nature structural & molecular biology

Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases

Woo Suk Choi[1], Byung-Cheon Jeong[1], Yoo Jin Joo[1], Myeong-Ryeol Lee[1], Joon Kim[1], Michael J Eck[2,3] & Hyun Kyu Song[1]

The N-end rule pathway is a regulated proteolytic system that targets proteins containing destabilizing N-terminal residues (N-degrons) for ubiquitination and proteasomal degradation in eukaryotes. The N-degrons of type 1 substrates contain an N-terminal basic residue that is recognized by the UBR box domain of the E3 ubiquitin ligase UBR1. We describe structures of the UBR box of *Saccharomyces cerevisiae* UBR1 alone and in complex with N-degron peptides, including that of the cohesin subunit Scc1, which is cleaved and targeted for degradation at the metaphase-anaphase transition. The structures reveal a previously unknown protein fold that is stabilized by a novel binuclear zinc center. N-terminal arginine, lysine or histidine side chains of the N-degron are coordinated in a multispecific binding pocket. Unexpectedly, the structures together with our *in vitro* biochemical and *in vivo* pulse-chase analyses reveal a previously unknown modulation of binding specificity by the residue at position 2 of the N-degron.

The N-end rule pathway is a regulated protein-degradation system that has a crucial role in all kingdoms of life[1–3]. In eukaryotes, substrates in the pathway that contain an N-degron are destined for ubiquitination and proteasomal degradation[2,4]. N-end rule–mediated degradation underlies many biological processes such as chromosome segregation[5], the import of short peptides[6], neurogenesis and cardiovascular development[7], mental retardation[8], apoptosis[9], the sensing of heme, nitric oxide and oxygen[10–12], plant development[13,14] and DNA repair[15].

The E3 ubiquitin ligase UBR1 is a component of the N-end rule pathway in all eukaryotes[2]. A deficiency of E3 in humans is associated with Johanson-Blizzard syndrome, an autosomal recessive disorder associated with pancreatic dysfunction, mental retardation and physical malformations[8]. Two classes of N-end rule substrates are recognized by two distinct domains of UBR1 (**Fig. 1a**). Type 1 substrates contain a basic N-terminal residue and are recognized by the cysteine-rich UBR box domain (**Supplementary Fig. 1**), whereas type 2 substrates contain a large hydrophobic N-terminal residue and are recognized by the ClpS-homology domain of UBR1 (refs. 2,16–19). The UBR box and ClpS-homology domains are unrelated in primary sequence.

A functional N-degron is created by proteolytic scission of a pro–N-degron, resulting in the exposure of a destabilizing N-terminal residue. For example, the cleavage of the *S. cerevisiae* cohesin subunit Scc1 by the protease separase during the metaphase-anaphase transition yields a 33-kDa fragment of Scc1 with an N-terminal arginine residue[5]. This fragment is recognized and targeted for degradation by UBR1, which binds the newly formed N-degron via the UBR domain. Because it is directly recognized by UBR1, the N-terminal arginine of this substrate is referred to as a 'primary' destabilizing residue. Such residues are directly recognized by a cognate ubiquitin ligase such as UBR1. Other destabilizing N-terminal residues such as asparagine, glutamine, aspartate and glutamate are referred to as 'secondary' or 'tertiary' destabilizing residues, as they undergo enzymatic deamidation and/or arginylation before the corresponding N-end rule substrates can be recognized by the UBR box[20,21].

Although the UBR box was the first substrate-binding site of a ubiquitin ligase to be characterized[22], delineation of its structure has remained elusive. Here we describe crystal structures of the UBR box of *S. cerevisiae* E3 ligase UBR1 alone and in complex with N-degron peptides, including that of Scc1. The structures reveal a previously undescribed protein fold that is stabilized by coordination of three zinc ions, two of which form an atypical binuclear zinc center. N-degron peptides bind in a shallow cleft on the surface of the domain. Comparison of diverse N-degron complexes shows that basic N-terminal residues arginine, lysine or histidine are coordinated by distinct sets of interactions with the domain. Additionally, the structures reveal specificity-determining interactions with the residue at position 2 of the N-degron. Our biochemical and functional analyses including peptide array, isothermal titration calorimetry (ITC) and *in vivo* pulse-chase studies of N-degron recognition confirm a previously unknown modulation of binding specificity by the residue at position 2 of the N-degron.

RESULTS

Overall structure of the UBR box

The crystal structure of the UBR box reveals a compact, heart-shaped domain with three zinc-coordination sites and little regular secondary structure (**Fig. 1b**). The V-shaped base of the domain is formed by

[1]School of Life Sciences and Biotechnology, Korea University, Anam-Dong, Korea. [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, USA. [3]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. Correspondence should be addressed to H.K.S. (hksong@korea.ac.kr).

Received 11 May; accepted 20 July; published online 12 September 2010; doi:10.1038/nsmb.1907

