Crystallization and preliminary X-ray diffraction analysis of Xaa-Pro dipeptidase from *Xanthomonas campestris*

Ashwani Kumar,^a
Venkata Narayana Are,^a
Biplab Ghosh,^a Utsavi Agrawal,^a
Sahayog N. Jamdar,^{b*}
Ravindra D. Makde^{a*} and
Surinder M. Sharma^a

^aHigh Pressure and Synchrotron Radiation Physics Division, Bhabha Atomic Research Centre, Trombay, Mumbai, Maharashtra 400 085, India, and ^bFood Technology Division, Bhabha Atomic Research Centre, Trombay, Mumbai, Maharashtra 400 085, India

Correspondence e-mail: snjam@barc.gov.in, ravimakde@rrcat.gov.in

Received 26 June 2014
Accepted 28 July 2014

Xaa-Pro dipeptidase (XPD; prolidase; EC 3.4.13.9) specifically hydrolyzes dipeptides with a prolyl residue at the carboxy-terminus. *Xanthomonas* spp. possess two different isoforms of XPD (48 and 43 kDa) which share ~24% sequence identity. The XPD of 43 kDa in size (XPD43) from *Xanthomonas* spp. is unusual as it lacks the strictly conserved tyrosine residue (equivalent to Tyr387 in *Escherichia coli* aminopeptidase P) that is suggested to be important in the proton-shuttle transfer required for catalysis in the M24B (MEROPS) family. Here, the crystallization and preliminary X-ray analysis of XPD43 from *X. campestris* (GenBank accession No. NP_637763) are reported. Recombinant XPD43 was crystallized using the microbatch-under-oil technique. Diffraction data were collected on the recently commissioned protein crystallography beamline (PX-BL21) at the Indian synchrotron (Indus-2, 2.5 GeV) to 1.83 Å resolution with 100% completeness. The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 84.32$, $b = 105.51$, $c = 111.35$ Å. Two monomers are expected to be present in the asymmetric unit of the crystal, corresponding to a solvent content of 58%. Structural analysis of XPD43 will provide new insights into the role of the conserved residues in catalysis in the M24B family.

1. Introduction

Xaa-Pro dipeptidase (XPD; EC 3.4.13.9) specifically cleaves a *trans* Xaa-Pro peptide bond in a dipeptide with a prolyl residue at the carboxy-terminus (MEROPS database; Rawlings *et al.*, 2014). The enzyme is also referred to as proline dipeptidase, prolidase or peptidase-Q (PepQ). XPD belongs to the M24B family of metallo-enzymes, which also includes aminopeptidase P, which is capable of hydrolyzing a *trans* Xaa-Pro peptide bond at the N-terminus of a polypeptide. XPD is ubiquitous in nature and has been isolated from mammals, bacteria and archaea (Theriot *et al.*, 2009). In archaea and bacteria, although the function of XPD has not been well elucidated, it has been suggested to be involved in the recycling of proline (Ghosh *et al.*, 1998). On the other hand, in humans the role of XPD in different pathophysiological conditions has been well established. For instance, the enzyme has been shown to be involved in the final stages of the degradation of dietary as well as endogenous proteins, particularly collagen. Moreover, mutations in the XPD gene in humans causes a rare autosomal recessive disorder, prolidase deficiency (PD), which exhibits phenotypes such as skin ulceration, mental retardation and recurrent infections (Endo *et al.*, 1982; Forlino *et al.*, 2002). The enzyme is also important commercially in the food and dairy industries for improving the flavour and texture of food (Theriot *et al.*, 2009). For example, during cheese ripening the action of this enzyme on proline dipeptides leads to a reduction in the bitterness of the product. In addition to the peptidase activity, XPD and the other peptidases of the M24B family display fortuitous activity against toxic organophosphorus (OP) compounds, which include pesticides and nerve agents, by cleaving P–F and P–O bonds (Cheng *et al.*, 1993). Thus, it has huge potential applications in detoxifying OP nerve agents and in biosensors for the detection of OP compounds.

© 2014 International Union of Crystallography
All rights reserved

Table 1
Macromolecule-production information.

Source organism	<i>X. campestris</i>
DNA source	Genomic DNA
Forward primer†	CGTGGATCCAGACGCGAGATCGGCGGATG
Reverse primer‡	CGTAAGCTTATGCAACGGTTGATCGATCGCCAC
Cloning vector	None
Expression vector	pST50Tr
Expression host	<i>E. coli</i> BL21(DE3)pLysS
Complete amino-acid sequence of the construct produced	MASWSHPQFEKGSSHHHHSSGGGGGGENLYFQGSSTQIG-GMSLDQARTQLAPWTQRAAPIGADEYQQRIERARVLMRAQ-GVDALLIGAGTSLRYFSGVPWGASERLVALLLTTEGDPVL-ICPAFEESGLDAVLQLPVRKRLWEEHEDPYALVVQAMDEQ-HAHALALDPGIAFAVHTGLRAHLGTAIRDAGAIIDGCRMC-KSPAELALMQACDMLLVQLRAAGIAHEGIGTDQLVRFI-DEAHRALGADNGSTFCIVQFGHATAFPHGIPGVQHLRAGE-LVLIDTGTCTVQGYHSDITRTWIYGTSPDAQRIWELELAA-QAAFAAVRPGVACEAVDQAARAVLQAAGLPGDYRLPLGP-HRTGHGCGLAIEAPYLVGRNRPQLPQGMCSANPEMIVVP-GAFGVRLEDHFYVTDGTAQWFTTPSVAIDQPFPA

† The *Bam*HI site is underlined. ‡ The *Hind*III site is underlined.

Genome analysis of *X. campestris* showed the presence of three different genes encoding peptidases of the M24B family; two of them are annotated as XPD, while the third is annotated as aminopeptidase P. The sizes of the polypeptides encoded by the two XPD genes are 441 (48.25 kDa; XPD48; GenBank accession No. NP_638603) and 399 (42.80 kDa; XPD43; GenBank accession No. NP_637763) residues, and the proteins share ~24% sequence identity. It was further shown that individual orthologues of all three genes are also present in other *Xanthomonas* spp. as well as in some families of gamma-proteobacteria, which include many pathogenic bacteria. In the case of XPD48 orthologues, detailed structural and functional studies have only been reported for the enzyme from the gammaproteobacterium *Alteromonas* sp. strain JD6.5 (PDB entry 3l24; Vyas *et al.*, 2010), while the orthologues from *Escherichia coli* and *X. maltophilia* have also been functionally characterized (Park *et al.*, 2004; Suga *et al.*, 1995). Structural and functional characterization of XPD43 orthologues has only been reported for the enzymes from the archaea *Pyrococcus furiosus* and *P. horikoshii* (PDB entries 1pv9 and 2how; Maher *et al.*, 2004; Jeyakanthan *et al.*, 2009). Although the structures of XPD43 orthologues from the eubacteria *Mycobacterium ulcerans*, *Bacillus anthracis* and *Thermotoga maritima* (PDB entries 4ege, 3q6d and 2zsg) have been deposited in the PDB by structural genomics consortia, none of these enzymes have been functionally characterized so far. Moreover, detailed information on both the structural and functional aspects of XPD43 proteins from the entire proteobacteria phylum is lacking. The closest sequence homologues of the XPD48 and XPD43 proteins of *X. campestris* in the PDB are the XPDs from *Alteromonas* sp. strain JD6.5 (PDB entry 3l24; 47% identity) and *Thermococcus sibiricus* (PDB entry 4fk; 29% identity; Trofimov *et al.*, 2012), respectively.

A bioinformatics analysis showed that XPD43 from *Xanthomonas* spp. is quite unique among the M24B family of peptidases owing to a change of the strictly conserved glycine and tyrosine (Gly385 and Tyr387; numbering according to *E. coli* aminopeptidase P) to methionine and valine, respectively (Fig. 1). In aminopeptidase P from *E. coli* Tyr387 has been proposed to be involved in the conserved hydrogen-bond network in the Asp260–Arg404–Tyr387 motif, which is able to shuttle a proton from the bulk solvent to the leaving peptide (Jao *et al.*, 2006). Moreover, it was also suggested that Tyr387, Arg404 and His350 are important residues for proline specificity (Wilce *et al.*, 1998; Graham *et al.*, 2006). Site-directed mutagenesis of Tyr387 and Arg404 resulted in a severalfold reduction in the activity of aminopeptidase P from *E. coli* (Jao *et al.*, 2006). From this viewpoint, structural and functional characterization of XPD43 from *X. campestris* is necessary. Structural analysis of XPD43 will provide new insights into the role of the conserved residues in catalysis. Further structure–function analysis may shed light on the significance of the two isoforms of XPD in *Xanthomonas* spp. and other members of the gammaproteobacteria.

