Structure and function of the yeast U-box-containing ubiquitin ligase Ufd2p

Daqi Tu†, Wei Li‡, Yihong Ye‡, and Axel T. Brunger†§

†Departments of Molecular and Cellular Physiology, Neurology and Neurological Sciences, Structural Biology, and Photon Science, Stanford University and Howard Hughes Medical Institute, Stanford, CA 94305; and ‡Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on May 3, 2005.

Edited by Brenda A. Schulman, St. Jude Children’s Research Hospital, Memphis, TN, and accepted by the Editorial Board August 15, 2007 (received for review February 13, 2007)

Proteins conjugated by Lys-48-linked polyubiquitin chains are preferred substrates of the eukaryotic proteasome. Polyubiquitination requires an activating enzyme (E1), a conjugating enzyme (E2), and a ligase (E3). Occasionally, these enzymes only assemble short ubiquitin oligomers, and their extension to full length involves a ubiquitin elongating factor termed E4. Ufd2p, as the first E4 identified to date, is involved in the degradation of misfolded proteins of the endoplasmic reticulum and of a ubiquitin–β-GAL fusion substrate in Saccharomyces cerevisiae. The mechanism of action of Ufd2p is unknown. Here we describe the crystal structure of the full-length yeast Ufd2p protein. Ufd2p has an elongated shape consisting of several irregular Armadillo-like repeats with two helical hairpins protruding from it and a U-box domain flexibly attached to its C terminus. The U-box of Ufd2p has a fold similar to that of the RING (Really Interesting New Gene) domain that is present in certain ubiquitin ligases. Accordingly, Ufd2p has all of the hallmarks of a RING finger-containing ubiquitin ligase: it associates with its cognate E2 Ubc4p via its U-box domain and catalyzes the transfer of ubiquitin from the E2 active site to Ufd2p itself or to an acceptor ubiquitin molecule to form unanchored diubiquitin oligomers. Thus, Ufd2p can function as a bona fide E3 ubiquitin ligase to promote ubiquitin chain elongation on a substrate.

E4 elongation factor | endoplasmic reticulum-associated degradation | Cdc48p/p97 | polyubiquitination | Armadillo repeats

In eukaryotes, short-lived proteins are degraded by the ubiquitin (Ub) proteasome system (reviewed in ref. 1). Substrates of the proteasome include transcription factors, cell cycle regulators, signal transducers, and misfolded proteins generated under stress conditions. Protein ubiquitination is achieved by a multistep mechanism involving a cascade of enzymes. Ub-activating enzyme (E1) hydrolyzes ATP to form a high-energy thioester bond between its catalytic cysteine residue and the C terminus of Ub. Activated Ub is subsequently transferred to a distinct Ub-conjugating enzyme (E2) by transthiolation. Finally, Ub is transferred to the ε-amino group of an internal lysine residue of a target protein by a Ub ligase (E3). Polyubiquitin chains can be assembled when additional Ub molecules are transferred, one at a time, to a lysine residue in the substrate-bound Ub molecule via an isopeptide bond linkage. In addition, with the assistance of an E3 ligase, certain Ub-conjugating enzymes can form Lys-48-linked Ub chains linked to its catalytic cysteine before transferring the assembled Ub chains to a substrate (2, 3).

The essential components involved in the degradation of certain short-lived proteins have been identified in Saccharomyces cerevisiae by using a model proteasomal substrate consisting of a Ub moiety fused to the N terminus of a reporter protein. These components, which were designated as UFD1–5, represent the so-called UFD (Ub fusion degradation) pathway (4). Interestingly, with the exception of Ufd5p, a transcriptional regulator of the proteasome (5), and of Ufd4p, a HECT (homologous to E6-associated protein C terminus) domain Ub ligase (6), the other UFD proteins all interact with a conserved AAA ATPase (ATPase associated with various activities) named Cdc48p in yeast or p97 in mammals to regulate its activities in a subset of proteasome-dependent degradation pathways (7). The role of Cdc48p/p97 in protein turnover is best characterized for the degradation of misfolded endoplasmic reticulum (ER) proteins, which occurs through a pathway termed ER-associated protein degradation or retrotranslocation. During this process, Cdc48p/p97 associates with the ER membrane to extract its client proteins (misfolded polypeptides undergoing retrotranslocation) out of the ER membrane and subsequently target them for degradation by the proteasome. The Cdc48p/p97-dependent retrotranslocation requires the function of at least two UFD proteins, Ufd1p and Ufd2p. Ufd1p associates with Npl4p to form a heterodimeric cofactor complex, which promotes substrate recognition by Cdc48p/p97 (8–10). Ufd2p appears to act downstream of Cdc48p/p97 to facilitate the transfer of ER-associated protein degradation substrates to Rad23p, a proteasome-associated Ub receptor (11). Although Ufd2p is not essential for cell viability under normal conditions, its activity becomes critical under stress conditions in yeast (6). Likewise, the Ufd2a+/− mice lacking one functional copy of the Ufd2a gene, a homologue of Ufd2, develop a neurological disorder as a result of axonal dystrophy induced by ER stress. Mice deficient in UFD2a die in utero because of marked apoptosis in the developing heart (12). Taken together, the major function of Ufd2 is probably to cooperate with Cdc48/p97 to maintain a stress-free environment for cells.

