Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the autophagic E2 enzyme Atg10

Seung Beom Hong, Byeong-Won Kim, Jun Hoe Kim and Hyun Kyu Song*

School of Life Sciences and Biotechnology, Korea University, Seoul 136-701, Republic of Korea

Correspondence e-mail: hksong@korea.ac.kr

Received 11 June 2012
Accepted 31 July 2012

PDB Reference: Atg10, 4ebr

Autophagy is a regulated degradation pathway that plays a critical role in all eukaryotic life cycles. One interesting feature of the core autophagic process, autophagosome formation, is similar to ubiquitination. One of two autophagic E2 enzymes, Atg10, interacts with Atg7 to receive Atg12, a ubiquitin-like molecule, and is also involved in the Atg12–Atg5 conjugation reaction. To date, no information on the interaction between Atg10 and Atg7 has been reported, although structural information is available pertaining to the individual components. Here, the crystal structure of Atg10 from *Saccharomyces cerevisiae* is described at 2.7 Å resolution. A significant improvement of the diffraction limit by heavy-atom derivatization was essential for structure determination. The core fold of yeast Atg10 is well conserved compared with those of Atg3 and other E2 enzymes. In contrast to other E2 enzymes, however, the autophagic E2 enzymes Atg3 and Atg10 possess insertion regions in the middle of the core fold and may be involved in protein function. The missing segment, which was termed the 'FR-region', in Atg10 may be important for interaction with the E1 enzyme Atg7. This study provides a framework for understanding the E2 conjugation reaction in autophagy.

1. Introduction

Autophagy represents the bulk degradation of a cell's own components by the lysosomal machinery (or the vacuole in yeast) and is tightly regulated to ensure a balance between the synthesis and degradation of cellular products (Nakatogawa *et al.*, 2009; Klionsky & Emr, 2000). It is a major mechanism for recycling nutrients from unnecessary to essential cellular processes in eukaryotes (Klionsky & Emr, 2000). Recently, a wealth of molecular details concerning autophagy have been reported such as the identification of new *atg* genes, the mechanism of autophagosome formation, disease progression related to autophagy and biochemical and structural studies of Atg proteins (Klionsky *et al.*, 2003; Noda *et al.*, 2009, 2011; Nakatogawa *et al.*, 2009; Suzuki & Ohsumi, 2010; Shintani & Klionsky, 2004; Taherbhoy *et al.*, 2011, 2012; Hong *et al.*, 2011).

One interesting molecular feature of the early stage of the autophagic process is that it is analogous to ubiquitination (Taherbhoy *et al.*, 2012; Nakatogawa *et al.*, 2009; Geng & Klionsky, 2008). Many known ubiquitin-like (Ubl) modifiers are conjugated to their targets through sequential enzymatic reactions catalyzed by E1 activating, E2 conjugating and E3 ligase enzymes (Varshavsky, 2005; Schulman & Harper, 2009).

Two distant cousins of ubiquitin, Atg8 and Atg12, provide a carboxyl-terminal glycine to the target and act as tagging proteins in autophagy (Noda *et al.*, 2008; Geng & Klionsky, 2008). The E1-like enzyme Atg7 activates Atg8 and Atg12 in an ATP-dependent manner and these are then transferred to the E2-like enzymes Atg3 and Atg10, respectively (Noda *et al.*, 2009). It is intriguing that two different Ubl molecules share the same activating enzyme, Atg7, and are then separately transferred to their respective partner E2 enzymes. In the pathway involving Atg10 as an E2-like enzyme, an isopeptide linkage is formed between the carboxyl-terminal glycine of Atg12 and the side chain of Lys149 of Atg5 (Mizushima *et al.*, 1998) and Atg10 plays a critical role in Atg12–Atg5 conjugation. However, Atg3, another E2-like enzyme, catalyzes the covalent modification of the carboxyl-terminal glycine of Atg8 using the lipid phosphatidylethanolamine (PE; Ichimura *et al.*, 2000).

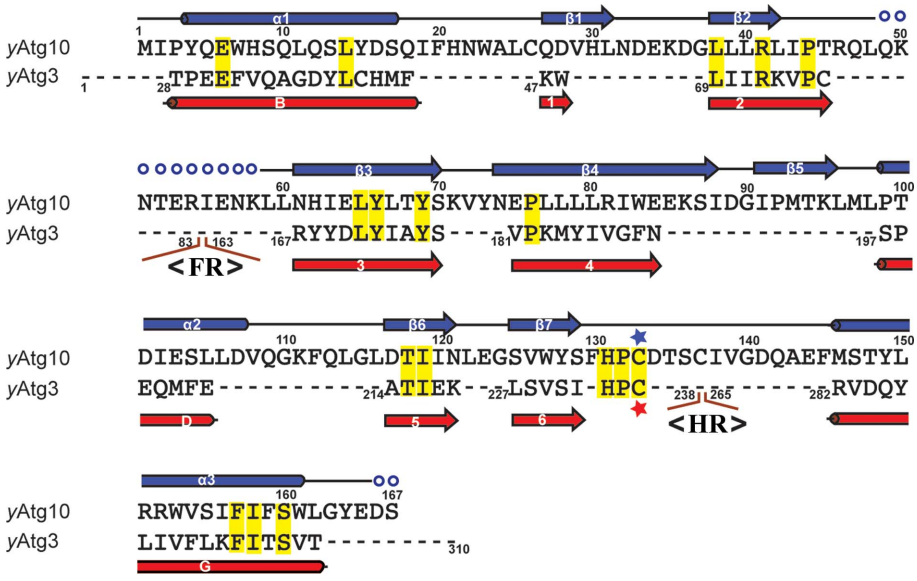

Figure 1
Structure-based sequence alignment between Atg10 from yeast (yAtg10) and Atg3 from yeast (yAtg3). The catalytic residues of yAtg10 (Cys133) and yAtg3 (Cys234) are marked by a blue and a brown star, respectively. Conserved residues are shaded in yellow. Secondary-structure elements are indicated above the sequence for yAtg10 (blue) and below for yAtg3 (brown). Invisible residues in the yAtg10 structure are shown as blue open circles. Seven β-strands (arrows) and three α-helices (cylinders) are numbered sequentially for yAtg10, whereas β-strands are numbered sequentially and α-helices are labelled alphabetically for yAtg3 according to a previous report (Yamada *et al.*, 2007). Structurally deviating residues in yAtg3 are omitted for clarity and the positions of the unique FR-region and HR-region in yAtg3 are indicated.