2. Materials and methods

2.1. Macromolecule production

The coding DNA sequence (CDS) region of the gene encoding the XPD43 protein (GenBank accession No. NP_637763) was from the genomic DNA of *X. campestris* ATCC 33913 (NCBI taxonomy ID 190485). This CDS region was PCR-amplified by thermostable *Pfu* DNA polymerase and was cloned into a T7-promoter-based expression plasmid (Tan *et al.*, 2005) to form an in-frame translational fusion protein with a Strep-tag II–hexahistidine–*Tobacco etch virus* protease site (STRHISTEV) tag (Table 1). The construct was verified by restriction-digestion analysis and DNA sequencing using T7 and T7-terminator oligos. The clone expressing the gene was grown in 3 l 2×TY broth at 310 K to an OD₆₀₀ of ~0.4; the culture was then shifted to a lower temperature (291 K) for further growth. Subsequently, it was induced at an OD₆₀₀ of ~0.8 for 18 h by the addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was harvested at 7000 rev min^{−1} (Sorvall GSA rotor) at ambient temperature. The cell pellet was suspended in P300 buffer (50 mM sodium phosphate pH 7.0, 300 mM NaCl) and the desired protein was purified from the soluble fraction of the cell lysate using Ni metal-ion affinity and gel-filtration chromatography techniques. Typically, the soluble lysate containing the desired protein (STRHISTEV–XPD43) was incubated with 3 ml pre-equilibrated (P300 containing 20 mM imidazole) Ni Sepharose resin (GE Healthcare) and the bound protein was eluted with P300 buffer containing 500 mM imidazole by the batch method at 277 K. The eluted protein STRHISTEV–XPD43

Figure 1
Multiple sequence alignment of peptidases from the M24B (MEROPS) family (only relevant blocks of sequence are shown). UniProt accession numbers are given. The conserved residues important in the hydrogen-bonding network and in proline specificity are marked in boxes. The substitutions in XPD43 (UniProt accession No. Q8P839) are marked by arrows. The numbering at the top is for *E. coli* aminopeptidase P (UniProt accession No. P15034).

Table 2
Crystallization.

Method	Microbatch-under-oil
Plate type	96-well U-bottom plate
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	~12
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0, 200 mM NaCl
Composition of reservoir solution	40 mM KH ₂ PO ₄ , 15% glycerol, 12% (w/v) polyethylene glycol (PEG) 8000
Volume and ratio of drop	4 µl (1:1)
Volume of reservoir	0 (microbatch)

was subsequently subjected to digestion by *Tobacco etch virus* protease (TEV; 1:50 molar ratio) at 293 K for 24 h while dialyzing against T200 buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl). The desired protein (XPD43) was separated from the STRHISTEV peptide and undigested protein (STRHISTEV-XPD43) using a 5 ml HisTrap column on an ÄKTA system (GE Healthcare) at room temperature. The protein was concentrated to ~15 mg ml⁻¹ using a 10 kDa cutoff ultrafiltration unit (Sartorius). Glycerol was added to 20% (v/v) to the protein solution, which was then stored at 203 K until further use.

2.2. Crystallization

For crystallization, the required amount of XPD43 protein was retrieved from the freezer at 203 K and subjected to size-exclusion column chromatography. A Superdex 200 10/300 GL (GE Healthcare) column was used with T200 buffer to remove impurities and heterogeneity and also for buffer equilibration (Fig. 2). The eluted protein was concentrated to ~12 mg ml⁻¹ and was used for setting up crystallization at 294 K employing the microbatch-under-oil method (Chayen *et al.*, 1992). Typically, 2 µl protein solution was mixed with 2 µl crystallization solution in a 96-well U-bottom plate and was overlaid with 50 µl Al's oil (Hampton Research). Initial crystallization screening was performed using the Index (Hampton Research) and JCSG-*plus* (Molecular Dimensions) crystal screens. Crystal hits were observed in several conditions of the JCSG-*plus* and Index screens, but all of them were restricted to pH ≤ 5. One of the

Figure 2
Size-exclusion chromatography profile of purified XPD43 (marked by an arrow) on pre-calibrated Superdex 200 using an ÄKTA system (GE Healthcare) with 20 mM Tris-HCl pH 8.0, 200 mM NaCl. Inset, a 15% SDS-PAGE showing molecular-mass markers (lane 1; labelled in kDa) and purified XPD43 (lane 2).