Ufd2p contains a U-box domain that is structurally related to the RING finger domain that is found in certain E3 Ub ligases (13, 14). Ufd2 is capable of elongating Ub chains, an activity that is essential for its function in ER-associated protein degradation. In the presence of an E1, Ubc4p (E2), and Ufd4p (E3), a short Ub chain attached to a substrate can be further extended by Ufd2p. For this reason, Ufd2p is often referred to as an E4 enzyme (6). It is unclear how Ufd2p promotes Ub chain elongation in conjunction with the HECT domain containing E3 ligase Ufd4p. A HECT domain E3 ligase usually accepts Ub

Author contributions: D.T., Y.Y., and A.T.B. designed research; D.T. and W.L. performed research; D.T. and A.T.B. analyzed data; and D.T., Y.Y., and A.T.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. B.A.S. is a guest editor invited by the Editorial Board.

Abbreviations: Ub, ubiquitin; ER, endoplasmic reticulum; SeMet, selenomethionine; MAD, multiple-wavelength anomalous dispersion.

Data deposition: The atomic coordinates and diffraction data have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 2QIZ (Native1) and 2QJ0 (Native2)].

†To whom correspondence should be addressed at: Stanford University, 318 Campus Drive, Room E300C, Stanford, CA 94305. E-mail: brunger@stanford.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0701369104DC1.

© 2007 by The National Academy of Sciences of the USA
from its cognate E2 and forms a thioester-linked Ub–E3 intermediate before transferring the attached Ub to a substrate. In contrast, RING finger E3 ligases usually associate directly with cognate E2 enzymes to promote the transfer of Ub directly from E2 to a substrate. RING finger E3 ligases can also mediate the discharge of Ub from the E2 active cysteine in the absence of a substrate. When excess free Ub molecules are present, the released Ub molecules can be linked to a lysine residue in these free Ub molecules, forming an isopeptide bond-linked diubiquitin molecule (15). Whether Ufd2p also contains these activities is unknown.

Here we present the crystal structure of the full-length yeast Ufd2p enzyme. The structure has an elongated shape with a flexible U-box domain attached to it. The core region of Ufd2p consists of multiple irregular Armadillo-like repeats with two pronounced helical hairpin protrusions. The core region of Ufd2p is structurally reminiscent of importin α, a nuclear transport protein. Based on the CHIP/E2 complex crystal structure (16), we built a model of the Ufd2p/E2 complex that predicts interactions between E2 and the U-box domain as well as the core of Ufd2p. Biochemical experiments confirmed that Ufd2p binds directly to the E2 enzyme Ubc4p and furthermore revealed that its U-box domain is involved in the transfer of an E2-conjugated Ub to a free Ub and to Ufd2p itself. Thus, Ufd2p has the characteristics of a typical RING finger E3 ligase.

Results

Structure Determination. Crystals of the yeast Ufd2p, a 110-kDa protein consisting of 961 residues, were formed in space group P2_12_1 containing one molecule per asymmetric unit. The structure was solved by multiple-wavelength anomalous dispersion (MAD) phasing to d_min = 3.0 Å using selenomethionine substitution. Density modification and phase extension to d_min = 2.74 Å using data set Native2(Se) (Tables 1 and 2, Native1) resulted in a readily interpretable electron density map (Fig. 1). Iterative model building and refinement produced a model with good statistics and geometry (Table 1). A higher-resolution data set was subsequently collected on a different crystal form in the same space group, but with a significantly different c cell axis (Tables 1 and 2, Native1). The final refined model comprising 933 residues produced R_free = 27.6% and R = 22.1% to d_min = 2.56 Å. A representative 2F_e – F_o map contoured at 1 σ is shown in supporting information (SI) Fig. 10. The quality of the electron density map was excellent except for residues 27–35 and residues 707–718.

Overall Structure. Ufd2p (residues 1–879) has an elongated shape, ~146 Å in length, ~84 Å in height, and ~70 Å in width (Fig. 2). Based on sequence conservation (SI Fig. 11), the major body of Ufd2p can be approximately divided into two regions, consisting of residues 1–187 and 188–879, respectively. The highly variable N-terminal region consists of a short β-hairpin and eight α-helices. Helices α1 to α4 form a four-helix bundle, whereas P2_12_1 containing one molecule per asymmetric unit. The structure was solved by multiple-wavelength anomalous dispersion (MAD) phasing to d_min = 3.0 Å using selenomethionine substitution. Density modification and phase extension to d_min = 2.74 Å using data set Native2(Se) (Tables 1 and 2, Native1) resulted in a readily interpretable electron density map (Fig. 1). Iterative model building and refinement produced a model with good statistics and geometry (Table 1). A higher-resolution data set was subsequently collected on a different crystal form in the same space group, but with a significantly different c cell axis (Tables 1 and 2, Native1). The final refined model comprising 933 residues produced R_free = 27.6% and R = 22.1% to d_min = 2.56 Å. A representative 2F_e – F_o map contoured at 1 σ is shown in supporting information (SI) Fig. 10. The quality of the electron density map was excellent except for residues 27–35 and residues 707–718.