Previous structural investigations of Atg3 showed that it has significant structural homology to E2 ubiquitin-conjugating enzymes, although two characteristic regions referred to as the FR-region and the HR-region in Atg3 have been proposed to account for its unique reaction (Yamada *et al.*, 2007). In particular, the isolated FR-region has been shown to bind the N-terminal domain of Atg7 (Hong *et al.*, 2011; Taherbhoy *et al.*, 2011). Hydrophobic residues in the FR-region of Atg3 interact with residues in the specific groove of Atg7 (Taherbhoy *et al.*, 2011). Furthermore, Atg3 recognizes Atg8 using the Trp-Xaa-Xaa-Leu sequence motif which is located in the HR-region of Atg3 (Yamada *et al.*, 2007) and is found in various autophagy-related molecules (Noda *et al.*, 2010). Intriguingly, Atg10 lacks the unique FR-region and the Trp-Xaa-Xaa-Leu motif in its primary structure (Fig. 1 and Supplementary Fig. S1[1]) and therefore the mechanism by which Atg10 interacts with its interacting partners Atg7 and Atg12 has remained elusive. While this paper was in preparation, a solution NMR structure of Atg10 from *Kluyveromyces marxianus* was reported (Yamaguchi *et al.*, 2012), in which the authors focused on the interaction between Atg10 and an Atg12–Atg5 conjugate but not that between Atg10 and Atg7.

Here, we describe the structure of Atg10 from *Saccharomyces cerevisiae* (yAtg10) at 2.7 Å resolution. The structure determination was not straightforward owing to the poor quality of the initial crystals and was overcome by the use of biochemical characterization and heavy-atom derivatization. Although the core fold of yAtg10 is well conserved compared with that of Atg3, we identified unique features in the structure of yAtg10. Our biochemical data confirmed that the autophagic E2 enzymes Atg3 and Atg10 share the same binding pocket of E1 enzyme Atg7. Taken together, this study provides a framework for understanding the E2 conjugation reaction in autophagy.

2. Materials and methods

2.1. Overexpression and purification

Full-length yAtg10 was cloned using standard PCR techniques. The amplified products were treated with the restriction enzymes *Bam*HI and *Eco*RI and inserted into a modified pET vector for the construction of GST-tagged protein. The resultant plasmids were transformed into *Escherichia coli* BL21 (DE3) cells. The integrity of the cloned *atg10* gene was verified by DNA sequencing. Expression of yAtg10 was induced by the addition of 0.5 m*M* IPTG at 291 K for 24 h. Following sonication, the cell lysate was loaded onto a glutathione-Sepharose 4B affinity column (GE Healthcare) and the GST tag was cleaved using tobacco etch virus (TEV) protease at a molar ratio of 1:10. The resultant yAtg10 protein (starting with Gly-Ser amino-acid residues ahead of the first Met) was

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MH5068). Services for accessing this material are described at the back of the journal.

further purified by anion-exchange column chromatography using 5 ml Q FF (GE Healthcare) and a second passage through a GST affinity column to remove any GST contaminants. Finally, yAtg10 was loaded onto a Superdex 75 gel-filtration column (GE Healthcare) pre-equilibrated with 50 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 2 m*M* Tris(2-carboxyethyl)phosphine hydrochloride (TCEP–HCl). Selenomethionine-substituted yAtg10 was expressed in methionine-auxotrophic *E. coli* B834 (DE3) cells in defined medium. Selenomethionyl yAtg10 and yAtg10ΔFR mutant (with a 13-residue deletion from Gln47 to Leu59 and containing a three-glycine-residue linker) were purified in a similar manner to the wild-type protein.

2.2. Size-exclusion chromatography with multi-angle light scattering

Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) experiments were performed using an FPLC system (GE Healthcare) coupled to a Wyatt MiniDAWN TREOS MALS instrument and a Wyatt Optilab rEX differential refractometer. For chromatographic separation, a Superose 12 or Superdex 200 10/300 GL size-exclusion column (GE Healthcare) with a 500 µl sample loop was used at a flow rate of 0.5 ml min^{-1} in running buffer comprising 50 m*M* Tris–HCl pH 8.0, 500 m*M* (or 50 m*M*) NaCl, 2 m*M* (or 1 m*M*) TCEP. The outputs were analyzed using the *ASTRA V* software (Wyatt). QELS signals, together with the protein concentration determined from the refractive index, were used to calculate the molecular mass of the complex. QELS data were collected by TREOS simultaneously to measure the hydrodynamic radius of yAtg10. The solvent viscosity was calculated using the *SEDNTERP* 1.09 program (http://bitc-wiki.sr.unh.edu/).

2.3. Heavy-atom derivatization screening and native gel electrophoresis

Purified yAtg10 protein was screened for derivatization with various heavy-atom solutions (Boggon & Shapiro, 2000). 10 µl protein solution (2 mg ml^{-1}) was mixed with 10 µl heavy-atom solution (10 m*M* of each heavy-atom compound in 10 m*M* bis-Tris pH 6.0) and left on ice for 10 min. 10 µl of the reaction solution was then loaded onto a polyacrylamide gel in the absence of SDS and reducing agent (Lee *et al.*, 2010). The derivatization was visualized by Coomassie Blue staining following gel electrophoresis (Supplementary Fig. S2*a*). Potential heavy-atom derivatives were used in crystallization experiments.

Table 1
Data-collection and refinement statistics for Atg10.

Values in parentheses are for the highest resolution shell.

		MAD		
	SeMet†	Peak	Inflection	Remote
Data collection				
X-ray source‡	NW12, PF	MX4A, PAL		
Wavelength (Å)	1.0	0.97951	0.97966	0.96
Resolution (Å)	2.70 (2.75–2.70)	3.50 (3.63–3.50)		
Space group	$P6_122$	$P6_122$		
Unit-cell parameters (Å)	a = 127.8, c = 169.6	a = 128.0, c = 169.8	a = 128.1, c = 169.9	a = 128.2, c = 169.8
R_{merge}§	0.060 (0.473)	0.095 (0.299)	0.094 (0.313)	0.095 (0.360)
$\langle I/\sigma(I)\rangle$	85.7 (9.88)	48.9 (13.1)	47.7 (11.7)	48.8 (13.2)
Completeness (%)	99.5 (100)	100 (99.9)	100 (99.9)	100 (100)
No. of reflections	569556	262728	263138	264149
Multiplicity	24.8 (26.6)	24.1 (25.1)	24.1 (25.1)	24.1 (25.2)
Figure of merit		0.65 [after DM]		
Refinement				
Resolution (Å)	27.8–2.70			
No. of reflections	21751			
R_{work}/R_{free}¶	0.223/0.251			
Asymmetric unit	2 monomers			
No. of atoms				
Protein	2590 [8 Se]			
Heteroatoms	2 [Hg]			
Water	31			
Average B factor (Å^2)				
Protein	67.5			
Heteroatoms	79.9			
Water	51.5			
R.m.s. deviations				
Bond lengths (Å)	0.012			
Bond angles (°)	1.44			
Ramachandran favoured (%)	96.4			
Ramachandran outliers (%)	0.0			
PDB code	4ebr			

† The highest resolution data obtained using a selenomethionyl-derivatized crystal. ‡ PF, Photon Factory; PAL, Pohang Accelerator Laboratory. § $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of hkl and $\langle I(hkl)\rangle$ is the corresponding average value for all i measurements. ¶ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$; R_{free} was calculated using 5% of the data.

