crystallization communications

Crystallization and preliminary X-ray diffraction analysis of human endoplasmic reticulum aminopeptidase 2

David B. Ascher,[a]* Galina Polekhina[a,b] and Michael W. Parker[a,c]*

[a]Biota Structural Biology Laboratory, St Vincent's Institute of Medical Research, 9 Princes Street, Fitzroy, Victoria 3065, Australia, [b]Centre for Cancer Research, Monash Institute of Medical Research, Monash University, 27-31 Wright Street, Clayton, Victoria 3168, Australia, and [c]Bio21 Molecular Science and Biotechnology Institute and Department of Biochemistry and Molecular Biology, The University of Melbourne, 30 Flemington Road, Parkville, Victoria 3010, Australia

Correspondence e-mail: dascher@svi.edu.au, mparker@svi.edu.au

Endoplasmic reticulum aminopeptidase 2 (ERAP2) is a critical enzyme involved in the final processing of MHC class I antigens. Peptide trimming by ERAP2 and the other members of the oxytocinase subfamily is essential to customize longer precursor peptides in order to fit them to the correct length required for presentation on major histocompatibility complex class I molecules. While recent structures of ERAP1 have provided an understanding of the 'molecular-ruler' mechanism of substrate selection, little is known about the complementary activities of its homologue ERAP2 despite their sharing 49% sequence identity. In order to gain insights into the structure–function relationship of the oxytocinase subfamily, and in particular ERAP2, the luminal region of human ERAP2 has been crystallized in the presence of the inhibitor bestatin. The crystals belonged to an orthorhombic space group and diffracted anisotropically to 3.3 Å resolution in the best direction on an in-house X-ray source. A molecular-replacement solution suggested that the enzyme has adopted the closed state as has been observed in other inhibitor-bound aminopeptidase structures.

1. Introduction

Endoplasmic reticulum aminopeptidase 2 (ERAP2), a type 1 membrane protein, is a member of the oxytocinase subfamily of zinc-dependent M1 aminopeptidases (Tanioka *et al.*, 2003). ERAP2 shares approximately 40–50% identity with the other members of the family: endoplasmic reticulum aminopeptidase 1 (ERAP1) and insulin-regulated aminopeptidase (IRAP) (Tsujimoto & Hattori, 2005). All three members of the oxytocinase family have been shown to be essential for the correct processing of antigens for presentation on major histocompatibility complex (MHC) class I molecules (Saric *et al.*, 2002; Serwold *et al.*, 2002; York *et al.*, 2002; Saveanu *et al.*, 2005, 2009). The MHC class I system is responsible for the presentation of cytosolic peptides on the cell surface for recognition by T cells. This allows the recognition of cells containing foreign proteins by the immune system. Antigens are produced by proteosomal degradation and then require N-terminal trimming by the oxytocinase family to enable correct presentation (Roelse *et al.*, 1994; Serwold *et al.*, 2001).

Given their role in the recognition of abnormal cells, both ERAP1 and ERAP2 have been shown to be down-regulated in a number of tumours (Fruci *et al.*, 2006, 2008). They have also both been implicated in regulating angiogenesis and blood pressure through the trimming of angiotensins II and III, respectively (Hattori *et al.*, 2000; Yamamoto *et al.*, 2002; Tanioka *et al.*, 2003; Watanabe *et al.*, 2003). ERAP2 expression levels have also been linked to the development of pre-eclampsia, a potentially dangerous complication of pregnancy (Founds *et al.*, 2009; Johnson *et al.*, 2009; Hill *et al.*, 2011).

Recently, the crystal structure of ERAP1 has been solved in a number of different conformational states (Kochan *et al.*, 2011; Nguyen *et al.*, 2011). These studies, together with biochemical studies focused on IRAP (Ascher *et al.*, 2011), have shed light upon the domain architecture and the large conformational changes that these proteins undergo upon the binding of a substrate or inhibitor. The oxytocinase family conserved domain structure consists of, from the N-terminus to the C-terminus, a short cytoplasmic unstructured tail,