Results

Structure Determination. Crystals of the yeast Ufd2p, a 110-kDa protein consisting of 961 residues, were formed in space group P2_12_1 containing one molecule per asymmetric unit. The structure was solved by multiple-wavelength anomalous dispersion (MAD) phasing to d_min = 3.0 Å using selenomethionine substitution. Density modification and phase extension to d_min = 2.74 Å using data set Native2(Se) (Tables 1 and 2, Native1) resulted in a readily interpretable electron density map (Fig. 1). Iterative model building and refinement produced a model with good statistics and geometry (Table 1). A higher-resolution data set was subsequently collected on a different crystal form in the same space group, but with a significantly different c cell axis (Tables 1 and 2, Native1). The final refined model comprising 933 residues produced R_free = 27.6% and R = 22.1% to d_min = 2.56 Å. A representative 2F_e – F_o map contoured at 1 σ is shown in supporting information (SI) Fig. 10. The quality of the electron density map was excellent except for residues 27–35 and residues 707–718.

Overall Structure. Ufd2p (residues 1–879) has an elongated shape, ~146 Å in length, ~84 Å in height, and ~70 Å in width (Fig. 2). Based on sequence conservation (SI Fig. 11), the major body of Ufd2p can be approximately divided into two regions, consisting of residues 1–187 and 188–879, respectively. The highly variable N-terminal region consists of a short β-hairpin and eight α-helices. Helices α1 to α4 form a four-helix bundle, whereas P2_12_1 containing one molecule per asymmetric unit. The structure was solved by multiple-wavelength anomalous dispersion (MAD) phasing to d_min = 3.0 Å using selenomethionine substitution. Density modification and phase extension to d_min = 2.74 Å using data set Native2(Se) (Tables 1 and 2, Native1) resulted in a readily interpretable electron density map (Fig. 1). Iterative model building and refinement produced a model with good statistics and geometry (Table 1). A higher-resolution data set was subsequently collected on a different crystal form in the same space group, but with a significantly different c cell axis (Tables 1 and 2, Native1). The final refined model comprising 933 residues produced R_free = 27.6% and R = 22.1% to d_min = 2.56 Å. A representative 2F_e – F_o map contoured at 1 σ is shown in supporting information (SI) Fig. 10. The quality of the electron density map was excellent except for residues 27–35 and residues 707–718.

Results

Structure Determination. Crystals of the yeast Ufd2p, a 110-kDa protein consisting of 961 residues, were formed in space group

Table 1. X-ray data collection and refinement

<table>
<thead>
<tr>
<th></th>
<th>Single wavelength</th>
<th>MAD (Se)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Native1</td>
<td>Native2 (Se)</td>
</tr>
<tr>
<td></td>
<td>SSRL/11-1</td>
<td>SSRL/11-1</td>
</tr>
<tr>
<td><strong>Crystallographic data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Beamline</strong></td>
<td>SSRL/11-1</td>
<td>ALS/8.2.2</td>
</tr>
<tr>
<td><strong>Wavelength, Å</strong></td>
<td>0.97945</td>
<td>1.00000</td>
</tr>
<tr>
<td><strong>a, b, c, Å</strong></td>
<td>65.5, 123.0, 176.4</td>
<td>65.8, 122.8, 178.9</td>
</tr>
<tr>
<td><strong>Resolution, Å</strong></td>
<td>2.56</td>
<td>2.74</td>
</tr>
<tr>
<td><strong>Unique reflections</strong></td>
<td>46,444</td>
<td>37,483</td>
</tr>
<tr>
<td><strong>Completeness, %</strong></td>
<td>99.2 (95.4)</td>
<td>96.3 (74.2)</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>4.2</td>
<td>3.8</td>
</tr>
<tr>
<td><strong>Bond length deviation, Å</strong></td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td><strong>Bond angle deviation, °</strong></td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td><strong>Average B factor, Å²</strong></td>
<td>63.0</td>
<td>77.7</td>
</tr>
<tr>
<td><strong>Maximum B factor, Å²</strong></td>
<td>23.6</td>
<td>31.4</td>
</tr>
<tr>
<td><strong>Residues in core region, %</strong></td>
<td>87.0</td>
<td>84.4</td>
</tr>
<tr>
<td><strong>Residues in disallowed regions, %</strong></td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Overall Structure. Ufd2p (residues 1–879) has an elongated shape, ~146 Å in length, ~84 Å in height, and ~70 Å in width (Fig. 2). Based on sequence conservation (SI Fig. 11), the major body of Ufd2p can be approximately divided into two regions, consisting of residues 1–187 and 188–879, respectively. The highly variable N-terminal region consists of a short β-hairpin and eight α-helices. Helices α1 to α4 form a four-helix bundle, whereas P2_12_1 containing one molecule per asymmetric unit. The structure was solved by multiple-wavelength anomalous dispersion (MAD) phasing to d_min = 3.0 Å using selenomethionine substitution. Density modification and phase extension to d_min = 2.74 Å using data set Native2(Se) (Tables 1 and 2, Native1) resulted in a readily interpretable electron density map (Fig. 1). Iterative model building and refinement produced a model with good statistics and geometry (Table 1). A higher-resolution data set was subsequently collected on a different crystal form in the same space group, but with a significantly different c cell axis (Tables 1 and 2, Native1). The final refined model comprising 933 residues produced R_free = 27.6% and R = 22.1% to d_min = 2.56 Å. A representative 2F_e – F_o map contoured at 1 σ is shown in supporting information (SI) Fig. 10. The quality of the electron density map was excellent except for residues 27–35 and residues 707–718.

Results

Structure Determination. Crystals of the yeast Ufd2p, a 110-kDa protein consisting of 961 residues, were formed in space group

Table 2. Phasing power and figure of merit

<table>
<thead>
<tr>
<th>Phasing power</th>
<th>Figure of merit</th>
</tr>
</thead>
<tbody>
<tr>
<td>48.67–2.99 Å</td>
<td>1.75</td>
</tr>
<tr>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>3.13–2.99 Å</td>
<td>0.44</td>
</tr>
<tr>
<td>0.14</td>
<td></td>
</tr>
</tbody>
</table>

Overall figure of merit after density modification and phase extension: 0.94.
helices α5 and α6 interact with α3 and α4 through hydrophobic contacts that are partly mediated by their connecting loops. The more conserved core region, residues 188–879, consists of a series of interacting helices with two pronounced helical hairpin protrusions (Fig. 2). The C-terminal residues (884–947) form a U-box domain consisting of a short three-stranded β-sheet and two α-helices (Fig. 2), which is connected to the major body by a five-residue loop. The structure of this domain superimposes well on that of the U-box of CHIP (16) with a rmsd of 1.1 Å out of 63 equivalent Cα atoms. The U-box domain forms no direct contact with the major body of Ufd2p. We observed two conformations of the U-box domain in the two different crystal forms (Fig. 3). Whereas the majority of the atomic coordinates between the two crystal forms are very similar (rmsd = 0.6 Å for 829 Cα atoms within residues 1–850), the U-box domain and helix α39 have rotated with a maximum displacement of 4.2 Å between equivalent Cα atoms, indicating flexibility of the linker connecting the U-box domain to the core region. Clearly, the movement of the U-box domain is restricted by crystal packing contacts, so we expect an increased amount of conformational variability in solution.

The Conserved Core Region. The core region of Ufd2p contains 31 α-helices of variable lengths that are connected by loops of different size (Fig. 2). Most of these helices form a compact unit that resembles a shallow-grooved, right-handed spiral, apart from the α19/20 and α32/33 helical hairpin protrusions, and helices α16 and α22. The helical packing pattern of the compact unit consists of five structurally repeating units that resemble tandem Armadillo (ARM) repeats, each of which comprises three α-helices formed by ~40 aa (α23–α24, α25–α28, α29–α31, α33–α34, and α36–α38) (Fig. 4). However, the Ufd2p repeating units are highly irregular, comprising multiple short helices and/or loops whose structural equivalent in ARM motif would be composed of a single helix.

Conserved Residues and Surface Features. As mentioned above, the N-terminal region of Ufd2p exhibits primary sequence variability, indicating that other Ufd2p homologues may adopt different structures in that region. The side of Ufd2p from which the U-box domain protrudes is more conserved compared with the other side (Fig. 5). There is a large continuous patch of conserved residues covering one side of the core region and part of one of the helical hairpins (α32/α33). In contrast, the surface residues of the N-terminal region of Ufd2p are not conserved apart from the spot that is formed by the termini of helices α14 and α17 and by the loop region around residue Lys-333 (white arrow in Fig. 5B).

One of the two sides of Ufd2p is mostly covered by negatively charged amino acids, especially around the linker region between the U-box and the core region (SI Fig. 12 Upper). The other side has a mostly neutral charge distribution, except for two positively charged regions: the abovementioned conserved spot near Lys-333 and the area between the two helical hairpins (SI Fig. 12 Lower).

Structural Comparisons with Other Proteins. A search for structurally related proteins of Ufd2p was performed with DALI (18). The core region of Ufd2p shares significant structural similarity with the nuclear import factor importin α, which recognizes proteins that carry a nuclear localization signals to mediate their transport from the cytosol into the nucleus (19, 20). However, in contrast to Ufd2p’s irregular ARM-like motifs, importin α consists of a series of regular ARM repeats. These repeating units form a superhelical structure with a shallow groove on one side that binds a nuclear localization signal-containing substrate.

---

Fig. 2. Overall structure of Ufd2p. The two views are related by a 90° rotation as specified. The N-terminal region is colored blue, the core region is colored green, and the U-box domain is colored red. Selected α-helices are labeled.

Fig. 3. Conformational variability of the U-box domain. The structures from the two crystal forms were superimposed by using the atoms of the major body. Green, structure from Native1; red, structure from Native2(Se).

Fig. 4. Repeating folding motifs of the core region of Ufd2p. Each repeating motif is colored differently. For clarity, helices α32 and α35 and part of helix α33 are not shown because they are not involved in the folding motifs.
Fig. 5. Molecular surface conservation of Ufd2p. Ufd2p homologues from 12 species (S. cerevisiae, Schizosaccharomyces pombe, Aspergillus oryzae, Drosophila melanogaster, Apis mellifera, Mus musculus, Rattus norvegicus, Homo sapiens, Danio rerio, Arabidopsis thaliana, Dictyostelium discoideum, and Caenorhabditis elegans) were aligned by using ClustalW (17). Residues were each assigned a conservation score by using the ConSurf server 3.0 (http://consurf.tau.ac.il), and scores were mapped onto the surface of the molecule. The colors range from 100% (red) to 50% (white) to 0% (cyan) sequence identity. Two views are shown. (A) View of the side of Ufd2p from where the U-box domain protrudes. (B) The view in A is rotated 200° away from the reader. The white arrow marks the only conserved spot in the region of residues 1–380.

Fig. 6. Superposition of residues 196–857 of the Ufd2p core region (green) and residues 141–509 of importin α (Protein Data Bank ID code 1EE4) (magenta) using the overlay matrix determined by DALI.

Fig. 7. A model of the Ufd2p/E2 complex. The Ufd2p core region is shown in green, its U-box domain is shown in red, and the docked E2 (Ub4cp) is shown in cyan. For clarity, the U-box domain of the CHIP–E2 complex is not shown.

We tested the model of the Ufd2p/E2 complex by demonstrating a direct interaction between Ufd2p and its cognate E2 Ub4cp (Fig. 8A and B) and a reduced interaction between the U-box-deleted Ufd2p mutant and Ub4cp (Fig. 8C). Thus, Ufd2p may act similarly to an E3 Ub ligase to mediate the transfer of Ub from Ub4cp to a substrate. To test this hypothesis, we developed a Ub turnover assay to monitor the single-round transfer of Ub from the active cysteine in Ub4cp to free Ub acceptors (Fig. 8D). His-tagged Ub4cp purified from Escherichia coli was incubated with a low concentration of Flag-Ub in the presence of E1 and ATP. This charging reaction generated Ub4cp carrying a Flag-Ub moiety on its catalytic cysteine residue (Fig. 8E, labeled Ub4cp·F-Ub; confirmed by mass spectrometry, Fig. 8F, lane 1, labeled U8) as well as some Ub4cp molecules that are linked to Ub via a lysine residue (Fig. 8E, labeled Ub4cp·F-Ub; confirmed by mass spectrometry, Fig. 8F, lane 1, labeled U7; note that this species is resistant to reducing agent; see Fig. 8F). Mass spectrometry experiments showed that the higher-molecular-weight species in the range of 35–37 kDa (marked by asterisks) were also ubiquitinated Ub4cp that contain extra Ub molecules (Fig. 8H, lane 1, labeled U5 and U6) attached to either a lysine or a noncatalytic cysteine residue, which was likely generated by Ub4cp autoubiquitination. The charging reaction was quenched by addition of EDTA and N-ethylmaleimide to prevent additional rounds of charging. The quenched reaction was further incubated in the presence of Ufd2p and excess untagged acceptor Ub. After incubation, a fraction of Flag-Ub molecules attached to the catalytic cysteine in Ub4cp was transferred to the untagged Ub acceptor Ub. After incubation, a fraction of Flag-Ub molecules attached to the catalytic cysteine in Ub4cp was transferred to the untagged Ub acceptor, leading to the formation of unanchored diubiquitin molecules (Fig. 8E, lanes 5 and 6, labeled F-Ub·Ub). The two Ub molecules are likely linked via an isopeptide bond because the F-Ub·Ub band is resistant to treatment by a reducing agent (Fig. 8F). As expected, no such diubiquitin molecules were formed in the absence of Ub acceptor or when a methylated Ub mutant was used as acceptor (Fig. 8E). Furthermore, the Ufd2p ΔBox mutant did not yield the formation of unanchored diubiquitin (Fig. 8G, lanes 4–6). These data suggest that the U-box of Ufd2p promotes the discharge of Ub from the catalytic cysteine of Ub4cp, a property shared by RING domain E3 ligases.

Interestingly, a fraction of Ub4cp-conjugated Flag-Ub was transferred to a protein of ~110 kDa to form a Ub ladder (Fig. 8E and G, labeled Ufd2p·F-Ubα; Fig. 8H, lane 1, labeled U1–U4). Because E1 and Ufd2p were the only two proteins that had the molecular mass of ~110 kDa, we performed immunoblotting experiments to rule out that E1 was ubiquitinated (Fig. 8I). Mass spectrometry analyses further confirmed that the Ub
Fig. 8. Ufd2p has E3 ligase activities. (A) Association of Ubc4p with Ufd2p. Biotinylated Ufd2p or the Ufd2p/H9004 box mutant immobilized on streptavidin-coated beads was incubated with purified Ubc4p. The precipitated material was analyzed by SDS/PAGE and Coomassie staining. Two salt conditions were used for binding: either 50 mM sodium chloride (lanes 4–6) or 150 mM potassium chloride (lanes 7–9). Each input protein was also analyzed directly (lanes 1–3). (B) Another independent binding experiment using the 50 mM salt condition as in A. Ten percent of the reaction was also probed by an anti-T7 antibody immunoblotting (lanes 4–6). (C) Amount of Ubc4p bound to Ufd2p or to Ufd2p/H9004 box quantified from Western blots of three independent binding experiments (using the 50 mM salt condition). Results were normalized based on the amount of Ufd2p or Ufd2p/H9004 box present in the precipitated materials. Error bars show the standard deviation. (D) Schematic representation of the single-round turnover experiment. (E) Transfer of Ub from Ubc4p to an acceptor Ub. Ubc4p charged with Flag-Ub was incubated with excess untagged Ub in the presence of Ufd2p. Where indicated, the reaction was carried out with no Ub acceptor or with methylated Ub (Me-Ub). The double bands marked by asterisks are Ubc4p molecules conjugated with two Ub molecules with the upper one likely containing lysine-linked Ub and the bottom band carrying only cysteine-linked Ub. (F) The sample in E, lane 6, was analyzed under both reducing and nonreducing conditions. (G) The formation of diubiquitin molecules and Ufd2p autoubiquitination require the Ufd2p U-box. Experimental details are as in E. Where indicated, the Ufd2p/H9004 box mutant was used. (H) Gel for mass spectrometry analysis. Experimental details similar to E, except that the reaction time was 1 h (lane 1) and the acceptor Ub was omitted from the reaction. Lane 2 was a mock reaction that contained everything, but not E1. The reactions were incubated with glutathione beads to remove most GST-E1. Ninety-five percent of the reaction was analyzed by Coomassie staining (Right), whereas the rest was analyzed by immunoblotting (IB) (Left). The bands labeled as U1–U8 were subjected to mass spectrometry analyses. (I) Experimental details are as in H. The Western blots show that GST-E1 is not ubiquitinated.
molecules were indeed conjugated to Ufd2p (data not shown). As expected, Ufd2p autoubiquitination required its U-box (Fig. 8G). Taken together, we conclude that Ufd2p has several hallmarks of a RING E3 ligase: direct association with its cognate E2, promotion of the transfer of E2-conjugated Ub to a free Ub molecule, and autoubiquitination.

**Prediction of Interactions Between Ufd2, Cdc48p, and Rad23p.** Biochemical studies suggest that Ufd2p binds Cdc48p via charged residues because the interaction between the two molecules is highly sensitive to salt concentration (24). Yeast two-hybrid analyses showed that a segment between Thr-808 and Lys-856 in Ufd2p may be critical for its interaction with Cdc48p (11). Within that region there are several conserved surface side chains: Asp-826, Glu-827, and Arg-828; Arg-844; and Glu-855 (SI Fig. 13). These residues are located near the C terminus of the core region of Ufd2p. The two proximal residues Arg-844 and Glu-855 are likely candidates for binding to Cdc48p because they are unobstructed by other domains and fully exposed, which may be required to accommodate the interaction with the 600-kDa Cdc48p.

The region in Ufd2p that is involved in Rad23p binding (1–380) (11) is less conserved. However, a conserved and positively charged spot stands out as a potential binding site for Rad23p, which includes residues Phe-326, Asp-328, Lys-333, Asn-380, and Phe-381 (white arrow in Fig. 5B).

**Discussion**

Ufd2p is the first discovered member of the U-box domain-containing protein family that is involved in polyubiquitination (22). Ufd2p was initially termed E4 because it can cooperate with an E3 to extend a Ub chain on an oligoubiquitinated artificial substrate. Subsequent studies showed that Ufd2p may be also involved in escorting ubiquitinated proteins from Cdc48p to the proteasome for degradation (11). During this process, Ufd2p appears to interact first with Cdc48p to upload polyubiquitinated substrates. Subsequently, the Ufd2p–substrate complex may dissociate from Cdc48p, allowing Ufd2p to bind and hand over substrates to Rad23p, a proteasome-associated factor.

The structure of Ufd2p provides some clues on how it can have such distinct functions. The large elongated body of Ufd2p is well suited for interactions with multiple binding partners. Furthermore, we have found an unexpected structural similarity between the core region of Ufd2p and the nuclear transporter protein, importin α, which may imply a transport function for Ufd2p. With respect to the polyubiquitination function of Ufd2p, Ufd2p also contains an U-box domain that is both structurally and functionally related to the RING domain in certain Ub E3 ligases. Based on the crystal structure of the complex between U-box domain of CHIP and E2 Ubc13 (16), we predicted a complex between Ufd2p and its cognate E2 (Fig. 7). The modeled interaction between the U-box domain and E2 is similar to that between a RING domain and an E2 (25); in both cases a hydrophobic ridge on the surface of E2 inserts into a hydrophobic groove formed by a short α-helix and the tips of two hairpin turns on the U-box or RING finger surface.

The ability of Ufd2p to bind to Ubc4p (Fig. 8A and B) and the structural similarity between the U-box and RING domains suggested that Ufd2p can act as an E3 Ub ligase. We tested this hypothesis by a single-round Ub turnover assay and found that Ufd2p functions by a mechanism similar to a RING domain Ub ligase. In particular, Ufd2p catalyzes the formation of diubiquitin by transferring Ub from an E2 to free Ub acceptors, a hallmark of a RING domain-containing E3 ligase. Furthermore, we found that Ufd2p can ubiquitinate itself in the presence of just E1 and E2. This is consistent with previous findings that mammalian UFD2 and other U-box proteins can catalyze the assembly of long polyubiquitin chains on themselves and heterologous substrates (26). We therefore propose that Ufd2p can function as a bona fide E3 ligase.

In the work of Koepl et al. (6), Ufd2p was classified as an E4 enzyme because Ufd2p cannot cooperate directly with an E2 in the ubiquitination reaction of a particular substrate: its ability to ubiquitinate a model UFD substrate (Ub-ProTα) requires the presence of a HECT domain E3. Our results do not contradict their conclusion because they only examined the fate of Ub-ProTα instead of following the flow of the Ub molecules. Thus, autoubiquitination of Ufd2p may have escaped detection. It is likely that the nature of the substrate prevents Ufd2p from directly using E2 to ubiquitinate Ub-ProTα. For example, the Ufd2p may have a low affinity to this substrate and thus would require another E3 to initiate the ubiquitination process.

How Ufd2p assists Ufd4p, a HECT domain E3, to extend oligoubiquitin chains remains to be determined. Based on our data we propose the following two possibilities. First, Ufd2p may act downstream of Ufd4p (sequential model in Fig. 9A). In this model, Ub would be first transferred via Ubc4p to the HECT domain of Ufd4p before being transferred to a substrate. Once a few Ub moieties have been added to the substrate, Ufd4p may become inactive and its role is taken over by Ufd2p. Ufd2p would then transfer Ub directly from an E2 to the substrate to extend the polyubiquitin chain. However, this sequential model does not explain the observation that Ufd2p combined with an E2 is unable to polyubiquitinate a purified oligoubiquitinylated substrate Ub-ProTα (6). In alternative models (cooperative models in Fig. 9B), Ufd2p would work in conjunction with Ufd4p. The two proteins may form a complex together with the E2 Ubc4p. Within this complex, Ufd2p may promote the transfer of Ub from Ubc4p to a substrate protein directly or via Ufd4p. Alternatively, Ufd4p may serve as a cofactor to activate Ufd2p, allowing it to efficiently polymerize Ub molecules on a substrate once a few Ub molecules have been added to it by Ufd4p. Our crystal structure of the yeast Ufd2p provides a platform for
phase separation when drops were exposed to air at room temperature. Addition of 20% PEG-3350 and 300 mM triammonium citrate quickly produced a precipitate. Combination of 30% PEG-3350, 200 mM tripotassium citrate, and 5 mM DTT. This crystal was cryoprotected as described above, was diffracted to d_{min} = 2.5 Å, and was less anisotropic. Integration, scaling, and merging of the diffraction data were performed with HKL2000 (28).

Structure Determination. The Se substructure was found by using dual-space direct methods with Patterson seeding as implemented in SHELXD (29). The structure factors, F_{o} – F_{c}, of the substructure were calculated with SHELXC. The correlation coefficient between the signed anomalous difference ΔF at the peak and high energy remote wavelengths fell below 25% past 3.5-Å resolution; hence, the MAD data were truncated at that resolution. The number of trials was set to 10, and seven of them produced solutions with correlation coefficients >63.0%. The top solution (correlation coefficient = 64.0) produced 18 selenium sites, and there was a large drop in occupancy for an additional site (0.31 vs. 0.17). The correct hand was determined by observing a much higher contrast value between the protein and solvent densities compared with that of the wrong hand (0.766 vs. 0.145) after density modification in SHELXE. All subsequent phasing, density modification, and refinement calculations were performed by using CNS 1.2 (30). Experimental MAD phase probability distributions were obtained from the Se substructure. The positions, thermal factors, and anomalous f{partial} and f{partial}^o factors were refined individually for each atom and wavelength. Phases were further improved by density modification and extended to d_{min} = 2.74 Å by using data set Native2.

