Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

The expression, purification and crystallization of a ubiquitin-conjugating enzyme E2 from *Agrocybe aegerita* underscore the impact of His-tag location on recombinant protein properties

De-Feng Li,[a] Lei Feng,[a] Yan-Jie Hou[a] and Wei Liu[b]*

[a]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, and [b]Institute of Immunology, The Third Military Medical University, Chongqing 400038, People's Republic of China

Correspondence e-mail: wei.liu.2005@gmail.com

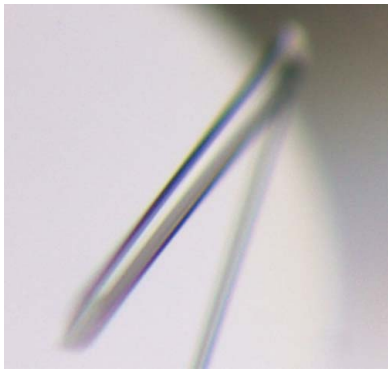
Received 5 November 2012
Accepted 23 December 2012

Ubiquitination is a post-translational modification involved in myriad cell regulation and disease pathways. The ubiquitin-conjugating (E2) enzyme is the central player in the ubiquitin-transfer pathway. Although a large array of E2 structures are available, not all E2 families have known structures and three-dimensional structures from fungal organisms other than yeast are lacking. Here, the expression, purification, crystallization and preliminary X-ray analysis of UbcA1, a novel ubiquitin-conjugating enzyme identified from the medicinal mushroom *Agrocybe aegerita*, which shows antitumour properties, are reported. As a potential anticancer drug candidate, the protein was expressed in either a C-terminally or an N-terminally His-tagged form. In the process of purification and crystallization, the location of the His tag seemed to play a crucial role in protein stability. In contrast to unsuccessful crystallization trials for the protein with a C-terminal tag, a crystal of N-terminally His-tagged UbcA1 grown under optimal conditions diffracted X-rays to 1.7 Å resolution. The crystal belonged to space group *C*2, with unit-cell parameters a = 84.93, b = 34.76, c = 128.10 Å, β = 118.57°. An X-ray data set was collected that was suitable for structure determination, showing satisfactory completeness, $\langle I/\sigma(I)\rangle$ and R factors. All of these results underscore the non-negligible impact of His-tag location on protein behaviour during the process of purification and crystallization.

1. Introduction

The ubiquitin-proteasome pathway (UPP) involves highly selective proteolytic machinery that plays crucial roles in protein quality control, cell-cycle control, proliferation, development, signal transduction, transcriptional regulation, receptor down-regulation and synaptic plasticity (Geng *et al.*, 2012; Haglund & Dikic, 2012; Mocciaro & Rape, 2012; Shang & Taylor, 2012*b*). The UPP begins with the attachment of a ubiquitin (Ub), a highly conserved 76-amino-acid polypeptide, to a substrate protein (Glickman & Ciechanover, 2002). The Ub or ubiquitin-like protein (UBL) transfer cascade requires three enzyme activities: a Ub-activating (E1) enzyme, a Ub-conjugating (E2) enzyme and a Ub ligase (E3). The concerted action of these enzymes results in substrate modification by mono-ubiquitylation or poly-ubiquitylation. Recent studies revealed that the E2 enyzme functions at the heart of the Ub-transfer pathway by governing ubiquitin-chain initiation and elongation, regulating the processivity of chain formation and establishing the topology of assembled chains (Wenzel *et al.*, 2011; Ye & Rape, 2009).

The family of E2s is characterized by the presence of a highly conserved ubiquitin-conjugating (Ubc) domain that accommodates an ATP-activated Ub or UBL *via* a covalently linked thioester onto its active-site residue (van Wijk & Timmers, 2010). E2s that contain variable N- or C-terminal extensions appended to the Ubc domain are more common, and the E2 family has accordingly been divided into four classes: class I, Ubc domain only; class II, Ubc plus a C-terminal extension; class III, Ubc plus an N-terminal extension; class IV, Ubc plus both N- and C-terminal extensions. Even the Ubc fold alone can be divided into at least 17 subfamilies (Michelle *et al.*, 2009). To date, more than 100 structures of E2 enzymes have been made available in the Protein Data Bank (PDB), most of which are of an individual Ubc domain. Nonetheless, not all E2 families have



known structures and no crystal structure from fungi other than yeast has yet been reported (Michelle *et al.*, 2009).