2.4. Crystallization and data collection

The purified protein was concentrated to 10 mg ml^{-1} for crystallization. Crystallization was performed by the hanging-drop vapour-diffusion method at 295 K, mixing equal volumes of yAtg10 protein solution and reservoir solution. Crystals were obtained within 3 d using a reservoir solution consisting of 50 m*M* bis-Tris–HCl pH 6.15, 3.0 *M* NaCl, 1 m*M* 1,4-di-acetoxymercury-2,3-dimethoxybutane (Supplementary Fig. S2). For cryocooling, a crystal was transferred to reservoir solution containing 25%(*v*/*v*) glycerol prior to flash-cooling in a nitrogen stream at 100 K. The crystals belonged to the hexagonal space group $P6_122$, with unit-cell parameters $a = b = 127.8$, $c = 169.6$ Å, and contained two molecules of yAtg10 in the asymmetric unit. Multi-wavelength anomalous diffraction (MAD) data were collected on beamline 4A of Pohang Accelerator Laboratory, Pohang, Republic of Korea and high-resolution data were collected on the NW12 beamline of Photon Factory, Tsukuba, Japan. Diffraction data were indexed, integrated and scaled using the *HKL*-2000 software

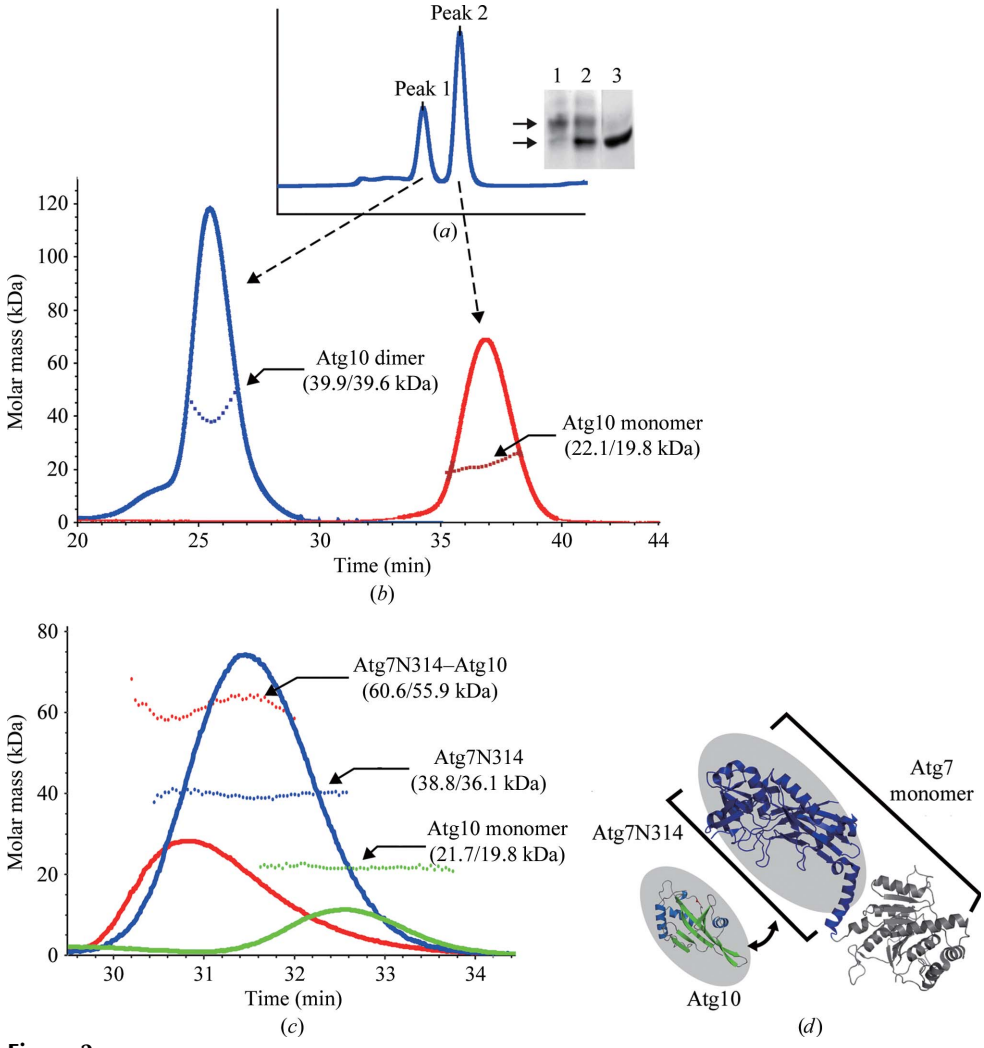

Figure 2
(*a*) Gel-filtration profile and native PAGE results of yAtg10. The protein exists in two different oligomeric states in solution (peaks 1 and 2). Lanes 1 and 2 represent yAtg10 protein from peaks 1 and 2, respectively. Lane 3 represents protein from peak 2 with SDS denaturing agent. (*b*) SEC-MALS data of yAtg10 in the presence of 2 m*M* TCEP and 500 m*M* NaCl. The gel-filtration elution profiles (OD_{280}, 0.5 ml min^{-1}) are shown as solid lines. The averaged molar mass determined with the *ASTRA V* program from UV data obtained using a UPC-9 UV detector and QELS data obtained using a Wyatt MiniDAWN TREOS instrument are shown as dotted lines. Each species is indicated by an arrow with experimental (MALS) and theoretically calculated (calc.) molar-mass values shown in parentheses (MALS/calc.) (*c*) SEC-MALS result confirming a 1:1 complex between Atg7N314 (1–314) and Atg10. Red, blue and green profiles represent Atg7N314–Atg10, Atg7N314 and Atg10, respectively. (*d*) Atg7N314 drawn as a purple ribbon and highlighted by a transparent grey oval in the full-length Atg7 structure, which is shown as a monomeric chain for clarity. This domain is known as the E2 enzyme-binding region; one of the E2 enzymes, Atg10, is shown in a ribbon-diagram form highlighted by a transparent oval.

package (Otwinowski & Minor, 1997). The statistics of the collected data are summarized in Table 1.

2.5. Structure determination and refinement

Eight possible selenium sites in the asymmetric unit were located using the three-wavelength MAD data. Initial phases were calculated and an initial model was partially built using the *PHENIX* software package (Adams *et al.*, 2010). The model was rebuilt manually using *Coot* and *O* (Emsley & Cowtan, 2004; Jones *et al.*, 1991). Refinement was also carried out using the *PHENIX* software package. Twofold noncrystallographic symmetry (NCS) restraints were maintained during initial refinement and were relaxed in the final stage. *MolProbity* (Chen *et al.*, 2010) and *STRIDE* (Heinig & Frishman, 2004) were used to assess model geometry and to assign secondary-structure elements. The *DALI* server (Holm & Sander, 1993; http://ekhidna.biocenter.helsinki.fi/dali_server/) was used for structural comparisons. The data-collection, phasing and refinement statistics are summarized in Table 1. All structural images were produced using *PyMOL* (http://www.pymol.org/).

