J Biomol NMR (2013) 56:387–392
DOI 10.1007/s10858-013-9746-6

NMR STRUCTURE NOTE

Solution structure of monomeric human FAM96A

Bingjie Ouyang · Lei Wang · Shuo Wan · Yang Luo · Lu Wang · Jian Lin · Bin Xia

Received: 26 February 2013 / Accepted: 14 May 2013 / Published online: 22 June 2013

Biological context

Domain of unknown function 59 (DUF59) is a protein family found in bacteria, archaea and some eukaryotes, which is highly conserved during evolution. Members of this family from prokaryotes are reported to be involved in physiological functions such as metabolism of phenylacetic acid (Ferrandez et al. 1998; Olivera et al. 1998) and metal-sulfur cluster synthesis (Lezhneva et al. 2004; Luo et al. 2012; Schwenkert et al. 2010). Three structures of bacterial DUF59s have been determined (Almeida et al. 2005) and all of them are in monomeric form.

There are only two DUF59 proteins in mammals, FAM96A and FAM96B (family with sequence similarity 96 member A and B). FAM96A contains 160 amino acid residues, and its N-terminal 27 residues are predicted to be a signal peptide and residues 28–160 are highly homologous to those of DUF59 family. Previous study reveals that *FAM96A* mRNA is enriched in macrophages, indicating the potential importance of FAM96A as regulator of inflammation and target for anti-inflammatory design (Chen et al. 2012). FAM96A was also found to interact both in vitro and in vivo with Ciao 1 (Chen et al. 2012), which functions as a cytoplasmic FeS assembly (CIA) protein (Srinivasan et al. 2007) and regulates the physiological function of WT1 (Wilms tumor suppressor protein) in cell growth and differentiation (Johnstone et al. 1998).

It was reported recently that there are monomeric, dimeric and oligomeric forms of FAM96A when expressed in *Escherichia coli* (Chen et al. 2012; Mas et al. 2012). Crystal structure of dimeric FAM96A is a domain-swapped form, but the lack of associated electron density for flexible loop (residues: T122–E127) region leads to uncertain domain swapping mode. Intriguingly, the monomeric FAM96A was converted into a distinct domain-swapped dimer during crystallization. Therefore, the closed conformational structure of the monomeric FAM96A has not been obtained.

In this work, we determine the solution structure of the monomeric FAM96A using NMR spectroscopy, which provides structural insight for the formation of domain-swapped dimer. In addition, we investigate the temperature-dependent interconversion between the monomeric and dimeric FAM96A, and the formation of oligomer.

Methods

Protein expression and purification

The cDNA fragment encoding *FAM96A* (residues: 28–160) without N-terminal signal peptide was subcloned into pGEX vector with a GST-tag at N-terminus and a 6His-tag at C-terminus. The recombinant plasmid was transformed into *E. coli* strain BL21 (DE3) for protein

B. Ouyang · L. Wang · S. Wan · B. Xia (✉)
Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China
e-mail: binxia@pku.edu.cn

B. Ouyang · L. Wang · J. Lin · B. Xia
College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

S. Wan · B. Xia
School of Life Sciences, Peking University, Beijing 100871, China

Y. Luo · L. Wang
Laboratory of Medical Immunology, School of Basic Medical Science, Peking University Health Science Center, Beijing 100191, China

