Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of human ubiquitin-conjugating enzyme E2 G2 (UBE2G2/UBC7)

Ryoichi Arai,[a,b]‡ Seiko Yoshikawa,[a]‡ Kazutaka Murayama,[a,c] Yuzuru Imai,[d] Ryosuke Takahashi,[d,e] Mikako Shirouzu[a,b] and Shigeyuki Yokoyama[a,b,f]*

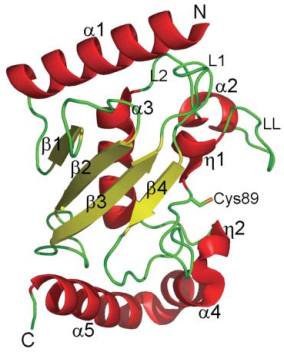
[a]Protein Research Group, RIKEN Genomic Sciences Center, Tsurumi, Yokohama 230-0045, Japan, [b]RIKEN SPring-8 Center, Harima Institute, Sayo, Hyogo 679-5148, Japan, [c]Tohoku University Biomedical Engineering Research Organization, Aoba, Sendai 980-8575, Japan, [d]RIKEN Brain Science Institute, Wako, Saitama 351-0198, Japan, [e]Department of Neurology, Graduate School of Medicine, Kyoto University, Sakyo, Kyoto 606-8507, Japan, and [f]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Bunkyo, Tokyo 113-0033, Japan

‡ These authors contributed equally to this work.

Correspondence e-mail: yokoyama@biochem.s.u-tokyo.ac.jp

Received 22 December 2005
Accepted 10 March 2006

PDB Reference: human UBE2G2/UBC7, 2cyx, r2cyxsf.



The human ubiquitin-conjugating enzyme E2 G2 (UBE2G2/UBC7) is involved in protein degradation, including a process known as endoplasmic reticulum-associated degradation (ERAD). The crystal structure of human UBE2G2/UBC7 was solved at 2.56 Å resolution. The UBE2G2 structure comprises a single domain consisting of an antiparallel β-sheet with four strands, five α-helices and two 3_{10}-helices. Structural comparison of human UBE2G2 with yeast Ubc7 indicated that the overall structures are similar except for the long loop region and the C-terminal helix. Superimposition of UBE2G2 on UbcH7 in a c-Cbl–UbcH7–ZAP70 ternary complex suggested that the two loop regions of UBE2G2 interact with the RING domain in a similar way to UbcH7. In addition, the extra loop region of UBE2G2 may interact with the RING domain or its neighbouring region and may be involved in the binding specificity and stability.

1. Introduction

Ubiquitin-dependent protein degradation plays an important role in the regulation of various cellular processes, including cell-cycle progression, signal transduction, transcription, DNA repair and protein quality control (Koepp *et al.*, 1999; Laney & Hochstrasser, 1999). Ubiquitination involves the successive actions of the ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-protein ligase enzymes (E3) (Hershko & Ciechanover, 1998; Pickart, 2001). The E1 enzyme activates free ubiquitin and transfers it to E2 through a thioester linkage between the ubiquitin C-terminus and an E2 active-site cysteine. The E3 enzyme recognizes its substrate and E2 and catalyzes the formation of an isopeptide bond between a lysine ε-amino group of the substrate (or ubiquitin) and the C-terminal carboxyl group of ubiquitin Gly76. Over 30 human E2s have been identified and they all contain a conserved ~150 amino-acid catalytic core. The E2 enzymes are grouped into four classes depending on the presence and the location of additional sequences (Jentsch, 1992). Some of these enzymes contain extra C-terminal and/or N-terminal extensions from the core domain. The class I enzymes are the smallest E2 enzymes and consist almost entirely of the conserved core domain. The class II enzymes contain an extra C-terminal extension from the core domain, while class III enzymes have an N-terminal extension. The class IV enzymes contain both N- and C- terminal extensions.

