Backbone resonance assignments of the monomeric DUF59 domain of human Fam96a

Caroline Mas[1], Kai-En Chen[2], Ian M. Brereton[1], Jennifer L. Martin[2,3] and Justine M. Hill[1,3*]

[1]*Centre for Advanced Imaging,* [2]*Institute for Molecular Bioscience, Division of Chemistry and Structural Biology, and* [3]*School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane QLD 4072, Australia*

[*]Corresponding author: justine.hill1@gmail.com; +61 7 3365 3925 (phone); +61 7 3365 4273 (fax)

Abstract

Proteins containing a domain of unknown function 59 (DUF59) appear to have a variety of physiological functions, ranging from iron-sulfur cluster assembly to DNA repair. DUF59 proteins have been found in bacteria, archaea and eukaryotes, however Fam96a and Fam96b are the only mammalian proteins predicted to contain a DUF59 domain. Fam96a is an 18 kDa protein comprised primarily of a DUF59 domain (residues 31-157) and an N-terminal signal peptide (residues 1-27). Interestingly, the DUF59 domain of Fam96a exists as monomeric and dimeric forms in solution, and X-ray crystallography studies of both forms unexpectedly revealed two different domain-swapped dimer structures. Here we report the backbone resonance assignments and secondary structure of the monomeric form of the 127 residue DUF59 domain of human Fam96a. This study provides the basis for further understanding the structural variability exhibited by Fam96a and the mechanism for domain swapping.

Keywords: DUF59 domain; Fam96a; NMR resonance assignments; secondary structure; domain-swapped dimer.

Biological context

Proteins containing a DUF59 domain have been found widely in bacteria, archaea and eukaryotes. While the function of many of these proteins remains to be identified, members of this family have been implicated in a range of physiological functions. Previous studies have reported that in aerobic bacteria, the DUF59-containing proteins PaaD and PaaJ are involved in phenylacetic acid degradation (Olivera et al. 1998). Moreover, the DUF59-containing protein HCF101 from *Arabidopsis thaliana* is likely to be responsible for metal-sulfur cluster assembly (Lezhneva et al. 2004). Structure and sequence analysis of the DUF59 protein TM0487 from *Thermotoga maritima* also suggested a conserved metal-binding site and a possible role for bacterial DUF59 proteins in binding iron-sulfur clusters (Almeida et al. 2005). To date, only three structures of prokaryotic DUF59-containing proteins have been reported: TM0487 from *Thermotoga maritima* (Almeida et al. 2005), DUF59 from *Bacillus anthracis*, and TTHB138 from *Thermus thermophilus*, revealing monomer structures with a similar secondary structure topology (PDB ID codes: 1WCJ, 3LNO and 3CQ1, respectively).

Fam96a (Family with sequence similarity 96 member a) and its homologue Fam96b (also known as MIP18) are the only two mammalian proteins predicted to contain a DUF59 domain. Yeast two-hybrid analysis suggested that both Fam96a and Fam96b interact with Ciao1 (cytosolic iron-sulfur protein assembly 1) (Rual et al. 2005). Ciao1 is a 38 kDa protein that is responsible for iron-sulfur cluster protein assembly in the cytoplasm and nucleus. Recently, Fam96b and Ciao1, together with ANT2 (adenine nucleotide translocator-2), MMS19 (nucleotide excision repair protein homolog) and the transcription subunit XPD, were shown to be part of the MMXD complex (MIP18-MMS19-XPD). This multiprotein complex is implicated in DNA repair and chromosome segregation (Ito et al. 2010). Indeed, knockdown of Fam96b in HCT116 colon cancer cells led to mis-localisation of the mitotic spindle and accumulation of abnormally shaped nuclei (Ito et al. 2010). Fam96a and Fam96b share 41% sequence identity and differ in that Fam96a has a predicted N-terminal signal sequence whereas Fam96b does not. According to size exclusion chromatography, chemical crosslinking and analytical ultracentrifugation experiments, the DUF59 domain of Fam96a can exist as monomeric and dimeric forms in solution (Chen et al, submitted). Furthermore, unlike the previously determined monomer structures of bacterial DUF59 proteins, crystal structures of the monomeric and dimeric forms of Fam96a were solved as two different domain-swapped dimers (Chen et al, submitted; PDB ID codes 3UX2 and 3UX3). In the

