Journal of Biomolecular NMR **29:** 423–424, 2004.

Letter to the Editor: ^{1}H, ^{15}N and ^{13}C resonance assignments of the ApaG protein of the phytopathogen *Xanthomonas axonopodis* pv. *citri*

Ângela M. Katsuyama[a,b], Daniel O. Cicero[c], Alberto Spisni[a,d], Maurizio Paci[c], Chuck S. Farah[b] & Thelma A. Pertinhez[a,*]

[a]*BioNMR Laboratory, Center for Structural Molecular Biology, LNLS, Campinas, Brazil;* [b]*Department of Biochemistry, Institute of Chemistry, University of São Paulo, São Paulo, Brazil;* [c]*Department of Chemical Science and Technology, University of Rome 'Tor Vergata', and INFM, section B, Rome, Italy;* [d]*Sect. Chemistry and Structural Biochemistry, Department of Experimental Medicine, University of Parma, Parma, Italy.*

Received 20 November 2003; Accepted 16 December 2003

Key words: ApaG protein, NMR, *Xanthomonas axonopodis* pv. *citri*

Biological context

Xanthomonas axonopodis pv. *citri* (*X. a.* pv. *citri*) is the agent of citrus canker, a severe disease responsible for substantial economic losses in the production of citrus fruit. The sequencing of the *X. a.* pv. *citri* genome (Da Silva et al., 2002) revealed the existence of several ORFs (Open Reading Frames) that encode conserved proteins of unknown function and structure. Considering that structural data for these proteins could provide information regarding their cellular function and could disclose new folds, a NMR screening of some *X.a.* pv. *citri* recombinant proteins was carried out to select the best proteins for further structural studies (Galvão-Botton et al., 2003).

In this context, ApaG, a 14 kDa protein encoded by ORF XAC0862 was over-expressed in *E. coli* and purified in order to determine its three-dimensional structure by NMR. The *apaG* gene is located in a multifunctional operon that encodes a protein involved in the pyridoxine biosynthetic pathway (PdxA) (Roa et al., 1989), a translational regulatory protein (KsgA) that confers sensitivity to the kasugamycin antibiotic and a diadenosine-tetraphosphates hydrolase (ApaH) probably involved in cellular protection process against stress. The fact that ApaG and ApaH expression is tightly linked (Blanchin-Roland et al., 1986) led to the hypothesis that the two proteins may be functionally correlated (Lévêque et al., 1990). ApaG has a highly conserved GXGXXG signature sequence (Liu et al., 2003), which is a pyrophosphate binding motif found in NAD- and FAD- binding proteins, suggesting that it may bind pyrophosphates or nucleotide phosphates. ApaG also shares sequence homology with CorD, a *Salmonella typhimurium* protein associated with Co^{2+} sensitivity and Mg^{2+} homeostasis (Gibson et al. 1991). While ApaG function is not clear at the moment, the high degree of sequence conservation among ApaG homologs in bacteria indicates that it must carry out some important biological function.

ApaG-like domains have also been found in the C-terminal regions of mammalian F-Box proteins (Liu et al., 2003), known to be involved in defining the selectivity of substrate targeting for protein ubiquitination and subsequent intracellular degradation. In this context, the ApaG domain may play some key role in favoring protein-protein interactions required for substrate specificity. Another eukaryotic ApaG homolog has been recently identified: PDIP38, a protein that interacts with the p50 DNA polymerase δ subunit and the proliferating cell nuclear antigen (Liu et al., 2003).

While the above considerations on ApaG functions in prokaryotes and eukaryotes remain speculative, the determination of its structure is of obvious interest. In fact, considering that ApaG shows no significant structural homology with any protein in the PDB (<18%), the resolution of the 3D solution structure is expected to provide clues to its biological functions and may also disclose some new protein fold. In this respect, the observation that the recombinant *X. a.* pv. *citri* ApaG can be expressed in large amounts in *E. coli*,

*To whom correspondence should be addressed. E-mail: thelma@lnls.br

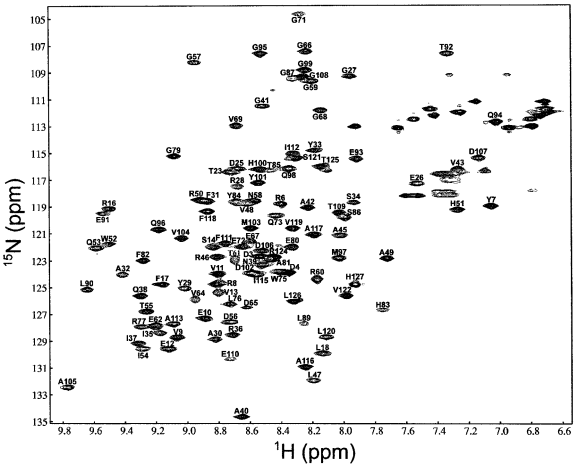

Figure 1. ^{1}H-^{15}N HSQC spectrum of 0.5 mM ApaG protein in phosphate buffer pH 6.8, collected at 600 MHz, at 303 K. Only the backbone cross peaks are indicated with residue numbers.