Agrocybe aegerita, an edible mushroom, is believed to have potential antitumour properties (Feng *et al.*, 2010; Yang *et al.*, 2009). A gene encoding a putative Ub-conjugating enzyme was identified in a recent transcriptomic analysis. The gene product, which comprises 152 amino acids and is named UbcA1, displayed remarkable pro-apoptosis activity on HeLa cells (Li *et al.*, unpublished data). A subsequent biochemical assay confirmed that UbcA1 possesses Ub-conjugating activity. Here, we describe the expression, purification, crystallization and preliminary crystallographic results of UbcA1, which is the first fungal Ub-conjugating enzyme other than yeast E2s crystallized to date. Determination of its structure would further expand our knowledge of the E2 superfamily.

2. Materials and methods

2.1. Expression and purification

The cDNA encoding *A. aegerita* UbcA1 was generously provided by Professor Hui Sun from Wuhan University. The coding sequence was first inserted into the expression plasmid pET22b (Novagen) at the *Nde*I and *Xho*I cleavage sites. A C-terminal His-tagged recombinant protein was overproduced from this construct, but the purified protein was not stable because of degradation occurring close to the C-terminus and was hence not suitable for crystallization (see §3). The ORF of UbcA1 was then subcloned into the pET28 plasmid (Novagen) at the same cleavage sites and a terminal codon juxtaposed to the *Xho*I site was introduced by PCR in order to produce an N-terminally His-tagged protein.

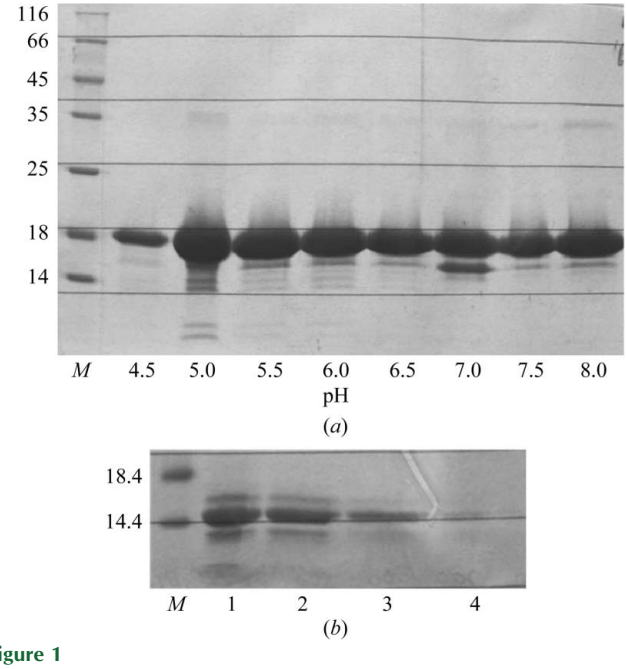


Figure 1
SDS–PAGE gels showing protein degradation occurring close to the C-terminus of recombinant UbcA1 produced from the pET22b-UbcA1 plasmid. (*a*) Protein degradation at different pH values during protein storage. Protein samples were prepared by adjusting the buffer pH after purification and were incubated for 3 d at 277 K before loading onto the gel. Lane *M* contains molecular-mass markers (labelled in kDa). (*b*) SDS–PAGE showing that the degraded fragment of UbcA1 (14 kDa) was unable to bind to an Ni–NTA column. Protein samples after a two-week incubation at 277 K were applied onto the column; the control represents the sample before loading. Lane *M*, molecular-mass markers (labelled in kDa); lane 1, control; lane 2, flowthrough; lane 3, wash; lane 4, elution.

An *Escherichia coli* BL21(DE3) host strain was used for the production of recombinant UbcA1. A small-scale culture was grown from these cells transformed overnight at 310 K in Luria Broth medium supplemented with kanamycin at a concentration of 34 µg ml^{-1}. Large-scale cultures were inoculated 1:100 with the starter culture and grown until the OD_{600} reached 0.6–0.8. The cultures were then transferred to 298 K for 15 min before 0.8 m*M* isopropyl β-D-1-thiogalactopyranoside was added to induce protein expression. Cells were harvested after 15 h incubation by centrifugation for 30 min at 4000*g* and 277 K.