3. Results and discussion

3.1. Biochemical characterization and crystal improvement

yAtg10 exists in at least two forms in solution based on the initial gel-filtration results and native gel analyses (Fig. 2*a*). The two major forms were confirmed as monomeric and dimeric yAtg10 by the SEC-MALS experiment (Fig. 2*b*) and given that a high concentration of salt (500 m*M* NaCl) in the buffer could reduce the number of oligomeric states. Crystallization efforts using monomeric yAtg10 yielded crystals, unlike when dimeric yAtg10 was used. One of the several crystallization conditions employed was very similar to that previously reported and the crystallographic parameters were basically the same (Yamaguti *et al.*, 2007). However, the quality of the data obtained from this tetragonal crystal form was insufficient to overcome the phase problem owing to a lack of data completeness along an axis, extremely weak diffraction beyond 4.0 Å resolution and poor merging statistics of the data (data not shown).

Another important factor in obtaining better crystals besides avoiding heterogeneous oligomeric states was to minimize the tendency of disulfide bridges to form during concentration and storage at 277 K. Disulfide-bridge formation was observed even in the presence of 1–2 m*M* dithiothreitol (DTT). A more homogeneous monomeric sample was obtained when TCEP was employed as an alternative reducing

agent. Crystallization of monomeric yAtg10 in the presence of a high concentration of salt and TCEP yielded a different crystal form. However, as the diffraction of this crystal was limited to only 4.5 Å resolution, heavy-atom derivatization was employed since several reports have shown that prior or post derivatization of protein crystals yielded a significant improvement in their diffraction limits (Xu *et al.*, 2004; Chang *et al.*, 1998; Green & Luo, 2006). Heavy-atom screening using native gel-shift assays indicated that yAtg10 reacts well with many mercury compounds (Boggon & Shapiro, 2000) and the best crystal was obtained on addition of 1 m*M* 1,4-diacetoxymercury-2,3-dimethoxybutane (Supplementary Fig. S2). This mercury-derivatized crystal diffracted to 2.7 Å resolution using synchrotron X-rays (Supplementary Fig. S2*c*) and all data statistics indicated that it was suitable for structure determination (Table 1).

3.2. Structure determination

The crystal structure of yAtg10 was determined by the MAD method and refined to an R_{work} and R_{free} of 0.223 and 0.251, respectively, at 2.7 Å resolution (Table 1). Two molecules of yAtg10 in the crystalline lattice suggested a potential homodimer (Supplementary Fig. S3), although MALS analysis indicated that yAtg10 exists as a monomer in solution (Fig. 2*b*). Furthermore, the yAtg10 monomer forms a 1:1 complex with the N-terminal domain of Atg7, which is a monomer in solution (Noda *et al.*, 2011; Taherbhoy *et al.*, 2011; Hong *et al.*, 2011), indicating that the yAtg10 monomer is a functional biological unit (Figs. 2*c* and 2*d*). The yAtg10 model in each chain accounts for 159 amino-acid residues. Two extra residues (Gly-Ser from the expression vector) at the N-terminus are clearly observed in the electron-density maps and have been built in the final model. However, ten residues (49–58) on the surface loop and two C-terminal residues (Asp166 and Ser167) could not be built owing to their absence from the electron-density map, perhaps as a consequence of their flexible nature.

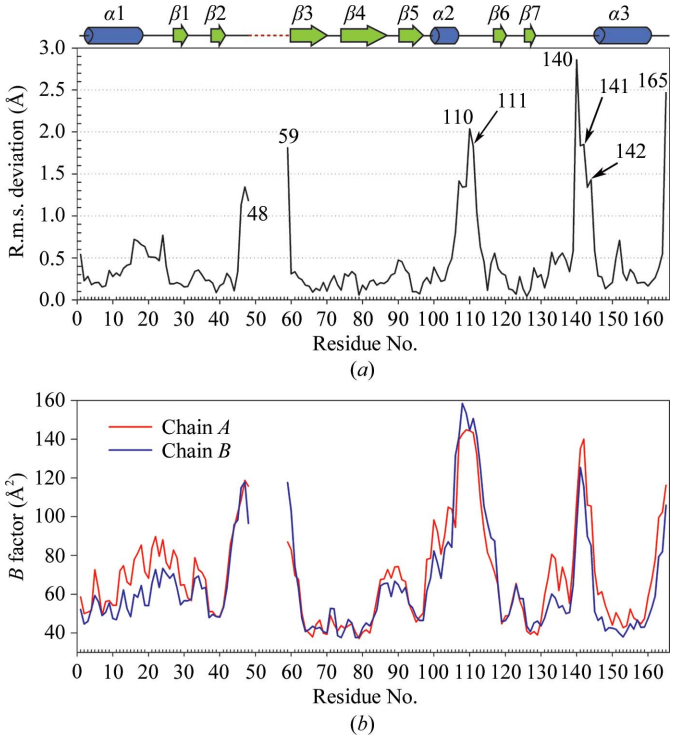

Figure 3
(*a*) Plot of the difference between each chain of yAtg10 in the asymmetric unit. The r.m.s. deviations for C^{α} atoms of each residue are plotted as a function of residue number. The residues with the largest deviations (>1.5 Å) are indicated. (*b*) *B*-factor plot for C^{α} atoms of each chain in the yAtg10 model. The *B*-factor distribution between chains *A* (red) and *B* (blue) is very similar. Residues 49–58 are missing in the structure and subsequently there are no calculations relating to r.m.s. deviations and *B*-factor distribution for these residues. α-Helices and β-strands are shown as blue cylinders and green arrows, respectively, and the missing region is indicated by red dots.

We first analyzed the structural differences between the two molecules in the asymmetric unit (Fig. 3*a*). When 157 matching C^{α} atoms (omitting the two extra residues at the N-terminus) of each yAtg10 chain were superposed, the root-mean-square (r.m.s.) deviation was 0.64 Å. As shown in Fig. 3(*a*), substantially deviating parts (>1.5 Å) were only found in loop regions between the β2 and β3 strands, between the α2 helix and the β6 strand and between the β7 strand and the α3 helix. The recalculated r.m.s. deviation after excluding the seven residues with the largest deviations (Leu59, Gly110, Lys111, Gly140, Asp141, Gln142 and Glu165) was reduced to only 0.45 Å. The electron-density maps of these regions were also weak and the temperature-factor distribution for each chain of yAtg10 was basically the same as that of the r.m.s. deviations (Fig. 3*b*), supporting the notion that the aforementioned loop regions are intrinsically flexible. Therefore, the two yAtg10 models in the asymmetric unit are essentially very similar except for the flexible loops, and one asymmetric unit was arbitrarily selected to describe the structure.