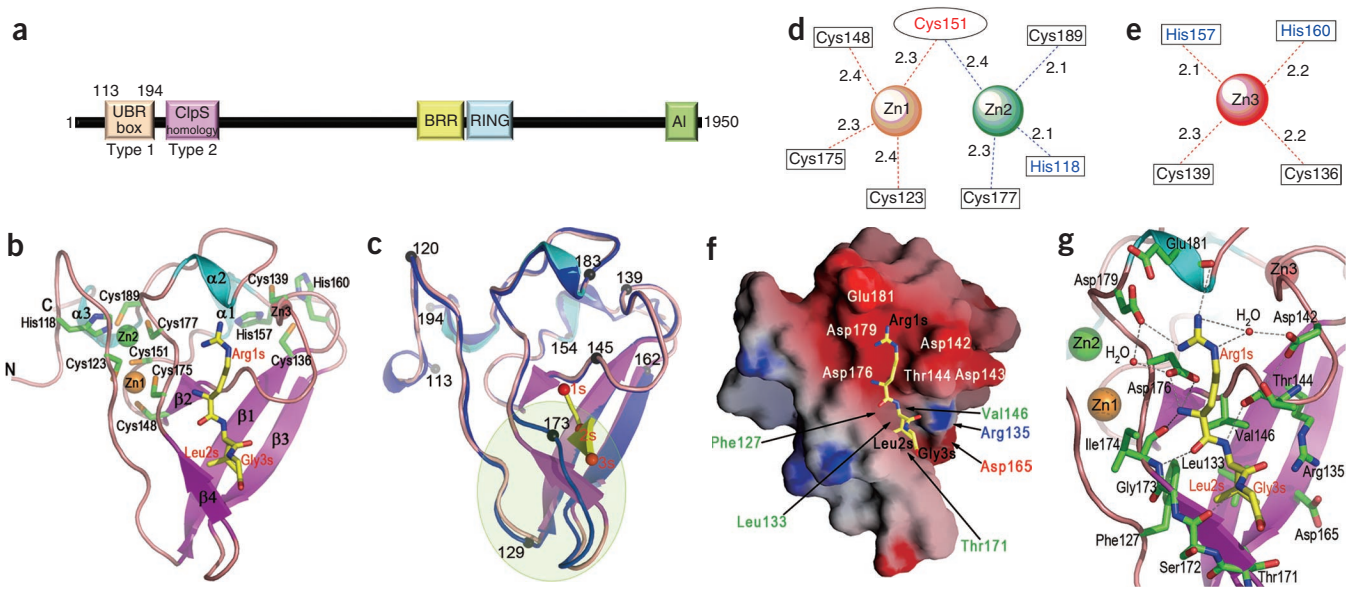

Figure 1 Structure of the UBR box. (**a**) Domain architecture of UBR1 from *S. cerevisiae*. The UBR box and ClpS-homology domain recognize type 1 and 2 N-end rule substrates, respectively. The basic residue–rich (BRR) region and RING domain are responsible for interaction with the E2 ubiquitin-conjugating enzyme Ubc2 and for E3 ligase activity. The C-terminal auto-inhibitory (AI) domain regulates ubiquitin ligase activity of UBR1. (**b**) Ribbon diagram showing the structure of the UBR box in complex with Scc1 peptide (RLGES). The bound peptide molecule is shown as a stick model, with carbon atoms in yellow. Residues in the peptide substrate are labeled in red and the letter 's' is appended to their sequence number. Side chains of residues in the UBR domain that participate in zinc coordination are shown in stick form, and the bound zinc ions are shown as orange, green and red spheres (for Zn1, Zn2 and Zn3, respectively). (**c**) Superposition of structures of the peptide-free UBR domain (blue ribbon) and the RLGES-complexed UBR domain (colored as in **b**). Selected residues in the peptide-free structure are marked with a black sphere and labeled. The region showing conformational change upon complex formation is indicated with a transparent green oval. (**d**,**e**) Schematic diagrams showing zinc coordination. Distances (Å) between coordinating residues and zinc ions are shown. UBR box cysteine and histidine residues are shown in open rectangles with black and blue labels, respectively. Colors for zinc atoms are the same as in **b**. (**d**) The binuclear zinc center. The cysteine residue (Cys151) bridging two zinc atoms is shown. (**e**) The third zinc-coordination site. The structure and corresponding electron density in these regions are shown in **Supplementary Figure 2a**,**b**. (**f**) Molecular surface showing the electrostatic potential of the UBR box. Negatively and positively charged surfaces are shaded red and blue, respectively. The bound peptide is shown in yellow. Selected residues that comprise the N-degron–binding cleft are labeled. (**g**) Interaction of the UBR domain with Scc1 peptide. Hydrogen bonds are indicated by dashed lines, and 'bridging' water molecules are shown as red spheres.

the intersection of a small β-sheet (formed by strands β1 and β3) and two irregular loops (**Fig. 1b**). The more bulbous upper aspect of the domain is formed by long connecting loops, which include three short segments of 3_{10}-helix, and by the zinc-coordination sites. As described below, N-degron peptides bind in the cleft of the V-shaped base, forming β-sheet interactions with strand β4 that are induced by complex formation (**Fig. 1c**). Strands β2 and β4 are foreshortened by participation of their C-terminal extensions in the formation of a binuclear zinc center (Zn1 and Zn2; **Fig. 1b**,**d**). A third zinc ion (Zn3) with typical Cys_2His_2 tetrahedral coordination is on the opposite side of the domain at the end of strands β1 and β3 (**Fig. 1b**,**e**). This zinc is expected to be important in stabilizing the fold of the UBR box. The overall fold of the UBR box has not previously been reported. Specifically, a search of protein structures in the Protein Data Bank using DALI and COPS structural-similarity servers did not reveal other proteins with similar three-dimensional folds[23,24].

The two zinc ions in the binuclear zinc center are tetrahedrally coordinated by a total of six cysteine residues and one histidine residue (**Fig. 1d** and **Supplementary Fig. 2**). One of the cysteine residues, Cys151, is bound to both zinc atoms, forming a bridge between them. This pattern of coordination differs from the well-known binuclear zinc clusters, such as that present in the DNA-binding domain of the Gal4 transcription factor, which forms two bridges between zinc atoms using two of the six cysteine residues and no histidine[25,26]. With the exception of His118, each of the residues involved in zinc binding is well-conserved in other UBR domain proteins[17]. His118 is near the N terminus of the UBR domain and is highly conserved among UBR1 E3 ligases of different organisms (**Supplementary Fig. 1**), but not in other UBR family proteins[17]. The apparent lack of conservation of this residue outside UBR1 E3 ligases is likely to stem from ambiguities in primary-sequence alignments in this region of the protein, and from the ability of other residues to participate in zinc coordination in lieu of this histidine. The imprecise definition of the domain boundaries of the UBR box (based on primary-sequence alignments) may have frustrated earlier efforts to elucidate the structure of this domain.

Recognition of the N-degron

Type 1 N-degrons bind in a relatively shallow, acidic cleft on the surface of the UBR domain (**Fig. 1f**), forming antiparallel β-sheet interactions with strand β4. The exposed nature of the substrate-binding site of UBR and its broad negatively charged surface account, at least in part, for the ability of this domain to recognize diverse type 1 N-degrons that include a basic N-terminal residue. The structure of the UBR box in complex with the Scc1 N-degron, which contains an N-terminal arginine and bears the sequence Arg-Leu-Gly-Glu-Ser (RLGES hereafter), is shown in **Figure 1b**. The first three residues of the peptide are ordered in the structure. The binding involves interactions with backbone atoms of each of the three residues, with side chains of the first and second residues, and with the protonated N-terminal NH_3^+ group (**Fig. 1g**). The participation of three residues is consistent with our binding data, where we find that a longer peptide ligand results in, at most, a marginally higher binding affinity (**Supplementary Fig. 3**).

Recognition of the positively charged N-terminal NH_3^+ group of the peptide seems to be particularly important. The side chain of Asp176 in UBR forms dual hydrogen bonds with the α-amino group, and the main chain carbonyl of Ile174 contributes a third hydrogen bond (**Fig. 1g**). This coordination is consistent with protonation of the α-amino group (to the positively charged, NH_3^+ state), as would be required for Asp176's participation as a donor in three hydrogen bonds. Consistent with its central structural role in recognizing the α-amino group of the N-degron, a recent mutagenesis study showed that Asp176 is essential for the function of the UBR box[27]. Asp176 is located between two zinc-coordinating cysteine residues (**Supplementary Fig. 1**), further highlighting the crucial role of zinc coordination in the structure and function of the domain. Additional substrate–UBR main chain hydrogen bond interactions are formed by Ser172 and Gly173 in strand β4, and by Thr144 on the opposite side of the binding cleft (**Fig. 1g**). The side chain of the N-terminal arginine residue forms a salt bridge with Asp179, a hydrogen bond with the main chain carbonyl of Glu181 and a water-mediated interaction with Asp142. This constellation of interactions is specific to the side chain of the N-terminal arginine, as each of its guanidinium nitrogen atoms participates in hydrogen bonds with the UBR domain. In contrast to the specific hydrogen bond interactions with the N-terminal arginine, the leucine side chain of the second residue is accommodated in a shallow hydrophobic pocket formed by residues Phe127, Leu133, Val146 and Thr171 (**Fig. 1f**). This pocket is partially exposed to solvent and is not completely filled by the leucine side chain. Constraints imposed by coordination of the peptide backbone orient the residue at position 3 so that its side chain extends away from the surface of the domain; thus, it is unlikely to form specificity-determining interactions. In the Scc1 peptide, this residue is glycine, which lacks a side chain, but examination of the alanine side chain in the corresponding position of the Arg-Ile-Ala-Ala-Ala (RIAAA) and other model peptides discussed below confirms this interpretation (**Supplementary Fig. 4**).