After discarding the supernatant, cell pellets were resuspended in lysis buffer (50 m*M* NaH_2PO_4/Na_2HPO_4 pH 8.0, 300 m*M* NaCl, 10 m*M* imidazole). Typically, bacteria from 1 g wet cell mass were resuspended in 25 ml lysis buffer immediately before the addition of phenylmethylsulfonyl fluoride (PMSF) to a final concentration of 1 m*M* in order to prevent nonspecific proteolysis. Cells were lysed by sonication on ice at 200 W using 3 s pulses with 6 s intervals for 10 min. Cell debris was removed by centrifugation for 30 min at 16 000*g* and 277 K.

The soluble lysate was loaded onto an Ni–NTA column (15 ml column volume) pre-equilibrated with the same lysis buffer with the addition of 5%(*v*/*v*) glycerol. After complete loading, weakly bound proteins were removed by washing the column with 12 column volumes of washing buffer [50 m*M* NaH_2PO_4/Na_2HPO_4 pH 8.0, 300 m*M* NaCl, 20 m*M* imidazole, 5%(*v*/*v*) glycerol] and the recombinant UbcA1 protein was subsequently eluted from the column with elution buffer [50 m*M* NaH_2PO_4/Na_2HPO_4 pH 8.0, 300 m*M* NaCl, 250 m*M* imidazole, 5%(*v*/*v*) glycerol]. The eluted protein was concentrated to 5 ml using ultrafiltration before applying it onto a HiLoad 16/60 Superdex 75 column (GE Healthcare) pre-equilibrated with 50 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 2 m*M* DTT, 1 m*M* EDTA. The size-exclusion chromatography was run at 1.0 ml min^{-1} and 279 K. The molecular weight of the eluted fraction was estimated using an online column-calibration service at the European Molecular Biology Laboratory website (http://www.embl.de/pepcore/pepcore_services/protein_purification/chromatography/hiload16-60_superdex75/index.html).

2.2. Crystallization

The purified UbcA1 protein was concentrated to approximately 20 mg ml^{-1} by ultrafiltration and stored in the same buffer. The protein concentration was determined using the bicinchoninic acid assay (Smith *et al.*, 1985). Crystallization trials were carried out by hand using the hanging-drop vapour-diffusion and microbatch methods at room temperature. The drop in each well was formed by mixing 1 µl protein solution with 1 µl screen solution. Initial conditions were screened from approximately 500 commercial crystallization conditions using the Crystal Screen, Crystal Screen 2, Index, SaltRx, PEG/Ion, PEG/Ion 2 and PEGRx kits (Hampton Research, California, USA). A number of crystallization conditions were found using the PEG/Ion kit (Hampton Research), but gave only needles. Rod-shaped single crystals were obtained after optimization of the crystallization conditions. In subsequent optimization, the microbatch method and a dynamic seeding technique implemented in our laboratory (Zhu *et al.*, 2005) were utilized. The method basically consists of two steps. Microseeds were first seeded into a new non-equilibrated drop; after being equilibrated for various times against the reservoir solution, macroseed drops were then used to prepare a dilution series with which the crystals with improved quality could be harvested using macroseeding.

2.3. Data collection and processing

UbcA1 crystals used for data collection were dipped into the cryoprotectant (reservoir solution supplemented with 10% glycerol) for approximately 10 s before being mounted in nylon cryoloops (Hampton Research) and flash-cooled in a stream of liquid nitrogen for optimal cryoprotection. X-ray diffraction data were collected at 100 K at a wavelength of 1.0 Å on beamline BL5A of KEK, the Photon Factory, Japan, using an ADSC Q315r CCD detector at a distance of 178 mm. The collected diffraction data were indexed, integrated and scaled using *iMOSFLM* (Battye *et al.*, 2011) and *SCALA* from the *CCP*4 program suite (Winn *et al.*, 2011).