3.3. Overall structure

yAtg10 has an elongated shape in one direction, with approximate dimensions of 65 × 41 × 40 Å (Fig. 4*a*). Two α-helices, α1 (Tyr4–Ser17) and α3 (Met146–Trp161), form one face of the protein, while an antiparallel β-sheet formed by five β-strands (β1–β5) is located on the back side followed by α-helix α2 (Pro99–Asp107) and another antiparallel β-sheet formed by the β6 and β7 strands to the side of the β4 strand (Fig. 4*a*). The active region near the key Cys133 residue is located in a loop region immediately after the β7 strand. Interestingly, this cysteine (Cys133) and a neighbouring cysteine residue (Cys137) are derivatized by Hg atoms (Fig. 4*b*) and may assist in stabilizing this flexible loop region and improving crystal formation. As expected, the key cysteine residue is exposed to the solvent (Fig. 4); this residue forms an intermediate covalent bond with the C-terminal glycine residue of Atg12 (Geng & Klionsky, 2008; Shintani *et al.*, 1999). The surface properties of Atg10 show clusters of hydrophobic residues which are not prominent (Fig. 4*c*), and the electrostatic surface potential shows an even distribution of positively and negatively charged residues (Fig. 4*d*). However, as described, ten residues (49–58) on the loop

connecting the β2 and β3 strands as well as two C-terminal residues (Asp166 and Ser167) are missing in the current model. The internal missing segment consists of four polar, three positively charged, two negatively charged and one hydrophobic residue (Fig. 1). This segment is located on the same side as the catalytic residue at some distance (Figs. 4*c* and 4*d*) and therefore the surface nature of yAtg10 must be quite different. Nonetheless, the structurally equivalent region to these missing ten residues in yAtg3 is the FR-region (Fig. 1), which also shows extreme flexibility (Yamada *et al.*, 2007). Only an α-helical segment in the FR-region has been structurally characterized and is responsible for interaction with Atg7 (Taherbhoy *et al.*, 2011; Noda *et al.*, 2011; Hong *et al.*, 2011). Interestingly, an α-helical structure in this region was also predicted using the *PSIPRED* secondary-structure server (http://bioinf.cs.ucl.ac.uk/psipred/; Jones, 1999).

3.4. Structural comparison with other E2 enzymes

Structural similarities of yAtg10 to yAtg3 and E2 ubiquitin-conjugating enzymes were anticipated on the basis of functional similarity, despite their low sequence homology (approximately 20% sequence identity or less, even with large alignment gaps). The proteins that were identified to show the highest structural similarity using the *DALI* server are Atg3 (Z = 11.6), F-actin capping protein subunit α1 (Z = 6.1), human frataxin (Z = 5.5), ubiquitin-conjugating enzyme E2 from *Plasmodium falciparum* (Z = 5.2), the UbcH7 E2 enzyme

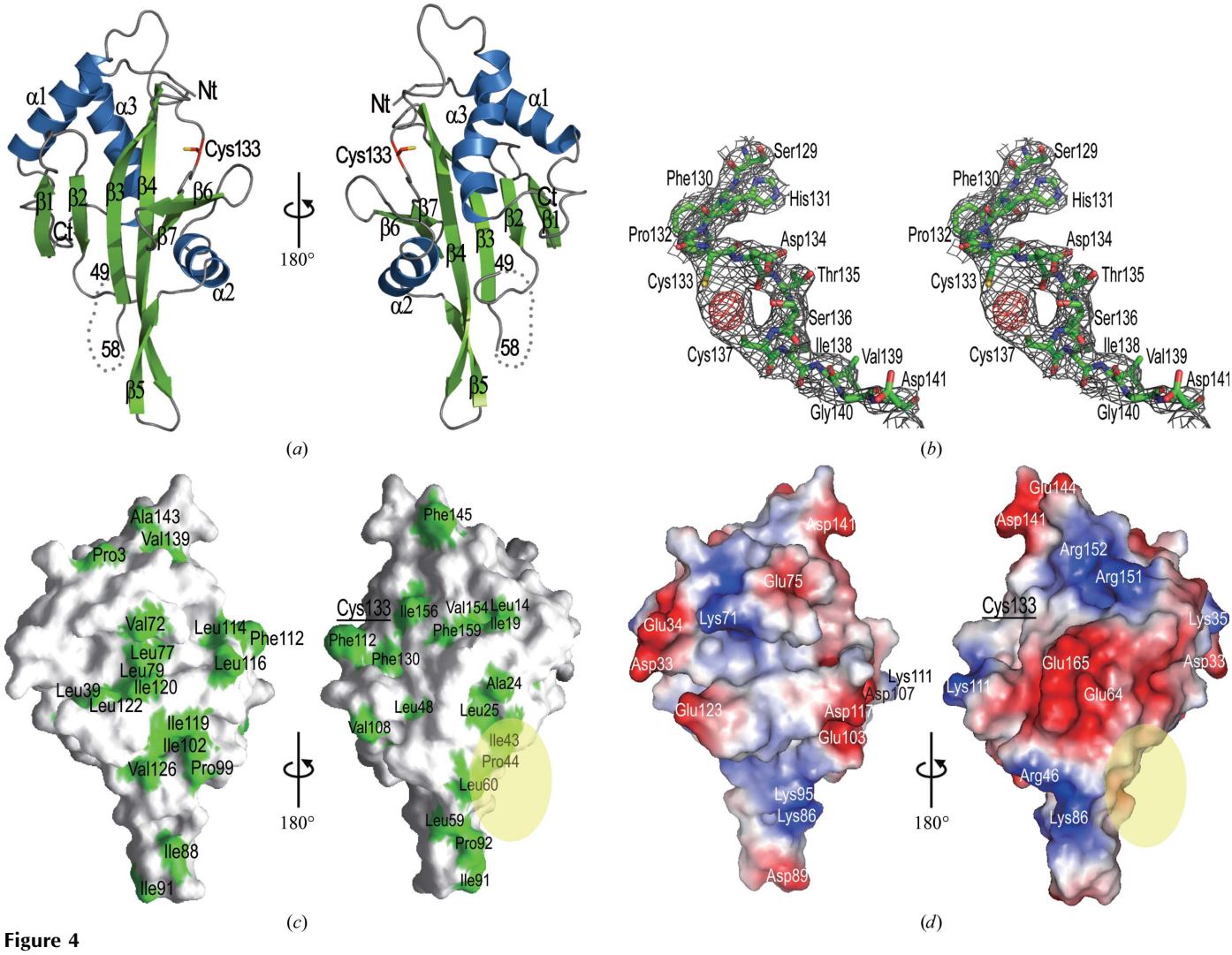

Figure 4

(*a*) Ribbon diagrams showing the secondary-structure elements of yAtg10. Front (left) and back (right) views showing the central β-strands of yAtg10. Seven β-strands (green arrows), three α-helices (blue ribbons) and connecting loops (grey) are shown. The catalytic residue Cys133 is represented in stick form and its C and S atoms are coloured red and yellow, respectively. The residues that are not visible (49–58) are indicated as grey dots. The N- and C-termini are labelled Nt and Ct, respectively. (*b*) Stereo electron-density map near the catalytic residue Cys133 of yAtg10. The final $2F_o - F_c$ map (grey) and the $F_o - F_c$ OMIT map for the bound Hg atom (red) were calculated using 27.8–2.7 Å resolution data and were contoured at 1.2σ and 6.0σ, respectively. The two free cysteine residues Cys133 and Cys137 are covalently modified by a heavy-metal atom. (*c*) Electrostatic potential surface of yAtg10 viewed as in (*a*). Positive and negative electrostatic potentials are coloured blue and red, respectively. The location of the missing ten residues is indicated as a transparent yellow oval and the key residue Cys133 is labelled and underlined. (*d*) Hydrophobic surface of yAtg10 viewed in the same orientation. Residues forming hydrophobic surfaces are coloured green and labelled.

