Biological Science

NW12A/2010G518

Crystal structures of UBR box in complex with N-degron peptides

Woo Suk CHOI and Hyun Kyu SONG*
School of Life Sciences and Biotechnology, Korea University, Anam-Dong
Seongbuk-Gu, Seoul 136-701, South Korea

Introduction

The N-end rule pathway is a regulated proteolytic system that targets proteins containing destabilizing N-terminal residues (N-degrons) for ubiquitylation and proteosomal 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 *S. 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 Arg, Lys or His side-chains of the N-degron are coordinated in a multi-specific 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.[1]

Methods

UBR box protein was concentrated to 20 mg/ml for crystallization screening. For the production of complexes, purified protein was incubated with a 3-fold molar excess of each peptide. The optimized crystallization condition for Scc1 peptide complex was 0.04 M sodium cacodylate trihydrate (pH 6.0), 0.04 M magnesium acetate tetrahydrate, 30% (v/v) MPD. For phasing, three wavelength MAD data sets were collected at the absorption edge and peak of the zinc atom and high energy remote with uncomplexed crystal. Three zinc atoms were located in the asymmetric unit using the SOLVE/RESOLVE programs. The partial model was built using the ARP/wARP program, and further model-building was performed using the program O. The protein model was refined using the programs CNS and REFMAC. The positions of the bound N-degron peptide were determined using a model-phased difference Fourier map contoured at 3.0σ. Model building and refinement were performed using COOT and CNS/REFMAC, respectively.

Results

The crystal structure of the UBR box reveals a compact, heart-shaped domain with three zinc coordination sites and little regular secondary structure (Fig 1a). The V-shaped base of the domain is formed by the intersection of a small β-sheet (formed by strands β1 and β3) and two irregular loops. 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. Strands β2 and β4 are foreshortened by participation of their C-terminal extensions in the formation of a binuclear zinc center (Zn1 and Zn2). 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. This zinc is expected to be important in stabilizing the fold of the UBR box. The unique 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 failed to reveal other proteins with similar three-dimensional folds.

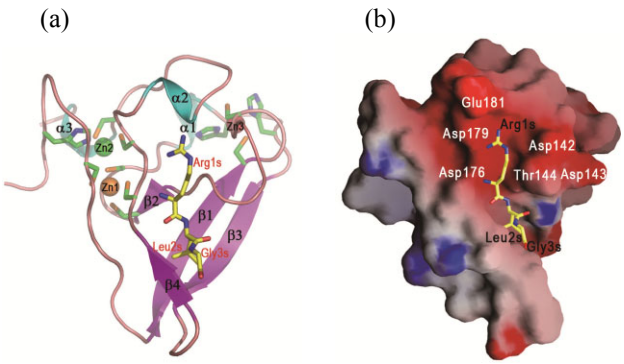

Figure 1. Structure of the UBR box:N-degron peptide complex. (a) Ribbon diagram of UBR box and stick model of bound N-degron peptide. (b) Electrostatic potential surface of UBR box. Highly negative cleft of UBR box accommodates N-terminal arginine residue of substrate.

Type-1 N-degrons bind in a relatively shallow, acidic cleft on the surface of the UBR domain (Fig 1b), 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 Scc1 N-degron, which contains an N-terminal arginine and bears the sequence RLGES, is shown in Figure 1b.

Reference

[1] W.S Choi et al., Nature Struct. Mol. Biol. 17, 1175 (2010)

* hksong@korea.ac.kr