The human UBE2G2 gene encodes the ubiquitin-conjugating enzyme E2 G2 (UBE2G2/UBC7), with a molecular weight of 18.6 kDa (165 amino acids). It was mapped to the region of human chromosome 21q22.3 and its transcripts are ubiquitously expressed in human tissues (Katsanis & Fisher, 1998; Rose *et al.*, 1998). Human UBE2G2 is a class I E2 enzyme. Recently, bacterial expression of His-tagged human UBE2G2 was reported (Reyes *et al.*, 2006). The human UBE2G2 protein shares 100, 62, 47 and 27% identities to murine UBE2G2/UBC7 (MmUBC7), yeast Ubc7, human UBE2G1 and human UbcH7, respectively (Fig. 1*a*). The crystal structures of yeast Ubc7 (Cook *et al.*, 1997), a human E6AP–UbcH7 complex (Huang *et*

al., 1999) and a human c-Cbl–UbcH7–ZAP-70 complex (Zheng *et al.*, 2000) have been reported. Functional studies have associated yeast Ubc7 and MmUBC7 with the degradation of endoplasmic reticulum (ER) substrates, a process known as ER-associated degradation (ERAD; Jungmann *et al.*, 1993; Fang *et al.*, 2001; Tiwari & Weissman, 2001). Parkin, a gene product responsible for autosomal recessive juvenile Parkinsonism (AR-JP), interacts with human UBE2G2/UBC7 and UBC6 through its RING domain and specifically ubiquitinates the Pael receptor in the presence of the E2s (Imai *et al.*, 2001). Furthermore, exogenous MmUBC7 mediates the ubiquitination and down regulation of both the inositol 1,4,5-triphosphate receptor in human neuroblastoma cells (Webster *et al.*, 2003) and the human type 2 iodothyronine selenodeiodinase (Kim *et al.*, 2003). Recently, the interactions of human UBE2G2/UBC7 with some RING-finger E3s, such as human HRD1 (Kikkert *et al.*, 2004) and TEB4 (Hassink *et al.*, 2005), have been reported. To analyze the structural and functional details of human UBE2G2/UBC7, which is involved in important cellular processes, its structure must be determined and compared with those of its homologues. Here, we report the crystal structure of human UBE2G2/UBC7 at 2.56 Å resolution and discuss its structural aspects.

2. Materials and methods

2.1. Protein expression and purification

The human UBE2G2 gene (Imai *et al.*, 2001) encoding human ubiquitin-conjugating enzyme E2 G2 (UBE2G2/UBC7) was cloned into a modified pENTR vector with a tobacco etch virus (TEV) protease cleavage site, derived from pENTR1A (Invitrogen). The expression vector pET/cMBP-UBE2G2 was constructed using Gateway technology (Invitrogen) with pENTR/TEV-UBE2G2 and pET/cMBP-GATEWAY bearing a T7 promoter, an N-terminal maltose-binding protein (MBP) tag and a Gateway reading frame cassette A (Invitrogen). The UBE2G2 protein was expressed as a fusion with an N-terminal MBP tag and a TEV protease cleavage site in *Escherichia coli* BL21(DE3). The protein was first purified on an amylose-resin column (New England Biolabs) and the MBP tag was then cleaved by His-tagged TEV protease, which was removed using a HisTrap column (GE Healthcare). The protein was purified further by Mono-Q and Superdex 75 column (GE Healthcare) chromatography steps. The yield of purified UBE2G2 protein was 8 mg per litre of culture. The construct that was used for crystallization contained the cloning artifact sequence GGSEF at the N-terminus.

