

# Structure of the FP domain of Fbxo7 reveals a novel mode of protein–protein interaction

Jinsai Shang,<sup>a</sup> Guan Wang,<sup>a</sup>  
Yang Yang,<sup>a</sup> Xiaolan Huang<sup>b</sup> and  
Zhihua Du<sup>a\*</sup>

<sup>a</sup>Chemistry and Biochemistry, Southern Illinois University, 1245 Lincoln Drive, Carbondale, IL 62901, USA, and <sup>b</sup>Computer Science, Southern Illinois University, 1000 Faner Drive, Carbondale, IL 62901, USA

Correspondence e-mail: zdu@chem.siu.edu

The FP (Fbxo7/PI31) domains found in the F-box protein Fbxo7 and the proteasome inhibitor PI31 mediate the homodimerization and heterodimerization of Fbxo7 and PI31. Fbxo7 is the substrate-recognition subunit of the SCF<sup>Fbxo7</sup> (Skp1-Cul1-F-box protein) E3 ubiquitin ligase that catalyzes the ubiquitination of hepatoma up-regulated protein (HURP) and inhibitor of apoptosis protein (IAP). Fbxo7 also interacts with proteins that are not substrates of the ubiquitin proteasome system, such as Cdk6 and PI31. Here, the crystal structure of the Fbxo7 FP domain is reported at 2.0 Å resolution. The Fbxo7 FP domain adopts an  $\alpha/\beta$ -fold similar to that of the PI31 FP domain. However, an  $\alpha$ -helix and three  $\beta$ -strands in the Fbxo7 FP domain are longer than their counterparts in the PI31 FP domain. The differences in these secondary-structural elements are spatially clustered to define a more structured and extended C-terminal end of the Fbxo7 FP domain. The two FP domains also differ substantially in the length and conformation of the longest connecting loop. More importantly, structural differences between the two FP domains lead to drastically different modes of inter-domain protein–protein interaction. The inter-domain interface of the Fbxo7 FP domain is defined by the  $\alpha$ -helical surface in one protomer and the  $\beta$ -sheet surface in the other protomer, whereas for the PI31 domain it is defined by either the  $\alpha$ -helical surfaces or the  $\beta$ -sheet surfaces in both protomers. The inter-domain interaction of the Fbxo7 FP domain is much more extensive, featuring a larger contact surface area, better shape complementarity and more hydrophobic and hydrogen-bonding interactions. The Fbxo7 FP domain also has the potential to bind two protein partners simultaneously using the  $\alpha$ -helical and  $\beta$ -sheet surfaces. The results of this structural study provide critical insights into how Fbxo7 may dimerize (or multimerize) and interact with other regulatory proteins *via* the FP domain.

Received 23 July 2013

Accepted 18 September 2013

**PDB References:** FP domain of Fbxo7, 4I9c; 4I9h

## 1. Introduction

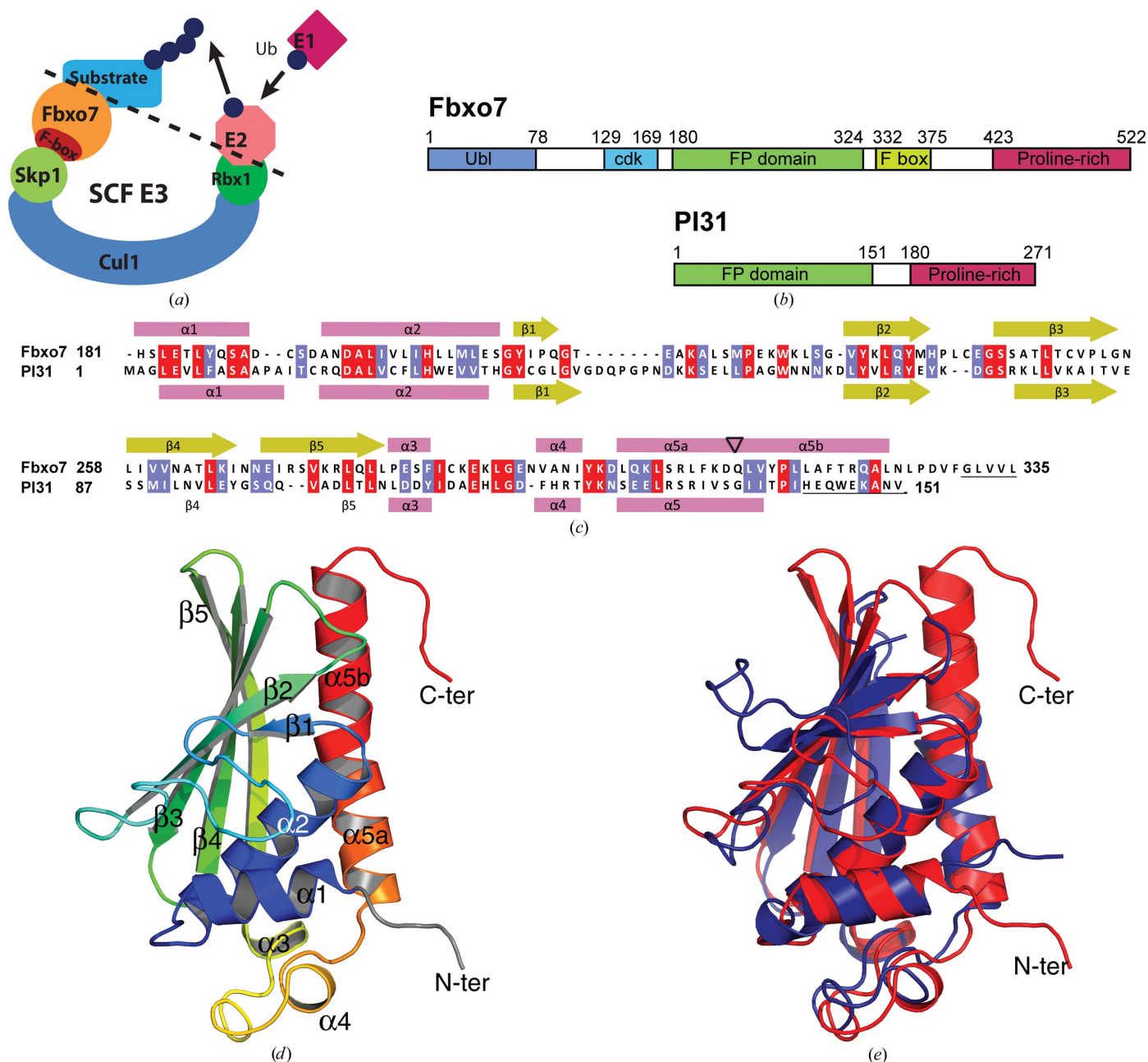
The F-box, which contains approximately 50 amino acids, is a widespread protein motif that mediates protein–protein interactions. The motif was named the F-box owing to its presence in cyclin F, one of the earliest identified F-box proteins (Kumar & Palletta, 1995; Bai *et al.*, 1996). There are a large number of F-box proteins. These proteins often contain some other protein–protein interaction motifs besides the F-box motif. Based on the additional motifs, F-box proteins are classified into several subfamilies: Fbxw for proteins containing WD40 repeats, Fbxl for proteins containing leucine-rich repeats and Fbxo for proteins containing other different or unidentified domains (Cenciarelli *et al.*, 1999; Winston *et al.*, 1999). This nomenclature is used for human and

## research papers

mouse F-box proteins. In other organisms, the presence of an F-box motif is not used as a basis for protein nomenclature.