expression. The bacteria were cultured overnight in 40 mL LB medium containing 100 μg/mL of ampicillin sodium at 35 °C and were transferred into 1 L LB medium for continuous growth till OD_{600} of the culture reached 1.0. The cells were collected by centrifugation and transferred into 500 mL M9 medium with 1 g/L $^{15}NH_4Cl$ and 4 g/L $^{13}C_6$-glucose for $^{13}C/^{15}N$ isotopic labeling. Then the bacterial culture was incubated at 35 °C for an hour and isopropyl-b-D-thiogalactoside (IPTG) was added to induce the protein production at a final concentration of 0.4 mM. After induced for 8 h at 25 °C, the cells were harvested by centrifugation and resuspended in the cell lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0). After the cells were lysed by freeze–thaw followed by sonication, the supernatant of cell lysate was applied onto the Ni–NTA affinity column. The GST-tag was removed by PreScission protease through on-column cleavage at 4 °C overnight. Oligomeric, dimeric and monomeric forms of FAM96A were separated by subsequent gel filtration (Superdex-75) with an ÄKTA fast protein liquid chromatography system (FPLC) (GE Healthcares). The monomeric and dimeric FAM96A were concentrated for further experiments.

NMR spectroscopy

The NMR samples containing 0.4 mM monomeric protein were in 50 mM sodium phosphate buffer (pH 7.0) with 95 % H_2O/5 % D_2O, along with 50 mM NaCl, 1 mM EDTA, 50 mM 1,4-dithiothreitol (DTT), 50 mM Arg, 50 mM Glu and 0.01 % DSS. The oxygen was removed from the samples in order to keep reducing condition and stabilize the protein. All NMR experiments were collected on Bruker Avance 600, 700 and 800 MHz NMR spectrometer at 298 K, each equipped with a cryoprobe. The backbone assignments of monomeric FAM96A were obtained based on 3D HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO experiments. The 3D HBHA(CO)NH, (H)CCH-COSY, (H)CCHTOCSY, H(C)CH-COSY, H(C)CH-TOCSY, TOCSY-HSQC experiments were performed for side-chain chemical shift assignments. Distance restraints for structure calculation were generated by using the 3D ^{15}N- and ^{13}C-edited NOESY-HSQC spectra, each measured with a mixing time of 100 ms. All NMR spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed with NMRView (Johnson and Blevins 1994). Proton chemical shifts were