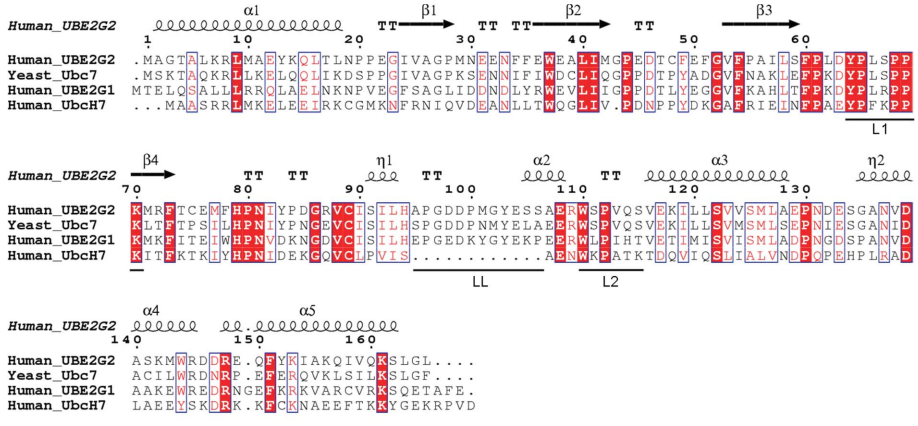


Figure 1
(*a*) Sequence alignment of homologues of human UBE2G2/UBC7. The alignment was generated by *ESPript* (Gouet *et al.*, 1999) with *CLUSTAL W* (Thompson *et al.*, 1994). The secondary structures of the human UBE2G2 protein, as determined by *DSSP* (Kabsch & Sander, 1983), are shown above the sequences (α, α-helix; β, β-strand; η, 3_{10}-helix; TT, β-turn). (*b*) Ribbon representation of the human UBE2G2/UBC7 structure (amino acids 1–165; stereoview). The helices and the β-strands are shown in red and yellow, respectively. The active-site residue (Cys89) is shown as a stick model.

2.2. Crystallization and data collection

Preliminary crystals of human UBE2G2 were obtained under condition No. 42 (0.1 *M* Tris–HCl buffer pH 8.5 containing 1.5 *M* ammonium sulfate and 12% glycerol) of the Crystal Screen 2 crystal screening kit (Hampton Research) using the 96-well sitting-drop vapour-diffusion method. The crystals of UBE2G2 used for structure determination were obtained in drops composed of 1 µl 8.5 mg ml^{-1} protein solution (20 m*M* Tris–HCl buffer pH 8.0 containing 120 m*M* NaCl, 2 m*M* DTT) and 1 µl reservoir solution (0.1 *M* Tris–HCl buffer pH 8.1 containing 1.45 *M* ammonium sulfate and 12% glycerol; Hampton Research) by the hanging-drop vapour-diffusion method against 500 µl reservoir solution. A rod-like crystal (~350 × 100 × 100 µm) was obtained within a few days and was used for data collection. The data collection was carried out at 100 K, with the reservoir solution containing 27.5% glycerol as a cryoprotectant. The diffraction data were collected at SPring-8 BL26B1 (Yamamoto *et al.*, 2002) and were recorded on a Jupiter 210 CCD detector (Rigaku). All diffraction data were processed with the *HKL*2000 program suite (Otwinowski & Minor, 1997).

2.3. Structure determination and refinement

The structure was solved by the molecular-replacement method using *MOLREP* (Vagin & Teplyakov, 1997) with the yeast Ubc7 structure (PDB code 2ucz; Cook *et al.*, 1997) as a search model. Data in the resolution range 50–3.0 Å were used in both rotation and translation calculations, which gave an obvious solution with significant contrast, resulting in three molecules in the asymmetric unit with a Matthews coefficient (V_M) of 3.83 Å^3 Da^{-1} and a solvent content of 67.91%. The model was corrected iteratively using *O* (Jones *et al.*, 1991) and was refined to 2.56 Å using *LAFIRE* (Yao *et al.*, 2006), *REFMAC*5 (Murshudov *et al.*, 1997) and *Crystallography & NMR System* (*CNS*; Brünger *et al.*, 1998). The crystallographic data and refinement statistics are presented in Table 1. Since there was additional electron density, four residues of the cloning artifact sequence at the N-terminus were also modelled. The final model includes 507 amino-acid residues of three UBE2G2 monomers and 23 water molecules in the asymmetric unit. In the loop regions (residues 100–106 and 131–133), the electron density corresponding to the side chains was ambiguous, which increased the *B* factor. In addition, relatively large areas of the molecular surface were exposed to the solvent in the crystal as the solvent content was high. These features resulted in the high average *B* factor. The quality of the model was inspected using *PROCHECK* (Laskowski *et al.*, 1993). The figures were created using *PyMOL* (DeLano, 2005).