Model Building and Refinement. The initial model was built by using the program COOT (31). The density-modified, phase-extended experimental map had clear electron density for most helices, loops, and side chains (Fig. 1). Thirty polyalanine α-helices were automatically placed into the electron density map by using the “Place Helix Here” feature in COOT. Guided by the 18 selenium sites, it was relatively straightforward to trace the loops and remaining helices and assign the entire sequence of the protein. The progress of model refinement was monitored by using the free R value computed from a randomly omitted set of 10% of the observed diffraction data at 50- to 2.65-Å resolution. The refinement consisted of alternating rounds of torsion angle molecular dynamics simulated annealing, individual restrained thermal factor refinement, and model building in COOT. All model refinement used the maximum likelihood target function using amplitudes and experimental phase probability distribution (option MLHL in CNS). The bulk solvent model refined to an electron density level ρ_{sol} = 0.3 e/Å^{3} with B_{sol} = 42.9 Å^{2}. The final model of data set Native2(Se) consists of residues 1–25, 34–706, and 719–951 and 94 water molecules.

The higher-resolution data set Native1 was solved by molecular replacement starting with the previous model using Phaser (32). Initially the entire refined structure of data set Native2(Se) was used as the search model. The electron density map calculated from molecular replacement phases showed that the C-terminal U-box had moved significantly. Consequently the structure was split into two fragments, residues 1–879 and 880–951 (U-box), which were then used as separate groups in rigid body refinement. The resulting model was subjected to molecular dynamics simulated annealing and restrained thermal factor refinement. Refinement used the maximum likelihood function of the experimental structure factors.
target function (option MLF in CNS) using amplitudes of Native1. The bulk solvent refined to an electron density level $\rho_{\text{sol}} = 0.3$ eÅ$^{-2}$ with $B_{\text{sol}} = 35.0$ Å$^2$. The final model consists of residues 1–27, 35–707, and 718–954, 145 water molecules, and one potassium ion. Final model statistics are shown in Tables 1 and 2. The atomic coordinates and diffraction data have been deposited in the Protein Data Bank [ID codes 2QIZ (Native1) and 2QJO (Native2)].

**Biochemical Experiments. Binding assay.** The biotinylation labeling was performed by using EZ-link sulfo-NHS-Biotin (Pierce) according to the manufacturer’s instruction. The E2 binding was performed by using EZ-link sulfo-NHS-Biotin (Pierce) and 2QJ0 (Native2).

**Binding assay.** The biotinylation labeling was performed by using EZ-link sulfo-NHS-Biotin (Pierce) according to the manufacturer’s instruction. The E2 binding was performed by using EZ-link sulfo-NHS-Biotin (Pierce) and 2QJ0 (Native2).

**Figure Preparation.** Figs. 1–7 and SI Figs. 10, 12, and 13 were prepared by using PyMol (33).

We thank Drs. Ron Kopito and Timothy Fenn for critical reading of the manuscript, Stefan Jentsch (Max Planck Institute of Biochemistry, Munich, Germany) for the Ufd2p plasmid, Rongsheng Jin for assistance with data collection, and Corie Ralston (Advanced Light Source Beamline 8.2.2) and Irimpan Mathews (Stanford Synchrotron Radiation Laboratory Beamline 11-1) for advice during data collection. The Stanford Synchrotron Radiation Laboratory is a national user facility operated by Stanford University on behalf of the U.S. Department of Energy (Office of Basic Energy Sciences). The Stanford Synchrotron Radiation Laboratory Beamline Structural Molecular Biology Program is supported by the Department of Energy (Office of Biological and Environmental Research) and by the National Institutes of Health (National Center for Research Resources, Biomedical Technology Program), and the National Institutes of General Medical Sciences. The Advanced Light Source is supported by the Office of Energy Research (Office of Basic Energy Sciences, Material Sciences Division) of the U.S. Department of Energy at Lawrence Berkeley National Laboratory. This work was supported by the Howard Hughes Medical Institute (A.T.B.) and by an intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (Y.Y.).

Supplemental Fig. S10. Electron density map of Ufd2p. Shown is a $\sigma_A$ weighted 2Fo-Fc map, contoured at 1 $\sigma$, computed at 2.56 Å resolution. Shown is the region around helix a23 (residues 526-544).

Supplemental Fig. S11. Sequence alignment of four Ufd2p homologues. Secondary structural elements are indicated above the alignment. Identical residues are indicated by the red background, conserved residues are indicated by red characters. This figure was produced using ESPript 2.2 (http://espript.ibcp.fr).

Supplemental Fig. S12. Electrostatic potential surface. The two views are identical to those in Figures 5A and 5B, respectively. The electrostatic potential surface is contoured at 10 kT/e with red and blue colors denoting negative and positive charge, respectively.

Supplemental Fig. S13. Potential interaction sites of Ufd2p with Cdc48p.
<table>
<thead>
<tr>
<th>Mouse Elegans</th>
<th>Yeast</th>
<th>Drosophila</th>
<th>Mouse Elegans</th>
<th>Yeast</th>
<th>Drosophila</th>
<th>Mouse Elegans</th>
<th>Yeast</th>
<th>Drosophila</th>
<th>Yeast</th>
<th>Drosophila</th>
<th>Yeast</th>
<th>Drosophila</th>
<th>Yeast</th>
<th>Drosophila</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure S11
Figure S13