Comparison of the complex structure with the unliganded UBR domain reveals conformational changes that are induced upon N-degron binding (**Fig. 1c**). These conformational changes are localized in the N-degron–binding site. In the absence of bound peptide, the β4 strand in the UBR box is not formed. Instead, the corresponding residues (172–176) form a loop that takes another path, partially occluding the peptide-binding cleft (**Fig. 1c**).

We also determined the structure of a longer UBR1 construct (residues 107–194) that includes the UBR domain in the unliganded state (**Table 1**). In this structure, the N terminus of an adjacent molecule in the crystal lattice occupies the N-degron–binding cleft, forming analogous backbone interactions, including key hydrogen bonds between its free (protonated) N-terminal amino group and Asp176 (**Supplementary Figs. 5** and **6**). Although the N-terminal segment does not correspond to an N-end rule sequence (Gly-Ser-Val-His-Lys-), it binds in the N-degron–recognition site apparently because of the high local concentration of the protein in the crystal. The main chain interactions of the bound segment after conformational change are identical to those in the N-degron peptide complex structures (**Fig. 1g** and **Supplementary Figs. 4** and **6**). This structure further highlights the central importance of recognition of the free N terminus and nonspecific main chain interactions by the UBR domain. Recent structural analysis of ClpS, an adaptor protein of the bacterial N-end rule pathway that recognizes large hydrophobic N-terminal residues, has also revealed a central role for coordination of the α-amino group in type 2 N-degrons[28–30] (**Supplementary Fig. 7**). However, the overall structure and N-degron recognition of ClpS are distinct from those of the UBR box and, furthermore, no conformational change was evident in ClpS upon complex formation[28], in contrast with the substantial conformational changes that we observed in the UBR box.

Recognition of other basic residues by the UBR box

Notably, three basic residues, arginine, lysine and histidine, are all specifically recognized by this small domain, despite their difference in structure. In an effort to investigate the recognition of other basic N-terminal residues by the UBR domain, we determined structures of the domain in complex with arginine-, lysine- and histidine-based model peptides (RIAAA, KIAA and HIAA, respectively) that have been previously used to study UBR box specificity. The N-terminal arginine, lysine and histidine side chains were found to be recognized

Table 1 Data collection, phasing and refinement statistics of the native UBR box (MAD)

	Zinc MAD			
	Native 1			Native 2
Data collection				
Space group		$P6_522$		$P3_2$
Cell dimensions				
a, *b*, *c* (Å)		39.2, 39.2, 274.5		44.6, 44.6, 140.0
α, β, γ (°)		90, 90, 120		90, 90, 120
	Peak	**Edge**	**Remote**	
Wavelength	1.28305	1.28325	1.12713	1.1
Resolution (Å)	2.88 (2.88–2.98)	2.86 (2.86–2.96)	2.80 (2.80–2.90)	1.68 (1.68–1.74)
R_{sym}	8.3 (47.4)	8.2 (48.9)	8.6 (58.5)	9.1 (52.1)
$I / \sigma I$	81.8 (6.5)	76.6 (5.8)	52.5 (4.5)	33.3 (3.3)
Completeness (%)	97.2 (75.2)	98.0 (82.8)	99.5 (100)	100 (99.9)
Redundancy	37.7	37.5	19.4	7.6
Refinement				
Resolution (Å)		45.8–2.60		29.8–1.68
No. reflections		4,194		33,688
R_{work}/R_{free}		24.4/28.4		19.3/22.1
No. atoms				
Protein		691		2,543
Ligand/ion		3 (zinc)		12 (zinc), 12 (acetate)
Water		21		262
B-factors				
Protein		60.5		18.1
Ligand/ion		54.9 (zinc)		18.2 (zinc)
				59.8 (acetate)
Water		59.5		31.7
R.m.s. deviations				
Bond lengths (Å)		0.007		0.007
Bond angles (°)		0.93		1.03

One crystal was used for each structure. Values in parentheses are for highest-resolution shell.

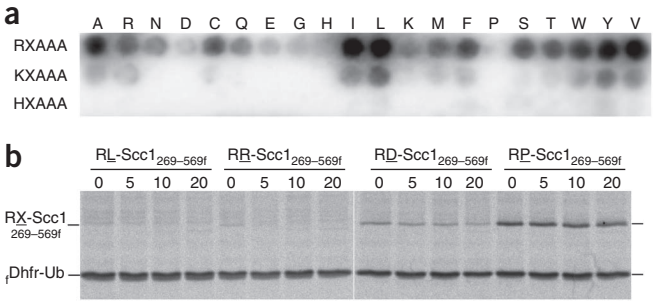

Figure 2 Effect of the residue at position 2 of N-degrons. (**a**) A SPOT peptide array in which each of the 20 naturally occurring amino acids was placed at position 2 of the peptides RXAAA, KXAAA or HXAAA. The substituted amino acid is indicated above the blot. Binding of the UBR1 UBR domain was detected using a horseradish peroxidase (HRP)-conjugated UBR box antibody. A parallel experiment using a GST–UBR box fusion protein and HRP-conjugated antibody to GST yielded similar results, but with increased detection sensitivity (**Supplementary Fig. 8**). (**b**) *In vivo* pulse-chase analysis of Scc1 variants (position 2 of the N-degron). Ub, ubiquitin. *S. cerevisiae* cells expressing $_f$Dhfr-Ub-Arg-Xaa-Scc1$_{269–566f}$ were labeled for 5 min with [^{35}S]methionine, then chased for 5 min, 10 min or 20 min, and anti-Flag immunoprecipitates were analyzed by SDS-PAGE. Bands corresponding to the Scc1 fragment (Arg-Xaa-Scc1$_{269–566f}$) and the reference protein $_f$Dhfr-Ub are indicated.

by distinct sets of interactions within the UBR domain, whereas the interactions with the backbone of these peptides are essentially identical. The coordination of the Arg-, Lys- and His-peptides is compared in **Supplementary Figure 4**. The arginine of the RIAAA peptide is bound in a manner identical to that in the Scc1 peptide described above. In contrast to the network of hydrogen bond interactions formed by the side chain of N-terminal arginine, the ε-nitrogen of the lysine side chain is positioned between Asp142 and Asp179, but is not close enough to form direct hydrogen bonds to either residue (**Supplementary Fig. 4b**). The side chain of the N-terminal histidine residue makes a hydrogen bond with the hydroxyl group of Thr144, and the imidazole nitrogen atoms are located ~4.0 Å and ~4.6 Å (that is, too far to form hydrogen bonds) from the side chain carboxylates of Asp179 and Asp142, respectively (**Supplementary Fig. 4c**).