dimer structure obtained from the monomeric fraction, the C-terminal region of the molecule is exchanged. In contrast, the structure determined for the dimeric protein fraction exhibits exchange of a single α-helix from the centre of the molecule, which is uncommon in the reported structures of domain-swapped dimers (Gronenborn 2009).

Here we report the backbone resonance assignments of the monomeric DUF59 domain of human Fam96a (residues 31-157). This represents an important step toward understanding the formation of domain-swapped dimers of Fam96a.

Methods and experiments

The DUF59 domain of Fam96a (residues 31-157) was subcloned into pMCSG7 vector, and the recombinant protein produced with a histidine tag at the N-terminus followed by a TEV protease recognition site. TEV protease cleavage of the His-tag results in an additional three residues (SNA) at the N-terminus of the protein. Uniformly ^{15}N-labelled and ^{13}C/^{15}N-labelled proteins were overexpressed in *E. coli* BL21(DE3) cells in minimal medium containing $^{15}NH_4Cl$ (1 g/L) and ^{13}C-glucose (2 g/L) as the sole nitrogen and carbon sources, respectively (Hill 2008). 2 L of cell culture was grown at 37°C to an optical density (A_{600} nm) of 1.0, then expression was induced with 1 mM IPTG and the cells incubated at 20°C overnight. Purification followed the same protocol reported for the native protein (Chen et al, submitted). Briefly, cell pellets were resuspended in lysis buffer and lysed by sonication. The protein was isolated using TALON® resin (Clontech), the His-tag subsequently cleaved using TEV protease and residual uncleaved protein removed using TALON® resin. The His-tag cleaved protein was further purified by gel filtration chromatography. Fam96a eluted as three peaks, consistent with a large oligomer, dimer and monomer, and the monomeric form was used to obtain backbone resonance assignments.

Samples for NMR contained 0.8 mM protein in 25 mM HEPES buffer at pH 7.0, 150 mM NaCl and 2 mM DTT in H_2O/D_2O (9:1). The NMR experiments were acquired at 25°C on either a Bruker Avance 750 MHz spectrometer equipped with a *z*-gradient triple resonance probe or a Bruker Avance 900 MHz spectrometer equipped with a *z*-gradient triple resonance cryoprobe. NMR spectra were processed with NMRPipe (Delaglio et al. 1995) and analysed using CcpNmr (Vranken et al. 2005). The backbone ^{1}H, ^{13}C and ^{15}N resonances of Fam96a monomer were assigned using a combination of standard triple resonance experiments

including 3D HNCO, HN(CA)CO, CBCA(CO)NH, HNCACB and C(CO)NH (Sattler et al. 1999).

Assignments and data deposition

The 2D ^{1}H-^{15}N HSQC spectrum of the DUF59 domain of Fam96a, with assignments labelled, is shown in Fig. 1. Backbone assignments were achieved for 86.8% of the amide resonances (105 of the 121 non-proline residues), 88.5% of the Cα and Cβ resonances, and 88.2% of the C′ resonances. Analysis of the Cα secondary shifts of Fam96a monomer revealed the presence of five α-helices and three β-strands (Fig. 2). The predicted secondary structure resembles the α/β-topology observed in both crystallographic forms of Fam96a (PDB ID codes 3UX2 and 3UX3) and the three prokaryotic DUF59 domain structures, TM0487 from *Thermotoga maritima* (Almeida et al. 2005; PDB ID 1WCJ), DUF59 from *Bacillus anthracis* (PDB ID 3LNO), and TTHB138 from *Thermus thermophilus* (PDB ID 3CQ1).