Table 3
Data-collection and processing statistics.

Values in parentheses are for the outer shell.	
Diffraction source	PX-BL21, Indus-2, India
Wavelength (Å)	0.97947
Temperature (K)	100
Detector	MAR225 CCD
Resolution (Å)	48–1.83 (1.86–1.83)
Space group	P2 ₁ 2 ₁ 2 ₁
Unit-cell parameters (Å, °)	a = 84.32, b = 105.51, c = 111.35, α = β = γ = 90
Total No. of reflections	617282 (19163)
No. of unique reflections	88109 (4412)
Multiplicity	7.0 (4.3)
⟨I/σ(I)⟩	18.2 (2.1)
Completeness (%)	99.9 (98.8)
R _{meas} †	8 (72)
Overall B factor from Wilson plot (Å ²)	16

† $R_{meas} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

crystallization conditions was subsequently optimized to produce crystals suitable for diffraction studies (Table 2).

2.3. Data collection and processing

The crystals were directly cooled in liquid nitrogen without any additional cryoprotectant. Single-crystal diffraction experiments were carried out on the protein crystals on the recently commissioned bending-magnet protein crystallography beamline (PX-BL21) at the 2.5 GeV Indus-2 synchrotron (India). Diffraction images were collected from a cryocooled crystal (100 K) with an oscillation of 1° at a wavelength of 0.97947 Å using a MAR225 CCD (Rayonix) detector. The data were collected to 1.83 Å resolution with a completeness of 100%. The data were indexed and integrated using XDS (Kabsch,

Figure 3
Diffraction pattern recorded on PX-BL21 at Indus-2 from a XPD43 crystal (inset) with an exposure time of 30 s and 1° oscillation. The crystal was grown in 45 d at 294 K from a microbatch drop made by mixing 2 µl each of crystallization solution (40 mM KH₂PO₄, 15% glycerol, 12% PEG 8000) and protein solution (12 mg ml⁻¹ XPD43 in 20 mM Tris-HCl pH 8.0, 200 mM NaCl).

2010) and subsequently scaled using *AIMLESS* (CCP4; Winn *et al.*, 2011). The data statistics are summarized in Table 3.

3. Results and discussion

The typical yield of recombinant XPD43 was found to be ~90 mg from 3 l bacterial culture. The purified protein showed a single symmetrical peak at a molecular weight corresponding to ~70 kDa on size-exclusion chromatography (Fig. 2), while the expected monomeric mass of the molecule deduced from the sequence is 42.8 kDa. This suggests that XPD43 exists as a dimer under the given conditions. The protein peak eluted somewhat later than the expected elution time, possibly owing to a more compact and globular nature of the dimer or owing to weak interaction of the protein with the Superdex 200 column. The recombinant protein was active and showed peptidase activity on several Xaa-Pro dipeptides (unpublished results). The initial crystallization screening produced crystal hits in more than ten different crystallization conditions containing different salts and precipitants at pH $\leq$ 5.5. However, the best diffraction-quality crystals were grown from the crystallization condition 40 mM KH_2PO_4 , 15% glycerol, 12% (*w/v*) polyethylene glycol (PEG) 8000. These crystals were rod-shaped and typically grew to dimensions of about $0.1 \times 0.1 \times 0.5$ mm in 45 d (Fig. 3, inset). One of the crystals diffracted to 1.83 Å resolution (Fig. 3) on the recently installed PX-BL21 beamline at the Indus-2 synchrotron. It belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 84.32$, $b = 105.51$, $c = 111.35$ Å (Table 3). The Matthews coefficient (V_M ; Matthews, 1968) was calculated to be $2.9 \text{ Å}^3 \text{ Da}^{-1}$, with a corresponding solvent content of 58%. Two monomers were expected to be present in the asymmetric unit of the cell. A model based on the monomer of PDB entry 4fkf was used for molecular replacement using *Phaser* (McCoy *et al.*, 2007) in the CCP4 suite (Winn *et al.*, 2011). A convincing molecular-replacement solution with a final log-likelihood gain (LLG) score of 191 and with two monomers in the asymmetric unit was obtained. Initial structure refinement carried out using *PHENIX* (Adams *et al.*, 2010) resulted in a model with an R_{work} and R_{free} of 26 and 30%, respectively. Further model building and structure refinement are in progress. Detailed structural analysis of XPD43 will provide new insights into the role of the conserved residues in catalysis in the M24B family of peptidases.