The best known biological function of F-box proteins is their role as the substrate-recognition component of the ubiquitin proteasome pathway. In this pathway, ubiquitin, a highly conserved 76-amino-acid protein, is conjugated to target proteins *via* the sequential action of three enzymes

(Fig. 1a): the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating enzyme and the E3 ubiquitin–protein ligase. Polyubiquitinated proteins are subsequently degraded by the 26S proteasome (Hershko, 2005). The largest and best characterized group of E3 ligases is the SCF (Skp1, cullin 1, F-box protein) family. The multicomponent SCF complexes consist of three common components, Skp1 (S-phase kinase-



**Figure 1**

Function of Fbxo7 and structure of its FP domain. (a) Function of Fbxo7 as the substrate-recognition component of the SCF<sup>Fbxo7</sup> E3 ligase. (b) Schematic representation of the putative domains of human Fbxo7. Ubl, ubiquitin-like domain; cdk, Cdk6-interacting motif; FP, Fbxo7/PI31 domain; PRR, proline-rich region. (c) Structure-based sequence alignment of the sequences of the FP domains of Fbxo7 and PI31. Identical and similar residues are highlighted with red and blue backgrounds, respectively. The secondary structures are indicated above and below the sequence for Fbxo7 and PI31, respectively. The  $\alpha 5$  helix in the FP domain of Fbxo7 has a kink in the middle denoted by a triangle. The underlined C-terminal sequences are present in the protein constructs used for crystallization, but no electron density was observed for these residues. Note that the protein constructs contain more residues than the putative FP domain as indicated in (b). (d) Overall structure of the FP domain of Fbxo7 colour-ramped from blue at the N-terminus to red at the C-terminus. The crystallization protein contains an artificial N-terminal sequence GPSSP as the result of a cloning artefact (grey). (e) Superimposition of the structures of the FP domains of Fbxo7 (red) and PI31 (blue).

associated protein 1), cullin 1 and Rbx1 (RING-box protein 1), as well as an F-box protein as the variable component. Structural and biochemical studies of SCF complexes reveal that cullin 1 serves as a rigid scaffold to position the Skp1 and Rbx1 subunits. In turn, Skp1 binds the F-box protein subunit by interacting with the F-box motif, while Rbx1 binds a ubiquitin-charged E2 enzyme. The F-box protein binds specific substrates and recruits the substrates to the SCF complex (Schulman *et al.*, 2000; Zheng *et al.*, 2002; Wu *et al.*, 2003; Goldenberg *et al.*, 2004; Hao *et al.*, 2007; Mizushima *et al.*, 2007; Zeng *et al.*, 2010). Numerous SCF-type E3 ligases exist. Each can only recognize a subset of protein substrates. The large number of diverse F-box proteins provides recognition specificities for various substrates. F-box proteins in each of the three subfamilies (Fbxw, Fbxl and Fbxo) have been shown to function in the SCF complexes (Cenciarelli *et al.*, 1999).

F-box-only protein 7 (Fbxo7) is a member of the Fbxo subfamily of F-box proteins (Hsu *et al.*, 2004; Laman *et al.*, 2005). In humans, two protein isoforms have been identified, with isoform 1 (more abundant) and isoform 2 containing 522 and 443 amino acids, respectively (Zhao *et al.*, 2011). Isoform 2 lacks the N-terminal 91 amino acids of isoform 1, replacing them with a different 12-amino-acid N-terminal sequence. Here, the term Fbxo7 refers to isoform 1 unless indicated otherwise.

As shown in Fig. 1(b), Fbxo7 contains a ubiquitin-like domain (Ubl; amino acids 1–78), a Cdk6-interacting motif (cdk; amino acids 129–169), an Fbxo7/PI31 domain (FP; amino acids 180–324), an F-box motif (amino acids 332–375) and a proline-rich region (PRR; amino acids 423–522) (Kirk *et al.*, 2008). Sequence homology between the Ubl domain and ubiquitin (20% identity and 38% similarity) suggests that the Ubl domain might interact with the ubiquitin-binding domains (UBDs) of ubiquitin-receptor proteins, including those in the proteasome. The cdk motif binds CDK6. The FP domain was first identified on the basis of its sequence homology to the N-terminal domain of the proteasome inhibitor PI31. It has been shown that the FP domains of Fbxo7 and PI31 mediate homodimerization and heterodimerization of the proteins. The F-box motif is required for interaction with Skp1. The sequence between the F-box motif and PRR also contributes to Skp1 binding, suggesting that the F-box motif of Fbxo7 might have a broader boundary (Hsu *et al.*, 2004). The PRR is most likely to be unstructured. It mediates the specific binding of protein substrates of the SCF<sup>Fbxo7</sup> ligases (Hsu *et al.*, 2004; Chang *et al.*, 2006).

Fbxo7 is one of the Fbxo subfamily proteins that have been shown to form functional SCF complexes. Specific protein substrates of the SCF<sup>Fbxo7</sup> ligase complex include HURP (hepatoma up-regulated protein), cIAP1 (cellular inhibitor of apoptosis protein 1) and TRAF2 (TNF receptor-associated factor 2) (Kuiken *et al.*, 2012). Fbxo7 also interacts with other proteins that are not substrates for ubiquitin-mediated degradation, including Cdk6, PI31 and p27 (Laman *et al.*, 2005; Kirk *et al.*, 2008). Selective binding of Fbxo7 to Cdk6 enhances the assembly and activity of the cyclin D–Cdk6 complex,

which is responsible for the ability of Fbxo7 to transform murine fibroblasts that are tumourigenic in athymic nude mice.

The crystal structure of the FP domain of PI31 has been determined (Kirk *et al.*, 2008). The FP domain adopts a novel  $\alpha/\beta$ -fold with five  $\alpha$ -helices and five  $\beta$ -strands. The five  $\beta$ -strands form an antiparallel  $\beta$ -sheet located on one side of the domain. The other side of the domain is mainly defined by the  $\alpha$ -helices. Inter-domain protein–protein contacts in the crystal are mediated by two different molecular interfaces: the  $\alpha$ -helical interface and the  $\beta$ -sheet interface. It was proposed that the  $\alpha$ -helical interface mediates homodimerization of the FP domain of PI31 and the  $\beta$ -sheet interface mediates heterodimerization of the FP domains of PI31 and Fbxo7.

The precise physiological role of the Fbxo7–PI31 interaction is not known. However, studies of homologous proteins in *Drosophila* suggest that Fbxo7 and PI31 may form a conserved regulatory complex that stimulates proteasome activity (Bader *et al.*, 2011). The *Drosophila* F-box protein Nutcracker shares some degree of homology with Fbxo7. Nutcracker also contains an FP domain that is located N-terminal to the F-box motif. It has been shown that Nutcracker interacts with DmPI31, the *Drosophila* homologue of human PI31, through heterodimerization of their FP domains (Bader *et al.*, 2011). Nutcracker–DmPI31 interaction stabilized DmPI31. The two proteins function together as a complex to stimulate proteasome activity, which is required for non-apoptotic caspase activation and sperm differentiation in *Drosophila*.

Fbxo7 has been implicated in cancers and neurodegenerative diseases. It has been shown that Fbxo7 is overexpressed in tumours of the lung and colon (Laman *et al.*, 2005). It is also well known that mutations in the Fbxo7 gene (*FBXO7*) cause PARK15, an early-onset Parkinsonian disease (Shojaee *et al.*, 2008; Di Fonzo *et al.*, 2009).