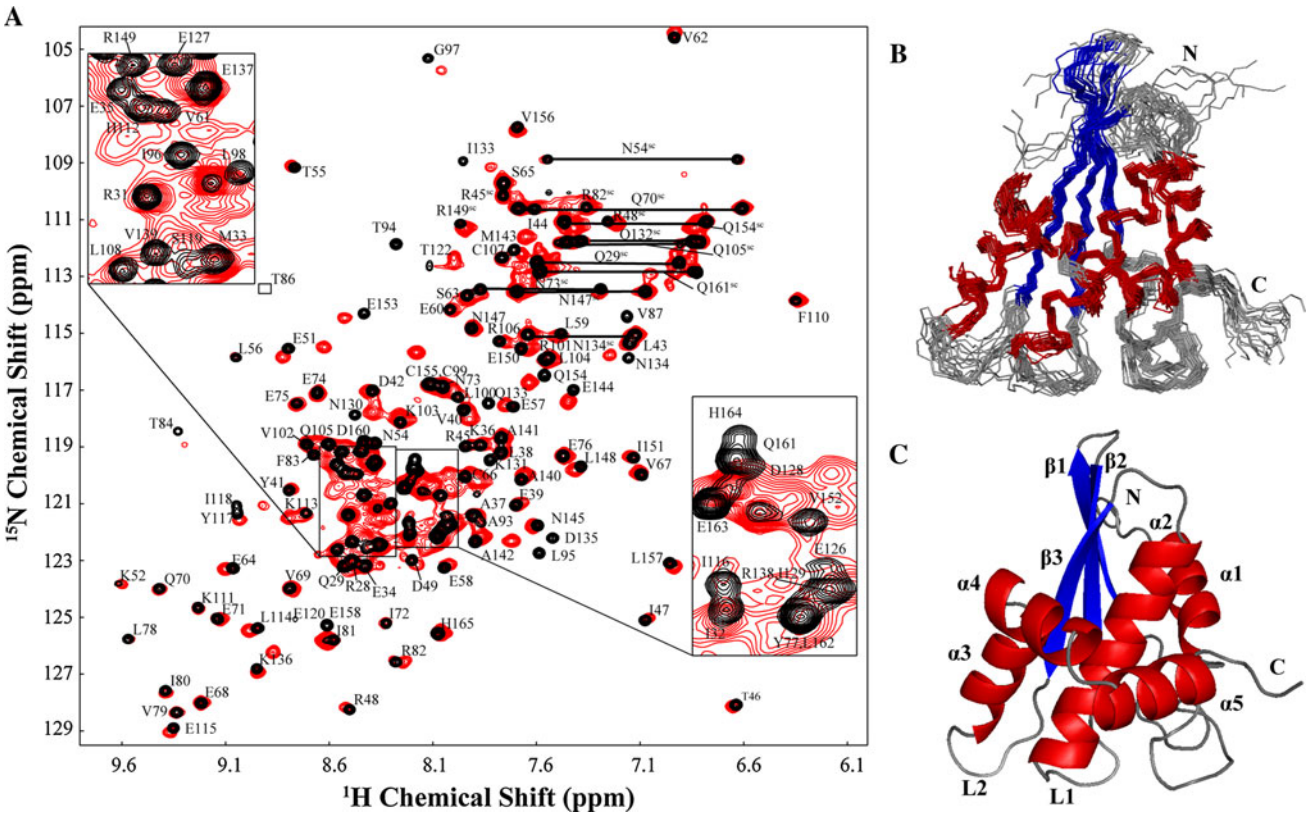

Fig. 1 **a** Overlay of 2D ^{1}H–^{15}N HSQC spectra of the monomeric (*black*) and dimeric (*red*) FAM96A. Resonance assignments of the monomeric FAM96A are indicated. **b** Superimposition of the backbone trace of the 20 representative structures of the monomeric FAM96A. **c** *Ribbon diagram* of the mean structure with secondary structural elements labeled

referenced to internal DSS, ^{15}N and ^{13}C chemical shifts were referenced indirectly to DSS (Markley et al. 1998).

Structure calculation

The initial structure of monomeric FAM96A was built by CANDID (Herrmann et al. 2002) based on the chemical shift assignments and NOE distance restraints. The distance restraints, dihedral angle restraints and H-bond restraints were determined by SANE (Duggan et al. 2001), TALOS (Cornilescu et al. 1999) and secondary structure elements in initial structure, respectively. The structure calculations were performed using CYANA 2.1 (Guntert et al. 1997) with standard CYANA simulated annealing schedule with 10,000 torsion angle dynamics steps per conformer. A total of 200 structures were calculated and 100 structures with the lowest target function values were selected. Then the structure refinement was carried out using ff03 force field and the generalized Born (GB) solvent model in AMBER9 (Pearlman et al. 1995). Finally, 20 lowest energy conformers that most consistent with experimental restraints were selected for representation. The final structures were analyzed using PROCHECK_NMR (Laskowski et al. 1996) and MOLMOL (Koradi et al. 1996).

Monomer–dimer interconversion

The samples for interconversion experiments containing 1.0 mM monomeric protein or 0.2 mM dimeric protein were in 50 mM sodium phosphate buffer (pH 7.0) with 50 mM NaCl, 1 mM EDTA, 50 mM 1,4-dithiothreitol (DTT). The oxygen was removed from the samples in order to keep reducing condition and stabilize the protein. The interconversion was detected by monitoring the 2D ^{1}H–^{15}N HSQC spectra of monomeric or dimeric FAM96A along with time using Bruker Avance 700 and 800 MHz NMR spectrometer at 298 or 303 K.