a single-pass transmembrane domain, an ~60 kDa catalytic domain composed of a β-sheet and then α-helical subdomains, a short β-sheet hinge domain and finally an ~40 kDa C-terminal α-helical domain. The structures also shed light upon a 'molecular-ruler' mechanism within ERAP1 that allows the correct processing of antigens to a certain length: ERAP1 has a preference for peptides of between nine and 16 residues in length (Chang *et al.*, 2005). However, little is known about the complementary partner ERAP2 as it has been less well studied, in part because of its absence in rodent models. Despite human ERAP1 and ERAP2 sharing 49% sequence identity, which is significantly higher in the catalytic domain, they show quite distinct substrate preferences and activities. ERAP2 has been shown to have an almost inverse substrate preference to that of ERAP1, preferring N-terminal Lys and Arg residues and possessing minimal affinity for ERAP1 substrates containing N-terminal Leu and Met residues (Hattori *et al.*, 2000; Tanioka *et al.*, 2003). It is not yet clear whether ERAP2 has any substrate-length preference. ERAP2 has been shown to co-locate with ERAP1 in the endoplasmic reticulum and to co-immunoprecipitate with ERAP1 during purification, which has led to the suggestion that they may form heterodimers, in contrast to the other members of the M1 family, which form homodimers (Saveanu *et al.*, 2005). Currently, there is no structural information to explain the molecular basis of the dimerization of any M1 aminopeptidase. All of these studies raise questions about how ERAP2 functions in concert with ERAP1 in the endoplasmic reticulum to process antigens to the correct length.

Given the pivotal role of ERAP2 in the correct processing of MHC class I antigens, we have initiated structural studies of ERAP2 and its domains in order to shed light on structure–function relationships and ultimately to determine the high-resolution atomic structure of ERAP2. Here, we report the crystallization and preliminary X-ray

analysis of the luminal region of ERAP2 bound to the inhibitor bestatin.

2. Experimental procedures and results

2.1. Expression and purification

The luminal region of human ERAP2 (UniProt sequence Q6P179) encompassing residues 50–960 was cloned using ligation-independent cloning into a pFastBac1 vector (Invitrogen) that had been modified to include an N-terminal honey bee melittin signal peptide followed by a maltose-binding protein (MBP), hexahistidine tag and TEV protease cleavage site. Expression in insect cells was chosen as ERAP2 was predicted to contain eight N-linked glycosylation sites. Recombinant bacmid was produced and identified following the manufacturer's instructions and recombinant virus was produced by transfection of Sf9 cells with Cellfectin (Invitrogen). Expression was performed in 12 l batches by infecting Sf21 cells at 2×10^6 cells ml^{-1} with recombinant virus at an MOI of 5.0. After 60 h the cells were pelleted by centrifugation at 1000g and the media were collected for purification.

The media were filtered, concentrated and buffer-exchanged into chilled 25 mM Tris pH 7.2, 500 mM NaCl, 100 mM maltose, 20 mM imidazole by tangental flow filtration using a Millipore ProFlux M12. The buffer-exchanged media were then clarified by centrifugation at 20 000g before being loaded onto a 5 ml HisTrap column (GE Healthcare) previously equilibrated with 25 mM Tris pH 7.2, 500 mM NaCl, 100 mM maltose, 20 mM imidazole. The column was washed with the same buffer and the protein of interest was eluted by a continuous increasing concentration of imidazole to 500 mM. All purification steps were performed at room temperature. The fractions containing MBP-6His-TEV-ERAP2$_{50–960}$ were pooled and dialyzed overnight against 25 mM Tris pH 7.2, 500 mM NaCl, 100 mM maltose, 20 mM imidazole at 277 K.

His-tagged TEV protease was added to the dialyzed sample at a 1:50–100 ratio to cleave the purification and solubility tags from ERAP2$_{50–960}$ and the mixture was incubated for 6–8 h at 295 K followed by a further 10–12 h at 277 K. A non-native serine-residue overhang remained following TEV protease cleavage. ERAP2$_{50–960}$ was separated from MBP-6His-TEV-ERAP2$_{50–960}$ and 6His-TEV by applying it onto a HisTrap column and collecting the flowthrough fractions. ERAP2$_{50–960}$ was then concentrated and purified further on a HiLoad Superdex 200 26/60 column (GE Healthcare) using a buffer consisting of 25 mM Tris pH 7.2, 150 mM NaCl. The fractions containing ERAP2$_{50–960}$ were pooled and concentrated using an Amicon Ultra-15 30K MWCO concentrator. NaCl was removed from the sample using a NAP-25 column (GE Healthcare). The purified protein was concentrated using an Amicon Ultra-4 10K MWCO concentrator to a final concentration of between 4 and 10 mg ml^{-1} for crystallization. The protein concentration was determined from the absorbance at 280 nm using an extinction coefficient calculated using the *ProtParam* tool (http://web.expasy.org/protparam/). While several species at approximately 110 kDa were observed by mass spectrometry, potentially owing to heterogenous glycosylation, the protein was judged to be sufficiently pure by SDS–PAGE (Fig. 1). The protein was stored at 277 K and crystallization screens were set up within a week of purification.