Table 1
X-ray data-collection and refinement statistics.

Values in parentheses are for the outer shell (2.65–2.56 Å).

Data collection	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 63.52, b = 87.61, c = 157.41
Wavelength (Å)	1.000
Resolution (Å)	50–2.56
Total reflections	117935
Unique reflections	28705
Redundancy	4.1 (3.7)
Completeness (%)	97.5 (84.6)
$I/\sigma(I)$	22.2 (4.2)
R_{sym}† (%)	5.5 (29.3)
Refinement	
Resolution (Å)	49.43–2.56
No. of reflections	28395
No. of protein atoms	3996
No. of water molecules	23
R_{work} (%)	22.8
R_{free}‡ (%)	26.2
R.m.s.d. bond lengths (Å)	0.009
R.m.s.d. bond angles (°)	1.6
Average B factor (Å^2)	75.7
Ramachandran plot	
Most favoured regions (%)	85.9
Additional allowed regions (%)	14.1
Generously allowed regions (%)	0.0
Disallowed regions (%)	0.0

† $R_{sym} = \sum|I_i - I_{avg}|/\sum I_i$, where I_i is the observed intensity and I_{avg} is the average intensity. ‡ R_{free} is calculated for 10% of randomly selected reflections excluded from refinement.

3. Results and discussion

The crystal structure of human UBE2G2 comprises a single domain consisting of an antiparallel β-sheet with four strands (β1–β4), five α-helices (α1–α5) and two 3_{10}-helices (η1 and η2; Fig. 1*b*). The ubiquitin-accepting residue Cys89 is located near η1. The overall folding of UBE2G2 corresponds to the typical fold of ubiquitin-conjugating enzymes. According to analytical ultracentrifugation, the

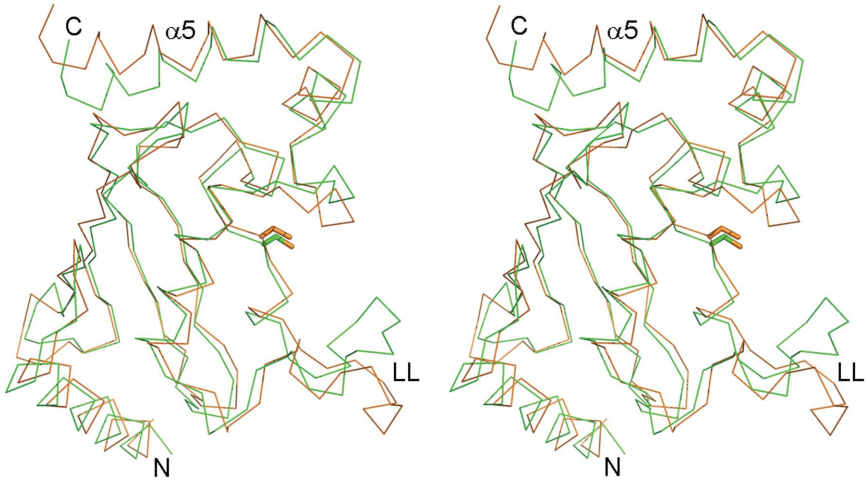


Figure 2
Superimposition of the main-chain structures of human UBE2G2 (green) and yeast Ubc7 (orange) (PDB code 2ucz; Cook *et al.*, 1997) (stereoview). The active-site cysteine residues are shown as stick models. The superimposition was carried out with *LSQKAB* (Kabsch, 1976).

molecular weight of UBE2G2 was ~18 kDa (data not shown), indicating that the UBE2G2 protein exists as a monomer in solution.

