The human aminopeptidase XPNPEP3 is associated with cystic kidney disease and TNF-TNFR2 cellular signaling. Its yeast and plant homolog Icp55 processes several imported mitochondrial matrix proteins leading to their stabilization. However, the molecular basis for the diverse roles of these enzymes in the cell is unknown. Here, we report the crystal structure of human XPNPEP3 with bound apstatin product at 1.65 Å resolution, and we compare its in vitro substrate specificity with those of fungal Icp55 enzymes. In contrast to the suggestions by earlier in vivo studies of mitochondrial processing, we found that these enzymes are genuine Xaa-Pro aminopeptidases, which hydrolyze peptides with proline at the second position (P1'). The mitochondrial processing activity involving cleavage of peptides lacking P1' proline was also detected in the purified enzymes. A wide proline pocket as well as molecular complementarity and capping at the S1 substrate site of XPNPEP3 provide the necessary structural features for processing the mitochondrial substrates. However, this activity was found to be significantly lower as compared with Xaa-Pro aminopeptidase activity. Because of similar activity profiles of Icp55 and XPNPEP3, we propose that XPNPEP3 plays the same mitochondrial role in humans as Icp55 does in yeast. Both Xaa-Pro aminopeptidase and mitochondrial processing activities of XPNPEP3 have implications toward mitochondrial fitness and cystic kidney disease. Furthermore, the presence of both these activities in Icp55 elucidates the unexplained processing of the mitochondrial cysteine desulfurase Nfs1 in yeast. The enzymatic and structural analyses reported here provide a valuable molecular framework for understanding the diverse cellular roles of XPNPEP3.

Human Xaa-Pro aminopeptidase 3 (XPNPEP3)3 (Xaa denotes any amino acid) and its yeast and plant homologs, intermediate-cleavage peptidase (Icp55), are metallopeptidases (EC 3.4.11.9) of the M24B (MEROPS) subfamily (1). Enzymes of this subfamily specifically remove the N-terminal amino acid from oligopeptides or polypeptides with Pro residue at the second position (P1') (supplemental Fig. S1). These proteins are strictly conserved in eukaryotes and are associated with several important biological processes (2–5). Homozygous mutations in gene encoding human XPNPEP3 protein have been implicated in the development of “nephronophthisis-like ciliopathy,” a cystic kidney disease (2). It has been proposed that XPNPEP3 enzyme processes several ciliary cystogenic proteins in cytoplasm into their functional form by removing the N-terminal residue adjacent to proline (Xaa-Pro aminopeptidase activity). Attenuated functions of few of these putative substrates of XPNPEP3 were earlier recognized to be the cause of cystic kidney disease. Moreover, XPNPEP3 is also reported to act as an adaptor molecule in the TNF-TNF2 signaling pathway (3). Upon stimulation by TNF, XPNPEP3 is recruited at the cytosolic domain of TNF2 to regulate phosphorylation of JNK kinases and exert anti-apoptotic functions (3). Human XPNPEP3 has two isoforms resulting from alternative splicing, one located in the mitochondrial matrix (XPNPEP3) and the other in the cytosol (XPNPEP3c) (6). It has been reported that an N-terminal region of XPNPEP3 is essential for its binding to the TNF-TNFR2 complex. Because of the absence of this region in XPNPEP3c, it lacks the ability to bind the complex. