Characterization, crystallization and preliminary X-ray crystallographic analysis of the Uba5 fragment necessary for high-efficiency activation of Ufm1

Shutao Xie

MOE Key Laboratory of Protein Science, School of Life Sciences, Tsinghua University, Beijing 100084, People's Republic of China

Correspondence e-mail: xieshutao@sina.com

Received 13 March 2014
Accepted 17 April 2014

Uba5 is the smallest ubiquitin-like molecule-activating enzyme and contains an adenylation domain and a C-terminal region. This enzyme only exists in multicellular organisms. The mechanism through which the enzyme recognizes and activates ubiquitin-fold modifier 1 (Ufm1) remains unknown. In this study, Uba5 adenylation domains with different C-terminal region lengths were cloned, expressed and purified. The results of an *in vitro* truncation assay suggest that Uba5 residues 57–363 comprise the minimal fragment required for the high-efficiency activation of Ufm1. Crystallization of Uba5 residues 57–363 was performed at 277 K using PEG 3350 as the precipitant, and crystals optimized by microseeding diffracted to 2.95 Å resolution, with unit-cell parameters $a = b = 97.66$, $c = 144.83$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. There is one molecule in the asymmetric unit; the Matthews coefficient and the solvent content were calculated to be 2.93 Å³ Da⁻¹ and 58.1%, respectively.

1. Introduction

Post-translational modification of protein by ubiquitin and ubiquitin-like molecules (UBLs) is essential in various physiological processes such as protein quality control, innate and adaptive immune responses, DNA repair and programmed cell death (Goldberg, 2007). Three different kinds of enzymes, namely UBL-activating enzymes (E1), UBL-conjugating enzymes (E2) and UBL ligases (E3), are found in this activation cascade (Hershko, 2005; Schulman & Harper, 2009). Different UBLs have distinct E1s. Among humans, eight E1s can initiate UBL conjugation. These E1s are divided into canonical and noncanonical E1s based on their domain architecture. Canonical E1s include Uba1, Uba6, Uba7, Sae1-Uba2 and Nae2-Uba3, and noncanonical E1s include Uba4, Uba5 and Atg7. Canonical E1s contain an inactive adenylation domain, an active adenylation domain, the first conserved catalytic cysteine domain (FCCH), the second conserved catalytic cysteine domain (SCCH) and a C-terminal ubiquitin-fold domain. Noncanonical E1s differ in molecular weight, domain architecture, structure and mechanisms underlying the activation of their corresponding UBLs.

Ubiquitin-fold modifier 1 (Ufm1) is among the different UBLs with a dedicated E1 (Uba5), E2-conjugating enzyme (Ufc1) and E3 ligase (Ufl1) (Komatsu *et al.*, 2004; Dou *et al.*, 2005; Tatsumi *et al.*, 2010). Uba5, which comprises ~400 amino-acid residues, is comparatively smaller than other E1s, and is weakly reactive towards SUMO2 (Zheng *et al.*, 2008). The Uba5–Ufm1 system is conserved in metazoans and plants, but not in yeasts, which suggests its significance in multicellular organisms (Komatsu *et al.*, 2004; Gannavaram *et al.*, 2011; Zhang *et al.*, 2012; Hertel *et al.*, 2013). Previous studies have shown that Uba5 is necessary in cell autonomous erythroid differentiation and is involved in the regulation of haematopoiesis (Tatsumi *et al.*, 2011). Sequence-alignment analysis revealed that Uba5 has no canonical FCCH and SCCH domains, but only contains an adenylation domain followed by a C-terminal region sequence (Fig. 1a). In 2010, Bacik and coworkers determined the crystal structure of the adenylation domain of Uba5 (residues 57–329; PDB entry 3h8v; Bacik *et al.*, 2010) and demonstrated that the Uba5 adenylation domain was very similar to those of its E1 homologues, and contains eight β -strands surrounded by seven α -helices. The location of the active-site cysteine (Cys250) is near the N-terminus of

the long α 6-helix in the adenylation domain (Bacik *et al.*, 2010). However, information on the underlying molecular mechanism of Uba5-catalyzed adenylation and transthiolation reactions of Ufm1 is still insufficient because of the absence of the C-terminal region that is required for Ufm1 activation.