In spite of the biological and physiological importance of Fbxo7, no structural studies of this protein have been reported. In this paper, we present the crystal structure of the FP domain of Fbxo7. Our results show that the FP domain of Fbxo7 shares the same  $\alpha/\beta$  protein fold as the FP domain of PI31. However, the FP domain of Fbxo7 utilizes an  $\alpha$ -helical– $\beta$ -sheet interface to mediate inter-domain protein–protein interactions. This interface is drastically different from and much more extensive than the  $\alpha$ -helical or  $\beta$ -sheet interfaces observed in the structure of the FP domain of PI31. Moreover, the FP domain of Fbxo7 has the potential to simultaneously interact with two protein partners *via* different molecular surfaces. Our structural study provides critical insights into how Fbxo7 may dimerize (or multimerize) and interact with other regulatory proteins *via* the FP domain.

## 2. Materials and methods

### 2.1. Cloning

A cDNA clone for full-length Fbxo7 isoform 1 (GenBank BC008361) was purchased from the DNASU Plasmid

Repository. The DNA encoding the FP domain and an extra C-terminal sequence (residues 181–335) was amplified by PCR from the plasmid. The DNA was inserted into an in-house-modified ligation-independent cloning (LIC) vector that contains DNA sequences encoding the HaloTag (Los *et al.*, 2008), a His tag and an eight-amino-acid recognition sequence for HRV (*Human rhinovirus*) 3C protease followed by a specific LIC sequence. Protein expression was carried out using NiCo21 (DE3) *Escherichia coli* cells (New England Biolabs).

## 2.2. Protein expression, purification and crystallization

The FP domain of Fbxo7 was expressed in *E. coli* as a fusion protein with an N-terminal HaloTag and His tag. Protein expression was induced using a 0.4 mM final concentration of IPTG when the cell culture reached an OD<sub>600</sub> of 0.6–0.8. After induction, the temperature was lowered to 288 K and the culture was allowed to grow overnight. Expressed proteins were purified using NTA affinity resin. The fusion tags were then cleaved by HRV 3C protease. The cleaved tags were separated from the target protein by a reverse IMAC process. The purified proteins were concentrated to a concentration of ~10 mg ml<sup>-1</sup> in a buffer consisting of 25 mM Tris pH 7.5, 200 mM NaCl, 1 mM DTT.

To obtain SeMet-labelled protein, the cells were grown in M9 minimal culture until an OD<sub>600</sub> of 0.6–0.8 was reached. Six amino acids (leucine, isoleucine, lysine, phenylalanine, threonine and valine) were added to the culture at final concentrations of 50–100 mg l<sup>-1</sup> to inhibit the methionine-biosynthesis pathway. After 30 min of continuous incubation, L-selenomethionine was added to a final concentration of 50 mg l<sup>-1</sup>. Protein expression was induced using a 0.4 mM final concentration of IPTG at 288 K overnight.

Crystals of the protein construct were obtained at 295 K by sitting-drop vapour diffusion against 50 µl well solution using 96-well format crystallization plates. Two crystal forms were obtained. Crystals belonging to space group *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> were obtained with a well solution consisting of 20% PEG 3350, 50 mM MOPS buffer pH 7.0, 100 mM KBr, 10% (v/v) glycerol. Crystals belonging to space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub> were obtained with a well solution consisting of 20% PEG 3350, 50 mM bis-tris buffer pH 7.2, 100 mM KBr, 10% (v/v) glycerol. Crystals were flash-cooled in liquid nitrogen.

## 2.3. Data collection, data processing and structure determination

Data collection was carried out on beamlines 21-ID-D, 21-ID-F and 21-ID-G of LS-CAT at the Advanced Photon Source (Argonne National Laboratory). Data were processed, integrated and scaled with the programs *MOSFLM* and *SCALA* in *CCP4* (Battye *et al.*, 2011; Winn *et al.*, 2011). The structure in space group *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> was solved by SAD using a data set collected at the peak wavelength of Se (0.97872 Å) from a single crystal of SeMet-labelled protein. The *PHENIX* package (Adams *et al.*, 2011) was used to locate the Se atoms and to build the initial structure. Interactive model building

**Table 1**

Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	Native	SeMet
Data collection		
Space group	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>4</sub> <sub>3</sub> <sub>2</sub> <sub>1</sub> <sub>2</sub>
Unit-cell parameters (Å)	<i>a</i> = 44.22, <i>b</i> = 64.41, <i>c</i> = 89.31	<i>a</i> = <i>b</i> = 57.65, <i>c</i> = 89.32
Resolution (Å)	44.2–2.1 (2.20–2.10)	48.4–2.0 (2.10–2.00)
Completeness (%)	99.8 (99.2)	100 (100)
<i>R</i> <sub>merge</sub> (%)	10.0 (48.8)	12.6 (68.4)
<i>I</i> / <i>σ</i> ( <i>I</i> )	10.9 (4.0)	15.0 (4.9)
Multiplicity	5.7 (5.7)	13.9 (14.2)
Refinement		
Resolution (Å)	44.2–2.1	48.4–2.0
No. of reflections	15476	10767
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> (%)	19.09/24.66	20.14/23.45
No. of atoms		
Protein	2362	1211
Water	96	56
<i>B</i> factors (Å <sup>2</sup> )		
Protein	32.3	28.4
Water	32.9	29.6
R.m.s.d., bonds (Å)	0.004	0.004
R.m.s.d., angles (°)	0.860	0.880
Ramachandran plot statistics (%)		
Residues in favoured region	96.0	98.0
Outliers	0	0

was carried out with *Coot* (Emsley *et al.*, 2010). The structure was refined using *PHENIX*. The structure in space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub> was solved by molecular replacement using *Phaser* (Storoni *et al.*, 2004) as implemented in the *PHENIX* package. The structure in space group *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> was used as the search model for molecular replacement. All plausible space groups of the *P*<sub>222</sub> family were tried. A single solution was found in space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>. Structure refinement was carried out by *PHENIX*. In *PHENIX* refinement, the initial isotropic *B* factors were all set to 20 Å<sup>2</sup> and individual *B* factors were refined. No restraints were used in the refinement except that weightings between X-ray and stereochemistry/ADP terms were optimized using the default values. No H atoms were included in the refinement. The input coordinates for refinement did not contain water molecules. Water molecules were automatically updated during the refinement process. The refinement was carried out in real space using simulated annealing (Cartesian, from 5000 to 300 K). The structure-determination statistics are given in Table 1. The figures were prepared with *PyMOL* (The *PyMOL* Molecular Graphics System v.1.5.0.4; Schrödinger). The structure in space group *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> was used to generate all of the figures.

Atomic coordinates and diffraction data for the structures in space groups *P*<sub>4</sub><sub>3</sub><sub>2</sub><sub>1</sub><sub>2</sub> and *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub> have been deposited in the Protein Data Bank with accession codes 4I9h and 4I9c, respectively.

## 3. Results

### 3.1. Overall structure

The FP domain of Fbxo7 was predicted to include residues 180–324 of the protein (Kirk *et al.*, 2008). In our study, the

FP domain construct contains residues 181–335, including an extra C-terminal sequence of 11 amino acids outside the putative FP domain boundary. The protein construct also contains an artificial sequence GPSSP at the N-terminus as a cloning artefact.