2.2. Protein crystallization

Initial crystallization screening was performed on ERAP2$_{50–960}$ with and without the aminopeptidase inhibitor bestatin at the Bio21 Collaborative Crystallization Centre (C³; http://www.csiro.au/c3).

Figure 1
4–12% SDS–PAGE under reducing conditions showing the purified ERAP2$_{50–960}$ running at approximately 95 kDa next to SeeBlue Plus 2 molecular-weight markers from Invitrogen (the orange marker is running at approximately 98 kDa).

crystallization communications

Over 1500 conditions were tried. A number of crystal forms grew in the presence of bestatin and in a range of PEG-based conditions. These conditions were optimized in-house, which included varying the precipitant concentration and the pH, trying PEGs of different molecular weights and types, varying the protein and inhibitor concentrations, the protein:reservoir ratio in the drops and crystallization temperature and using additive screening. The optimal crystallization conditions for ERAP2$_{50-960}$ in the presence of 5 mM bestatin were achieved using the hanging-drop vapour-diffusion method by mixing 2 µl 6 mg ml^{-1} protein solution in 25 mM Tris pH 7.2 with an equal volume of reservoir solution [0.5 M MES pH 7.0, 0.2 M KSCN, 0.02 M CaCl$_2$, 12%(w/v) PEG monomethylether 5000] and equilibrating against 1 ml reservoir solution. The crystals grew at 295 K and reached dimensions of 0.05 × 0.3 × 0.3 mm after several months (Fig. 2).

2.3. Data collection and preliminary X-ray analysis

Crystals were flash-cooled after stepwise transfer into the final cryobuffer. Ethylene glycol was used as a cryoprotectant and was added to the mother liquor in steps of 5%(v/v) up to 25%(v/v). The soaking time for each step was 2–3 min. A data set was collected on our in-house X-ray generator facility: a Rikagu MicroMax-007 HF with Cu $K\alpha$ X-rays and an R-AXIS IV^{++} imaging-plate detector. The crystal-to-detector distance was set to 240 mm and each image was exposed for 15 min with 0.5° oscillation (total rotation range 95°). However, the diffraction became very weak after about 50° rotation, possibly owing to radiation damage. The diffraction extended to approximately 3.3 Å resolution (Fig. 3), but was anisotropic along the c^* axis as the crystals were very thin plates. The crystals belonged to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 77.6, b = 191.3, c = 90.6 Å.

The data were processed using d^*TREK (Pflugrath, 1999) and imported into CCP4 (Winn $et\ al.$, 2011). Because of the anisotropic diffraction, many reflections were rejected in the 3.8–3.3 Å resolution shell, leading to low completeness values and generally poor signal to noise. Thus, for molecular replacement a low-resolution cutoff of 3.8 Å was chosen based on a completeness of 80%. The statistics for the X-ray data that were subsequently used in the molecular-replacement calculations are shown in Table 1. The calculated Matthews coefficient (V_M) of 3.37 Å^3 Da^{-1} suggested the presence of one molecule per asymmetric unit, corresponding to a solvent content of 60%, or (less likely) two molecules per asymmetric unit and a solvent content of 27% (Matthews, 1968). The Wilson B factor was

Table 1
Diffraction data statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12$
Unit-cell parameters (Å)	a = 77.6, b = 191.3, c = 90.6
Wavelength (Å)	1.54182
Resolution range (Å)	25–3.8 (3.94–3.80)
Total observations	29289
Unique reflections	11126 (1087)
Multiplicity	2.6 (2.6)
Completeness (%)	80.0 (84.5)
Mean $I/\sigma(I)$	4.1 (2.7)
R_{meas}†	0.168 (0.299)
$R_{p.i.m.}$†	0.075 (0.136)

† $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl} \sum_i I_i(hkl)$ and $R_{p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of a symmetry-related reflection with indices hkl.

calculated to be 35 Å^2 in $TRUNCATE$ (French & Wilson, 1978) and the mosaicity refined to 0.67°.