To investigate the contribution of Asp142, Asp179, Thr144 and Glu181 of UBR to the recognition of N-degron peptides, we prepared mutants of UBR1 by replacing each of these residues with alanine and measured the affinity for RIAA, KIAA and HIAA peptides using ITC (**Supplementary Table 1**). Mutation of either Asp142 or Asp179 of the UBR domain abolished detectable binding of UBR1 to the Lys-peptide and had a weaker effect on the binding of His-peptide and Arg-peptide, decreasing binding affinity by approximately an order of magnitude. Consistent with conclusions drawn from the structure of the UBR box, mutation of Thr144 to alanine had the strongest effect on binding of the His-peptide. Despite the lack of direct hydrogen bonds between Asp142 and Asp179 and the N-terminal lysine and histidine residues of N-degron substrates, the negative electrostatic environment created by Asp142 and Asp179 is clearly important for the binding of these N-terminal residues.

Importance of residue at position 2 of N-degrons

Whereas the crucial role of the substrate N-terminal residue in promoting degradation has been extensively studied, the importance of the residue at position 2 has not been systematically examined. The UBR domain structures suggest that the residue located at position 2 may affect binding of type 1 substrates. To investigate this possibility, we prepared a peptide library arrayed in a spot format in which each of the 20 amino acids was placed at position 2 of the peptides to generate sequences RXAAA, KXAAA or HXAAA, where X denotes the substituted position. Binding of the purified UBR domain to peptides of the SPOT array is shown in **Figure 2a**, and a parallel experiment showing binding of a glutathione S-transferase (GST)-UBR domain fusion protein is shown in **Supplementary Figure 8**. Any amino acid except proline can be accommodated at position 2 of an Arg-peptide, but hydrophobic residues (isoleucine, leucine, valine, tyrosine, alanine and phenylalanine) at position 2 bind particularly strongly. Proline at this position can be expected to alter the main chain conformation and cannot form the backbone hydrogen bond with the carbonyl of Thr144, thus precluding high-affinity binding, irrespective of the type of N-terminal residue (**Fig. 1g**). Apart from proline and hydrophobic residues located at position 2 of a substrate, there is modest discrimination (hydrophobic > positively charged (except histidine) > polar > negatively charged) in terms of affinity among residues at this position in the Arg-peptides (**Fig. 2a** and **Supplementary Fig. 8**). However, greater discrimination is observed in the Lys- and His-peptides, which bind with lower affinity to the UBR domain (**Supplementary Fig. 8**).

To explore the effect of position 2 on the rate of degradation *in vivo*, we designed mutants at position 2 of an engineered *in vivo* N-end rule

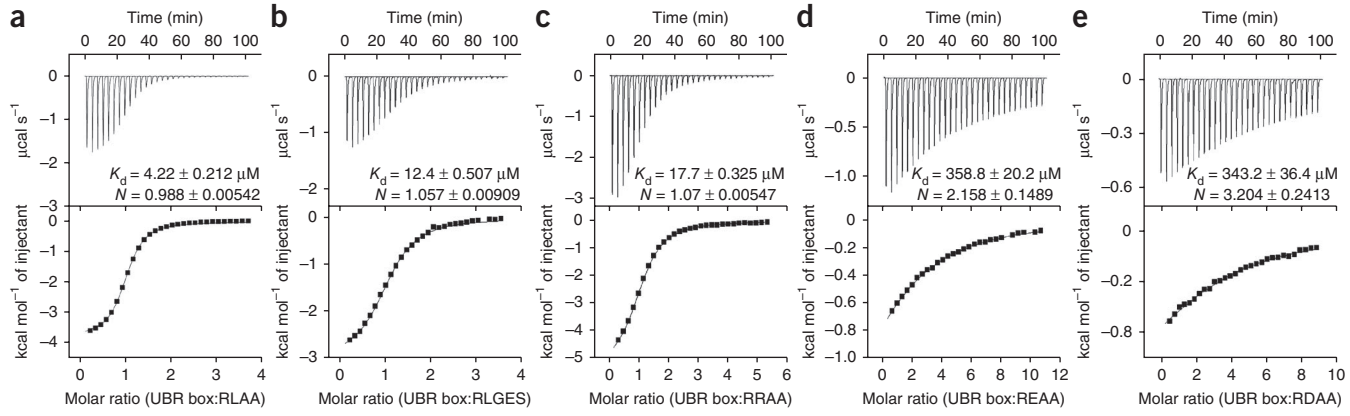

Figure 3 Binding affinities of the UBR box to peptides of different residues at position 2. (**a**–**e**) Measurement of the affinity of the UBR box for the RLAA peptide (**a**), RLGES peptide (**b**), RRAA peptide (**c**), REAA peptide (**d**) and RDAA peptide were (**e**) using ITC. The K_d values of the RIAA and RAAA peptides as measured by ITC were 7.30 ± 0.136 μM and 22.4 ± 0.49 μM, respectively (see **Supplementary Table 1** and **Supplementary Fig. 3**).

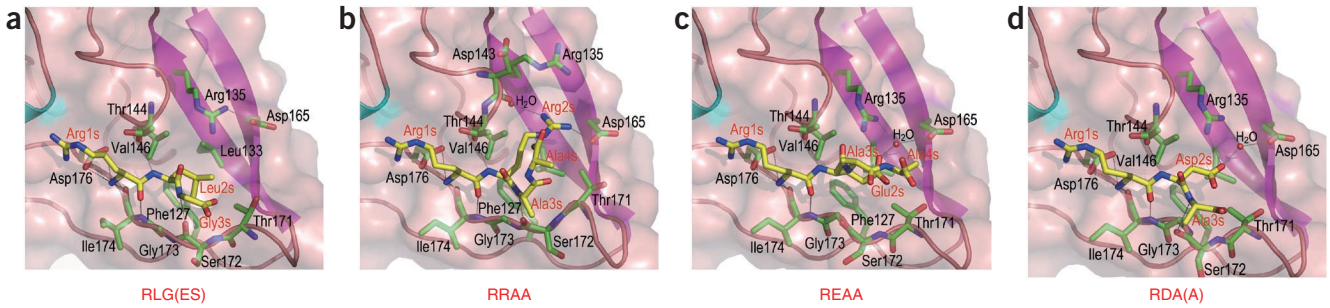

Figure 4 Details of UBR domain recognition of residues at position 2. (**a**–**d**) Interaction of the UBR box with the RLGES peptide (**a**), RRAA peptide (**b**), REAA peptide (**c**) and RDAA peptide (**d**). Transparent molecular surface of the UBR box is shown, and detailed interactions between the first arginine (Arg1s) in N-degron peptides and the UBR box are omitted for clarity. See **Supplementary Figure 4a** for a comparison with the RIAAA peptide.