3. Results

3.1. Expression and purification

In the first trial of protein production, UbcA1 with a C-terminal His tag was expressed from the reconstituted pET22b-UbcA1 plasmid as a soluble species in the bacterial lysate. A yield of approximately 10 mg was obtained from 1 l bacterial culture and the protein purity reached approximately 95% after two steps of chromatography. However, crystallization of the purified protein was unsuccessful as only spherical crystalline precipitates were observed under a single condition. More seriously, the protein was found to be quite unstable: degradation occurred during protein storage at 277 K even in the presence of 1 m*M* PMSF as a serine protease inhibitor. A 14 kDa band appeared on SDS–PAGE within 3 d (Fig. 1*a*) and the intact protein was completely degraded after two weeks of storage (Fig. 1*b*, lane 2). The fully degraded protein was then applied onto an Ni–NTA column, but the 14 kDa species failed to bind to the affinity resin (Fig. 1*b*), indicating that a possible proteolytic site is located close to the C-terminus of the intact protein. Consistently, the C-terminal sequence (residues 120–130) was predicted as a highly flexible loop region by the *DISOPRED*2 server (Ward *et al.*, 2004). Since histidine composition at the protein termini is believed to be a good predictor of local disorder (Li *et al.*, 1999), it is tempting to speculate that the C-terminal His tag present on this construct very likely renders the neighbouring loop region more exposed and thus more vulnerable to protease attack.

Based on this analysis, we decided to produce an N-terminally His-tagged recombinant protein in order to avoid increasing C-terminal flexibility, and a pET28a-UbcA1 plasmid was made to this end. The expressed product was soluble with a comparable yield to the other construct and also purified to ∼95% (Fig. 2). The single absorption peak apart from the flowthrough fraction on the elution profile of size-exclusion chromatography contained homogeneous UbcA1 with an estimated molecular weight of 42 kDa (Fig. 2*a*), indicating a dimeric form of the recombinant protein. A single band at 18 kDa observed on SDS–PAGE suggested that no detectable degradation had occurred in the protein expressed from this construct (Fig. 2*b*).

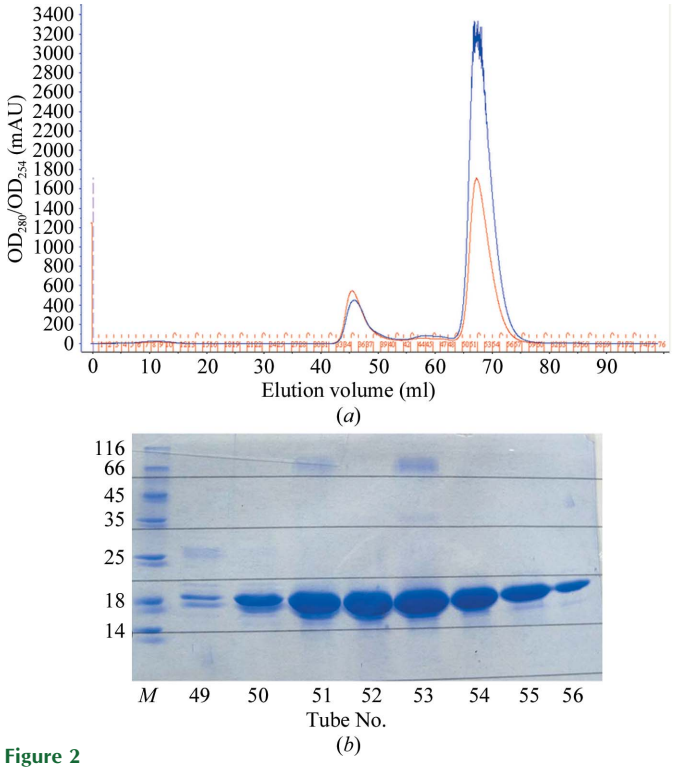


Figure 2
The purification results of N-terminally His-tagged UbcA1 expressed from the reconstituted pET28a-UbcA1 plasmid. (*a*) Chromatogram of the second round of gel filtration on a HiLoad 16/600 Superdex 75 column. The single absorption peak apart from the flowthrough fraction (the minor peak) contained the homogeneous UbcAl protein with an estimated molecular weight of 42 kDa. (*b*) The eluted fraction from size-exclusion chromatography detected by SDS–PAGE. The major band at 17 kDa indicated no detectable degradation of the expressed product from this construct. Lane *M* contains molecular-mass markers (labelled in kDa).