Results and discussion

Resonance assignments of the monomeric FAM96A

Consistent with previous reports (Chen et al. 2012; Mas et al. 2012), FAM96A also existed as monomer, dimer and oligomer when expressed in *E. coli* in our study, even though we used a different expression construct. We purified both the monomeric and dimeric protein and carried NMR resonance assignments for the monomer. NH signals of 117 residues were observed in 2D ^{1}H–^{15}N HSQC spectra of the monomeric FAM96A, while residues H89, C90, S91, L92, G121, H123, S124 and T125 did not show a NH signal in addition to 8 prolines (Fig. 1a). Overall, more than 96 % of backbone and side chain chemical shift assignments of the monomeric FAM96A were obtained. The ^{1}H, ^{13}C and ^{15}N chemical shifts have been deposited at BioMagResBank database under accession number 18548.

Table 1 Structural statistics for the monomeric FAM96A

NOE restraints	4225
Intraresidue	1780
Sequential	653
Medium-range	311
Long-range	376
Ambiguous	1105
Dihedral angel restraints	
φ	80
ψ	79
Hydrogen bond restrains	50
Chirality restraints	385
Omega angel	138
Structure statistics of final 20 conformers	
Restraints violations	
Distance (>0.2 Å)	0
Dihedral angle (>5°)	0
RMSD from mean structure (Å)	
Secondary-structure backbone atoms	0.62 ± 0.17
Secondary-structure heavy atoms	0.92 ± 0.16
All backbone atoms	2.13 ± 0.44
All heavy atoms	2.48 ± 0.42
AMBER Energy (kcal/mol)	
Mean AMBER energy	−5591.29
NOE distance restraints violation energy	4.62
Torsion angle restraints violation energy	0.28
Ramachandran statistics (%)	
Most favored regions	90.1
Additional allowed regions	7.8
Generously allowed regions	1.1
Disallowed regions	0.9

Solution structure of the monomeric FAM96A

Solution structure of the monomeric FAM96A was finally calculated using 4225 NOE distance restraints, 159 dihedral angle restraints and 50 H-bond restraints. The 20 lowest energy conformers for the monomeric FAM96A have a root mean square deviation (RMSD) of 0.6 Å for backbone heavy atoms of secondary structure elements (Fig. 1b). Detailed restraints information and structural statistics are summarized in Table 1. The structure has been deposited into Protein Data Bank (PDB ID: 2M5H).

Unlike the crystal structure, the solution structure of the monomeric FAM96A shows a closed monomeric conformation, which makes FAM96A a bona fide example of 3D