substrate, a mimic of the separase-produced Scc1 fragment[5]. The plasmid expressed dihydrofolate reductase (Dhfr)-ubiquitin fused with Scc1 fragment ($_f$Arg-Leu-Scc1$_{269–566f}$, where position 2 is underlined and 'f' denotes the N-terminal and C-terminal Flag epitopes). We replaced the original leucine residue at position 2 with either arginine, glutamate or proline and performed *in vivo* pulse-chase experiments (**Fig. 2b**). Deubiquitinating enzymes co-translationally cleave this fusion protein at the C terminus of the ubiquitin moiety, producing a target Arg-Xaa-Scc1$_{269–566f}$ fragment and the reference Dhfr-ubiquitin protein. Wild-type Arg-Leu-Scc1$_{269–566f}$ was short-lived ($t_{1/2}$ < 2 min), as reported[5], and the arginine mutant (Arg-Arg-Scc1$_{269–566f}$) was also short-lived, consistent with our SPOT array data, which show clear binding to peptides with arginine at the second position (**Fig. 2a** and **Supplementary Fig. 8**). In contrast, the proline mutant (Arg-Pro-Scc1$_{269–566f}$) was long-lived, showing little if any degradation over the 20-min time course. Also consistent with the peptide array data, the aspartate mutant (Arg-Asp-Scc1$_{269–566f}$) showed an intermediate rate of degradation. This rate of degradation is expected to be physiologically relevant, as acidic aspartate or glutamate residues are present at position 2 of N-degrons derived by deamidation and/or arginylation of substrates with secondary destabilizing residues (aspartate or glutamate) or tertiary destabilizing residues (asparagine or glutamine) at their N terminus[20,31]. These data clearly show that the degradation of Scc1

Table 2 Data collection and refinement statistics of the UBR box:peptide complex (molecular replacement)

	RLGES	RIAAA	KIAA	HIAA	RRAA	REAA	RDAA
Data collection							
Space group	$P3_2$	$P6_122$	$P6_122$	$P6_122$	$P3_2$	$P3_2$	$P3_2$
Cell dimensions							
a, *b*, *c* (Å)	45.8, 45.8, 87.5	58.3, 58.3, 111.0	58.2, 58.2, 110.9	58.3, 58.3, 111.2	44.9, 44.9, 140.1	44.5, 44.5, 139.6	44.3, 44.3, 139.5
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	2.1 (2.1–2.14)	2.1 (2.1–2.18)	2.1 (2.1–2.18)	2.1 (2.1–2.18)	2.0 (2.0–2.07)	1.85 (1.85–1.88)	1.75 (1.75–1.78)
R_{sym}	6.1 (49.0)	7.7 (51.0)	6.9 (46.2)	8.2 (52.2)	8.0 (59.6)	7.8 (47.3)	6.4 (49.8)
$I / \sigma I$	48.3 (4.5)	83.2 (7.0)	83.4 (6.5)	78.0 (6.6)	31.2 (3.7)	32.9 (3.5)	34.4 (2.7)
Completeness (%)	99.8 (99.8)	99.8 (100)	99.8 (100)	99.6 (100)	92.8 (91.3)	100 (100)	99.2 (96.8)
Redundancy	6.2	26.3	23.7	21.0	4.9	6.9	6.5
Refinement							
Resolution (Å)	22.9–2.1	37.4–2.1	24.6–2.1	24.6–2.1	22.7–2.0	38.6–1.85	38.4–1.75
No. reflections	11,396	6,664	6,628	6,647	18,786	25045	29067
R_{work}/R_{free}	21.4/27.3	25.3/28.5	25.1/28.1	26.1/29.4	19.1/24.0	19.1/24.3	20.8/26.2
No. atoms							
Protein	1,256	618	618	618	2,538	2542	2525
Ligand/ion	6 (zinc)	3 (zinc)	3 (zinc)	3 (zinc)	12 (zinc)	12 (zinc)	12 (zinc)
	57 (peptide)	24 (peptide)	22 (peptide)	23 (peptide)	33 (peptide)	31 (peptide)	25 (peptide)
							8 (acetate)
Water	41	22	22	19	255	466	463
B-factors							
Protein	18.2	45.5	32.7	48.2	38.8	29.3	29.5
Ligand/ion	45.4 (zinc)	45.0 (zinc)	39.1 (zinc)	48.0 (zinc)	31.6 (zinc)	22.5 (zinc)	22.8 (zinc)
	55.7 (peptide)	60.7 (peptide)	70.0 (peptide)	78.0 (peptide)	57.3 (peptide)	32.1 (peptide)	44.9 (peptide)
							56.4 (acetate)
Water	23.1	53.4	50.3	53.4	48.1	43.2	41.2
R.m.s. deviations							
Bond lengths (Å)	0.008	0.008	0.007	0.007	0.009	0.008	0.009
Bond angles (°)	1.00	1.03	0.93	0.94	1.21	1.08	1.13

One crystal was used for each structure. Values in parentheses are for highest-resolution shell.

in *S. cerevisiae* is affected by the type of residue at position 2, as indicated by the peptide array experiment and our structural results.

To quantify the effect of different residues in position 2 on the affinity of N-degron peptides for the UBR box, we performed ITC experiments using a series of model peptides (**Fig. 3a–e**). Peptides bearing hydrophobic residues (leucine or isoleucine) at position 2 showed tight binding to the UBR box with dissociation constants (K_d) ranging from 4 μM to 13 μM (**Fig. 3a**, **Supplementary Figs. 9** and **10** and **Supplementary Table 1**). We also found relatively tight binding (~18 μM) for the RRAA peptide. However, REAA and RDAA peptides, bearing an acidic residue at position 2, bound the UBR box with an affinity that was reduced to approximately 3.5 % that of the RLGES peptide, consistent with our SPOT array results.

Structural plasticity of position 2 recognition by the UBR box

The structural basis for the preference of hydrophobic residues at position 2 is clear: the second residue is shown to extend into a hydrophobic pocket of UBR, as we observed with both the Scc1 (RLGES) and RIAAA peptides (**Fig. 4a** and **Supplementary Fig. 4a**). The residues forming this hydrophobic pocket (Phe127, Leu133, Val146 and Thr171) are fairly well conserved in UBR1 E3 ligase sequences (**Supplementary Fig. 1**), and in particular Val146 was previously identified as a crucial residue for type 1 substrate binding by extensive genetic screening[27].

To reconcile the binding of basic and acidic residues in this position with the apparently hydrophobic environment of the cognate pocket in the UBR domain, we determined the structure of the UBR domain in complex with RRAA, REAA and RDAA peptides (**Table 2**). The RRAA complex shows that the side chain guanidinium nitrogen of Arg2 of the peptide forms a salt bridge with Asp165, which lies at the outside of the hydrophobic pocket (**Fig. 4b**). A water molecule bridges the other guanidinium nitrogen of Arg2 to the main chain carbonyl oxygen of Asp143. Notably, the guanidinium moiety of Arg2 of the N-degron displaces the side chain of Arg135 in the UBR domain to form a new salt bridge with Asp165 (**Fig. 4b**).

