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

Crystallization and preliminary X-ray diffraction studies of the ubiquitin-like (UbL) domain of the human homologue A of Rad23 (hHR23A) protein

Yu Wai Chen,* Toshitaka Tajima,‡ Martin Rees and Mitla Garcia-Maya

Randall Division of Cell and Molecular Biophysics, King's College London, New Hunt's House, Guy's Campus, London SE1 1UL, England

‡ Current address: Department of Structural and Molecular Biology, University College London, Gower Street, London WC1E 6BT, England.

Correspondence e-mail: yu-wai.chen@kcl.ac.uk

Received 26 May 2009
Accepted 8 August 2009

Human homologue A of Rad23 (hHR23A) plays dual roles in DNA repair as well as serving as a shuttle vehicle targeting polyubiquitinated proteins for degradation. Its N-terminal ubiquitin-like (UbL) domain interacts with the 19S proteasomal cap and provides the docking mechanism for protein delivery. Pyramidal crystals of the UbL domain of hHR23A were obtained by the hanging-drop vapour-diffusion method with ammonium sulfate as the crystallizing agent. The crystals diffracted to beyond 2 Å resolution and belonged to the hexagonal space group $P6_522$, with unit-cell parameters $a = b = 78.48$, $c = 63.57$ Å. The structure was solved by molecular replacement using the UbL domain of yeast Dsk2 as the search model.

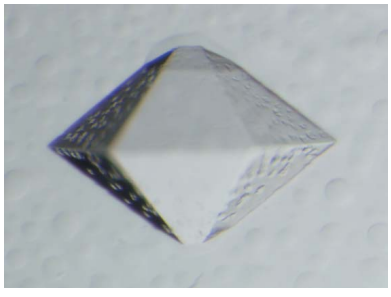
1. Introduction

Rad23 proteins are essential cellular components that are conserved in a wide range of organisms from yeast to humans. They play dual roles and bridge between two important cellular processes: DNA repair and protein recycling. When UV-induced DNA lesion occurs, Rad23 proteins are recruited to the damage site and are involved in nucleotide excision (for a review, see Dantuma *et al.*, 2009). In humans, there are two homologues of the yeast Rad23 protein, hHR23A and hHR23B, which have overlapping DNA-repair functions.

Rad23 proteins also serve to deliver polyubiquitinated proteins to the 26S proteasome for degradation (reviewed in Grabbe & Dikic, 2009). Although their DNA-repair functions are similar, the two human homologues exhibit different characteristics of binding to the proteasome and to polyubiquitinated proteins (Chen & Madura, 2006).

The Rad23 proteins have a modular architecture and consist of four well defined folded domains: from the N-terminus, a ubiquitin-like domain (UbL), a ubiquitin-associated domain (UbA1), an XPC-binding domain and a second ubiquitin-associated domain (UbA2). The domains are interconnected by unstructured linker regions. Study of full-length hHR23A revealed its intrinsic structural flexibility, with individual domains interacting with each other in a highly dynamic manner. This has been argued to be an essential feature that allows hHR23A to carry out its protein-delivery functions (Walters *et al.*, 2003; Goh *et al.*, 2008). By a complex interplay of competition for binding sites among sets of ubiquitin or UbL moieties and their receptors (*e.g.* the UbA domains), the polyubiquitinated proteins are loaded, transported and unloaded by hHR23A into the 26S proteasome for processing (Kang, Chen *et al.*, 2007; Grabbe & Dikic, 2009).

After hHR23A picks up a polyubiquitinated protein *via* its UbA domains (Raasi *et al.*, 2004), the UbL domain then binds strongly to proteasomal subunit 5a (S5a; Mueller & Feigon, 2003) and translocates its cargo into the proteasome. Interestingly, the UbL domain of hHR23A interacts in a highly dynamic intramolecular manner with both of its UbA domains (Walters *et al.*, 2003). In a similar way, this



UbL domain also binds to the UbA domain of another ubiquitin-receptor molecule, hHLIC2 (Kang, Zhang *et al.*, 2007). The hHR23A UbL domain also binds the catalytic domain of ataxin-3, a cysteine protease enzyme that carries out de-ubiquitination (Wang *et al.*, 2000).