Fig. 2 shows the superimposition of the main-chain structures of human UBE2G2 and yeast Ubc7 (Cook *et al.*, 1997). The overall structure of UBE2G2 is remarkably similar to that of yeast Ubc7 (r.m.s.d. = 2.15 Å over 164 C^{α} atoms). The major differences between human UBE2G2 and yeast Ubc7 are the structure of the long loop (LL) region (95–106) and the angle of the C-terminal helix. The C-terminal helix (α5) of UBE2G2 is closer to the β-sheet core region than that of yeast Ubc7. The important interactions of UBE2G2 in the contact region of the C-terminal helix and the core region are the hydrophobic interactions among Phe54, Met77, Phe78, Ile154 and Ile158 and the salt bridge between Glu76 and Lys161. The residues Glu76, Ile154 and Ile158 are replaced with Ser76, Gln154 and Ser158 in yeast Ubc7, respectively, suggesting that the interactions of yeast Ubc7 are weaker than those of UBE2G2. Consequently, the angle of the C-terminal helix (α5) may change. Recently, the crystal structure of the human ubiquitin-conjugating enzyme E2 G1 (UBE2G1), which is another human homologue of yeast Ubc7, was deposited in the PDB (PDB code 2awf). A structural comparison of UBE2G2 with UBE2G1 revealed that the overall folding of UBE2G2 is similar to that of UBE2G1 (r.m.s.d. = 1.12 Å over 115 C^{α} atoms), but in UBE2G1 the residues 98–106 within the long loop (LL) region and the C-terminal helices (η2, α4 and α5) were not located in the model owing to disorder.

Zheng and coworkers reported the crystal structure of a c-Cbl–UbcH7–ZAP70 peptide ternary complex (PDB code 1fbv; Zheng *et al.*, 2000). It revealed how the RING domain of c-Cbl recruits the ubiquitin-conjugating enzyme UbcH7. Fig. 3(*a*) shows the superimposition of the main-chain structures of UBE2G2 and UbcH7 in the ternary complex. The overall folding of UBE2G2 and UbcH7 overlaps roughly (r.m.s.d. = 2.95 Å over 143 C^{α} atoms). Fig. 3(*b*) shows a close-up view of the interface between the RING domain and the E2s. The critical residues of UbcH7 for the interaction with the RING domain, Pro62, Phe63, Lys96, Pro97 and Ala98 (Zheng *et al.*, 2000), and the corresponding residues of UBE2G2, Pro65, Leu66,