3.2. Crystallization

As mentioned above, no suitable crystallization conditions for C-terminally His-tagged Ubc-A1 were found from the initial screening because of protein instability. In sharp contrast, a number of conditions for the N-terminally His-tagged protein were established, in particular from the PEG/Ion kit (Hampton Research). All of these conditions led to the growth of single needles or needle clusters. After optimization, rod-shaped single crystals were obtained within 24 h using a condition consisting of 20%(*w*/*v*) PEG 3350, 0.2 *M* sodium acetate, 0.1 *M* bis-tris pH 6.5. However, the crystals grown under this condition seemed to be quite unstable as they were prone to redissolve 1 week after the crystal reached maximal size, leaving visible dents on the crystal surface and resulting in poor quality crystals that were not suitable for X-ray data collection. This adverse effect might be attributed to rapid crystal growth, which could be overcome by decreasing the diffusion rate. The microbatch method with sample drops sitting under paraffin oil was then used for this

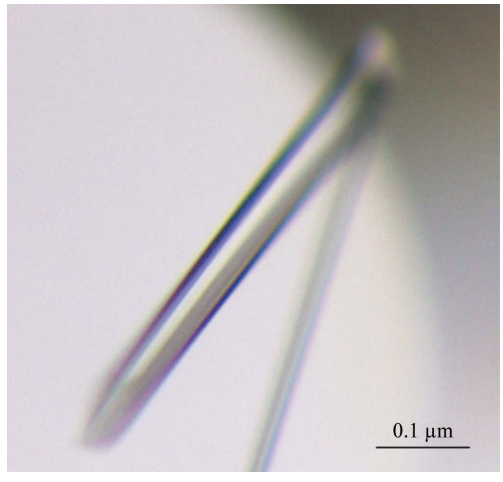


Figure 3
An optimized crystal of UbcA1 grown under the condition 20%(*w*/*v*) PEG 3350, 0.2 *M* sodium acetate, 0.1 *M* bis-tris pH 6.5 using the microbatch method and the dynamic seeding technique (Zhu *et al.*, 2005).

Table 1
Summary of native X-ray data collection.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.0
Space group	*C*2
Unit-cell parameters (Å, °)	a = 84.93, b = 34.76, c = 128.10, β = 118.57
Resolution range (Å)	42.45–1.70 (1.79–1.70)
No. of unique reflections	37008 (3958)
Multiplicity	3.4 (2.3)
Completeness (%)	94.1 (72.2)
Mean $I/\sigma(I)$	9.7 (3.4)
Solvent content (%)	41.7%
$R_{\rm merge}$† (%)	4.4 (20.4)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

purpose. As expected, crystals grown by these means were much more stable. A dynamic seeding technique implemented in our laboratory (Zhu *et al.*, 2005) was also utilized to grow larger crystals. Diffraction-quality crystals with approximate dimensions of 500 × 50 × 40 µm were grown using these techniques (Fig. 3).