(Z = 5.2), the UEV domain of tumour susceptibility gene 101 protein (Z = 5.1) and UFM1-conjugating enzyme (Z = 5.1). As shown in Fig. 5, the folding pattern of Atg10 shows similarity to Atg3 and other E2 enzymes. Unexpectedly, a subunit of F-actin capping protein shares structural homology in the central β-strand but differs greatly in other regions, and frataxin (or CyaY), an iron-homeostasis protein, has a similar folding pattern (Fig. 5; Takeda *et al.*, 2010; Dhe-Paganon *et al.*, 2000; Cho *et al.*, 2000). Although the significance of the structural homology between yAtg10 and the aforementioned functionally unrelated proteins is unclear, these proteins and yAtg10 share a common task in mediating protein–protein interactions. The E2 enzymes interact with Ub (or Ubl molecules), E1 activating enzymes and E3 ligases to shuttle Ub (or Ubl molecules) from an E1 to an E3–substrate complex (Pruneda *et al.*, 2011; Varshavsky, 2005; Schulman & Harper, 2009). The E2 enzymes have a relatively small size and share an α/β folding pattern (Fig. 5) and a conserved catalytic core. Despite the wealth of structural information on the E2 enzyme alone, information pertaining to the interaction mode of the E2 enzyme with its binding partners is limited owing to the determination of only a few complex structures. Furthermore, detailed analysis of the binding interface between E2 enzymes and their binding proteins has shown that E2 enzymes have a dynamic property that allows the interaction to adopt diverse orientations (Pruneda *et al.*, 2011; Ko *et al.*, 2010). Therefore, attempts at formulating a general binding mode of yAtg10 to its binding partners using structural comparison with other E2 enzymes is problematic. However, employing the most similar protein, the autophagic E2 enzyme Atg3, may be a good foundation from which to delineate the characteristics of yAtg10. Both yAtg10 and Atg3 interact with the same E1 activating enzyme, Atg7, for the delivery of activated Ubl molecules. Atg10 has been studied less extensively than Atg3 in many respects, especially given the difficulty in obtaining purified Atg12 protein. Furthermore, Atg10 lacks two unique insertions present in Atg3, the FR-region and the HR-region, which play critical roles in binding to the E1 enzyme Atg7 and Ubl Atg8, respectively (Yamada *et al.*, 2007; Hong *et al.*, 2011; Taherbhoy *et al.*, 2011; Noda *et al.*, 2010). These regions protrude from the shared core structure with yAtg10 (Fig. 6*a*), suggesting that there are no structurally equivalent regions in Atg10. The Trp-Xaa-Xaa-Leu sequence motif at the end of the long HR helix in Atg3 is recognized by hydrophobic residues of Atg8 (Yamaguchi *et al.*, 2010). Similarly, hydrophobic residues are found in the structurally equivalent region of Atg12 from *A. thaliana*, although the hydrophobic binding pocket seems to be much smaller in Atg12 (Suzuki *et al.*, 2005). Since the Trp-Xaa-Xaa-Leu sequence motif has not been detected in Atg10, it is tempting to speculate that the hydrophobic residues with smaller side-chain atoms participate in interactions with Atg10