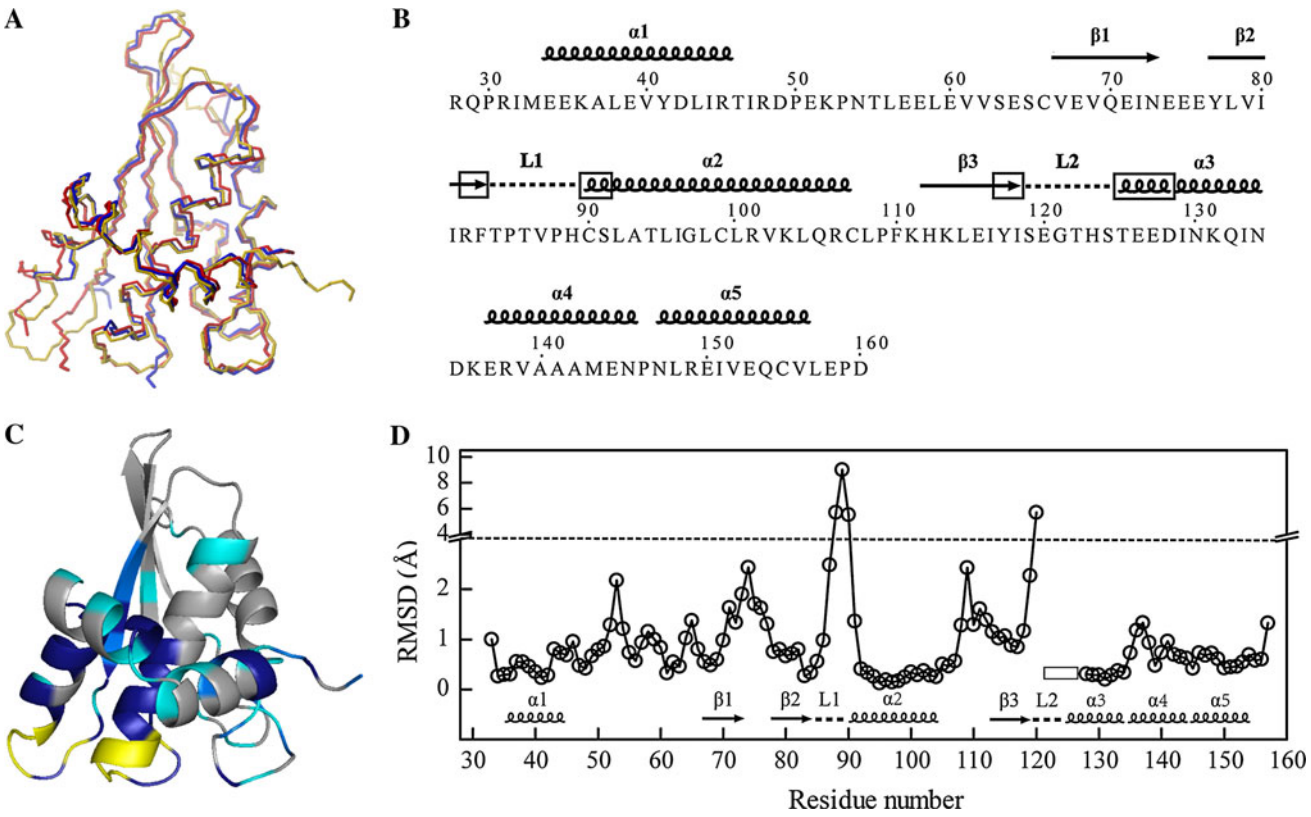

Fig. 2 **a** Superimposition of the solution structure of the monomeric FAM96A and monomer-like functional units (MFUs) of the two domain-swapped crystal structures. The solution structure of the monomeric FAM96A is shown in *yellow*, MFU of 3UX2 is shown in *blue* and MFU of 3UX3 is shown in *red*. **b** Primary sequence of FAM96A and secondary structure elements. Differences of secondary structure elements between the monomeric and dimeric FAM96A are indicated by *black box*. **c** Mapping of the residues with significant NH chemical shift differences between the monomer and dimer onto the structure of the monomeric FAM96A. Residues of which NH peaks do not overlap at all are shown in *dark blue*, residues with partially overlapped NH peaks are shown in *light blue*, residues with mostly overlapped HN peaks are shown in *cyan* and residues with unchanged NH peaks are shown in *gray*. Prolines and residues without NH assignment are shown in *yellow*. **d** Per-residue RMSD between the monomeric FAM96A and MFU of 3UX2. Residues (T122–E127) that are lack of electron density in 3UX2 are indicated by the *black box*. Flexible residues (R28–I32, G121–S124 and E158–D160) with RMSD over 1.5 Å in the ensemble of the monomer structures are omitted from comparison

domain swapping. It consists of five α helices (α1: residues 34–45; α2: residues 90–107; α3: residues 125–134; α4: residues 136–144; α5: residues 146–156) and a three-strand mixed β-sheet (β1: residues 67–73; β2: residues 76–83; β3: residues 112–118). The connectivity of secondary structure elements is α1–β1–β2–α2–β3–α3–α4–α5. The α2 helix is located in the center of the structure and other secondary structure elements surround it. There are two flexible loops: L1 loop (residues: 85–89) that links β2 strand and α2 helix and L2 loop (residues: 119–124) that links β3 strand and α3 helix, which have a RMSD of 1.1 and 2.1 Å for backbone heavy atoms, respectively.