The crystal structures of several M1 aminopeptidases in both apo and inhibitor-bound forms were available for use in molecular replacement. Using the open apo structure of the homologue ERAP1 (PDB entry 3qnf; Kochan $et\ al.$, 2011) as a search model for molecular replacement did not yield a solution. However, using bestatin-bound closed-state ERAP1 crystal structures (PDB entries 2yd0 and 3mdj; M. Vollmar, G. Kochan, T. Krojer, E. Ugochukwu, J. R. C. Muniz, J. Raynor, A. Chaikuad, C. Allerston, F. Von Delft, C. Bountra, C. H. Arrowsmith, J. Weigelt, A. Edwards & S. Knapp, unpublished work; Nguyen $et\ al.$, 2011) as search models with the program $Phaser$ (McCoy $et\ al.$, 2007) yielded favourable statistics (RFZ and TFZ scores of >8). The main difference between the three ERAP1 crystal structures is the relative positioning of the C-terminal domain with respect to the catalytic domain. Following rigid-body refinement using $REFMAC$ [ten cycles in which the molecule in the asymmetric unit was refined as a rigid body and ten cycles in which the catalytic,

Figure 2
Crystals of ERAP2$_{50-960}$ grown by vapour diffusion at 295 K in a mother liquor consisting of 0.5 M MES pH 7.0, 0.2 M KSCN, 0.02 M CaCl$_2$, 12%(w/v) PEG monomethylether 5000. The crystal in the centre of the picture is approximately 0.03 × 0.1 × 0.1 mm in size.
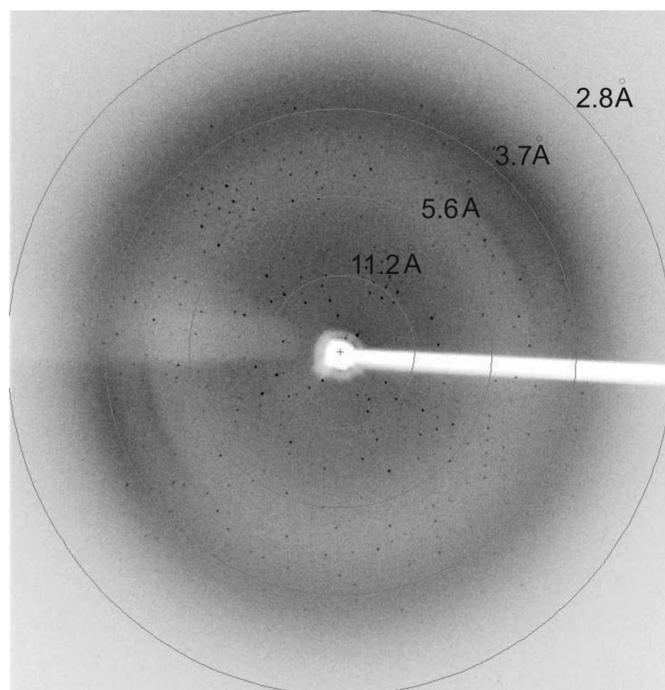
Figure 3
X-ray diffraction pattern recorded in-house from an ERAP2$_{50-960}$ crystal (0.5° oscillation) with visible diffraction to 3.3 Å resolution.

hinge and C-terminal domains (residues 50–550, 551–644 and 645–960, respectively) were treated separately; Murshudov *et al.*, 2011] followed by inspection in *Coot* (Emsley & Cowtan, 2004), the solution provided by 2yd0 showed the greatest improvement in statistics ($R = 0.43$ and $R_{free} = 0.47$, with the model reflecting the sequence of ERAP1), packed well in the unit cell and yielded sensible electron density for the C-terminal region. This suggests that we have crystallized the tightly closed inhibitor-bound form of human ERAP2. Higher resolution data at a synchrotron source will now be sought and the ERAP2 model will be refined to completion.

References

Ascher, D. B., Cromer, B. A., Morton, C. J., Volitakis, I., Cherny, R. A., Albiston, A. L., Chai, S. Y. & Parker, M. W. (2011). *Biochemistry*, **50**, 2611–2622.