In this study, a minimal fragment of Uba5 that is required for the high-efficiency activation of Ufm1 was identified. This fragment includes the adenylation domain and part of the C-terminal region, ranging from residues 330 to 363. The purification, characterization, crystallization and preliminary X-ray crystallographic analysis of the fragment are also reported. These results provide insights into the activation mechanism of Ufm1 by Uba5.

2. Materials and methods

2.1. Cloning

Homo sapiens Uba5 (EC 6.3.2.19, NCBI accession No. NM_024818.3) fragments consisting of residues 57–329, 57–363 and 57–404 were amplified from the cDNA plasmid by PCR using the same forward primer with the sequence 5'-gagcatgtccatag-GCATTGAAACGAATGGGAA-3' and reverse primers with sequences 5'-gtgacgtcactcgagtcATTGTATAACCTCTTGTTA-3', 5'-gtgacgtcactcgagtcGTCTGGAAGTGGACCTGAAA-3' and 5'-tgacgtcactcgagtcCATATTCTTCATTTTGGCCA-3', respectively. The PCR product was digested with the restriction endonucleases *Nde*I and *Xho*I (New England Biolabs) and ligated into the pETDuet-1 vector (Novagen). Ligation products were then transformed into *Escherichia coli* DH5 α competent cells by heat shock and

the positive transformants were then selected from agar plates containing 100 μ g ml⁻¹ ampicillin. The authenticity of the recombinant plasmids was confirmed by DNA sequencing.

2.2. Expression and purification

The recombinant plasmid was transformed into *E. coli* BL21 (DE3) competent cells *via* a heat-shock procedure. The positive transformants were selected from agar plates containing 100 μ g ml⁻¹ ampicillin, inoculated into 10 ml LB (Luria–Bertani) medium and grown overnight at 310 K with agitation. The overnight cell culture was diluted into 1 l LB and grown under the same condition to an OD₆₀₀ of 0.8. Recombinant proteins were induced with 0.2 mM IPTG (isopropyl β -D-1-thiogalactopyranoside) at 293 K for 12 h.

The bacterial cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris–HCl pH 8.0, 300 mM NaCl) supplemented with 10 mM MgCl₂, 200 U ml⁻¹ DNaseI, 1 mM phenylmethanesulfonyl fluoride (PMSF) and were then disrupted by sonication. The cell lysate was centrifuged at 8000g at 277 K for 45 min and the soluble fractions were loaded onto Ni–NTA beads (Qiagen). After washing with buffer consisting of 50 mM Tris–HCl pH 8.0, 500 mM NaCl, the recombinant protein was eluted with elution buffer (50 mM Tris–HCl pH 8.0, 250 mM imidazole). The partially purified proteins were loaded onto a Source 15Q column (GE Healthcare Biosciences) pre-equilibrated with 50 mM Tris–HCl pH 8.0 and then eluted using a linear gradient of 0–1 M NaCl. The 6 $\times$ His tags were cleaved by 3C protease (derived from rhinovirus) and removed by reloading the mixture onto an Ni–NTA affinity column. Gly-Pro-His residues were left at the N-terminus of the

Figure 1

Uba5 domain architecture and *in vitro* assay. (a) Summary of the Uba5 constructs used for activation assay. (b) Secondary-structure prediction of the C-terminal region of Uba5 using the PHD method. (c) *In vitro* activation of Ufm1 by Uba5. Different Uba5 recombinant proteins at 20 μ M and Ufm1 at 100 μ M were incubated in 100 μ l reaction buffer consisting of 50 mM Tris–HCl pH 7.5, 150 mM NaCl, 5 mM MgCl₂, 1 mM ATP at 303 K for different times. Reaction products at 10 μ l were resolved by nonreducing 12% SDS–PAGE. The reaction time was 0 min for lanes 1, 5 and 9, 0.5 h for lanes 2, 6 and 10, 1 h for lanes 3, 7 and 11, and 2 h for lanes 4, 8 and 12. The arrows indicate the position of Uba5 fragments before reacting with Ufm1. Lane M contains molecular-mass markers (labelled in kDa).