Structural comparison of the monomeric and dimeric FAM96A

We compared the structure of the monomeric FAM96A to monomer-like folding units (MFUs) of the two domain-swapped dimeric structures (Fig. 2a). The RMSD between FAM96A monomer and MFU of the domain-swapped dimer from the crystallized monomeric FAM96A (PDB ID: 3UX3) is 1.2 Å for backbone heavy atoms of secondary structure elements, while it is only 0.7 Å between the monomer and MFU of the dimeric FAM96A (PDB ID: 3UX2), which is even smaller than the RMSD of 0.9 Å between MFUs of the two domain-swapped dimers. Further comparison between the monomer and 3UX2 indicates that β2, β3 strands and α2 helix are each 2 residues longer in the monomer (Fig. 2b). The α3 helix is also longer in the monomer due to that three residues (T125–E127) are not visible and the residue D128 does not adopt α helix conformation in 3UX2. As a result, L1 loop linking β2 strand and α2 helix and L2 loop linking β3 strand and α3 helix are both shorter in the monomeric FAM96A.

In the structure of dimeric FAM96A (PDB ID: 3UX2), due to the lack of associated electron density for L2 loop (residues T122–E127), there are two possible domain swapping modes (Chen et al. 2012). One mode is to swap

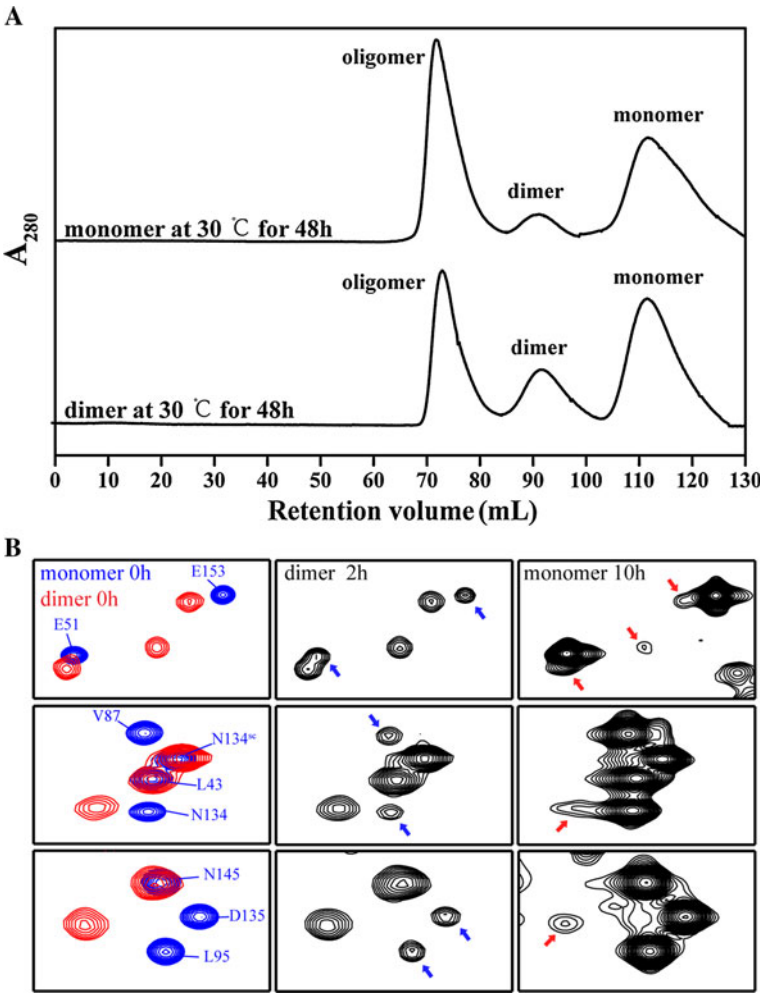

Fig. 3 Interconversion between the monomeric and dimeric FAM96A. **a** Gel filtration chromatography of the monomeric and dimeric FAM96A at 30 °C for 48 h. **b** Changes of 2D ^{1}H–^{15}N HSQC spectra of the monomeric and dimeric FAM96A along with time. *Left* initial spectra of the monomeric (*blue*) and dimeric (*red*) FAM96A; *Middle* spectra of the dimeric FAM96A at 30 °C for 2 h; *Right* spectra of the monomeric FAM96A at 30 °C for 10 h. The changes of signature NH signals of the monomer and dimer are indicated by the *blue* and *red arrows*, respectively

