Jinsai Shang,^a Guan Wang,^a Yang Yang,^a Xiaolan Huang^b and Zhihua Du^{a*}

^aChemistry and Biochemistry, Southern Illinois University, 1245 Lincoln Drive, Carbondale, IL 62901, USA, and ^bComputer Science, Southern Illinois University, 1000 Faner Drive, Carbondale, IL 62901, USA

Correspondence e-mail: zdu@chem.siu.edu

Received 19 June 2013

Accepted 17 August 2013

Expression, purification and crystallization of the FP domain of the human F-box protein Fbxo7

Fbxo7 is a conserved protein in higher eukaryotes that belongs to the F-box protein family. Fbxo7 is the substrate-recognition component of the SCF^{Fbxo7} (Skp1-Cul1-Fbox protein) E3 ubiquitin ligase. Besides the F-box motif, Fbxo7 also contains a C-terminal proline-rich region, an N-terminal ubiquitin-like domain and a novel FP (Fbxo7/PI31) domain preceding the F-box motif. The FP domains of Fbxo7 and the PI31 proteasome inhibitor mediate interaction between the two proteins. For structure determination of the FP domain of Fbxo7, a protein construct (amino acids 181–335) corresponding to the FP domain was expressed, purified and crystallized. The native and selenomethionine-labeled proteins crystallized in different crystal forms. Native and single-wavelength anomalous dispersion data sets with diffraction to 2.1 and 2.0 Å resolution, respectively, have been collected and structure determination is in progress.

1. Introduction

F-box proteins belong to a large protein family that is characterized by the presence of the F-box motif, a loosely conserved ~40–50 amino-acid motif involved in protein–protein interaction in eukaryotes (Kumar & Paietta, 1995; Bai *et al.*, 1996). In most F-box proteins, the F-box motif accounts for only a small fragment of the sequence; the majority of the sequence harbors a wide variety of other motifs. Two of the most common protein–protein interaction motifs besides the F-box are WD40 repeats and the leucine-rich repeats. F-box proteins containing these two motifs are classified into the Fbxw and Fbwl subfamilies, while the other F-box proteins are classified into the Fbxo subfamily (Cenciarelli *et al.*, 1999; Winston *et al.*, 1999).

F-box proteins are best known for their critical involvement in the ubiquitin–proteasome system (Hershko, 2005). This system mediates the selective destruction of many regulatory and misfolded/unfolded proteins. In the ubiquitin–proteasome pathway, proteins targeted for destruction are conjugated with multiple copies of the small protein ubiquitin. This ubiquitination process is carried out by the successive action of three key enzymes: the ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-ligating (E3) enzymes. The poly-ubiquitinated proteins are subsequently degraded by the 26S proteasome.

In the ubiquitin–proteasome system, the E3 ligases are responsible for recruiting specific target proteins and catalyzing the transfer of ubiquitin from the E2 enzymes to the targets (Jackson & Eldridge, 2002). The E3 ligases are very diverse. They can be either single proteins or multi-component protein complexes. The largest group of E3 ligases is the SCF complex that contains four components: Skp1 (S-phase kinase-associated protein 1), cullin1, Rbx1 (RING-box protein 1) and an F-box protein. Cullin1 is a rigid and elongated molecule that binds Skp1 and Rbx1 at its N-terminus and C-terminus, respectively, keeping these two subunits ~100 Å apart. Rbx1 in turn binds a ubiquitin-charged E2 enzyme while Skp1 binds the F-box protein subunit through interaction with the F-box. The F-box proteins recruit the substrates to the SCF complexes (Schulman *et al.*, 2000; Zheng *et al.*, 2002; Wu *et al.*, 2003; Goldenberg *et al.*, 2004; Hao *et al.*, 2007; Mizushima *et al.*, 2007; Zeng *et al.*, 2010). The F-box proteins are the specificity components of the SCF-type E3 ligase

© 2013 International Union of Crystallography
All rights reserved

complexes. The existence of a large number of diverse F-box proteins provides the recognition specificities for various substrates.

F-box only protein 7 (Fbxo7) belongs to the Fbxo subfamily of the F-box proteins (Hsu *et al.*, 2004; Laman *et al.*, 2005). It has been shown that Fbxo7 can serve as a component of functional SCF complexes. The ubiquitination and proteasomal degradation of HURP (hepatoma up-regulated protein), cIAP1 (cellular inhibitor of apoptosis protein 1) and TRAF2 (TNF receptor-associated factor 2) are mediated by SCF^{Fbxo7} (the F-box protein component is indicated by the superscript; Hsu *et al.*, 2004; Chang *et al.*, 2006; Kuiken *et al.*, 2012).