The protein construct crystallizes in two different crystal forms in space groups  $P4_32_12$  and  $P2_12_12_1$ . The structure in space group  $P4_32_12$  was solved by the SAD phasing method using a data set collected from a single crystal of SeMet-labelled protein. The structure in space group  $P2_12_12_1$  was solved by the molecular-replacement method (Table 1).

In both structures, electron density was observed for residues 181–330. Electron density was also observed for the N-terminal artificial sequence in the structure in space group  $P4_32_12$ . No electron density was observed for residues 331–335 in both structures. The two structures are very similar (pairwise r.m.s.d. of 0.340 Å). The structures in space groups  $P4_32_12$  and  $P2_12_12_1$  contained two molecules and one molecule in the asymmetric unit, respectively. Despite this difference, the inter-domain protein–protein interactions in both structures were virtually identical. In the following sections, we will mainly discuss the structure in space group  $P4_32_12$ .

The FP domain of Fbxo7 adopts an  $\alpha/\beta$ -fold that is similar to the FP domain of PI31 (Figs. 1c and 1e). Both FP domains contain five  $\alpha$ -helices and five  $\beta$ -strands. The five  $\beta$ -strands form an antiparallel  $\beta$ -sheet located on one side of the molecule. The five  $\alpha$ -helices pack against the  $\beta$ -sheet on the other side (Fig. 1d). Most of the connecting loops are short. Two relatively long connecting loops are present between  $\alpha$ -helices  $\alpha_3$  and  $\alpha_4$  (referred to as the  $\alpha_3$ – $\alpha_4$  loop; nine amino acids) and  $\beta$ -strands  $\beta_1$  and  $\beta_2$  (referred to as the  $\beta_1$ – $\beta_2$  loop; 17 amino acids). These loops seem to be well structured. All of the residues within the loops yield good electron density. Comparing the structures in the two different space groups, the longer  $\beta_1$ – $\beta_2$  loop has basically the same conformation, while the  $\alpha_3$ – $\alpha_4$  loop shows a slight difference, most likely owing to different crystal packing.

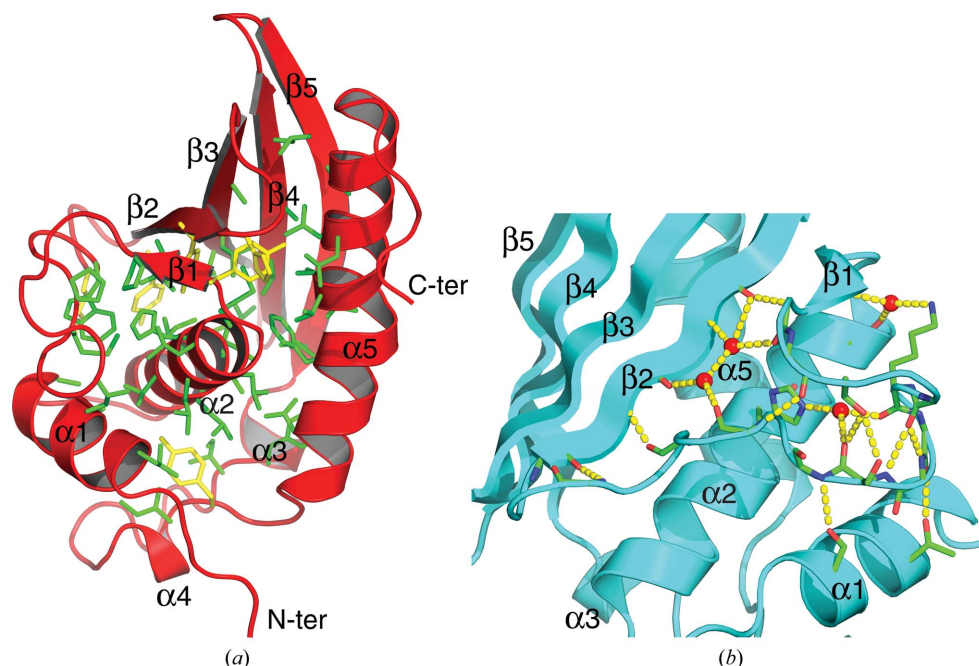
Most of the secondary-structural elements are comparable in the two structures. However, helix  $\alpha_5$  and strands  $\beta_3$ ,  $\beta_4$  and  $\beta_5$  are longer in the FP domain of Fbxo7 (Fig. 1c). The differences in these secondary-structural elements are spatially clustered to define a more extended and structured C-terminal end of the FP domain of Fbxo7 (Fig. 1e). Most noticeably, the FP domain of Fbxo7 features a long helix  $\alpha_5$  with seven helical

turns and a pronounced kink in the middle. Helix  $\alpha_5$  in the FP domain of PI31 is only about half as long (comparable to  $\alpha_5a$  in the FP domain of Fbxo7). The crystallization construct of the PI31 FP domain contains more C-terminal residues that correspond to residues within helix  $\alpha_5b$  of the Fbxo7 FP domain (Fig. 1c). However, no electron density was observed for the C-terminal residues in the PI31 FP domain, indicating that these residues are flexible in the structure. The two FP domains also show significant differences in the length and conformation of the  $\beta_1$ – $\beta_2$  loop (Figs. 1c and 1e).

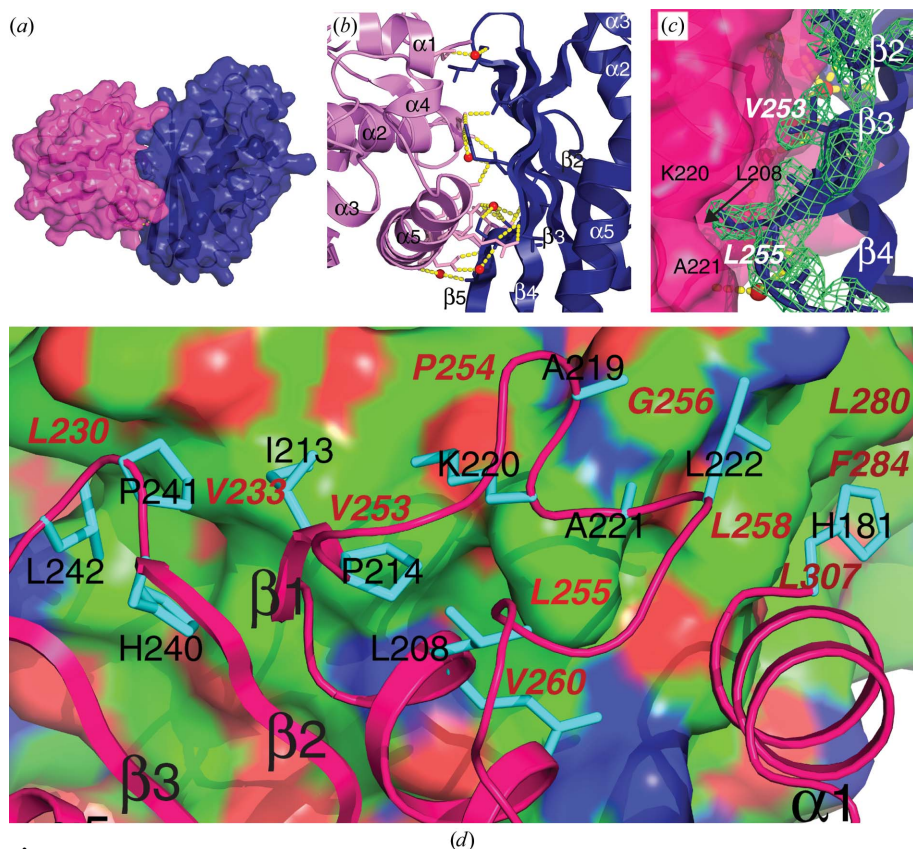
In short, although the FP domains of Fbxo7 and PI31 share the same protein fold, they show significant differences in many of the structural elements. These differences define very different molecular surfaces of the two FP domains that mediate different modes of protein–protein interaction (see below for details).