Disruption of the intradomain Arg135–Asp165 salt bridge also has a role in binding of the REAA and RDAA peptides. In these structures the side chain of glutamate or aspartate in the peptide salt bridges with Arg135 at the edge of the hydrophobic pocket (**Fig. 4c**,**d**). A water molecule occupies a central position in a hydrogen bonding network among Arg135, Asp165 and glutamate or aspartate of the N-degron (**Fig. 4c**,**d**). This binding mode alters the path of the peptide backbone, (as compared with the 'canonical' Scc1 binding mode) such that the hydrogen bond to the backbone amide of the third residue is not formed in the REAA peptide complex structure. A similar but less dramatic distortion is observed in the RDAA peptide complex structure. Quantitative measurement of binding of these peptides reveals considerably lower affinity than that observed for the favored RLAA peptide, which binds in the canonical mode (**Fig. 3**).

DISCUSSION

Our results define for the first time an essential structural aspect of ubiquitin ligases involved in the N-end rule pathway, thus providing detailed insights into the recognition of type 1 N-degrons by the UBR domain. The core fold of the UBR box, stabilized in part by the unique binuclear zinc center, provides the structural plasticity required for interaction with diverse N-end rule substrates (**Fig. 5**). Three important elements contribute to the recognition of type 1 N-degrons by the UBR box: (i) basic side chains of residues at position 1 in N-degrons are recognized by distinct residues in the acidic environment of the UBR box, which involves sequence-specific ionic interaction; (ii) the polypeptide backbone of the first three residues of N-end rule substrates form hydrogen bonds with β4 in the UBR box in an induced β-sheet interaction; (iii) the side chain of the residue at position 2 is recognized in a multispecific pocket that can alternately accommodate hydrophobic, acidic or basic residues.

In addition to its role in recognizing type 1 substrates, the UBR box is also actively involved in allosteric activation of full-length UBR1 for CUP9 degradation[32]. CUP9 is a transcriptional repressor of the peptide transporter PTR2, and it contains an internal (non–N-terminal) degron. The affinity between UBR1 and CUP9 is markedly increased by binding of both type 1 and type 2 N-degron peptides to UBR1. The following sequential process has been proposed[32]. Binding of type 1 Arg-Ala peptide to the UBR box increases the accessibility of type 2 Leu-Ala peptide to the adjacent ClpS-homology domain. The later binding event in turn induces the open conformation of UBR1 allowing CUP9 to bind. In the absence of ligation of the ClpS-homology domain, the CUP9-binding site might be masked by the C-terminal autoinhibitory domain (**Fig. 1a**). The only conformational change we observed in the UBR box upon peptide binding was localized within the N-degron–binding site; thus, this region of the UBR

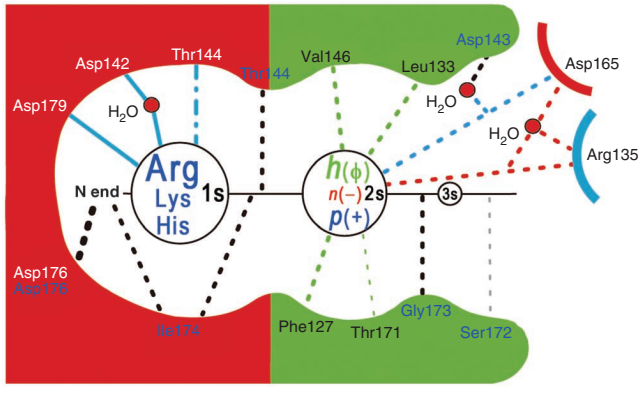

Figure 5 A schematic view of the interactions between the UBR box and the N-degron peptide. Type 1 substrate bearing an N-terminal basic residue, arginine, lysine or histidine (1s), is recognized in a sequence-specific manner, and the nature of the residue at position 2 (2s) modulates the binding affinity. Nonspecific main chain substrate binding contributes to stabilization of the complex. Red and green colors in the UBR box–binding site represent acidic and hydrophobic surface properties, respectively. Basic residues (blue) at position 1 (1s) are distinctly recognized by several acidic residues (white) in the UBR box. The blue dashed-and-dotted line represents the unique interaction between Thr144 and His1s. The α-amino group, labeled as 'N end', forms an ionic interaction with the carboxyl group of Asp176 and a hydrogen bond with the main chain carbonyl oxygen of Ile174. Nonspecific main chain–main chain interactions are depicted by black dotted lines, with the residues involved labeled in blue. Ser172, Gly173 and Ile174 in the loop region of the UBR box form hydrogen bonds in a short antiparallel β-sheet following conformational change. Different classes—*h*(φ), hydrophobic; *n*(–), negatively charged; *p*(+), positively charged—of amino acid residues at position 2 are recognized by different binding modes of the UBR box. Hydrophobic interactions are depicted by a green dotted line, and the contribution of the side chain methyl group of Thr171 by a thinner green dotted line. The ionic interactions between *p*(+) residue (2s) and Asp165 and that between *n*(–) residue (2s) and Arg135 are depicted by a blue and a red dotted line, respectively. Note that a water molecule is involved in both ionic interactions. The size of the labels denoting the N-degron is roughly proportional to the binding affinity at each position.

box may interact with the ClpS-homology domain to impede the accessibility of type 2 substrates. If so, the structural change that we observed upon complex formation might represent the initiating conformational switch in the sequential activation process of UBR1.

The potential importance of the residue at position 2 has now been systematically examined, through both peptide array experiments and *in vivo* pulse-chase experiments (**Fig. 2**). The number of confirmed *in vivo* N-end rule substrates is still small (fewer than ten in total from diverse organisms), making it difficult to expound on the importance of the residue at position 2 in physiologically relevant terms. However, in one clear example, cleavage of the yeast cohesin subunit Scc1 by Esp1 separase produces a C-terminal fragment with N-terminal arginine and either leucine or phenylalanine at position 2 (x_1 in the recognition motif, SxExGR|Rx_1 of separase), in agreement with our finding that hydrophobic residues are preferred[5]. We do anticipate that more physiological N-end rule substrates will be identified and our findings will have a crucial role in rational interpretation and generalization.