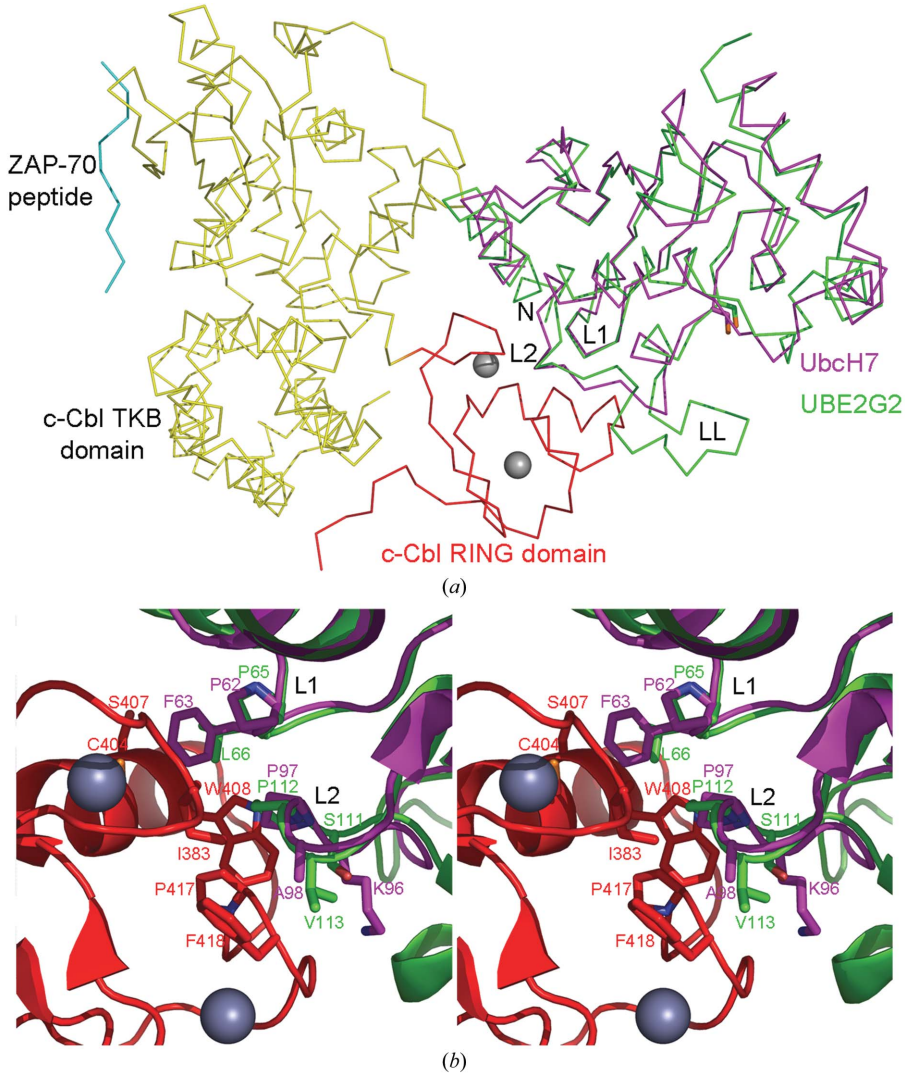


Figure 3
(*a*) Superimposition of the main-chain structures of human UBE2G2 and UbcH7 in the c-Cbl–UbcH7–ZAP70 peptide ternary complex (PDB code 1fbv; Zheng *et al.*, 2000). The TKB domain and linker sequence of c-Cbl, the RING domain of c-Cbl, the ZAP-70 peptide and human UbcH7 are coloured yellow, red, cyan and magenta, respectively. The zinc ions are indicated by grey spheres. The human UBE2G2 protein is coloured green. The active-site cysteine residues are shown as stick models. (*b*) Close-up view of the ribbon representation of the interface between the RING domain and UbcH7 in the c-Cbl–UbcH7–ZAP70 complex and the superimposition of UBE2G2 on UbcH7 (stereoview). The colouring is the same as that in Fig. 3(*a*). The critical residues for the interaction of UbcH7 with the *RING* domain and the corresponding residues of UBE2G2 are shown as stick models. All superimpositions were carried out with *LSQKAB* (Kabsch, 1976).

Ser111, Pro112 and Val113, overlap remarkably well (r.m.s.d. = 0.768 Å over five C^{α} atoms), suggesting that the L1 (64–70) and L2 (110–115) loops of UBE2G2 are involved in the interaction with the RING domain in a similar way as UbcH7. This is consistent with the previous results that Parkin binds to UBE2G2 as well as UbcH7 and ubiquitinates substrates (Imai *et al.*, 2000, 2001). In addition, UBE2G2 has the extra long loop (LL) region (95–106), which is probably located on the side near the RING domain. The *B* factor of the LL region is relatively high, implying the possibility of conformational flexibility. The LL region may interact with the RING domain or its neighbouring region and may be involved in the binding specificity and stability.