Fbxo7 also interacts with other proteins that are not substrates for ubiquitin-mediated degradation, including p27, Cdk6 and PI31 (Laman *et al.*, 2005; Kirk *et al.*, 2008). Selective binding of Fbxo7 to Cdk6 enhances the assembly and activity of cyclin D–Cdk6 complexes, which are responsible for the ability of Fbxo7 to transform murine fibroblasts. The biological function of Fbxo7–PI31 interaction is not clear at the present time. However, a similar interaction was observed between the DmPI31 and Nutcracker (an F-box protein that is homologous to Fbxo7) proteins in *Drosophila*. Nutcracker binding stabilizes DmPI31 and they form a regulatory complex that controls proteasome activity (Bader *et al.*, 2011).

Fbxo7 has been implicated in cancers and neurodegenerative disease. It has been shown that Fbxo7 is overexpressed in tumors of the lung and colon (Laman *et al.*, 2005). It has also been shown that mutations in the Fbxo7 gene (*fbxo7*) caused PARK15, an early-onset Parkinsonian disease (Shojaee *et al.*, 2008; Di Fonzo *et al.*, 2009).

Besides the F-box motif (amino acids 329–379), Fbxo7 also contains a ubiquitin-like domain (Ubl; amino acids 1–78), a Cdk6-interacting motif (Cdk; amino acids 129–169), an Fbxo7/PI31 domain (FP; amino acids 180–324) and a proline-rich region (PRR; amino acids 423–522) (Kirk *et al.*, 2008). The PRR is most likely to be unstructured. It mediated specific binding and recruitment of HURP and cIAP1 to the SCF^{Fbxo7} ligase complex (Hsu *et al.*, 2004; Chang *et al.*, 2006). The FP domain was identified in Fbxo7 and the proteasome inhibitor PI31 based on sequence homology between corresponding regions of the two proteins. It was shown that the FP domains in Fbxo7 and PI31 mediated homodimerization and heterodimerization of the proteins. FP domains have also been identified in Nutcracker and DmPI31 proteins in *Drosophila*. The crystal structure of the FP domain of PI31 has been determined, which features a novel α/β fold (Kirk *et al.*, 2008). In this study, we aim to determine the crystal structure of the FP domain of Fbxo7. In spite of the emerging biological and physiological significance of Fbxo7, no structural study on this protein has been reported. The structure of the FP domain of Fbxo7 will provide insights into the protein–protein interaction mechanism of this important protein.

2. Materials and methods

2.1. Cloning

A cDNA clone for the full-length human Fbxo7 (GenBank BC008361) was purchased from DNASU Plasmid Repository. The protein product of this cDNA corresponds to the Fbxo7 isoform 1 with GenBank accession No. NP_036311. The DNA encoding the FP domain (residues 181–335) was amplified by PCR from the plasmid with the forward primer 5'-CAAGGACCGAGCAGCCCCCTCA-CATTCATTAGAGACCTTGTATC-3' and the reverse primer 5'-ACCACGGGGAACCAACCCTTAGAGGACGACCAACCCAAATAC-3'. Both of these primers contain 21 nt at the 5'-end as a specific sequence for ligation-independent cloning (LIC). The PCR product

was purified by agarose gel electrophoresis and processed by T4 DNA polymerase in the presence of 2.5 mM dATP to generate 5'-overhangs at both ends. The cloning vector is an in-house modified LIC vector that contains DNA sequences encoding the HaloTag (Los *et al.*, 2008), His tag and an eight amino-acid recognition sequence for HRV 3C protease in front of a specific LIC sequence. The vector was linearized by *Sma*I digestion and subsequently processed by T4 DNA polymerase in the presence of 2.5 mM dTTP to generate 5'-overhangs at both ends. The processed PCR insert and vector were mixed and incubated at room temperature for 20 min. The mixture was used to transform DH5 α *Escherichia coli* competent cells. Protein expression was carried out using NiCo21 (DE3) *E. coli* cells (New England Biolabs).

2.2. Protein expression and purification

The FP domain of Fbxo7 was expressed in *E. coli* as a fusion protein with an N-terminal HaloTag and His tag. After removal of the tags by HRV 3C protease, the target protein contains an artificial sequence GPSSP at the N-terminus. Protein expression was induced by a 0.4 mM final concentration of IPTG when the cell culture reached an OD₆₀₀ of 0.6–0.8. After induction, the temperature was lowered to 288 K and the culture was allowed to grow overnight. Expressed proteins were purified by NTA affinity resin. The fusion tags were then cleaved by HRV 3C protease. The cleaved tags were separated from the target protein by a reverse IMAC (immobilized metal-affinity chromatography) process with NTA resin. The purified proteins were concentrated to ~10 mg ml⁻¹ in a buffer consisting of 25 mM Tris pH 7.5, 200 mM NaCl, 1 mM DTT.