The UBR domain is present in at least seven distinct proteins in mammalian systems that are proven or putative E3 ligases[17]. Notably, residues required for accommodating hydrophobic (Phe127, Leu133 and Val146), acidic (Asp165) or basic (Arg135) residues in the position 2–binding pocket are distinctly conserved within a different subset of mammalian UBR family proteins[17]. It is tempting to speculate that distinct UBR family proteins in mammals have different roles depending on the type of residue occupying position 2 of the N-degron. The results of the present study can provide insights into structural investigations of UBR box domains in other families of proteins, and of uncharacterized domains within UBR1. Consequently, the current atomic-resolution details pertaining to the recognition of type 1 N-degrons broaden our understanding of the functional repertoire of the N-end rule pathway.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Accession codes. Protein Data Bank: Atomic coordinates and structure factors have been deposited under the following accession codes: 3NIT (native1) and 3NIS (native2) for free UBR box; 3NIH (RIAAA), 3NII (KIAA), 3NIJ (HIAA), 3NIL (RDAA), 3NIK (REAA), 3NIM (RRAA) and 3NIN (RLGES) for UBR box–N-degron peptide complexes.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank the staff at the 4A and 6B beamlines, Pohang Accelerator Laboratory, Korea, and the NW12 beamline, Photon Factory, Japan, for help with the data collection. We also thank Y.T. Kwon (University of Pittsburgh) for providing us with unpublished data of UBR fragments and A. Varshavsky (California Institute of Technology) for yeast strains and plasmids used for the *in vivo* study. This work was supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF 2007-0055395), the 21C Frontier Functional Proteomics Project (FPR08B2-270), World-Class University Project (R33-10108), the Plant Signaling Network Research Center and the Korea Healthcare Technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea (A092006).

AUTHOR CONTRIBUTIONS

W.S.C., B.-C.J., M.-R.L. and H.K.S. performed structural studies; W.S.C. and B.-C.J. performed biochemical studies; Y.J.J. and J.K. performed *in vivo* assay; W.S.C., M.J.E. & H.K.S. designed experiments, analyzed data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Bachmair, A., Finley, D. & Varshavsky, A. *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186 (1986).
2. Varshavsky, A. The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. USA* **93**, 12142–12149 (1996).
3. Tobias, J.W., Shrader, T.E., Rocap, G. & Varshavsky, A. The N-end rule in bacteria. *Science* **254**, 1374–1377 (1991).
4. Bachmair, A. & Varshavsky, A. The degradation signal in a short-lived protein. *Cell* **56**, 1019–1032 (1989).
5. Rao, H., Uhlmann, F., Nasmyth, K. & Varshavsky, A. Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* **410**, 955–959 (2001).
6. Turner, G.C., Du, F. & Varshavsky, A. Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature* **405**, 579–583 (2000).
7. An, J.Y. *et al.* Impaired neurogenesis and cardiovascular development in mice lacking the E3 ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway. *Proc. Natl. Acad. Sci. USA* **103**, 6212–6217 (2006).
8. Zenker, M. *et al.* Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). *Nat. Genet.* **37**, 1345–1350 (2005).
9. Ditzel, M. *et al.* Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nat. Cell Biol.* **5**, 467–473 (2003).
10. Hu, R.G., Wang, H., Xia, Z. & Varshavsky, A. The N-end rule pathway is a sensor of heme. *Proc. Natl. Acad. Sci. USA* **105**, 76–81 (2008).
11. Hu, R.G. *et al.* The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature* **437**, 981–986 (2005).
12. Lee, M.J. *et al.* RGS4 and RGS5 are *in vivo* substrates of the N-end rule pathway. *Proc. Natl. Acad. Sci. USA* **102**, 15030–15035 (2005).
13. Holman, T.J. *et al.* The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **106**, 4549–4554 (2009).
14. Graciet, E. *et al.* The N-end rule pathway controls multiple functions during *Arabidopsis* shoot and leaf development. *Proc. Natl. Acad. Sci. USA* **106**, 13618–13623 (2009).
15. Hwang, C.S., Shemorry, A. & Varshavsky, A. Two proteolytic pathways regulate DNA repair by cotargeting the Mgt1 alkylguanine transferase. *Proc. Natl. Acad. Sci. USA* **106**, 2142–2147 (2009).
16. Tasaki, T. *et al.* A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol. Cell. Biol.* **25**, 7120–7136 (2005).
17. Tasaki, T. *et al.* The substrate recognition domains of the N-end rule pathway. *J. Biol. Chem.* **284**, 1884–1895 (2009).
18. Lupas, A.N. & Koretke, K.K. Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. *J. Struct. Biol.* **141**, 77–83 (2003).
19. Dougan, D.A., Reid, B.G., Horwich, A.L. & Bukau, B. ClpS, a substrate modulator of the ClpAP machine. *Mol. Cell* **9**, 673–683 (2002).
20. Baker, R.T. & Varshavsky, A. Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway. *J. Biol. Chem.* **270**, 12065–12074 (1995).
21. Kwon, Y.T. *et al.* Altered activity, social behavior, and spatial memory in mice lacking the NTAN1p amidase and the asparagine branch of the N-end rule pathway. *Mol. Cell. Biol.* **20**, 4135–4148 (2000).
22. Bartel, B., Wunning, I. & Varshavsky, A. The recognition component of the N-end rule pathway. *EMBO J.* **9**, 3179–3189 (1990).
23. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780–2781 (2008).
24. Suhrer, S.J., Wiederstein, M., Gruber, M. & Sippl, M.J. COPS–a novel workbench for explorations in fold space. *Nucleic Acids Res.* **37**, W539–W544 (2009).
25. Marmorstein, R., Carey, M., Ptashne, M. & Harrison, S.C. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356**, 408–414 (1992).
26. Baleja, J.D., Marmorstein, R., Harrison, S.C. & Wagner, G. Solution structure of the DNA-binding domain of Cd2–GAL4 from *S. cerevisiae*. *Nature* **356**, 450–453 (1992).
27. Xia, Z. *et al.* Substrate-binding sites of UBR1, the ubiquitin ligase of the N-end rule pathway. *J. Biol. Chem.* **283**, 24011–24028 (2008).
28. Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T. & Baker, T.A. The molecular basis of N-end rule recognition. *Mol. Cell* **32**, 406–414 (2008).
29. Roman-Hernandez, G., Grant, R.A., Sauer, R.T. & Baker, T.A. Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. *Proc. Natl. Acad. Sci. USA* **106**, 8888–8893 (2009).
30. Schuenemann, V.J. *et al.* Structural basis of N-end rule substrate recognition in *Escherichia coli* by the ClpAP adaptor protein ClpS. *EMBO Rep.* **10**, 508–514 (2009).
31. Wang, H., Piatkov, K.I., Brower, C.S. & Varshavsky, A. Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Mol. Cell* **34**, 686–695 (2009).
32. Du, F., Navarro-Garcia, F., Xia, Z., Tasaki, T. & Varshavsky, A. Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain. *Proc. Natl. Acad. Sci. USA* **99**, 14110–14115 (2002).

ONLINE METHODS

Sample preparation. Of the more than 20 different constructs, two (native1, residues 107–194, and native2, residues 113–194) of the UBR box in UBR1 E3 ligase from *S. cerevisiae* were chosen for the final structural study. All experimental procedures were the same for both constructs except for the crystallization conditions. The constructs were amplified from genomic DNA of *S. cerevisiae* using standard PCR techniques and cloned into the BamHI and XhoI sites of modified a pET vector for the construction of GST-tagged protein. Protein was expressed in BL21(DE3)RIL cells at 25 °C for 12 h. Cells were lysed in PBS and subsequently subjected to GST-affinity chromatography. The GST tag was cleaved by TEV protease and separated from the UBR box by ion-exchange chromatography (Mono Q, GE Healthcare). Protein was further purified by size-exclusion chromatography (Superdex 75, GE Healthcare) in a final buffer containing 50 mM Tris-HCl, pH 8.0, 200 mM NaCl and 2 mM DTT. Binding-site mutants (D142A, T144A, D179A and E181A) were generated using QuikChange mutagenesis (Stratagene), and sample preparation of the mutants was performed in the same way as that of the wild-type UBR box.