We thank Mr R. Akasaka and Dr M. Kukimoto-Niino for the analytical ultracentrifugation, Mr S. Kamo for computer maintenance and Ms A. Ishii, Ms K. Yajima, Ms M. Sunada and Ms T. Nakayama for clerical assistance. We also thank Dr M. Yamamoto for data collection at the RIKEN Structural Genomics beamline BL26B1 at SPring-8. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cook, W. J., Martin, P. D., Edwards, B. F., Yamazaki, R. K. & Chau, V. (1997). *Biochemistry*, **36**, 1621–1627.
DeLano, W. L. (2005). *PyMOL* v.0.98. DeLano Scientific, South San Francisco, CA, USA.
Fang, S., Ferrone, M., Yang, C., Jensen, J. P., Tiwari, S. & Weissman, A. M. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 14422–14427.
Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Hassink, G., Kikkert, M., van Voorden, S., Lee, S. J., Spaapen, R., van Laar, T., Coleman, C. S., Bartee, E., Fruh, K., Chau, V. & Wiertz, E. (2005). *Biochem. J.* **388**, 647–655.
Hershko, A. & Ciechanover, A. (1998). *Annu. Rev. Biochem.* **67**, 425–479.
Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P. M., Huibregtse, J. M. & Pavletich, N. P. (1999). *Science*, **286**, 1321–1326.
Imai, Y., Soda, M., Inoue, H., Hattori, N., Mizuno, Y. & Takahashi, R. (2001). *Cell*, **105**, 891–902.
Imai, Y., Soda, M. & Takahashi, R. (2000). *J. Biol. Chem.* **275**, 35661–35664.
Jentsch, S. (1992). *Annu. Rev. Genet.* **26**, 179–207.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Jungmann, J., Reins, H. A., Schobert, C. & Jentsch, S. (1993). *Nature (London)*, **361**, 369–371.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Katsanis, N. & Fisher, E. M. (1998). *Genomics*, **51**, 128–131.
Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thanedar, S., Roitelman, J., Chau, V. & Wiertz, E. (2004). *J. Biol. Chem.* **279**, 3525–3534.
Kim, B. W., Zavacki, A. M., Curcio-Morelli, C., Dentice, M., Harney, J. W., Larsen, P. R. & Bianco, A. C. (2003). *Mol. Endocrinol.* **17**, 2603–2612.
Koepp, D. M., Harper, J. W. & Elledge, S. J. (1999). *Cell*, **97**, 431–434.
Laney, J. D. & Hochstrasser, M. (1999). *Cell*, **97**, 427–430.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pickart, C. M. (2001). *Annu. Rev. Biochem.* **70**, 503–533.
Reyes, L. F., Sommer, C. A., Beltramini, L. M. & Henrique-Silva, F. (2006). *Protein Expr. Purif.* **45**, 324–328.
Rose, S. A., Leek, J. P., Moynihan, T. P., Ardley, H. C., Markham, A. F. & Robinson, P. A. (1998). *Cytogenet. Cell Genet.* **83**, 98–99.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Tiwari, S. & Weissman, A. M. (2001). *J. Biol. Chem.* **276**, 16193–16200.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Webster, J. M., Tiwari, S., Weissman, A. M. & Wojcikiewicz, R. J. (2003). *J. Biol. Chem.* **278**, 38238–38246.
Yamamoto, M., Kumasaka, T., Ueno, G., Ida, K., Kanda, H., Miyano, M. & Ishikawa, T. (2002). *Acta Cryst.* A**58**, C302.
Yao, M., Zhou, Y. & Tanaka, I. (2006). *Acta Cryst.* D**62**, 189–196.
Zheng, N., Wang, P., Jeffrey, P. D. & Pavletich, N. P. (2000). *Cell*, **102**, 533–539.