Previously, researchers have used nuclear magnetic resonance (NMR) spectroscopy to investigate the structure and interactions of the individual domains of hHR23A. The structure of hHR23A residues 1–78 has been determined and the important residues that mediate its binding to proteasomal subunit S5a have been identified (Mueller & Feigon, 2003). Here, we report the crystallization and structure solution of the 9.6 kDa UbL domain comprising residues 1–82. This is the first crystallographic study concerning this protein. It will be of great interest to compare the structures of this domain as determined independently by the two structural methods. The high-resolution crystal structure will reveal accurate structural details, especially concerning the conformations of the critical side chains that confer the ability to bind S5a and ataxin-3 upon hHR23A. Such information will be valuable in guiding future protein-engineering experiments and studies of the molecular binding interface.

2. Experimental procedures

2.1. Overproduction and purification

The plasmid pGEX-hHR23A(1–82) contains the cDNA sequence encoding residues 1–82 of hHR23A subcloned into the pGEX-5X-1 vector (GE Healthcare) for the production of an N-terminal glutathione *S*-transferase (GST) fusion protein (Wang *et al.*, 2000). The plasmid was transformed into *Escherichia coli* Rosetta 2 (DE3) cells (Novagen) and grown at 310 K until the absorbance at 600 nm (A_{600}) reached 0.4–0.6. Overproduction of hHR23A UbL was then induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to 0.5 m*M* followed by a further 3 h incubation at 310 K. The cells were lysed with a Constant cell-disruption system at 165 MPa at 277 K in a buffer containing 20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl with Complete (EDTA-free) protease-inhibitor cocktail (Roche). After centrifugation, the GST-fusion protein was extracted from the supernatant using glutathione-Sepharose HP resin (GE Healthcare) according to the standard batch protocol provided by the manufacturer, with a wash buffer containing 1× PBS (137 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 2 m*M* KH_2PO_4), 5% glycerol and 1 m*M* dithiothreitol. When the GST-fusion protein was bound to the resin, the buffer was replaced with factor Xa cleavage buffer (20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl and 1 m*M* $CaCl_2$). The GST tag was cleaved by incubation with factor Xa (New England Biolabs; 10 µg of 1 mg ml^{-1} enzyme per millilitre of reaction volume) for 13 h at room temperature. The unbound phase was collected and further purified by Superdex 75 16/60 size-exclusion chromatography (GE Healthcare) in a buffer containing 20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 1 m*M* EDTA and 1 m*M* dithiothreitol. The purified protein had three additional plasmid-encoded residues (Gly-Ile-Pro) N-terminal to the hHR23A UbL domain.

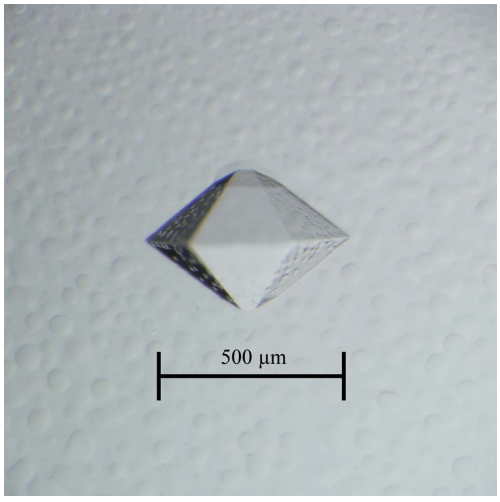


Figure 1
A hexagonal bipyramidal crystal of hHR23A UbL domain.

