Journal of Biomolecular NMR, **20:** 287–288, 2001.
KLUWER/ESCOM

Letter to the Editor: ^{1}H, ^{13}C and ^{15}N resonance assignment of YajQ, a protein of unknown structure and function from *Escherichia coli*

Simona Miron[a], Tudor Borza[b], Cosmin Saveanu[b], Anne-Marie Gilles[b], Octavian Bârzu[b] & Constantin T. Craescu[a,*]

[a]*Institut National de la Santé et de la Recherche Médicale U350 & Institut Curie-Recherche, Centre Universitaire, Bâtiments 110-112, F-91405 Orsay, France;*
[b]*Laboratoire de Chimie Structurale des Macromolécules (CNRS, URA 2185), Institut Pasteur, Paris, France*

Received 20 February 2001; Accepted 6 April 2001

Key words: resonance assignment, triple resonance NMR, YajQ

Biological context

A systematic search, performed in our laboratories, for bacterial nucleotide binding proteins with unknown function revealed several poorly represented species in *Escherichia coli* proteome (Blattner et al., 1997). These proteins were identified by ligand elution pseudo-affinity chromatography in combination with two-dimensional gel electrophoresis and mass spectrometry. One of them was identified to be YajQ, a polypeptide with unknown structure and no assigned function. Therefore it belongs to the 38% fraction of *E. coli* proteome designated as hypothetical proteins. Sequence comparison with BLAST (Altschul et al., 1997) allowed us to identify homologous proteins in many other bacteria (*H. influenzae*, *B. subtilis*, *M. tuberculosis*, etc.) but not in eukaryotic organisms. YajQ is a small size (163 residues) monomeric protein with a stable and regular structure, including both α-helix and β-sheet secondary elements, as indicated by the circular dichroism analysis. In the absence of any functional properties, determination of the 3D structure may open the way towards the annotation of the molecular function and understanding of the relationships between the sequence and biological role of the protein (Hwang et al., 1999; Eisenstein et al., 2000). In addition to this fundamental aspect, YajQ and the other bacterial homologous proteins may have a pharmaceutical relevance as potential new targets for antimicrobial compounds, provided that the members of this family constitute an essential gene product for a given pathogenic organism.

*To whom correspondence should be addressed.

We therefore initiated a structural NMR project dedicated to the solution structure determination of YajQ and to the study of its intermolecular interactions with potential relevance for the biological function. For the NMR analysis, the protein was overproduced and uniformly isotope labeled with ^{13}C/^{15}N in *E. coli.* Here we communicate the almost complete ^{1}H, $^{13}C^{\alpha}$ and ^{15}N backbone assignment and partial assignment for the majority of side-chain resonances of YajQ protein in the free form. These data will allow the solution structure determination and study of interactions with putative ligands or other biopolymers.

Methods and experiments

Uniformly labeled recombinant YajQ was overproduced in *E. coli* strain Bli5/pAOT7 using M9 minimal medium containing 1.5 g/l 99% (^{15}N)-ammonium sulfate and/or 3.0 g/l 99% (^{13}C)-glucose as the sole nitrogen and carbon sources, respectively, and supplemented with 70 μg/ml kanamycin and 30 μg/ml chloramphenicol. At OD_{600} of 1.2–1.4, the culture was induced by addition of isopropyl-β-D-thiogalactoside (1 mM final concentration) and further incubated at 37 °C for 16 h. The protein was purified by a two-step procedure including chromatography on Blue-Sepharose and Ultrogel AcA54. The molecular mass of YajQ, measured by electrospray ionization mass spectrometry (18 213.98±0.8 Da) was 130 Da less than that calculated from the sequence (18 343.89 Da), meaning that the N-terminal methionine residue is missing.