the α2 helix, β3 strand and α3–α5 helices with L1 loop as the only hinge loop and the other is only α2 helix and β3 strand swapped with both L1 and L2 loops serving as hinge loops. Comparison of 2D ^{1}H–^{15}N HSQC spectra of the monomeric and dimeric FAM96A shows that about 30 % of NH peaks show significant deviation between the monomer and dimer, while the other NH signals remain unchanged. Some NH peaks of the monomer do not overlap with those of the dimer at all, including residues T84, T86 and V87 in L1 loop, E120 and T122 in L2 loop, F83 in β2 strand, A93–L98 in α2 helix, I116 and I118 in β3 strand, E126, D128 and N130–D135 in α3 helix, along with residues E51, M143 and E144 (Fig. 2c). This should indicate that the chemical environment and even the conformation for L1 and L2 loop regions differ significantly between the domain-swapped dimer and the monomer. This is consistent with the fact that residues S119, E120 and G121 on L2 loop in 3UX2 show very large deviations (RMSD of 2.3, 5.7 and 9.6 Å) from those of the monomer (Fig. 2d), although RMSD of these three residues in the ensemble of the monomer structures are 1.0, 1.4 and 2.7 Å, respectively. Similarly, L1 loop in 3UX2 which is a confirmed hinge loop also deviate from that of the monomer with a RMSD of 4.5 Å, while RMSD of that in the ensemble of the monomer structures is only 1.1 Å. Meanwhile, L1 loop in 3UX3 is not a hinge loop, and it has a RMSD of only 1.0 Å to that in the monomer. Taken together, we believe that the dimeric FAM96A is formed through swapping α2 helix and β3 strand, and both L1 and L2 loops serve as hinge loops. Therefore, shortened secondary structure elements in the dimeric FAM96A may arise from the need for lengthening hinge loops to stabilize the domain-swapped dimer.

Interconversion between the monomeric and dimeric FAM96A

When either the pure monomeric or dimeric FAM96A protein was left at 30 °C for 2 days, gel filtration chromatography analysis showed that the samples were separated into three peaks of which retention volumes correspond to the oligomeric, dimeric and monomeric

forms of FAM96A (Fig. 3a). Thus, the monomeric and dimeric FAM96A can convert into each other, and oligomer is also formed during the interconversion. The same results were also observed when the samples were left at 25 °C and the conversion was much slower.

To further examine the interconversion between the monomeric and dimeric FAM96A, we monitored the change of 2D ^{1}H–^{15}N HSQC spectra along with time. For the dimeric FAM96A at 30 °C, signature NH signals of the dimer became weaker, while signature NH signals of the monomer appeared and became stronger, as the time was increased (blue arrows in Fig. 3b). We also performed the NMR conversion experiments from the pure dimer at 25 °C. The estimated half conversion time ($t_{1/2}$) of the dimeric FAM96A is about 6 h at 30 °C and about 48 h at 25 °C. This reveals that the rate of interconversion is highly dependent on the temperature and may explain why the conversion wasn't detected at 4 °C in the previous work (Chen et al. 2012). Meanwhile, some weak signature NH signals of the dimeric FAM96A could be observed in 2D ^{1}H–^{15}N HSQC spectra of a pure monomer sample (1 mM) after 10 h at 30 °C (red arrows in Fig. 3b), revealing that the dimeric FAM96A converted from the monomer is as the same as the one formed in expression.