2.2. Crystallization and data collection

Initial screening for crystallization conditions was performed with Hampton Research Crystal Screen and Crystal Screen 2 at room temperature. Crystals for diffraction studies were grown using the hanging-drop vapour-diffusion method at room temperature. Each drop was set up by mixing 2 µl 5.6 mg ml^{-1} protein solution with 2 µl reservoir solution (0.1 *M* HEPES pH 7.5, 2% PEG 400 and 2.0 *M* ammonium sulfate) on a cover slip, which was inverted and sealed over a well containing 1 ml reservoir solution. No crystals were observed for more than two months, but in many of these trials a single hexagonal bipyramidal crystal grew to maximum dimensions of 0.5 × 0.5 × 0.4 mm in one year (Fig. 1). The crystals were soaked in a cryoprotectant solution having the same composition as the reservoir solution but with 15% glycerol (Garman & Mitchell, 1996) for 3 min. The crystals were then mounted and flash-cooled in a nitrogen stream at 100 K for data collection.

One crystal (0.4 × 0.4 × 0.3 mm) was used to collect 257 frames (each of 60 s exposure and 1° oscillation; Fig. 2) in six separate runs using an in-house Oxford Diffraction Xcalibur Nova system with sealed-tube X-ray generator (Cu $K\alpha$; λ = 1.5418 Å) equipped with an Onyx CCD detector.

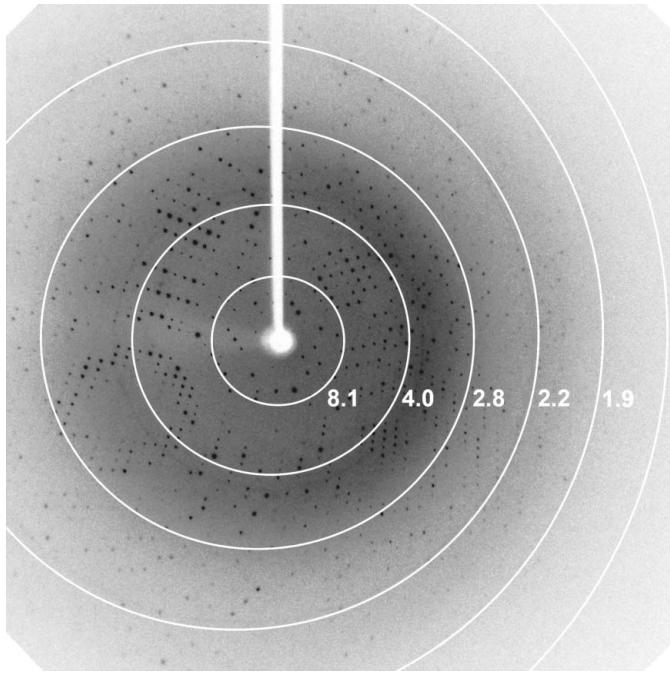


Figure 2
Typical X-ray diffraction image of an hHR23A UbL domain crystal. The rings indicate resolution (in Å).

Table 1
Data-processing statistics for hHR23A UbL domain.

Values in parentheses are for the highest resolution shell.

Space group	$P6_122/P6_522$
Unit-cell parameters (Å)	$a = b = 78.48$, $c = 63.57$
Resolution (Å)	20.2–1.97 (2.08–1.97)
No. of measurements (no cutoff)	200887
No. of unique reflections	8601
Multiplicity	23.3 (16.2)
Completeness (%)	99.9 (100)
$R_{\rm merge}$† (%)	5.6 (38.1)
Mean $I/\sigma(I)$	12.4 (2.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

3. Results and discussion

3.1. Data processing

The diffraction data were automatically indexed and integrated with the Oxford Diffraction *CrysAlis Pro* software and were then merged, scaled and reduced with programs (*SCALA* and *TRUNCATE*) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Pseudo-precession images generated with *HKLVIEW* (from *CCP*4) exhibited 6/*mmm* Laue symmetry and belonged to the hexagonal system. The only reflection condition observed was $l = 6n$, indicating that the space group was either $P6_122$ or $P6_522$. The statistics of data processing are summarized in Table 1. The most probable asymmetric unit content is one 9.6 kDa protein molecule, giving a Matthews coefficient ($V_{\rm M}$) of 3.04 Å^3 Da^{-1}, which is within the usually observed range (Matthews, 1968) and corresponds to 59.5% solvent content.