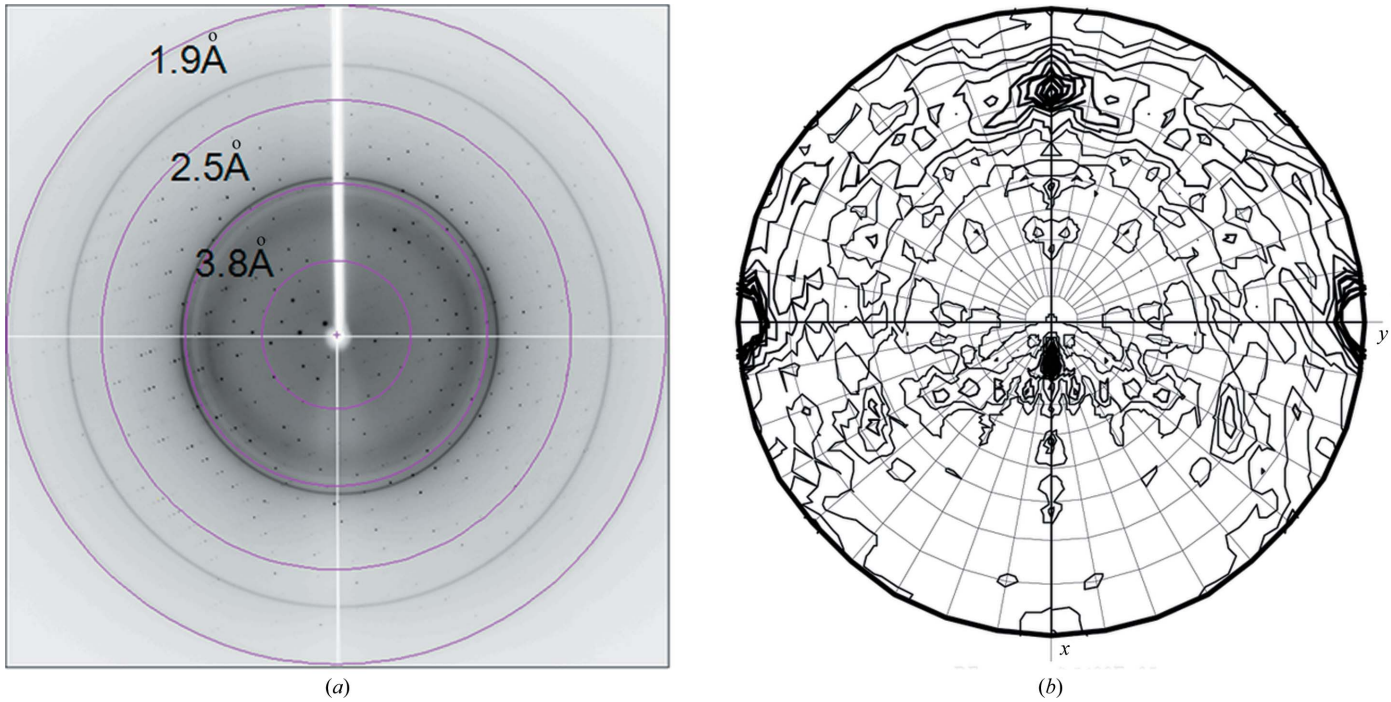
3.3. Data collection and evaluation

A crystal grown under the optimal condition using the microbatch method and the dynamic seeding technique was flash-cooled just before a complete diffraction data set was collected to 1.7 Å resolution on beamline BL5A at KEK, the Photon Factory, Japan. A total of 180 frames were recorded. Indexing results showed that the crystal was monoclinic, space group *C*2, with unit-cell parameters a = 84.93, b = 34.76, c = 128.10 Å, β = 118.57°, giving a value for the Matthews coefficient (Matthews, 1968) of 2.25 Å^3 Da^{-1} and a solvent content of 41.7%, assuming the presence of two protomers of UbcA1 in the asymmetric unit. Twofold noncrystallographic symmetry between the two monomers was clearly shown on a self-rotation map calculated using the diffraction data (Fig. 4*b*). All statistical results are given in Table 1 and a representative diffraction image is shown in Fig. 4(*a*).

4. Discussion

The His tag, which very often consists of six consecutive histidine residues, has gained great popularity over the past 15 years as a purification tool for recombinant proteins. Increasingly, determined structures are of protein constructs that contain a His tag at the N- or C-terminus, accounting for nearly 60% of the crystal structures deposited in the Protein Data Bank (Derewenda, 2004). The impact of the His tag on native protein structures has been systematically surveyed and well documented (Carson *et al.*, 2007). Statistical comparison of structures with or without His tags showed that the His tag had no significant effect on the structures of native proteins, but was protein-target specific (Carson *et al.*, 2007). Numerous cases have exemplified either advantageous or adverse effects of the His tag on protein crystallization, *i.e.* some proteins can only be crystallized without a His tag, while others could only be crystallized with a His tag. However, the impact of the location of the His tag has not been systematically analysed, although several reports have shown that the composition and position of His tags might influence protein expression (Doray *et al.*, 2001; Mast *et al.*, 2004; Woestenenk *et al.*, 2004).