Crystallization. The UBR box protein was concentrated to 20 mg ml^{-1} for crystallization screening. For the production of complexes, purified protein was incubated with a three-fold molar excess of each peptide. The optimized crystallization conditions for each crystal were as follows: 0.1 M HEPES, pH 7.7, 70% (v/v) 2-methyl-2,4-pentanediol (MPD) for native1; 0.02 M calcium chloride, 0.1 M sodium acetate, pH 4.8, 30% (v/v) MPD for native2; 0.16 M ammonium acetate, 0.01 M calcium chloride, 0.05 M sodium cacodylate, pH 6.5, 8% (w/v) PEG4000 for the RIAAA, KIAA and HIAA complexes; 0.04 M sodium cacodylate, pH 6.0, 0.04 M magnesium acetate, 30% (v/v) MPD for the RRAA and RLGES complexes; 0.1 M Bis-Tris, pH 6.5, 10% (w/v) PEG3350 for the REAA complex; 0.1 M Bis-Tris, pH 5.5, 0.2 M ammonium acetate, 25% (w/v) PEG3350 for the RDAA complex. Different crystallization conditions resulted in different crystal packing (**Tables 1** and **2**).

Structure determination. For phasing, three-wavelength MAD data sets were collected at the absorption edge and peak of the zinc atom and at a high-energy wavelength remote from these with native1 crystal. Three zinc atoms were located in the asymmetric unit using SOLVE/RESOLVE[33]. The partial model was built using ARP/wARP[34], and further model building was performed using the program O[35]. The protein model was refined using CNS[36] and REFMAC[37]. Phases of native2 and peptide complex crystals were obtained by molecular replacement using the programs MOLREP[38] or PHASER[39] with the refined structure of native1 as a search model. The positions of the bound N-degron peptide were determined using a model-phased difference Fourier map contoured at 3σ. Model building and refinement were performed using COOT[40] and CNS/REFMAC[36,37], respectively. Statistics for all collected data and refinement are described in **Tables 1** and **2**. The assessment of model geometry and assignment of secondary-structure elements were performed using the PROCHECK program[41]. No Ramachandran outliers were found in all refined structures. For structure comparison, the DALI (http://ekhidna.biocenter.helsinki.fi/dali_server/) and COPS (http://cops.services.came.sbg.ac.at/) servers were used[23,24].

Isothermal titration calorimetry. For ITC experiments, ITC1 buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl and 2 mM DTT) was used for all binding experiments except for that using the HIAA peptide, for which ITC2 buffer (20 mM MES, pH 6.5, 50 mM NaCl and 2 mM DTT) was used. UBR box protein was diluted to a concentration of 100 μM in ITC1 or ITC2 buffer and peptides were dissolved in the same buffers at a concentration of 2–3 mM. The experiment was performed at 22 °C using a VP-ITC microcalorimeter (MicroCal). Each peptide was injected 30 times (8 μl each) into 1.4-ml samples of UBR box protein. The experimental data were calculated using the ORIGIN software package provided with the instrument. At least three experiments were performed using varied peptide and protein concentrations.

Peptide array. The cellulose membrane containing 120 synthetic peptide spots was produced by JPT Peptide Technologies GmbH in Germany. Peptides in each spot were composed of 5 residues and had two common residues, Ala-Ala, at their C terminus covalently attached to the membrane. For peptide arrays, low-pH buffer (the same as ITC2 buffer) in the presence of 0.5% (v/v) Tween20 was used. The membrane was blocked for 2 h in western blocking reagent (Roche), incubated for 2 h in array buffer with excess purified UBR box protein or excess GST-fused UBR box protein, and finally incubated for 3 h with HRP-conjugated UBR box antibody (Peptron) or HRP-conjugated GST antibody (Amersham). In the case of GST detection, arrays were first analyzed by probing with purified GST alone, and no background signals were apparent (data not shown). In the interval between each procedure, extensive washing was performed using array buffer. The resultant membrane was developed by treatment with WEST-one solution (Intron Biotechnology) for 1 min. The binding was visualized and quantified using LAS-3000 (Fujifilm). Multiple experiments were performed using different membranes and detection procedures.

Pulse-chase analysis. To determine the *in vivo* half-life of $Scc1^{269–566}$ and its variants at position 2, pulse-chase analysis was performed as previously described with minor modifications[5]. *S. cerevisiae* JD52 strains harboring one of the plasmids expressing $_f$Dhfr-Ub-Arg-Leu-$Scc1_{269–566f}$, $_f$Dhfr-Ub-Arg-Arg-$Scc1_{269–566f}$, $_f$Dhfr-Ub-Arg-Asp-$Scc1_{269–566f}$ or $_f$Dhfr-Ub-Arg-Pro-$Scc1_{269–566f}$ from the GAL1 promoter were grown to mid-log phase in SR(–ura) medium (2% (w/v) raffinose, 0.67% (w/v) YNB (Difco) without amino acid with auxotrophic supplements). The GAL1 promoter was induced by incubating in SRG(–ura) medium (2% (w/v) raffinose, 2% (w/v) galactose, 0.67% (w/v) YNB without amino acid with auxotrophic supplements) for 2 h. Cells from a 40-ml culture were pelleted. The cell pellets were washed with 1 ml SRG(–ura –met), resuspended in 0.4 ml of SRG(–ura –met) and pulse-labeled for 5 min with 0.16 mCi of ^{35}S-Express (NEN), followed by centrifugation and resuspension of cells in SD(–ura) containing 4 mM L-methionine and 2 mM L-cysteine. Samples of cultured mixture of volume 0.1 ml were harvested at the indicated time and processed for immunoprecipitation with anti-Flag M2 agarose beads (Sigma-Aldrich) as described[42]. Immunoprecipitated complexes were subjected to 10% (w/v) SDS-PAGE and autoradiography signals were detected using a BAS-2500 image leader (Fujifilm).

33. Terwilliger, T.C. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972 (2000).
34. Lamzin, V.S. & Wilson, K.S. Automated refinement of protein models. *Acta Crystallogr. D Biol. Crystallogr.* **49**, 129–147 (1993).
35. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
36. Brunger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
37. Winn, M.D., Murshudov, G.N. & Papiz, M.Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374**, 300–321 (2003).
38. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).
39. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
40. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
41. Laskowski, R.A., Moss, D.S. & Thornton, J.M. Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067 (1993).
42. Ghislain, M., Dohmen, R.J., Levy, F. & Varshavsky, A. Cdc48p interacts with Ufd3p, a WD repeat protein required for ubiquitin-mediated proteolysis in *Saccharomyces cerevisiae*. *EMBO J.* **15**, 4884–4899 (1996).

doi:10.1038/nsmb.1907