recombinant protein from the vector cloning sites. Fractions were then condensed and loaded onto a Superdex 200 (GE Healthcare Biosciences) column equilibrated with 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM dithiothreitol (DTT). Purified protein was analyzed by 12% SDS-PAGE and the protein concentration was determined by ultraviolet absorption at 280 nm (for recombinant Uba5 57–363, $1 A_{280} = 1.70 \text{ mg ml}^{-1}$). The purified proteins concentrated to 15 mg ml^{-1} were aliquoted, frozen in liquid nitrogen and stored at 193 K for further experiments and crystallization.

2.3. *In vitro* activation of Ufm1 by Uba5

Activation and transfer of Ufm1 by Uba5 were analyzed by incubating the indicated concentrations of Uba5 and Ufm1 in a 200 μl reaction system. The reaction was performed at 303 K for different durations, stopped using SDS-PAGE sample buffer (80 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol) and then resolved by nonreducing 12% SDS-PAGE.

2.4. Crystallization

Crystallization of Uba5 57–363 was performed at 293 K *via* the hanging-drop method with different commercially available screening kits. Needle-like crystals appeared in drops comprising 1.5 μl protein solution and 1.5 μl reservoir solution consisting of 100 mM MES-KOH pH 6.0, 23–25% PEG 3350, 0.3–0.5 M ammo-

Table 1

Data-collection and processing statistics for Uba5 57–363.

Values in parentheses are for the outer shell.

Wavelength (Å)	0.97928
Temperature (K)	100
Space group	<i>P</i> 6322
Unit-cell parameters (Å, °)	$a = b = 97.66$, $c = 144.83$, $\alpha = \beta = 90$, $\gamma = 120$
Mosaicity (°)	0.516
Resolution range (Å)	50–2.95 (3.06–2.95)
Total No. of reflections	177240
No. of unique reflections	9068
Completeness (%)	99.7 (100)
Multiplicity	19.5 (19.2)
$\langle I/\sigma(I) \rangle$	27.98 (7.03)
$R_{\text{r.i.m.}}$	0.110 (0.532)
$R_{\text{p.i.m.}}$	0.025 (0.119)
$R_{\text{merge}}^{\dagger}$ (%)	10.7 (51.9)

$$^{\dagger} R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

niun tartrate dibasic salt. To improve the X-ray diffraction quality, the needle-like crystals were subsequently optimized by microseeding. Proteins were mixed in advance with the reservoir solution at a 1:1 ratio and equilibrated for 30 min. Several crystal clusters were scooped up using a 0.1 mm loop (Hampton Research), after which they were transferred into a tube containing 50 μl seed-stock solution. The crystals were broken up using a Seed Bead kit (Hampton Research) when vortex-mixed at $10 \times 5 \text{ s}$ bursts. The seed stock was diluted in 1:10, 1:100 and 1:1000 ratios with the seed-stock solution. Subsequently, 0.3 μl seeds with different concentrations were added to the equilibrated drops. The seed stocks were freshly prepared for each round of microseeding. Crystals suitable for data collection were obtained after microseeding in 100 mM MES-KOH pH 6.0, 24% PEG 3350, 0.2 M ammonium tartrate dibasic salt for 3 d of growth.

2.5. Data collection and processing

Fresh single crystals were harvested by transferring them with a cryoloop into the cryoprotectant solution containing the reservoir solution supplemented with 25% (v/v) ethylene glycol and were then flash-cooled in liquid nitrogen. The diffraction data were collected on Shanghai Synchrotron Radiation Facility (SSRF) beamline BL17U using a MAR CCD 225 detector. Data sets were collected at 100 K with an exposure time of 1 s per image. The best crystal diffracted to 2.95 Å resolution using a crystal-to-detector distance of 290 mm. The complete data set was subsequently processed using *HKL-2000* (Otwinowski & Minor, 1997).