For selenomethionine-labeled protein, the cells were grown in M9 minimal culture until reaching an OD₆₀₀ of 0.6–0.8. Six amino acids, leucine, isoleucine, lysine, phenylalanine, threonine and valine, were added to the culture at a final concentration of 50–100 mg l⁻¹ to inhibit the methionine-biosynthesis pathway. After 30 min of continuous incubation, L-selenomethionine (SeMet) was added to a final concentration of 50 mg l⁻¹. Protein expression was induced by 0.4 mM final concentration of IPTG at 288 K overnight.

2.3. Crystallization and data collection

Crystallization trials were carried out using 96-well-format plates. 6 × 96 in-house prepared screening solutions (using various PEGs as the precipitants) were used in the trials. Multiple crystallization conditions were identified within a few days. After optimizing the crystallization conditions, diffracting crystals of the Fbxo7 FP domain were obtained at 295 K by sitting-drop vapor diffusion against 50 μ l well solution using 96-well-format crystallization plates. The crystallization drops contained 1 μ l protein sample mixed with 1 μ l well solution. Two crystal forms were obtained. Crystals in space group *P*₄₃₂₁2 were obtained with a well solution consisting of 20% PEG 3350, 50 mM MOPS buffer pH 7.0, 100 mM KBr, 10% (v/v) glycerol. Crystals in space group *P*₂₁₂₁2₁ were obtained with a well solution consisting of 20% PEG 3350, 50 mM MES buffer pH 7.2, 100 mM KBr, 10% (v/v) glycerol. Crystals were flash-cooled in liquid nitrogen.

Data collection was carried out on beamline 21ID-D of LS-CAT at the Advanced Photon Source (Argonne National Laboratory). Data were processed, integrated and scaled with the programs *MOSFLM* and *SCALA* in the *CCP4* suite (Battye *et al.*, 2011).

3. Results and discussion

The FP domain of Fbxo7 was predicted to contain residues 180–324 of the protein. In this study, our FP-domain construct contains

Table 1

Data-collection and processing statistics.

Values in parentheses are for the outermost resolution shell.

	Native	SeMet
Space group	$P2_12_12_1$	$P4_32_12$
Unit-cell parameters (Å)	$a = 44.22$, $b = 64.41$, $c = 89.31$	$a = b = 57.65$, $c = 89.32$
Resolution (Å)	44.22–2.10 (2.21–2.10)	48.38–2.10 (2.21–2.10)
No. of observed reflections	88774	129484
No. of unique reflections	15476 (2202)	9341 (1342)
Completeness (%)	99.8 (99.2)	100 (100)
R_{merge} (%)	10.0 (48.8)	11.7 (49.2)
$\langle I/\sigma(I) \rangle$	10.9 (4.0)	15.7 (6.3)
Multiplicity	5.7 (5.7)	13.9 (14.2)

Figure 1

Crystals of the FP domain of the human F-box protein Fbxo7. The crystals were yielded by SeMet-labeled protein.

residues 181–335. The protein was expressed in bacteria as a fusion protein with a HaloTag and a His tag to facilitate expression and purification of the protein. After purification, the tags were removed by the very efficient and specific HRV 3C protease. The target protein for crystallization contains an artificial pentapeptide sequence GPSSP at the N-terminus as a cloning artefact resulting from the 3C protease-recognition sequence and the LIC sequence of the cloning vector. However, the presence of the artificial sequence did not seem to interfere with crystallization of the protein.

Native and SeMet-labeled protein samples were prepared. Both samples yielded diffraction-quality crystals under similar crystallization conditions. Although all of the crystals had an elongated rod-shaped appearance, two different crystal forms existed for each of the two protein preparations, in space groups $P4_32_12$ and $P2_12_12_1$. Many data sets were collected at the Se peak wavelength (0.97872 Å). Finally, the highest-resolution (2.1 Å) data set in space group $P2_12_12_1$ was collected from a single native crystal and the highest-resolution (2.1 Å) data set in space group $P4_32_12$ was collected from a single SeMet-labeled crystal (Fig. 1). Data-collection and processing statistics are presented in Table 1.