Chang, S.-C., Momburg, F., Bhutani, N. & Goldberg, A. L. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 17107–17112.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Founds, S. A., Conley, Y. P., Lyons-Weiler, J. F., Jeyabalan, A., Hogge, W. A. & Conrad, K. P. (2009). *Placenta*, **30**, 15–24.

French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.

Fruci, D., Ferracuti, S., Limongi, M. Z., Cunsolo, V., Giorda, E., Fraioli, R., Sibilio, L., Carroll, O., Hattori, A., van Endert, P. M. & Giacomini, P. (2006). *J. Immunol.* **176**, 4869–4879.

Fruci, D., Giacomini, P., Nicotra, M. R., Forloni, M., Fraioli, R., Saveanu, L., van Endert, P. & Natali, P. G. (2008). *J. Cell. Physiol.* **216**, 742–749.

Hattori, A., Kitatani, K., Matsumoto, H., Miyazawa, S., Rogi, T., Tsuruoka, N., Mizutani, S., Natori, Y. & Tsujimoto, M. (2000). *J. Biochem.* **128**, 755–762.

Hill, L. D., Hilliard, D. D., York, T. P., Srinivas, S., Kusanovic, J. P., Gomez, R., Elovitz, M. A., Romero, R. & Strauss, J. F. III (2011). *BMC Med. Genet.* **12**, 64.

Johnson, M. P., Roten, L. T., Dyer, T. D., East, C. E., Forsmo, S., Blangero, J., Brennecke, S. P., Austgulen, R. & Moses, E. K. (2009). *Hum. Genet.* **126**, 655–666.

Kochan, G., Krojer, T., Harvey, D., Fischer, R., Chen, L., Vollmar, M., von Delft, F., Kavanagh, K. L., Brown, M. A., Bowness, P., Wordsworth, P., Kessler, B. M. & Oppermann, U. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 7745–7750.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 499–501.

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.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nguyen, T. T., Chang, S.-C., Evnouchidou, I., York, I. A., Zikos, C., Rock, K. L., Goldberg, A. L., Stratikos, E. & Stern, L. J. (2011). *Nature Struct. Mol. Biol.* **18**, 604–613.

Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.

Roelse, J., Grommé, M., Momburg, F., Hämmerling, G. & Neefjes, J. (1994). *J. Exp. Med.* **180**, 1591–1597.

Saric, T., Chang, S.-C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M. & Goldberg, A. L. (2002). *Nature Immunol.* **3**, 1169–1176.

Saveanu, L., Carroll, O., Lindo, V., Del Val, M., Lopez, D., Lepelletier, Y., Greer, F., Schomburg, L., Fruci, D., Niedermann, G. & van Endert, P. M. (2005). *Nature Immunol.* **6**, 689–697.

Saveanu, L., Carroll, O., Weimershaus, M., Guermonprez, P., Firat, E., Lindo, V., Greer, F., Davoust, J., Kratzer, R., Keller, S. R., Niedermann, G. & van Endert, P. (2009). *Science*, **325**, 213–217.

Serwold, T., Gaw, S. & Shastri, N. (2001). *Nature Immunol.* **2**, 644–651.

Serwold, T., Gonzalez, F., Kim, J., Jacob, R. & Shastri, N. (2002). *Nature (London)*, **419**, 480–483.

Tanioka, T., Hattori, A., Masuda, S., Nomura, Y., Nakayama, H., Mizutani, S. & Tsujimoto, M. (2003). *J. Biol. Chem.* **278**, 32275–32283.

Tsujimoto, M. & Hattori, A. (2005). *Biochim. Biophys. Acta*, **1751**, 9–18.

Watanabe, Y., Shibata, K., Kikkawa, F., Kajiyama, H., Ino, K., Hattori, A., Tsujimoto, M. & Mizutani, S. (2003). *Clin. Cancer Res.* **9**, 6497–6503.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yamamoto, N., Nakayama, J., Yamakawa-Kobayashi, K., Hamaguchi, H., Miyazaki, R. & Arinami, T. (2002). *Hum. Mutat.* **19**, 251–257.

York, I. A., Chang, S.-C., Saric, T., Keys, J. A., Favreau, J. M., Goldberg, A. L. & Rock, K. L. (2002). *Nature Immunol.* **3**, 1177–1184.