3. Results and discussion

Uba5 is the smallest UBL-activating enzyme and comprises only an adenylation domain and a C-terminal extended region. Two distinct forms of Uba5 exist in humans, and these forms result from differential splicing of the full-length transcript (1–404 and 57–404). The N-terminal residues (1–56) of Uba5 are not conserved among different species, and are not necessary for the activation of Ufm1 (Zheng *et al.*, 2008). Although the crystal structure of the Uba5 adenylation domain has been determined, the molecular mechanism of the recognition and activation of Ufm1 and the physiological function of the Uba5 C-terminal region remain unknown. Residues 330–363 presumably function in the formation of an α -helix based on a secondary-structure prediction using the *PHD* method (Fig. 1*b*; Rost & Sander, 1993). In larger multi-domain E1s, the FCCH domain and other parts of the enzyme aid in the recognition and recruitment of

Figure 2

Crystals of recombinant Uba5 57–363. (a) Needle-like crystals in drops comprising 1.5 μl protein solution and 1.5 μl reservoir solution consisting of 100 mM MES-KOH pH 6.0, 23–25% PEG 3350, 0.3–0.5 M ammonium tartrate dibasic salt after 4 d growth. (b) Crystals obtained in reservoir solution consisting of 100 mM MES-KOH pH 6.0, 24% PEG 3350, 0.2 M ammonium tartrate dibasic salt after four rounds of microseeding and 3 d of growth.

Figure 3
A representative image of an X-ray diffraction pattern collected from the Uba5 57–363 crystal. The crystal diffracted to 2.95 Å resolution.

the UBL molecules. As in another noncanonical E1, Atg7, an α -helix tail follows the adenylation domain which can help in recognizing and recruiting the Atg8 molecule for activation (Hong *et al.*, 2011; Noda *et al.*, 2011; Taherbhoy *et al.*, 2011).

To identify the minimal fragment of Uba5 required for the activation of Ufm1, three different truncations containing the adenylation domain of Uba5 and different lengths of the C-terminal sequence were designed based on sequence-alignment analysis. The calculated molecular weights for Uba5 57–404, 57–363 and 57–329 were 38.8, 34.3 and 30.4 kDa, respectively, and the yields of the recombinant proteins were about 48, 60 and 40 mg per litre of culture, respectively. The ability of the truncations to activate Ufm1 was assessed using an *in vitro* activation assay. Uba5 57–404 or 57–363 can efficiently form a covalently linked complex with Ufm1 in the activation assay, as indicated by nonreducing SDS–PAGE (Fig. 1c). However, for Uba5 57–329 no obvious bands for the co-migrating complex were observed even after the reaction proceeded for 2 h. These results suggest that Uba5 57–363 is the minimal fragment for high-efficiency activation of Ufm1. The secondary structure of residues 330–363 was predicted to contain an α -helix, and a similar helix assignment was involved in Atg8 recruitment by Atg7 (Hong *et al.*, 2011; Noda *et al.*, 2011; Taherbhoy *et al.*, 2011). Thus, this C-terminal tail probably assists in the recognition and recruitment of Ufm1, similar to Atg7.

To elucidate the mechanism of recognition and activation of Ufm1 by Uba5, crystallographic studies on Uba5 57–363 were performed. High-quality crystals required for X-ray analysis were obtained after four rounds of repeated microseeding. Single rod-like crystals were visible after 2 d, and these crystals developed to their maximum

dimensions, typically $400 \times 80 \times 80 \mu\text{m}$, after 4 d (Fig. 2). Seeding beyond four rounds produced no obvious improvement in the crystal size and diffraction quality.