3.2. Molecular replacement

We attempted structure solution using the molecular-replacement method implemented in the program *Phaser* (McCoy *et al.*, 2007) by employing the UbL-domain structures deposited in the Protein Data Bank (Bernstein *et al.*, 1977; Berman *et al.*, 2000) as search models. Early attempts using the NMR ensemble structure of the same protein (Mueller & Feigon, 2003) as a search model did not yield any discriminating solution. Next, we used the high-resolution crystal structure of the UbL domain of the yeast Dsk2 protein (Lowe *et al.*, 2006; PDB code 2bwf) as a search model. The Dsk2 and hHR23A UbL domains share 39% amino-acid identity. The 2bwf coordinates were edited such that nonconserved residues had their side-chain atoms truncated beyond C^{β}.

The molecular-replacement calculations were performed using data to 4.0 Å resolution in the auto mode of *Phaser*. The solution showed up as the top peak in $P6_522$, with a modest RF (rotation function) Z score of 3.9 but a prominent TF (translation function) Z score of 12.2 that stood out from the second best peak, which had a TF Z score of 5.0. The corresponding refined log-likelihood gain of the solution was 106. In contrast, the only solution in the enantiomorphic space group $P6_122$ had an RF Z score of 3.4, a TF Z score of 4.2 and a refined log-likelihood gain of 18.

The molecular-replacement solution was examined graphically with the program *Coot* (Emsley & Cowtan, 2004) and found to agree convincingly well with the $2mF_{\rm o} - DF_{\rm c}$ electron-density map. This model was then subjected to molecular-dynamics refinement and grouped *B*-factor refinement using the program *CNS* (Brünger *et al.*, 1998) and the *R* factor improved to 0.46 with an $R_{\rm free}$ of 0.48. After two cycles of manual rebuilding and refinement, the *R* factor and $R_{\rm free}$ were 0.29 and 0.33, respectively. Further rebuilding and refinement is under way.

MR was supported by a studentship from the EPSRC, UK. The pGEX-hHR23A(1–82) plasmid was a gift from Dr Nobuyuki Nukina.

References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
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.
Chen, L. & Madura, K. (2006). *FEBS Lett.* **580**, 3401–3408.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dantuma, N. P., Heinen, C. & Hoogstraten, D. (2009). *DNA Repair*, **8**, 449–460.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Garman, E. F. & Mitchell, E. P. (1996). *J. Appl. Cryst.* **29**, 584–587.
Goh, A. M., Walters, K. J., Elsasser, S., Verma, R., Deshaies, R. J., Finley, D. & Howley, P. M. (2008). *BMC Biochem.* **9**, 4.
Grabbe, C. & Dikic, I. (2009). *Chem. Rev.* **109**, 1481–1494.
Kang, Y., Chen, X., Lary, J. W., Cole, J. L. & Walters, K. J. (2007). *J. Mol. Biol.* **369**, 168–176.
Kang, Y., Zhang, N., Koepp, D. M. & Walters, K. J. (2007). *J. Mol. Biol.* **365**, 1093–1101.
Lowe, E. D., Hasan, N., Trempe, J.-F., Fonso, L., Noble, M. E. M., Endicott, J. A., Johnson, L. N. & Brown, N. R. (2006). *Acta Cryst.* D**62**, 177–188.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mueller, T. D. & Feigon, J. (2003). *EMBO J.* **22**, 4634–4645.
Raasi, S., Orlov, I., Fleming, K. G. & Pickart, C. M. (2004). *J. Mol. Biol.* **341**, 1367–1379.
Walters, K. J., Lech, P. J., Goh, A. M., Wang, Q. & Howley, P. M. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12694–12699.
Wang, G., Sawai, N., Kotliarova, S., Kanazawa, I. & Nukina, N. (2000). *Hum. Mol. Genet.* **9**, 1795–1803.