We sincerely thank Drs P. D. Gupta, S. K. Deb and G. S. Lodha for providing the necessary support and infrastructure at Raja Ramanna Centre for Advanced Technology to carry out this research. We thank the staff at Indus-2 for providing the synchrotron beam for the present work. RDM and VNA are grateful to Dr Deepak Bhatnagar and Swapan K. Bhattacharjee for their support during this work. We thank Dr Venuka Goyal for editorial suggestions.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
 Chayen, N. E., Shaw Stewart, P. D. & Blow, D. M. (1992). *J. Cryst. Growth*, **122**, 176–180.
 Cheng, T.-C., Harvey, S. P. & Stroup, A. N. (1993). *Appl. Environ. Microbiol.* **59**, 3138–3140.
 Endo, F., Matsuda, I., Ogata, A. & Tanaka, S. (1982). *Pediatr. Res.* **16**, 227–231.
 Forlino, A., Lupi, A., Vaghi, P., Icaro Cornaglia, A., Calligaro, A., Campari, E. & Cetta, G. (2002). *Hum. Genet.* **111**, 314–322.
 Ghosh, M., Grunden, A. M., Dunn, D. M., Weiss, R. & Adams, M. W. W. (1998). *J. Bacteriol.* **180**, 4781–4789.
 Graham, S. C., Lilley, P. E., Lee, M., Schaeffer, P. M., Kralicek, A. V., Dixon, N. E. & Guss, J. M. (2006). *Biochemistry*, **45**, 964–975.
 Jao, S.-C., Huang, L.-F., Hwang, S.-M. & Li, W.-S. (2006). *Biochemistry*, **45**, 1547–1553.
 Jeyakanthan, J., Takada, K., Sawano, M., Ogasahara, K., Mizutani, H., Kunishima, N., Yokoyama, S. & Yutani, K. (2009). *J. Biophys.* **2009**, 434038.
 Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
 Maher, M. J., Ghosh, M., Grunden, A. M., Menon, A. L., Adams, M. W. W., Freeman, H. C. & Guss, J. M. (2004). *Biochemistry*, **43**, 2771–2783.
 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.
 Park, M.-S., Hill, C. M., Li, Y., Hardy, R. K., Khanna, H., Khang, Y.-H. & Rauschel, F. M. (2004). *Arch. Biochem. Biophys.* **429**, 224–230.
 Rawlings, N. D., Waller, M., Barrett, A. J. & Bateman, A. (2014). *Nucleic Acids Res.* **42**, D503–D509.
 Suga, K., Kabashima, T., Ito, K., Tsuru, D., Okamura, H., Kataoka, J. & Yoshimoto, T. (1995). *Biosci. Biotechnol. Biochem.* **59**, 2087–2090.
 Tan, S., Kern, R. C. & Selleck, W. (2005). *Protein Expr. Purif.* **40**, 385–395.
 Theriot, C. M., Tove, S. R. & Grunden, A. M. (2009). *Adv. Appl. Microbiol.* **68**, 99–132.
 Trofimov, A. A., Slutskaya, E. A., Polyakov, K. M., Dorovatovskii, P. V., Gumerov, V. M. & Popov, V. O. (2012). *Acta Cryst.* **F68**, 1275–1278.
 Vyas, N. K., Nickitenko, A., Rastogi, V. K., Shah, S. S. & Quiocho, F. A. (2010). *Biochemistry*, **49**, 547–559.
 Wilce, M. C., Bond, C. S., Dixon, N. E., Freeman, H. C., Guss, J. M., Lilley, P. E. & Wilce, J. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 3472–3477.
 Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.