The crystallization protein construct has a calculated molecular weight of 17 743.7 Da. For the crystal in space group $P4_32_12$ the

asymmetric unit can accommodate only one molecule, with a solvent content of 41.2%. For the crystal in space group $P2_12_12_1$ the asymmetric unit can accommodate either one or two molecules, with a solvent content of 65.7 or 31.4%, respectively. In summary, we have successfully expressed, purified and crystallized a protein construct corresponding to the FP domain of human Fbxo7. High-quality native and SAD (single-wavelength anomalous dispersion) data sets have been collected. The SeMet-labeled protein contains three SeMet residues, which should provide an adequately strong anomalous signal for solving the structure in space group $P4_32_12$ using the SAD phasing method. The structure in space group $P2_12_12_1$ can then be solved using the molecular-replacement method. The structures will provide the much needed structural information for Fbxo7.

We thank the staff of the LS-CAT at the Advanced Photon Source (Argonne National Laboratory) for assistance with data collection. The work was supported by the start-up fund and a seed grant from Southern Illinois University Carbondale.

References

- Bader, M., Benjamin, S., Wapinski, O. L., Smith, D. M., Goldberg, A. L. & Steller, H. (2011). *Cell*, **145**, 371–382.
- Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J. W. & Elledge, S. J. (1996). *Cell*, **86**, 263–274.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst. D*, **67**, 271–281.
- Cenciarelli, C., Chiaur, D. S., Guardavaccaro, D., Parks, W., Vidal, M. & Pagano, M. (1999). *Curr. Biol.* **9**, 1177–1179.
- Chang, Y.-F., Cheng, C.-M., Chang, L.-K., Jong, Y.-J. & Yuo, C.-Y. (2006). *Biochem. Biophys. Res. Commun.* **342**, 1022–1026.
- Di Fonzo, A. *et al.* (2009). *Neurology*, **72**, 240–245.
- Goldenberg, S. J., Cascio, T. C., Shumway, S. D., Garbutt, K. C., Liu, J., Xiong, Y. & Zheng, N. (2004). *Cell*, **119**, 517–528.
- Hao, B., Oehlmann, S., Sowa, M. E., Harper, J. W. & Pavletich, N. P. (2007). *Mol. Cell*, **26**, 131–143.
- Hershko, A. (2005). *Cell Death Differ.* **12**, 1191–1197.
- Hsu, J.-M., Lee, Y.-C. G., Yu, C.-T. R. & Huang, C.-Y. F. (2004). *J. Biol. Chem.* **279**, 32592–32602.
- Jackson, P. K. & Eldridge, A. G. (2002). *Mol. Cell*, **9**, 923–925.
- Kirk, R., Laman, H., Knowles, P. P., Murray-Rust, J., Lomonosov, M., Meziane, E. K. & McDonald, N. Q. (2008). *J. Biol. Chem.* **283**, 22325–22335.
- Kuiken, H. J., Egan, D. A., Laman, H., Bernards, R., Beijersbergen, R. L. & Dirac, A. M. (2012). *J. Cell. Mol. Med.* **16**, 2140–2149.
- Kumar, A. & Paietta, J. V. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 3343–3347.
- Laman, H., Funes, J. M., Ye, H., Henderson, S., Galinanes-Garcia, L., Hara, E., Knowles, P., McDonald, N. & Boshoff, C. (2005). *EMBO J.* **24**, 3104–3116.
- Los, G. V. *et al.* (2008). *ACS Chem. Biol.* **3**, 373–382.
- Mizushima, T., Yoshida, Y., Kumanomidou, T., Hasegawa, Y., Suzuki, A., Yamane, T. & Tanaka, K. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 5777–5781.
- Schulman, B. A., Carrano, A. C., Jeffrey, P. D., Bowen, Z., Kinnucan, E. R., Finnin, M. S., Elledge, S. J., Harper, J. W., Pagano, M. & Pavletich, N. P. (2000). *Nature (London)*, **408**, 381–386.
- Shojaee, S., Sina, F., Banihosseini, S. S., Kazemi, M. H., Kalhor, R., Shahidi, G. A., Fakhrirad, H., Ronaghi, M. & Elahi, E. (2008). *Am. J. Hum. Genet.* **82**, 1375–1384.
- Winston, J. T., Koeppe, D. M., Zhu, C., Elledge, S. J. & Harper, J. W. (1999). *Curr. Biol.* **9**, 1180–1182.
- Wu, G., Xu, G., Schulman, B. A., Jeffrey, P. D., Harper, J. W. & Pavletich, N. P. (2003). *Mol. Cell*, **11**, 1445–1456.
- Zeng, Z., Wang, W., Yang, Y., Chen, Y., Yang, X., Diehl, J. A., Liu, X. & Lei, M. (2010). *Dev. Cell*, **18**, 214–225.
- Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koeppe, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W. & Pavletich, N. P. (2002). *Nature (London)*, **416**, 703–709.