### 3.2. Intramolecular interactions of the Fbxo7 FP domain

The FP domain of Fbxo7 features a hydrophobic core that is centred around helix  $\alpha_2$  (Fig. 2a). Most of the residues of helix  $\alpha_2$  have a hydrophobic side chain, including a stretch of 11 amino acids ( $^{198}$ ALIVLIHLLML) in the middle of the helix (Fig. 1b). Helix  $\alpha_2$  is therefore mostly hydrophobic on all sides. A large number of residues from other structural elements pack around helix  $\alpha_2$  to form the hydrophobic core of the Fbxo7 FP domain. Most of these residues have a hydrophobic side chain, including Leu183, Leu186, Ala190, Cys192, Leu199, Ile200, Val201, Leu202, Ile203, His204, Leu205, Leu206, Met207, Leu208, Pro214, Met224, Pro225,



**Figure 2** Hydrophobic and hydrogen-bonding interactions within the structure of the FP domain of Fbxo7. (a) The hydrophobic core. The side chains of the residues that define the core are shown as sticks. The hydrophobic side chains are coloured green. A few Tyr, His and Lys residues are coloured yellow. (b) Hydrogen-bonding interactions that contribute to define the structure of the 17-amino-acid loop between strands  $\beta_1$  and  $\beta_2$ . Hydrogen bonds are represented as yellow dashed lines. Red spheres represent structured water molecules.


**Figure 3**

Inter-domain interactions of the FP domain of Fbxo7. (a) Two adjacent FP domains in the crystals interact with each other through the  $\alpha$ -helical interface in one domain (in purple) and the  $\beta$ -stranded interface (in blue). The shown structures do not contain the N- and C-terminal loops. The structures are rendered in surface mode (with 20% transparency) and cartoon mode. (b) Hydrogen bonds in the domain-domain interface. The residues involved in the hydrogen bonds are shown as sticks. The view is identical to that in (a). (c) The  $2F_o - F_c$  electron-density map contoured at  $1\sigma$  is shown for several residues (shown as sticks) in the domain-domain interface. (d) Hydrophobic interactions in the domain-domain interface. The two interacting FP domains are rendered in surface mode (coloured by element, with C, N and O atoms in green, blue and red, respectively) and cartoon mode (coloured magenta), respectively. The residues involved in the hydrophobic interactions are indicated in two different ways for the two protomers: on the molecular surface with red/italic residue names and as sticks (coloured in cyan) with black residue names. The inter-domain interface is the same in the two different crystal forms.

Trp228, Leu236, His240, Ala248, Leu250, Cys252, Pro254, Ile259, Val261, Ala263, Leu265, Ile267, Val274, Leu277, Leu279, Phe284, Ile285, Ile297, Ile304, Phe308, Leu312, Val313, Leu316 and Leu317 (Fig. 2a, green sticks). The hydrophobic moieties of the side chains of Tyr212, Lys229, Tyr234, Tyr238 and Tyr298 (Fig. 2a, yellow sticks) also participate in the formation of the hydrophobic core.

Besides the extensive hydrophobic interactions, other kinds of molecular interactions such as hydrogen-bonding interactions also contribute to the folding of the Fbxo7 FP domain. Most noticeably, hydrogen-bonding interactions play an important role in maintaining the conformation of the  $\beta_1$ – $\beta_2$  loop (Fig. 2b). A total of 22 hydrogen bonds mediate medium-range or long-range interactions within the loop or between the loop and neighbouring structural elements. Electron densities for four structured water molecules involved in the hydrogen-bonding network are also clearly observed.

Owing to the presence of extensive intramolecular interactions, the Fbxo7 FP domain has a tightly packed stable structure.

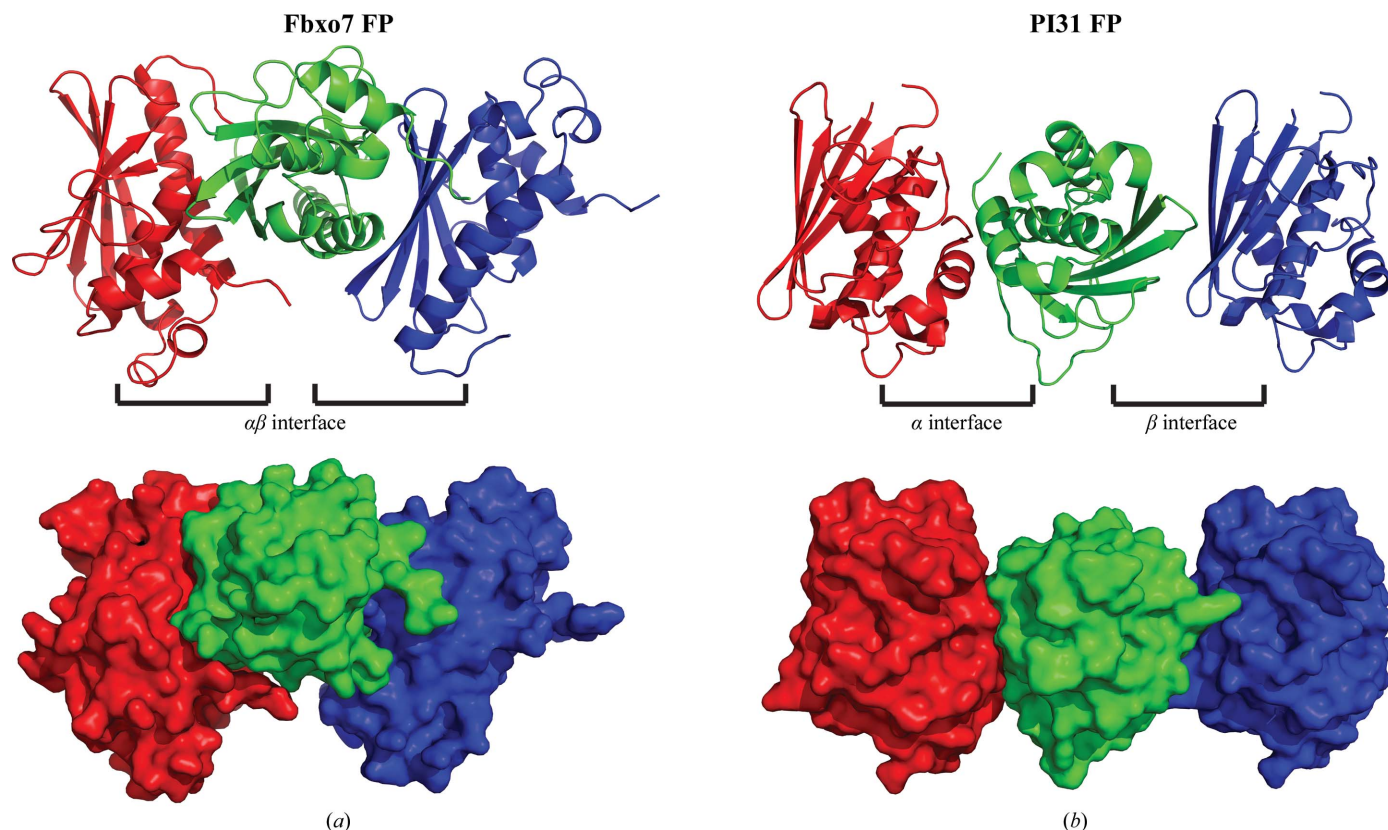
### 3.3. Intermolecular interactions of the Fbxo7 FP domain

For the crystal in space group  $P2_12_12_1$ , there are two FP domains in the asymmetric unit, which contact each other through substantial inter-domain protein–protein interactions. The asymmetric unit of the crystal in space group  $P4_32_12$  contains only one FP domain. However, two FP domains from adjacent asymmetric units show very similar inter-domain protein–protein interactions (Fig. 3).

