 125-megabase genome of *Arabidopsis* has been sequenced (The Arabidopsis Initiative, 2000). It comprises five chromosomes with a total of 25 498 genes encoding proteins from ∼11 000 families.

Arabidopsis was chosen by the RIKEN Structural Genomics/Proteomics Initiative (RSGI) as one of the target organisms for determining three-dimensional structures of representative members from all known protein families (Yokoyama et al., 2000).

A 121 amino acid domain was selected from the gene At4g01050 on chromosome 4, for which conventional sequence alignment did not reveal significant homology with proteins of known structure or function. Nevertheless, more sensitive algorithms for multiple sequence alignments and hidden Markov models such as Pfam (Bateman et al., 1997) provide tentative evidence that this domain may have a rhodanese fold. This paper reports the virtually complete NMR assignments of the hypothetical rhodanese domain At4g01050 which will provide the basis for future NMR studies of its structure, dynamics, and function. To our knowledge, no three-dimensional structure of a plant rhodanese domain is known yet.

*To whom correspondence should be addressed. E-mail: guentert@gsc.riken.go.jp

Methods and experiments

The Arabidopsis rhodanese domain At4g01050 was produced as a 167 amino acid recombinant protein with an N-terminal HAT affinity tag and a TEV protease cleavage site. The ^{13}C- and ^{15}N-labeled protein was produced by the *E. coli* cell-free synthesis system (Kigawa et al., 1999). The protein was first adsorbed to a TALON Superflow affinity column using 20 mM Tris-HCl (pH 8.0) containing 1 M NaCl, and eluted with 20 mM Tris-HCl (pH 8.0) containing 300 mM NaCl and 500 mM imidazole. The HAT-tag was then removed by incubation with TEV protease with an overnight dialysis at 4 °C against 20 mM Tris-HCl (pH 8.0) containing 1 M NaCl. To remove the HAT-tag and the TEV protease from the reaction mixture, the solution was added to a TALON Superflow affinity column. The flowthrough fraction was collected, desalted and loaded onto a HiTrap SP cation exchange column equilibrated with 5 mM Tris-HCl (pH 7.0). Finally, the purified protein was eluted with a gradient of 0–1 M NaCl.

The protein sample used for the NMR measurements comprises 134 amino acid residues, residues 8–128 constitute the hypothetical rhodanese domain which is surrounded by non-native flanking sequences of residues 1–7 and 129–134 that are related to the expression and purification system. A single sample of approximately 1.1 mM, uniformly ^{13}C and ^{15}N labeled protein was prepared in 20 mM phosphate buffer, pH 6.0, 100 mM NaCl, 1 mM dithiothreitol, 0.02% NaN_3, 10% D_2O (v/v). All NMR measurements were performed at 25 °C on Bruker DRX 600

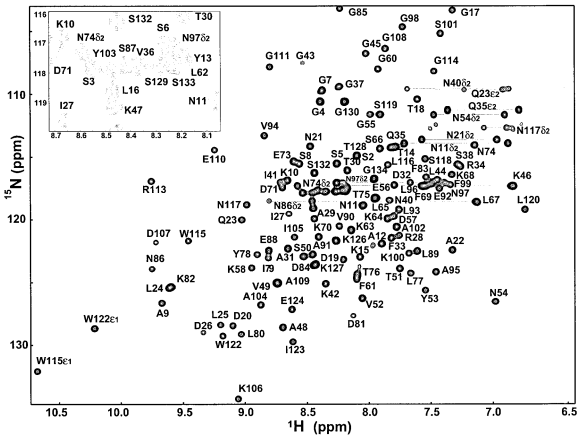

Figure 1. [^{1}H,^{15}N]-HSQC spectrum of the hypothetical rhodanese domain At4g01050 from *Arabidopsis thaliana* recorded at 600 MHz and 25 °C using 1.1 mM uniformly ^{13}C and ^{15}N labeled protein, 20 mM phosphate buffer, pH 6.0, 100 mM NaCl, 1 mM dithiothreitol, 0.02% NaN_3, 10% D_2O (v/v).

or, in the case of the NOESY experiments, Bruker AV 800 spectrometers. ^{1}H, ^{15}N and ^{13}C chemical shifts were referenced relative to the frequency of the ^{2}H lock resonance of water.

Sequence-specific assignments of the polypeptide backbone resonances were obtained by standard triple resonance techniques (Cavanagh, 1996) using 2D [^{1}H,^{15}N]-HSQC (Figure 1) and 3D HNCO, HN(CA)CO, HNCA, H^{α}-coupled HNCA, HN(CO)CA, HNCACB and CBCA(CO)NH spectra. 3D HNHA and HBHA(CO)NH spectra were used for the assignment of the H^{α} and H^{β} protons, respectively. ^{1}H and ^{13}C assignments of the non-aromatic side-chain CH_n moieties, including all prolines, were obtained from 2D [^{1}H,^{13}C]-HSQC, 3D (H)CC(CO)NH with 20 ms mixing time, 3D H(CCCO)NH with 20 ms mixing time and 3D HCCH-TOCSY with 14 ms mixing time. Assignments were confirmed using 3D ^{15}N-edited NOESY-HSQC and ^{13}C-edited NOESY-HSQC spectra with 80 ms mixing time. The ^{1}H and ^{13}C spin systems of the aromatic rings of Phe, Trp and Tyr were identified using a 3D HCCH-COSY experiment. Sequence-specific assignments of aromatic side chains were obtained using NOEs between the aromatic protons and the βCH_2 group or the α-proton in a 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY (80 ms mixing time) spectrum.

The programs NMRPipe (Delaglio et al., 1995) and NMRView (Johnson and Blevins, 1994) were used for spectral processing and data analysis, respectively.

Extent of assignments and data deposition

Ninety-nine percent of the backbone amide and nonlabile ^{1}H chemical shifts could be assigned. Backbone ^{1}H assignments are complete except for H^{N} of Gly 1 and H^{α} of Gly 134. The assignments of the nonlabile side-chain hydrogen atoms are complete except for H^{ζ} of Phe 33, H^{δ} of Lys 58, $H^{\varepsilon}/H^{\zeta}$ of Phe 83, and H^{δ} of Lys 127. All hydrogen-bound ^{15}N resonances are assigned except for Gly 1, N^{ε}/N^{η} of Arg 34, N^{ε}/N^{η} of Arg 113 and N^{ζ} of the 15 Lys residues. 99% of the hydrogen-bound ^{13}C resonances are assigned, i.e., all but C^{ζ} of Phe 33, $C^{\delta}/C^{\varepsilon}$ of Lys 58, $C^{\varepsilon}/C^{\zeta}$ of Phe 83 and C^{γ} of Lys 127. Except for Gly 1, all backbone carbonyl carbon resonances are assigned. The ^{1}H, ^{13}C and ^{15}N chemical shifts have been deposited in the BioMagResBank database (http://www.bmrb.wisc.edu) under BMRB accession number 5929.

Acknowledgements

D.P.-U. is supported by a pre-doctoral grant from the Spanish Ministry of Science and Technology. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

References

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M. and Sonnhammer, E.L. (2002) *Nucl. Acids Res.*, **30**, 276–280.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy, Principles and Practice*, Academic Press, Inc., San Diego.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Johnson, B.A. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.

Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. and Yokoyama, S. (1999) *FEBS Lett.*, **442**, 15–19.

The Arabidopsis Initiative (2000) *Nature*, **408**, 796–815.

Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R. and Kuramitsu, S. (2000) *Nat. Struct. Biol.*, **7**, 943–945.