In our study, the location of the His tag seemed to be crucial for protein stability and crystallization of UbcA1. Degradation of C-terminally His-tagged UbcA1 occurred, but did not for the protein with an N-terminal His tag; a number of crystallization conditions were established for the latter, compared with none for the former. All of these results underscore the important influence of the location of the His tag on protein behaviour in the process of purification and crystallization. Hence, the case of UbcA1 is a good example



(*a*) (*b*)

Figure 4
(*a*) A diffraction image of a crystal of UbcA1 obtained using an ADSC Q315r CCD detector on beamline BL5A of KEK, the Photon Factory, Japan. (*b*) A self-rotation function map calculated from the diffraction data displaying twofold noncrystallographic symmetry between the two protomers of UbcA1 in the asymmetric unit.

suggesting that optimization of the His tag location is a good strategy to increase the stability and crystallizability of a recombinant protein.

The ubiquitin-proteasome system is a highly selective proteolytic system that plays crucial roles in various cellular processes and is closely linked to a number of diseases including cancer, infectious diseases, cardiovascular diseases and neurodegenerative disorders (Shang & Taylor, 2012*a*). UbcA1 from *A. aegerita*, a medicinal mushroom that exhibits potential antitumour properties, showed remarkable pro-apoptotic activity towards HeLa cells, suggesting a possible role in cancer therapy. The successful crystallization of this Ub-conjugating enzyme has established a good starting point for the determination of the first mushroom E2 structure as a potential antitumour drug candidate.

This work was supported by the '973' (2006CB806502 and 2006CB910901), '863' (2006AA02A322) research grants and the National Natural Science Foundation of China (grant Nos. 30725001 and 31270788). The authors would like to thank the beamline scientists working at beamline BL5A of KEK, the Photon Factory, Japan, for their kind help with data collection.

References

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. (2007). *Acta Cryst.* D**63**, 295–301.
Derewenda, Z. S. (2004). *Methods*, **34**, 354–363.
Doray, B., Chen, C.-D. & Kemper, B. (2001). *Arch. Biochem. Biophys.* **393**, 143–153.
Feng, L., Sun, H., Zhang, Y., Li, D.-F. & Wang, D.-C. (2010). *FASEB J.* **24**, 3861–3868.
Geng, F., Wenzel, S. & Tansey, W. P. (2012). *Annu. Rev. Biochem.* **81**, 177–201.
Glickman, M. H. & Ciechanover, A. (2002). *Physiol. Rev.* **82**, 373–428.
Haglund, K. & Dikic, I. (2012). *J. Cell Sci.* **125**, 265–275.
Li, X., Romero, P., Rani, M., Dunker, A. K. & Obradovic, Z. (1999). *Genome Inform. Ser. Workshop Genome Inform.* **10**, 30–40.
Mast, N., Andersson, U., Nakayama, K., Bjorkhem, I. & Pikuleva, I. A. (2004). *Arch. Biochem. Biophys.* **428**, 99–108.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mocciaro, A. & Rape, M. (2012). *J. Cell Sci.* **125**, 255–263.
Michelle, C., Vourc'h, P., Mignon, L. & Andres, C. R. (2009). *J. Mol. Evol.* **68**, 616–628.
Shang, F. & Taylor, A. (2012*a*). *Mol. Aspects Med.* **33**, 446–466.
Shang, F. & Taylor, A. (2012*b*). *Prog. Mol. Biol. Transl. Sci.* **109**, 347–396.
Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985). *Anal. Biochem.* **150**, 76–85.
Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. & Jones, D. T. (2004). *Bioinformatics*, **20**, 2138–2139.
Wenzel, D. M., Stoll, K. E. & Klevit, R. E. (2011). *Biochem. J.* **433**, 31–42.
Wijk, S. J. van & Timmers, H. T. (2010). *FASEB J.* **24**, 981–993.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Woestenenk, E. A., Hammarström, M., van den Berg, S., Härd, T. & Berglund, H. (2004). *J. Struct. Funct. Genomics*, **5**, 217–229.
Yang, N., Li, D.-F., Feng, L., Xiang, Y., Liu, W., Sun, H. & Wang, D.-C. (2009). *J. Mol. Biol.* **387**, 694–705.
Ye, Y. & Rape, M. (2009). *Nature Rev. Mol. Cell Biol.* **10**, 755–764.
Zhu, D.-Y., Zhu, Y.-Q., Xiang, Y. & Wang, D.-C. (2005). *Acta Cryst.* D**61**, 772–775.