16 of the 121 non-proline residues of Fam96a have not been assigned, which are: E51, K52, T86, C90, S91, L92 and residues S119 to D128. These residues are located within three loops in the structure, between α-helix 1 and β-strand 1 (E51 and K52), β-strand 2 and α-helix 2 (T86-L92), and β-strand 3 and α-helix 3 (S119-D128), respectively (Fig. 2). The absence of signals from these residues due to line broadening suggests conformational exchange in these regions of the molecule. Consistent with this hypothesis, the latter two loops were also not observed in the crystal structures of Fam96a and correspond to hinge loops that connect the exchanged regions to form the domain-swapped dimers. In an attempt to detect the missing amide resonances, 2D ^{1}H-^{15}N HSQC spectra were recorded at various temperatures (10, 15, 25 and 30°C) and in different buffer compositions (25 mM HEPES or 25 mM potassium phosphate at pH 7.0); however, no additional peaks were observed under the conditions examined. Assignments for the ^{1}H, ^{13}C and ^{15}N backbone resonances of the DUF59 domain of Fam96a in its monomeric form have been deposited in the BioMagResBank database (http://www.bmrb.wisc.edu/) with the accession number 18137.

Acknowledgments

We thank the Queensland NMR Network for measurement time on the 750 MHz and 900 MHz NMR spectrometers. This work was supported by an Australian Research Council (ARC) grant (DP0770465) to JLM. CM is a University of Queensland Postdoctoral Research

Fellow, and JLM is an ARC Australian Laureate Fellow (FL0992138) and Honorary National Health and Medical Research Council (NHMRC) Research Fellow (455829).

References

Almeida MS, Herrmann T, Peti W, Wilson IA, Wüthrich K (2005) NMR structure of the conserved hypothetical protein TM0487 from Thermotoga maritima: implications for 216 homologous DUF59 proteins. Protein Sci 14:2880-2886. doi:10.1110/ps.051755805

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-293. doi:10.1007/BF00197809

Gronenborn AM (2009) Protein acrobatics in pairs – dimerization via domain swapping. Curr Opin Struct Biol 19:39-49. doi:10.1016/j.sbi.2008.12.002

Hill JM (2008) NMR screening for rapid protein characterization in structural proteomics. Methods Mol Biol 426:437-446. doi:10.1007/978-1-60327-058-8_29

Ito S, Tan LJ, Andoh D, Narita T, Seki M, Hirano Y, Narita K, Kuraoka I, Hiraoka Y, Tanaka K (2010) MMXD, a TFIIH-independent XPD-MMS19 protein complex involved in chromosome segregation. Mol Cell 39:632-640. doi:10.1016/j.molcel.2010.07.029

Lezhneva L, Amann K, Meurer J (2004) The universally conserved HCF101 protein is involved in assembly of [4Fe-4S]-cluster-containing complexes in *Arabidopsis thaliana* chloroplasts. Plant J 37:174-185. doi:10.1046/j.1365-313X.2003.01952.x

Olivera ER, Miñambres B, García B, Muñiz C, Moreno MA, Ferrández A, Díaz E, García JL, Luengo JM (1998) Molecular characterization of the phenylacetic acid catabolic pathway in *Pseudomonas putida* U: the phenylacetyl-CoA catabolon. Proc Natl Acad Sci U S A 95:6419-6424. doi:10.1073/pnas.95.11.6419

Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S, Albala JS, Lim J, Fraughton C, Llamosas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S, Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP, Vidal M (2005) Towards a proteome-scale map of the human protein-protein interaction network. Nature 437:1173-1178. doi:10.1038/nature04209

Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR

experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc 34:93–158. doi:10.1016/S0079-6565(98)00025-9

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59:687-696. doi:10.1002/prot.20449

Figure legends

Fig. 1 2D ^{1}H–^{15}N HSQC spectrum and backbone resonance assignments of the monomeric Fam96a DUF59 domain. The spectrum was recorded in 25 mM HEPES buffer (pH 7.0), 150 mM NaCl and 2 mM DTT at 900 MHz and 25°C. Assignments of residues are indicated with the one-letter amino acid code and the sequence number. *Horizontal lines* connect peaks corresponding to side chain NH_2 groups of Asn and Gln residues