X-ray diffraction data were collected to 2.95 Å resolution (Fig. 3). The space group was determined to be *P*6322. The unit-cell parameters were $a = b = 97.66$, $c = 144.83$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. If one molecule is assumed to be present in the crystallographic asymmetric unit, the calculated Matthews coefficient (V_M) is $2.93 \text{ Å}^3 \text{ Da}^{-1}$ and the solvent content is 58.1% (Matthews, 1968). The data statistics are presented in Table 1. Structure solution was performed by molecular replacement using *Phaser* (McCoy *et al.*, 2007) and the structure of Uba5 adenylation domain (PDB entry 3h8v) as a search model. Refinement of the structure is in progress. This structure will provide new insights into the molecular mechanisms of Ufm1 recognition and activation by Uba5.

We thank Dr Masaaki Komatsu for generously providing the cDNA plasmid of human Uba5 and Ufm1. We are also grateful to the staff members at the SSRF for the collection of the data set. This work was supported by grants 31070643 and 31130062 from the Natural Science Foundation of China and grants 2011CB910803 and 2013CB530603 from the China National Key Basic Research Program.

References

- Bacik, J. P., Walker, J. R., Ali, M., Schimmer, A. D. & Dhe-Paganon, S. (2010). *J. Biol. Chem.* **285**, 20273–20280.
- Dou, T., Gu, S., Liu, J., Chen, F., Zeng, L., Guo, L., Xie, Y. & Mao, Y. (2005). *Mol. Biol. Rep.* **32**, 265–271.
- Gannavaram, S., Sharma, P., Duncan, R. C., Salotra, P. & Nakhasi, H. L. (2011). *PLoS One*, **6**, e16156.
- Goldberg, A. L. (2007). *Biochem. Soc. Trans.* **35**, 12–17.
- Hershko, A. (2005). *Cell Death Differ.* **12**, 1191–1197.
- Hertel, P., Daniel, J., Stegehake, D., Vaupel, H., Kailayangiri, S., Gruel, C., Woltersdorf, C. & Liebau, E. (2013). *J. Biol. Chem.* **288**, 10661–10671.
- Hong, S. B., Kim, B.-W., Lee, K.-E., Kim, S. W., Jeon, H., Kim, J. & Song, H. K. (2011). *Nature Struct. Mol. Biol.* **18**, 1323–1330.
- Komatsu, M., Chiba, T., Tatsumi, K., Iemura, S., Tanida, I., Okazaki, N., Ueno, T., Kominami, E., Natsume, T. & Tanaka, K. (2004). *EMBO J.* **23**, 1977–1986.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Noda, N. N., Satoo, K., Fujioka, Y., Kumeta, H., Ogura, K., Nakatogawa, H., Ohsumi, Y. & Inagaki, F. (2011). *Mol. Cell*, **44**, 462–475.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rost, B. & Sander, C. (1993). *J. Mol. Biol.* **232**, 584–599.
- Schulman, B. A. & Harper, J. W. (2009). *Nature Rev. Mol. Cell Biol.* **10**, 319–331.
- Taherbhoy, A. M., Tait, S. W., Kaiser, S. E., Williams, A. H., Deng, A., Nourse, A., Hammel, M., Kurinov, I., Rock, C. O., Green, D. R. & Schulman, B. A. (2011). *Mol. Cell*, **44**, 451–461.
- Tatsumi, K., Sou, Y.-S., Tada, N., Nakamura, E., Iemura, S., Natsume, T., Kang, S. H., Chung, C. H., Kasahara, M., Kominami, E., Yamamoto, M., Tanaka, K. & Komatsu, M. (2010). *J. Biol. Chem.* **285**, 5417–5427.
- Tatsumi, K., Yamamoto-Mukai, H., Shimizu, R., Waguri, S., Sou, Y.-S., Sakamoto, A., Taya, C., Shitara, H., Hara, T., Chung, C. H., Tanaka, K., Yamamoto, M. & Komatsu, M. (2011). *Nature Commun.* **2**, 181.
- Zhang, Y., Zhang, M., Wu, J., Lei, G. & Li, H. (2012). *PLoS One*, **7**, e48587.
- Zheng, M., Gu, X., Zheng, D., Yang, Z., Li, F., Zhao, J., Xie, Y., Ji, C. & Mao, Y. (2008). *J. Cell. Biochem.* **104**, 2324–2334.