Taken together, the solution structure of the monomeric FAM96A appears in a closed monomeric conformation, significantly distinct to that of the crystal structure. FAM96A can consequently become a bona fide example of 3D domain swapping. Based on our study, we determine the domain swapping mode of the dimeric FAM96A. We also find that the monomeric and dimeric FAM96A can interconvert at physiological condition, while the physiological role of the conversion needs further study.

Acknowledgments All NMR experiments were carried out at Beijing NMR Center. This research was supported by Grants 2012CB910703 and 2009CB521703 (to BX) from the 973 Program and Grant 81172001 (to LW) from the National Science Foundation of China.

References

Almeida MS, Herrmann T, Peti W, Wilson IA, Wuthrich K (2005) NMR structure of the conserved hypothetical protein TM0487 from *Thermotoga maritima*: implications for 216 homologous DUF59 proteins. Protein Sci 14(11):2880–2886

Chen KE, Richards AA, Ariffin JK, Ross IL, Sweet MJ, Kellie S, Kobe B, Martin JL (2012) The mammalian DUF59 protein Fam96a forms two distinct types of domain-swapped dimer. Acta Crystallogr D Biol Crystallogr 68(Pt 6):637–648

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13(3):289–302

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(3):277–293

Duggan BM, Legge GB, Dyson HJ, Wright PE (2001) SANE (Structure Assisted NOE Evaluation): an automated model-based approach for NOE assignment. J Biomol NMR 19(4):321–329

Ferrandez A, Minambres B, Garcia B, Olivera ER, Luengo JM, Garcia JL, Diaz E (1998) Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway. J Biol Chem 273(40):25974–25986

Guntert P, Mumenthaler C, Wuthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 273(1):283–298

Herrmann T, Guntert P, Wuthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 319(1):209–227

Johnson BA, Blevins RA (1994) NMR View: a computer program for the visualization and analysis of NMR data. J Biomol NMR 4(5):603–614

Johnstone RW, Wang J, Tommerup N, Vissing H, Roberts T, Shi Y (1998) Ciao 1 is a novel WD40 protein that interacts with the tumor suppressor protein WT1. J Biol Chem 273(18):10880–10887

Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14(1):51–55, 29–32

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8(4):477–486

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(2):174–185

Luo D, Bernard DG, Balk J, Hai H, Cui X (2012) The DUF59 Family Gene AE7 Acts in the Cytosolic Iron–Sulfur Cluster Assembly Pathway to Maintain Nuclear Genome Integrity in *Arabidopsis*. Plant Cell 24(10):4135–4148

Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wuthrich K (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-union task group on the standardization of data bases of protein and nucleic acid structures determined by NMR spectroscopy. J Biomol NMR 12(1):1–23

Mas C, Chen KE, Brereton IM, Martin JL, Hill JM (2012) Backbone resonance assignments of the monomeric DUF59 domain of human Fam96a. Biomol NMR Assign. doi:10.1007/s12104-012-9390-1

Olivera ER, Minambres B, Garcia B, Muniz C, Moreno MA, Ferrandez A, Diaz E, Garcia JL, Luengo JM (1998) Molecular characterization of the phenylacetic acid catabolic pathway in *Pseudomonas putida* U: the phenylacetyl-CoA catabolon. Proc Natl Acad Sci USA 95(11):6419–6424

Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham III TE, DeBolt S, Ferguson D, Seibel G, Kollman P (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun 91(1–3):1–41

Schwenkert S, Netz DJ, Frazzon J, Pierik AJ, Bill E, Gross J, Lill R, Meurer J (2010) Chloroplast HCF101 is a scaffold protein for [4Fe-4S] cluster assembly. Biochem J 425(1):207–214

Srinivasan V, Netz DJ, Webert H, Mascarenhas J, Pierik AJ, Michel H, Lill R (2007) Structure of the yeast WD40 domain protein Cia1, a component acting late in iron–sulfur protein biogenesis. Structure 15(10):1246–1257