It should be noted that the crystallization protein construct contains an extra C-terminal sequence (amino acids 325–335) that should not be part of the FP domain (Fig. 1b). This sequence makes a reverse turn at the end of helix  $\alpha_5$  and runs antiparallel to the helix. No electron density was observed in the protein–interaction interface. The C-terminal end of the sequence points inwards, where there is no space to accommodate more residue(s). In view of this, the peculiar positioning of residues 325–330 is most likely to be driven by crystal packing. We therefore exclude these residues in the following analysis of the intermolecular interactions of the Fbxo7 FP domain.

The inter-domain protein–protein interaction interface of the Fbxo7 FP domain is best described as an  $\alpha\beta$  interface because the contact surfaces in the two interacting FP domains are mainly defined by  $\alpha$ -helices in one protomer and  $\beta$ -strands in the other protomer. Fig. 2(a) provides a good view of the  $\alpha$ -helical interacting surface that is constituted by the N-terminal end of  $\alpha_1$ , the C-terminal end of  $\alpha_2$ , most of  $\alpha_5$  and portions of the  $\beta_1$ – $\beta_2$  and  $\beta_2$ – $\beta_3$  connecting loops. The  $\beta$ -sheet interacting surface is defined by the four  $\beta$ -strands  $\beta_2$ – $\beta_5$ , as shown by the molecular surface facing left in Fig. 2(b). Interaction between two FP domains through the  $\alpha\beta$  interface buries  $1700 \text{ \AA}^2$  of solvent-accessible molecular surface area from the two domains ( $1934 \text{ \AA}^2$  if the C-terminal sequence of amino acids 325–330 is included in the calculation). The buried surface area is significantly larger than the estimated minimal area of  $1200 \text{ \AA}^2$  required for a stable protein–protein complex (Chothia & Janin, 1975; Jones & Thornton, 1995).

The inter-domain protein–protein interaction is mediated by several driving forces including shape complementarity and

**Figure 4**

Comparison of the inter-domain interactions in the structures of the FP domains of Fbxo7 and PI31. For each structure, three domains in the crystals are shown in cartoon and surface modes with different colouring in red, green and blue. For Fbxo7, inter-domain interaction is mediated by an  $\alpha/\beta$  interface. For PI31, two different interfaces exist: the  $\alpha$  interface and  $\beta$  interface.

hydrogen-bonding and hydrophobic interactions. As shown in Figs. 3(a) and 3(c), the overall shapes of the  $\alpha$ -helical and  $\beta$ -sheet molecular surfaces of the Fbxo7 FP domain have a high degree of complementarity. In the  $\alpha\beta$  interface, 23 inter-domain hydrogen bonds (some of which involve structured water molecules) are present (Fig. 3b). Inter-domain hydrophobic interactions are mediated by two hydrophobic patches on the surface of the FP domain. On the  $\alpha$ -helical surface, the hydrophobic patch is defined by the side chains of His181 (noncharged and hydrophobic at neutral pH), Leu208, Ile213, Pro214, Ala219, Lys220 (the aliphatic moiety), Ala221, Leu222, His240, Pro241 and Leu242 (Fig. 3d; the side chains are represented as cyan sticks with black residue names). On the  $\beta$ -sheet surface, the hydrophobic patch is defined by the side chains of Leu230, Val233, Val253, Pro254, Leu255, Gly256, Leu258, Val260, Leu280, Phe284 and Leu307 (Fig. 3d; the surface areas defined by these residues are indicated by italic residue names in red). Some of the hydrophobic interactions are augmented by shape complementarity, as shown in Fig. 3(c) for the side chains of Val253 and Leu255, for example (note that the shown electron-density meshes are smaller than the van der Waals radius).

In a previous study (Kirk *et al.*, 2008), it was shown that a V253E mutation in an Fbxo7 protein construct (amino acids 169–398) disrupted the homodimerization of Fbxo7 and the heterodimerization between Fbxo7 and PI31. Similarly, mutation of this conserved valine in Nutcracker (Val to Glu)

significantly weakens the Nutcracker–DmPI31 interaction in *Drosophila* (Bader *et al.*, 2011). As shown in Figs. 3(c) and 3(d), Val253 is located at the centre of the hydrophobic patch of the  $\beta$ -sheet surface. This central location explains the effect of the V253E mutation.

The presence of extensive molecular interactions and shape complementarity at the  $\alpha\beta$  interface of the Fbxo7 FP domain indicates that the interface mediates stable and most likely genuine protein–protein interaction. In solution, the protein construct used in this study exists as a dimer, based on results from size-exclusion chromatography (Supplementary Fig. S1<sup>†</sup>). In a previous study (Kirk *et al.*, 2008), it was shown that an Fbxo7 protein construct (amino acids 169–389) containing the FP domain homodimerized. In the crystals in both space groups  $P2_12_12_1$  and  $P4_32_12$  each FP domain interacts with two adjacent FP domains using the  $\alpha$ -helical and  $\beta$ -sheet interacting surfaces located on opposite sides of the domain. The same  $\alpha\beta$  interface is present at both of the inter-domain interacting sites (Fig. 4a).

### 3.4. Different modes of protein–protein interaction of the two FP domains

The crystal structures of the FP domains of Fbxo7 (this study) and PI31 (Kirk *et al.*, 2008) reveal drastically different

<sup>†</sup> Supporting information has been deposited in the IUCr electronic archive (Reference: DW5066).

modes of inter-domain protein–protein interaction (Figs. 4*a* and 4*b*). In the structure of the PI31 FP domain, two FP domains interact with each other through either the ‘ $\alpha$  interface’ or the ‘ $\beta$  interface’ mediated by the  $\alpha$ -helical or  $\beta$ -sheet surfaces of the FP domains, respectively (Fig. 4*b*). Each of these two interfaces is formed by the self-association of the same molecular surface in the two adjacent FP domains (either the  $\alpha$ -helical surfaces of the two domains in red and green or the  $\beta$ -sheet surfaces of the two domains in green and blue in Fig. 4*b*). Inter-domain protein–protein interaction through the  $\alpha$  interface buries 1114 Å<sup>2</sup> of solvent-accessible molecular surface area from the two domains. For the  $\beta$  interface, 560 Å<sup>2</sup> of surface area is buried. Obviously, the molecular surfaces involved in the inter-domain protein–protein interactions in the structure of the PI31 FP domain are much smaller than those in the structure of the Fbxo7 FP domain.