together with its limited molecular size, high solubility and remarkable stability (Galvão-Botton et al., 2003 and data not shown) clearly identify this protein as appropriate for being studied by NMR. In this paper we report the ApaG backbone and side-chain assignments.

Methods and experiments

The *X. a.* pv. *citri* ApaG gene was cloned into the pET-3a expression vector and over-expressed in *E. coli* BL21(DE3)pLys cells grown in M9 minimal media containing $^{15}NH_4Cl$ and $^{13}C_6H_{12}O_6$ at 37 °C. The doubly labeled soluble protein was purified using anion exchange chromatography, ammonium sulfate precipitation and gel filtration chromatography.

Protein samples were prepared for analysis by NMR using 40 mM NaH_2PO_4 pH 6.8, 50 mM NaCl, 0.05% NaN_3 to a 0.5 mM protein final concentration (in 100% or 5% D_2O). NMR spectra were collected at 303 K on either a Varian Inova® 600 MHz AS (CeBiME-LNLS) or on a Bruker Avance® 700 MHz (University of Rome II) spectrometers.

The ^{15}N edited HSQC spectrum and the HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB and CB-CACONH 3D triple resonance experiments were obtained with the Varian Inova 600 MHz, using the 5% D_2O sample for the backbone assignment process. For the side-chain assignments, we employed NMR data from HCCH-TOCSY, HCCH-COSY, HACACO, NOESY-CHSQC and CHSQC experiments, collected on the Bruker Avance 700 MHz instrument, using the 100% D_2O sample. Spectra were processed and analyzed using NMRPipe/NMRView software packages (Delaglio et al., 1995; Johnson and Blevins, 1994).

Extent of assignments and data deposition

Approximately 93% of the backbone has been assigned. Figure 1 shows the 110 assigned peaks from a total of 119 (127 amino acids minus 8 prolines). Briefly, the CO, Cα, Cβ and Hα assigned atoms corresponded to 98%, 99%, 99% and 100% of the total resonances of each nucleus. From the unassigned amide residues, two are located at the N-terminus (M1 and Q2) and four are situated probably in a highly flexible or solvent exposed loop (A19, H20, Q21 and S22). The remaining three unassigned residues correspond to residues Q63, V70 and V88.

Approximately 90% of the total side-chain signals have been completely assigned (115 from 127 residues). Some aromatic protons remain unassigned for F17, F31, F82, Y84, F111 and F118, as well as some aliphatic protons of R28, R50 and Q53. The forthcoming NOE analysis is expected to allow the determination of the remaining chemical shifts.

At this stage, the global chemical shift analysis carried out using the CSI method (Wishart et al., 1992) provided a preliminary description of the protein secondary structure characterized by the presence of several β-strands (data not shown). Circular dichroism spectroscopy and computational secondary structure prediction analyses are consistent with a high β-sheet content (data not shown).

Resonance assignments have been deposited in the BioMagRes Bank under accession number 5998.

Acknowledgements

This work has been partly supported by FAPESP and CNPq, Brazil and by MIUR, Italy.

References

Blanchin-Roland, S. et al. (1986) *Mol. Gen. Genet.*, **205**, 515–522.
Da Silva, A.C.R. et al. (2002) *Nature*, **417**, 459–463.
Delaglio, F. et al. (1995) *J. Biomol. NMR*, **6**, 277–293.
Galvão-Botton, L.M.P. et al. (2003) *FEBS Lett.*, **552**, 207–213.
Gibson, M.M. et al. (1991) *Mol. Microbiol.*, **5**, 2753–2762.
Johnson, B. and Blevins, R.A. (1994) *J. Biomol. NMR,* **4**, 603–614.
Lévêque, F. et al. (1990) *J. Mol. Biol.*, **212**, 319–329.
Liu, L. et al. (2003) *J. Biol. Chem.*, **278**, 10041–10047.
Roa, B.B. et al. (1989) *J. Bacteriol.,* **171**, 4767–4777.
Wishart, D.S. et al. (1992) *Biochemistry*, **31**, 1647–1651.