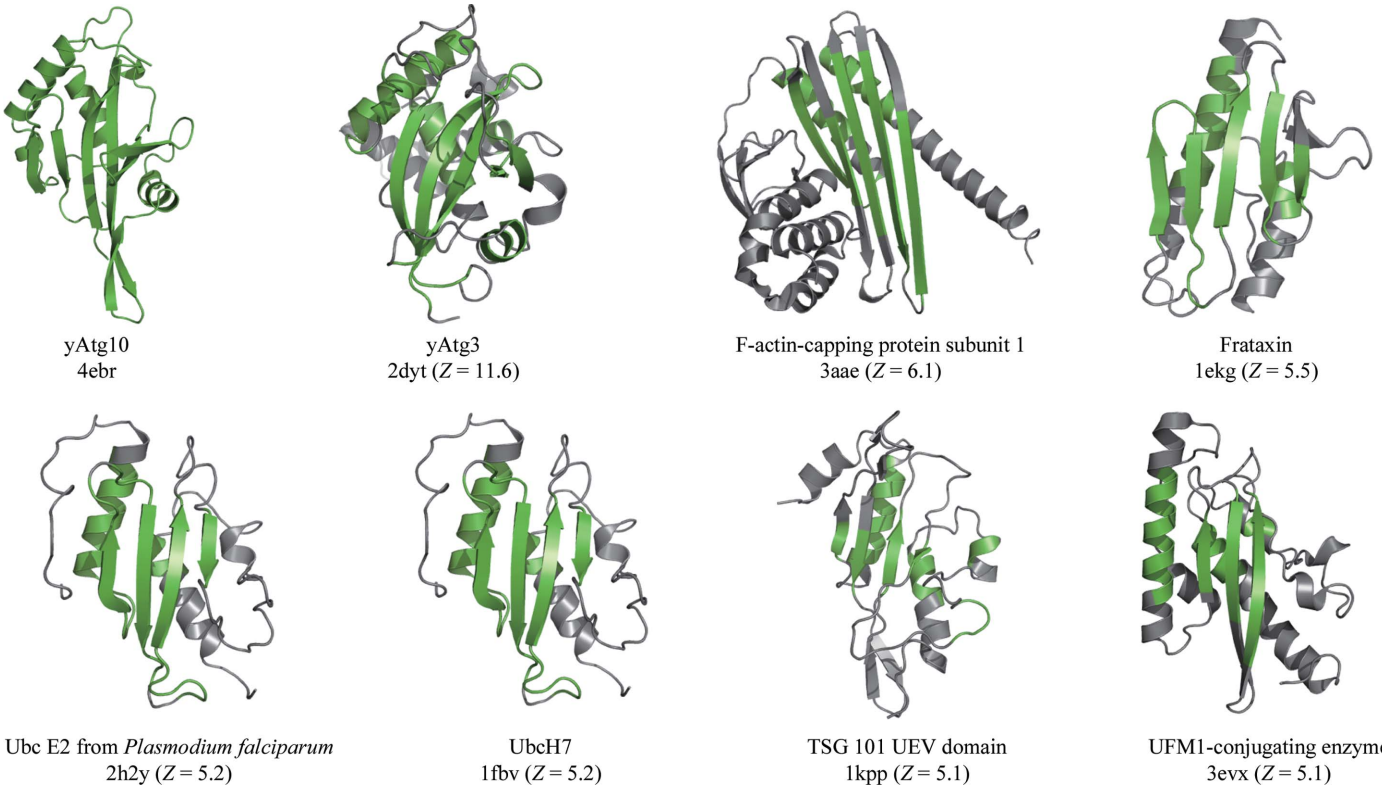

Figure 5
Ribbon diagrams comparing the overall structures of yAtg10, yAtg3, F-actin capping protein subunit α1, frataxin, Ubc E2 enzyme from *P. falciparum*, the UEV domain of tumour susceptibility gene 101 protein and UFM1-conjugating enzyme. The view is the same as that on the left in Fig. 4(*a*). PDB codes and Z-scores from the *DALI* server for each structure are shown below each structure. Structural regions that match corresponding regions of yAtg10 are coloured green in each structure.

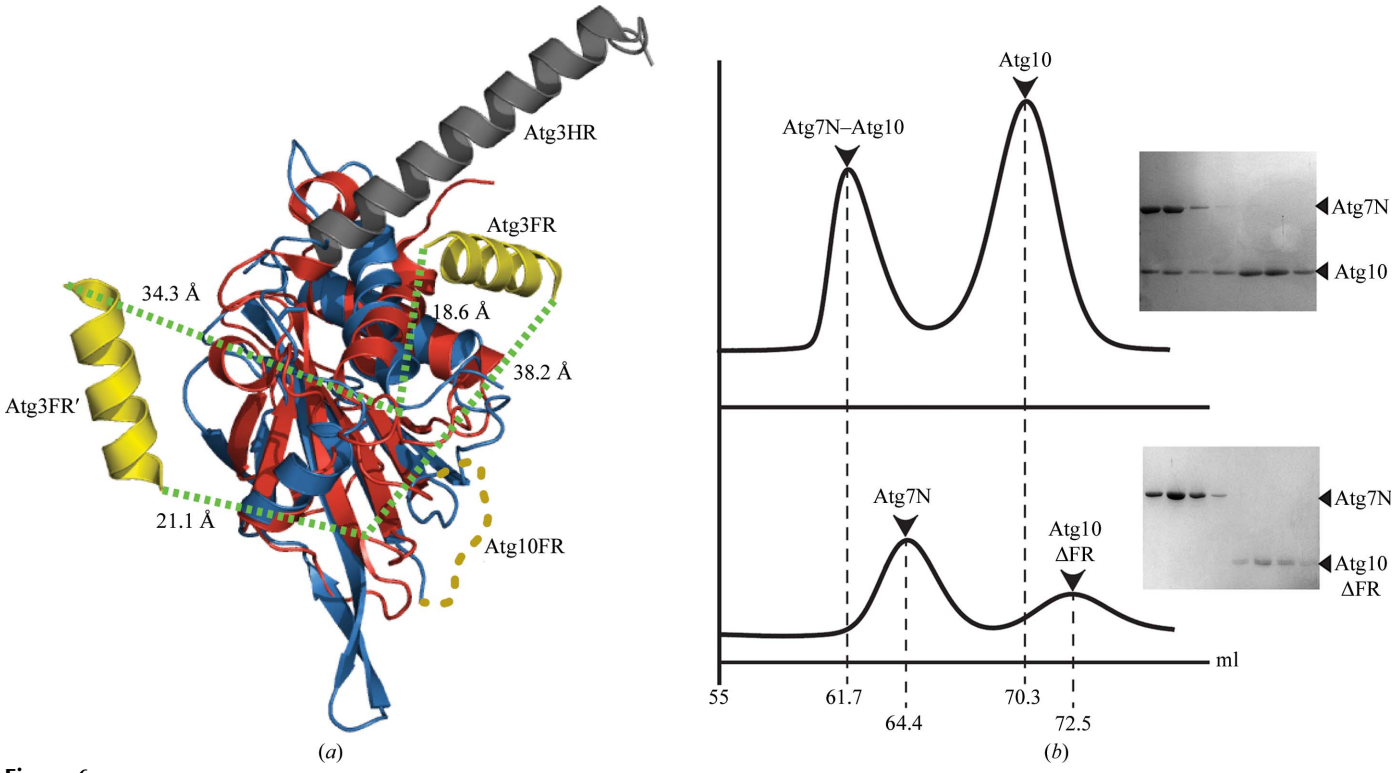

Figure 6
(*a*) Ribbon diagram showing structural superposition of yAtg10 (blue) and yAtg3 (red). The core region is relatively well superposed, whereas two uniquely protruding regions, FR and HR in yAtg3, are located in completely different positions. The FR-region and HR-region are coloured yellow and grey, respectively. Two possible positions of the FR-helix in yAtg3 are shown with distances indicated (Taherbhoy *et al.*, 2011). The ten missing flexible residues in yAtg10 are indicated by yellow dots. (*b*) Gel-filtration chromatography results using a mixture of yAtg7N and yAtg10 (or the yAtg10ΔFR mutant). The elution volume for each molecule is indicated. Atg7N co-migrates with wild-type Atg10, and SDS–PAGE confirms Atg7N–Atg10 complex formation (top). The Atg10ΔFR mutant migrates separately from Atg7N (bottom), suggesting that the FR-region in Atg10 plays a critical role in interaction with Atg7.

and Atg12. As shown in Fig. 4(*c*), several hydrophobic patches are exposed on the Atg10 surface, although it remains to be determined whether these hydrophobic residues are involved in the interaction with Atg12.

3.5. Atg10 FR-region

Atg10 does not possess an obvious protruding α-helical region for interaction with Atg7. While the FR-region in Atg3 refers to a flexible region, only 13 residues which form an α-helix have been modelled among the 80 amino-acid residues (Yamada *et al.*, 2007). Careful analysis based on the crystallographic symmetry of the deposited PDB coordinates for Atg3 (PDB entry 2dyt) has suggested another possibility for the location of the FR-region (Taherbhoy *et al.*, 2011), supporting extreme flexibility of this region (Fig. 6*a*). It has been reported that Atg10 and Atg3 share the same E1 activating enzyme, Atg7, and it is reasonable to speculate that a similar α-helical segment of FR in Atg10 binds to the N-terminal domain of Atg7. There are three α-helices in Atg10 (Fig. 4*a*), although none of these seem to be feasible for the interaction. We generated a GST-tagged N-terminal helix (α1) of Atg10 and performed pull-down assays, but no interaction was observed (data not shown). The ten residues (49–58) in the loop connecting the β2 and β3 strands in Atg10 are missing in the final model (Fig. 4*a*) and the temperature factors in this region are extremely high (Fig. 3*b*). For further analysis, we generated a mutant, Atg10ΔFR, by deleting the flexible region. We then examined the binding of Atg10ΔFR to Atg7N and failed to detect any interaction by gel filtration (Fig. 6*b*), suggesting the importance of the flexible region of yAtg10 for binding to yAtg7. The amino-acid sequence of the flexible region is not well conserved among Atg10 proteins (Supplementary Fig. S1), as is the case for the FR-region in Atg3. Interestingly, the relative orientation of these missing ten residues in yAtg10 is structurally equivalent to the FR-region in yAtg3, although the length of the matching residues is much shorter than in Atg3 (Figs. 1 and 6*a*). There is the possibility that the missing flexible region in Atg10 forms the same secondary structure (α-helix) as predicted by the server. Therefore, we defined this missing region as an FR-region (flexible region) in Atg10, although the functional equivalency should be further tested. For a conclusive answer, however, analysis of an Atg10–Atg7 complex structure is required.