NMR samples at a concentration of 1.0–1.3 mM (pH 6.5) were obtained by dissolving the lyophilized protein in potassium phosphate buffer (50 mM) in 95% $^{1}H_2O$/5% $^{2}H_2O$ or in 100% $^{2}H_2O$. Assignment was mainly performed from the analysis of 2D homonuclear, and double- and triple-resonance (HSQC, NOESY-HSQC, TOCSY-HSQC, HNCA and HN(CO)CA) NMR experiments (Wüthrich, 1986; Cavanagh et al., 1996) at 500 MHz (Varian Unity-500). Additional 2D ^{15}N HSQC and 3D ^{15}N NOESY-HSQC spectra were also acquired on a Bruker 800 MHz spectrometer (ICSN, Gif-sur-Yvette). Proton chemical shifts (in ppm) were referenced relative to internal DSS and ^{15}N and ^{13}C references were set indirectly relative to DSS using frequency ratios (Wishart et al., 1995). The NMR data were processed and analyzed using Felix98 software (MSI, San Diego, CA), running on a Silicon Graphics Indigo workstation.

Extent of assignments and data deposition

The ^{15}N-HSQC spectrum (Figure 1) of YajQ illustrates the excellent dispersion of the proton and nitrogen resonances in the amide groups. A total of 155 backbone amide cross peaks have been observed and assigned in the HSQC spectrum. The lacking ^{1}H/^{15}N cross peaks correspond to the Phe32, Gly89, Lys90, Gly127, and Lys135 residues. The sequential resonance assignment has been realized by HNCA/HN(CO)CA paired experiment analysis, which provides the intra- and interresidue connectivities, as well as by analysis of sequential NOEs in the ^{15}N NOESY-HSQC spectrum.

We assigned ^{1}H and ^{15}N backbone resonances of 157 amide groups from a total of 160 non-proline residues. The three unassigned residues (Gly89, Lys90, Lys135) were not observed, neither in the HSQC, nor in the ^{15}N NOESY-HSQC and triple resonance spectra. This may be explained by their position in loop segments between regular secondary structure elements (S.M. and C.T.C., unpublished results), which are usually characterized by an increased flexibility. ^{13}C resonances of α-carbons from 144 residues were also assigned from the triple-resonance experiments.

The present assignment data is now used for the secondary structure analysis and the 3D structure determination. The chemical shift values of the proton, nitrogen and carbon resonances have been deposited in the BioMagResBank in Madison, WI, U.S.A. (accession number: 4956).

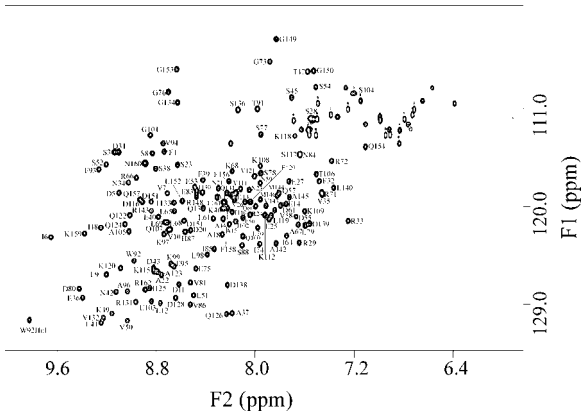

Figure 1. 2D ^{1}H-^{15}N HSQC spectrum at 800 MHz of *E. coli* YajQ at 308 K. Assignments of backbone amides and tryptophan imine resonances are indicated using one-letter codes for amino acids.

Acknowledgements

This work was supported by the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, the Institut Curie, and the Institut Pasteur. S.M. gratefully acknowledges a postdoctoral fellowship from Institut Curie.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) *Nucleic Acids Res.*, **25**, 3389–3402.

Blattner, F.R., Plunkett III, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. (1997) *Science*, **277**, 1453–1474.

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

Eisenstein, E., Gilliland, G.L., Herzberg, O., Moult, J., Orban, J., Poljak, R.J., Banerjei, L., Richardson, D. and Howard, A.J. (2000) *Curr. Opin. Biotechnol.*, **11**, 25–30.

Hwang, K.Y., Chung, J.H., Kim, S.-H., Han, Y.S. and Cho, Y. (1999) *Nat. Struct. Biol.*, **6**, 691–696.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.

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