More detailed comparisons in terms of the specific features that are present at the inter-domain interfaces are even more revealing. Probably owing to the self-associating nature of the  $\alpha$  and  $\beta$  interfaces of the PI31 FP domain, the degree of shape complementarity at these interfaces is low (Fig. 4*b*). The number of inter-domain hydrogen bonds at the  $\alpha$  and  $\beta$  interfaces (eight and two, respectively) of the PI31 FP domain is much smaller than the number (23) at the  $\alpha\beta$  interface of the Fbxo7 FP domain. Corresponding to the smaller molecular surface areas involved in inter-domain interaction, hydrophobic contacts at the  $\alpha$  and  $\beta$  interfaces of the PI31 FP domain are also less extensive than at the  $\alpha\beta$  interface of the Fbxo7 FP domain.

#### 4. Discussion

The results of our studies confirm the structural relationship between Fbxo7 and PI31: both proteins contain an FP domain with a similar overall fold. However, significant structural differences are observed between the two FP domains, especially regarding the modes of inter-domain protein–protein interaction. It should be noted that there is a difference in the length of the crystallization protein constructs used in the two structural studies. The protein construct for the crystallization of the Fbxo7 FP domain contains an extra C-terminal sequence that extends into the F-box motif (Fig. 1*c*). It is possible that the presence of this extra sequence stabilizes the structure of helix  $\alpha$ 5b in the Fbxo7 FP domain. Formation of helix  $\alpha$ 5b in turn stabilizes the  $\beta$ -strands ( $\beta$ 3,  $\beta$ 4 and  $\beta$ 5) that are in close proximity to helix  $\alpha$ 5b (Fig. 1*d*). The sequences of the two FP domains are 22.5% identical and 41.7% homologous (Fig. 1*c*). With this degree of sequence homology, it is also possible that the observed structural differences between the two FP domains are genuine and the structural differences may reflect functional differences.

Fbxo7 and PI31 have different biological functions and domain structures (Fig. 1*b*). The FP domain of PI31 is located at the N-terminus (amino acids 1–150) and is followed by an unstructured proline-rich region (amino acids 151–271). In Fbxo7, the FP domain is closely juxtaposed to the most likely

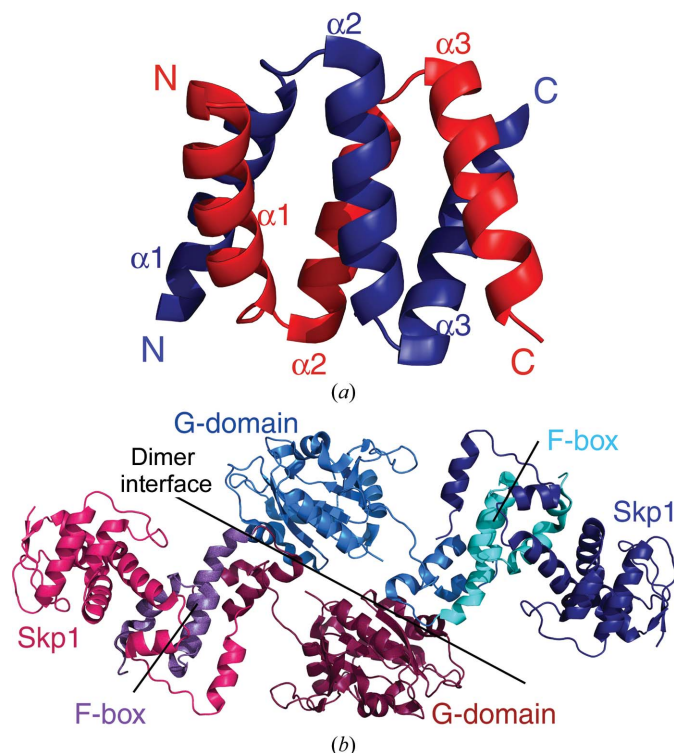
well structured F-box motif, suggesting that they may act together in certain function(s) of Fbxo7. This possibility is supported by the observation that the FP domain and F-box motif of Nutcracker are both required for positive regulation of DmPI31 stability (Bader *et al.*, 2011). In view of the different contexts of the FP domains of Fbxo7 and PI31, it is fully possible that the two FP domains may have some genuine structural differences.

The homodimerization of Fbxo7 and the heterodimerization of Fbxo7 and PI31 may be functionally important. Quite a number of F-box proteins are known to mediate dimerization of their corresponding SCF E3 ubiquitin ligases. These include proteins from all three F-box protein subfamilies, such as Fbw7, Cdc4,  $\beta$ -TrCP, Met30 and Pop1/Pop2 from the Fbxw subfamily, Skp2 from the Fbxl subfamily and Fbx4 from the Fbxo subfamily (Hao *et al.*, 2007; Welcker & Clurman, 2007; Tang *et al.*, 2007; Li & Hao, 2010). Dimerization of SCF ligases through their F-box protein subunits is emerging as an important feature for the ubiquitination activities of the ligases (Petroski & Deshaies, 2005). It is therefore key to understand the molecular basis of F-box protein association.

Homodimerization of F-box proteins of the Fbxw subfamily is mediated by a small (~45 amino acids) and highly conserved motif known as the D domain. Similar to the FP domain of Fbxo7, the D domain is closely juxtaposed to the F-box motif in Fbxw proteins. Structures of the D domains of Cdc4 and  $\beta$ -TrCP have been determined (Tang *et al.*, 2007). Each of the D domains contains three  $\alpha$ -helices that fold into similar overall structures. The D domains self-associate in a side-by-side manner to form a right-handed superhelical parallel dimer (Fig. 5*a*). The molecular details of the homodimerization of Fbx4 are revealed by the crystal structure of an Fbx4 core fragment (containing the F-box motif and the C-terminal substrate-binding G domain) in complex with Skp1 (Li & Hao, 2010). The Fbx4<sup>core</sup>–Skp1 complexes form an antiparallel homodimer mediated by Fbx4. Dimerization is achieved by interaction between a loop (within the linker region connecting the F-box motif and G domain) and the G domain (Fig. 5*b*). Apparently, the possible mode of FP domain-mediated dimerization of Fbxo7 as revealed in this study is different from those of other F-box proteins. More interestingly, as shown in Fig. 4*a*, the FP domain of Fbxo7 has the potential to mediate the simultaneous binding of two protein partners using different molecular surfaces. This unique potential may enable the FP domain to serve as a versatile protein–protein interaction module that plays important roles in various functions of Fbxo7.

It is not known whether the functional SCF<sup>Fbxo7</sup> complex is monomeric, dimeric or multimeric. The ability of the FP domain of Fbxo7 to dimerize (or multimerize as shown in Fig. 4*a*) suggests that the SCF<sup>Fbxo7</sup> complex may function as a dimer or a multimer. Fbxo7 is somewhat special in the F-box protein family because several of the partners of Fbxo7 (such as Cdk6, PI31 and p27) are not substrates of the ubiquitin proteasome system, indicating that Fbxo7 may participate in both SCF-dependent and SCF-independent activities. How Fbxo7 balances these activities is a central issue for under-





**Figure 5**

Structures of other F-box protein dimers. (a) The crystal structure of the D domain of  $\beta$ -TrCP (PDB entry 2p64; Tang *et al.*, 2007) shows a parallel superhelical dimer. The two protomers are coloured red and blue. (b) The crystal structure of the Fbx4<sup>core</sup>-Skp1 complex (PDB entry 3l2o; Li & Hao, 2010) shows an antiparallel dimer. The two protomers are roughly separated by the black line indicated as the ‘dimer interface’. The Fbx4 fragment contains the F-box motif, a linker region and the substrate-binding G domain. Dimerization is mediated by interaction between a loop within the linker region and the G domain.