Fig. 1

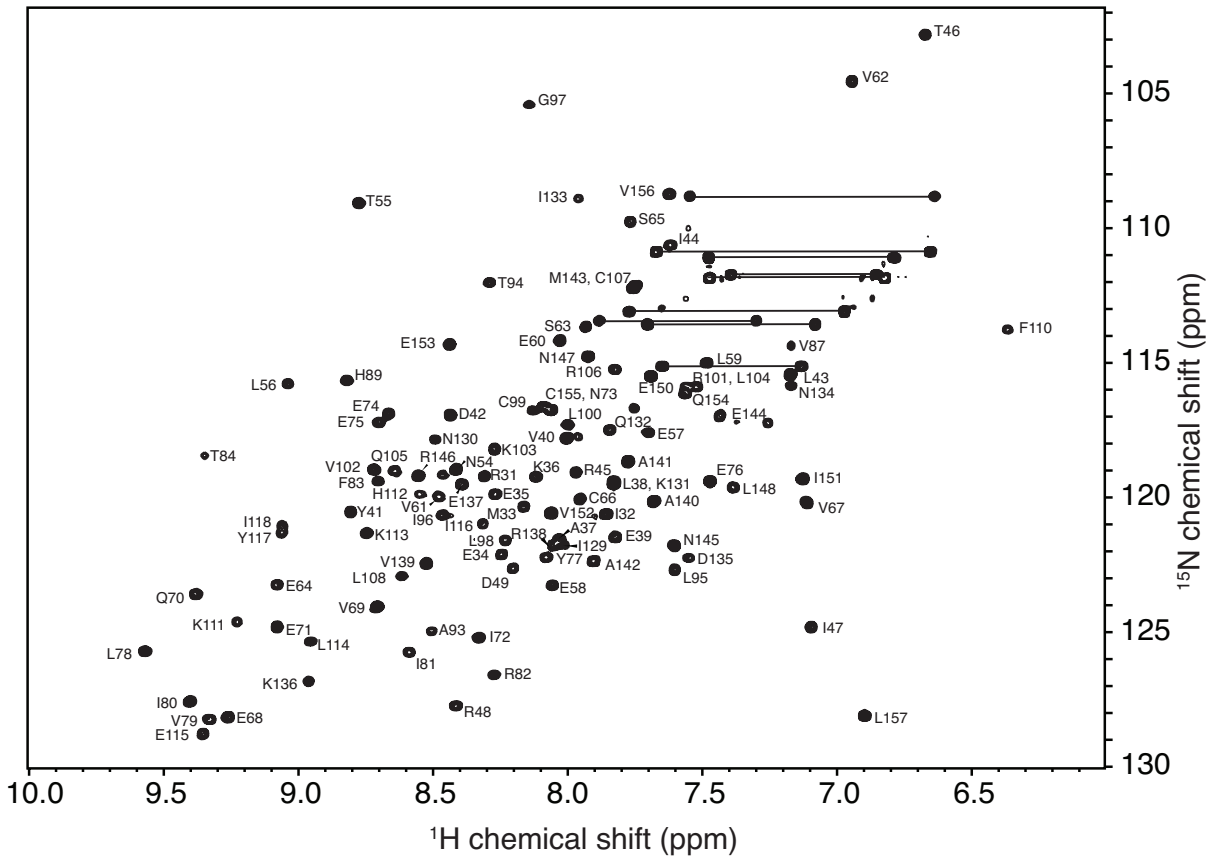

Fig. 2 Chemical shift deviations from random coil values for the Cα resonances of the monomeric Fam96a DUF59 domain. Unassigned residues are indicated by an *asterisk*. Secondary structure elements are indicated on the *top* of the figure by *pink boxes* for α-helices and *blue boxes* for β-strands

Fig. 2