4. Conclusions

The crystal structure of the autophagic E2 enzyme Atg10 from *S. cerevisiae* has been determined. The employment of heavy-atom derivatization techniques was essential for structure

determination. The core fold of yAtg10 is well conserved compared with those of Atg3 and other E2 enzymes. In contrast to other E2 enzymes, the autophagic E2 enzymes Atg3 and Atg10 possess insertion regions in the middle of the core fold and may be involved in protein function. The missing segment that we termed the 'FR-region' in Atg10 may be important for interaction with the Atg7 E1 enzyme.

We thank the staff of the 4A beamline, Pohang Light Source, Korea and the NW12 beamline, Photon Factory, Japan for help with data collection. This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST; 2011-0028168), a Korea University Grant and the Korea Healthcare Technology R&D Project, Ministry for Health, Welfare and Family Affairs, Republic of Korea (A092006).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Boggon, T. J. & Shapiro, L. (2000). *Structure*, **8**, R143–R149.
Chang, G., Spencer, R. H., Lee, A. T., Barclay, M. T. & Rees, D. C. (1998). *Science*, **282**, 2220–2226.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cho, S.-J., Lee, M. G., Yang, J. K., Lee, J. Y., Song, H. K. & Suh, S. W. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 8932–8937.
Dhe-Paganon, S., Shigeta, R., Chi, Y.-I., Ristow, M. & Shoelson, S. E. (2000). *J. Biol. Chem.* **275**, 30753–30756.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Geng, J. & Klionsky, D. J. (2008). *EMBO Rep.* **9**, 859–864.
Green, T. J. & Luo, M. (2006). *Acta Cryst.* D**62**, 498–504.
Heinig, M. & Frishman, D. (2004). *Nucleic Acids Res.* **32**, W500–W502.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
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.
Ichimura, Y., Kirisako, T., Takao, T., Satomi, Y., Shimonishi, Y., Ishihara, N., Mizushima, N., Tanida, I., Kominami, E., Ohsumi, M., Noda, T. & Ohsumi, Y. (2000). *Nature (London)*, **408**, 488–492.
Jones, D. T. (1999). *J. Mol. Biol.* **292**, 195–202.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Klionsky, D. J., Cregg, J. M., Dunn, W. A., Emr, S. D., Sakai, Y., Sandoval, I. V., Sibirny, A., Subramani, S., Thumm, M., Veenhuis, M. & Ohsumi, Y. (2003). *Dev. Cell*, **5**, 539–545.
Klionsky, D. J. & Emr, S. D. (2000). *Science*, **290**, 1717–1721.
Ko, S. *et al.* (2010). *J. Biol. Chem.* **285**, 36070–36080.
Lee, B.-G., Park, E. Y., Lee, K.-E., Jeon, H., Sung, K. H., Paulsen, H., Rübsamen-Schaeff, H., Brötz-Oesterhelt, H. & Song, H. K. (2010). *Nature Struct. Mol. Biol.* **17**, 471–478.
Mizushima, N., Noda, T., Yoshimori, T., Tanaka, Y., Ishii, T., George, M. D., Klionsky, D. J., Ohsumi, M. & Ohsumi, Y. (1998). *Nature (London)*, **395**, 395–398.
Nakatogawa, H., Suzuki, K., Kamada, Y. & Ohsumi, Y. (2009). *Nature Rev. Mol. Cell Biol.* **10**, 458–467.
Noda, T., Fujita, N. & Yoshimori, T. (2008). *Autophagy*, **4**, 540–541.
Noda, N. N., Ohsumi, Y. & Inagaki, F. (2009). *Chem. Rev.* **109**, 1587–1598.
Noda, N. N., Ohsumi, Y. & Inagaki, F. (2010). *FEBS Lett.* **584**, 1379–1385.
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.
Pruneda, J. N., Stoll, K. E., Bolton, L. J., Brzovic, P. S. & Klevit, R. E. (2011). *Biochemistry*, **50**, 1624–1633.
Schulman, B. A. & Harper, J. W. (2009). *Nature Rev. Mol. Cell Biol.* **10**, 319–331.
Shintani, T. & Klionsky, D. J. (2004). *Science*, **306**, 990–995.
Shintani, T., Mizushima, N., Ogawa, Y., Matsuura, A., Noda, T. & Ohsumi, Y. (1999). *EMBO J.* **18**, 5234–5241.
Suzuki, K. & Ohsumi, Y. (2010). *FEBS Lett.* **584**, 1280–1286.
Suzuki, N. N., Yoshimoto, K., Fujioka, Y., Ohsumi, Y. & Inagaki, F. (2005). *Autophagy*, **1**, 119–126.
Taherbhoy, A. M., Kaiser, S. E. & Schulman, B. A. (2012). *Cell Cycle*, **11**, 635–636.
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.
Takeda, S., Minakata, S., Koike, R., Kawahata, I., Narita, A., Kitazawa, M., Ota, M., Yamakuni, T., Maéda, Y. & Nitanai, Y. (2010). *PLoS Biol.* **8**, e1000416.
Varshavsky, A. (2005). *Trends Biochem. Sci.* **30**, 283–286.
Xu, Y., Moseley, J. B., Sagot, I., Poy, F., Pellman, D., Goode, B. L. & Eck, M. J. (2004). *Cell*, **116**, 711–723.
Yamada, Y., Suzuki, N. N., Hanada, T., Ichimura, Y., Kumeta, H., Fujioka, Y., Ohsumi, Y. & Inagaki, F. (2007). *J. Biol. Chem.* **282**, 8036–8043.
Yamaguchi, M., Noda, N. N., Nakatogawa, H., Kumeta, H., Ohsumi, Y. & Inagaki, F. (2010). *J. Biol. Chem.* **285**, 29599–29607.
Yamaguchi, M., Noda, N. N., Yamamoto, H., Shima, T., Kumeta, H., Kobashigawa, Y., Akada, R., Ohsumi, Y. & Inagaki, F. (2012). *Structure*, **20**, 1244–1254.
Yamaguti, M., Suzuki, N. N., Fujioka, Y., Ohsumi, Y. & Inagaki, F. (2007). *Acta Cryst.* F**63**, 443–445.