standing the functional mechanisms of Fbxo7. Because Fbxo7 and PI31 both contain an FP domain that has the ability to homodimerize and heterodimerize, it is possible that PI31 may antagonize the homodimerization of Fbxo7 and therefore modulate the function of the SCF<sup>Fbxo7</sup> ligase complex. However, since the Fbxo7 FP domain has the potential to simultaneously interact with two binding partners (Fig. 4a), the formation of an FP<sub>PI31</sub>-FP<sub>Fbxo7</sub>-FP<sub>Fbxo7</sub> trimeric complex (or a multimeric complex containing the two FP domains) is also possible. In the latter scenario, binding of PI31 to Fbxo7 via their FP domains may not disrupt the functional dimeric (or multimeric) SCF<sup>Fbxo7</sup> complex. It is known that some E3 ubiquitin ligases physically associate with the 26S proteasome. The substrate-bound E3 ubiquitin ligase may participate in the delivery of substrates to the proteasome (Xie & Varshavsky, 2000; Bloom *et al.*, 2006; Farrás *et al.*, 2001). PI31 is an established regulatory protein that controls proteasome activities (Zaiss *et al.*, 2002; McCutchen-Maloney *et al.*, 2000; Zaiss *et al.*, 1999; Bader *et al.*, 2011). Interaction of Fbxo7 and PI31 may play a functional role in bringing the SCF<sup>Fbxo7</sup> ligase complex to the proteasome.

## References

Adams, P. D. *et al.* (2011). *Methods*, **55**, 94–106.

- Bader, M., Benjamin, S., Wapinski, O. L., Smith, D. M., Goldberg, A. L. & Steller, H. (2011). *Cell*, **145**, 371–382.
- Bai, C., Sen, P., Hofmann, K., Ma, L., Goebel, M., Harper, J. W. & Elledge, S. J. (1996). *Cell*, **86**, 263–274.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Bloom, J., Peschiaroli, A., DeMartino, G. & Pagano, M. (2006). *Cell Div.* **1**, 5.
- Cenciarelli, C., Chiaur, D. S., Guardavaccaro, D., Parks, W., Vidal, M. & Pagano, M. (1999). *Curr. Biol.* **9**, 1177–1179.
- Chang, Y.-F., Cheng, C.-M., Chang, L.-K., Jong, Y.-J. & Yuo, C.-Y. (2006). *Biochem. Biophys. Res. Commun.* **342**, 1022–1026.
- Chothia, C. & Janin, J. (1975). *Nature (London)*, **256**, 705–708.
- Di Fonzo, A. *et al.* (2009). *Neurology*, **72**, 240–245.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Farrás, R., Ferrando, A., Jásik, J., Kleinow, T., Okrészl, L., Tiburcio, A., Salchert, K., del Pozo, C., Schell, J. & Koncz, C. (2001). *EMBO J.* **20**, 2742–2756.
- Goldenberg, S. J., Cascio, T. C., Shumway, S. D., Garbutt, K. C., Liu, J., Xiong, Y. & Zheng, N. (2004). *Cell*, **119**, 517–528.
- Hao, B., Oehlmann, S., Sowa, M. E., Harper, J. W. & Pavletich, N. P. (2007). *Mol. Cell*, **26**, 131–143.
- Hershko, A. (2005). *Cell Death Differ.* **12**, 1191–1197.
- Hsu, J.-M., Lee, Y.-C. G., Yu, C.-T. R. & Huang, C.-Y. F. (2004). *J. Biol. Chem.* **279**, 32592–32602.
- Jones, S. & Thornton, J. M. (1995). *Prog. Biophys. Mol. Biol.* **63**, 31–65.
- Kirk, R., Laman, H., Knowles, P. P., Murray-Rust, J., Lomonosov, M., Mezziane, el K. & McDonald, N. Q. (2008). *J. Biol. Chem.* **283**, 22325–22335.
- Kuiken, H. J., Egan, D. A., Laman, H., Bernards, R., Beijersbergen, R. L. & Dirac, A. M. (2012). *J. Cell. Mol. Med.* **16**, 2140–2149.
- Kumar, A. & Paietta, J. V. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 3343–3347.
- Laman, H., Funes, J. M., Ye, H., Henderson, S., Galinanes-Garcia, L., Hara, E., Knowles, P., McDonald, N. & Boshoff, C. (2005). *EMBO J.* **24**, 3104–3116.
- Li, Y. & Hao, B. (2010). *J. Biol. Chem.* **285**, 13896–13906.
- Los, G. V. *et al.* (2008). *ACS Chem. Biol.* **3**, 373–382.
- McCutchen-Maloney, S. L., Matsuda, K., Shimbara, N., Binns, D. D., Tanaka, K., Slaughter, C. A. & DeMartino, G. N. (2000). *J. Biol. Chem.* **275**, 18557–18565.
- Mizushima, T., Yoshida, Y., Kumanomidou, T., Hasegawa, Y., Suzuki, A., Yamane, T. & Tanaka, K. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 5777–5781.
- Petroski, M. D. & Deshaies, R. J. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 9–20.
- Schulman, B. A., Carrano, A. C., Jeffrey, P. D., Bowen, Z., Kinnucan, E. R., Finnin, M. S., Elledge, S. J., Harper, J. W., Pagano, M. & Pavletich, N. P. (2000). *Nature (London)*, **408**, 381–386.
- Shojaee, S., Sina, F., Banihosseini, S. S., Kazemi, M. H., Kalhor, R., Shahidi, G. A., Fakhrai-Rad, H., Ronaghi, M. & Elahi, E. (2008). *Am. J. Hum. Genet.* **82**, 1375–1384.
- Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* **D60**, 432–438.
- Tang, X., Orlicky, S., Lin, Z., Willems, A., Neculai, D., Ceccarelli, D., Mercurio, F., Shilton, B. H., Sicheri, F. & Tyers, M. (2007). *Cell*, **129**, 1165–1176.
- Welcker, M. & Clurman, B. E. (2007). *Cell Div.* **2**, 7.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Winston, J. T., Koeppe, D. M., Zhu, C., Elledge, S. J. & Harper, J. W. (1999). *Curr. Biol.* **9**, 1180–1182.
- Wu, G., Xu, G., Schulman, B. A., Jeffrey, P. D., Harper, J. W. & Pavletich, N. P. (2003). *Mol. Cell*, **11**, 1445–1456.
- Xie, Y. & Varshavsky, A. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 2497–2502.

- Zaiss, D. M. W., Standera, S., Holzhütter, H., Kloetzel, P.-M. & Sijts, A. J. A. M. (1999). *FEBS Lett.* **457**, 333–338.
- Zaiss, D. M. W., Standera, S., Kloetzel, P.-M. & Sijts, A. J. A. M. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 14344–14349.
- Zeng, Z., Wang, W., Yang, Y., Chen, Y., Yang, X., Diehl, J. A., Liu, X. & Lei, M. (2010). *Dev. Cell*, **18**, 214–225.
- Zhao, T., De Graaff, E., Breedveld, G. J., Loda, A., Severijnen, L. A., Wouters, C. H., Verheijen, F. W., Dekker, M. C. J., Montagna, P., Willemsen, R., Oostra, B. A. & Bonifati, V. (2011). *PLoS One*, **6**, e16983.
- Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W. & Pavletich, N. P. (2002). *Nature (London)*, **416**, 703–709.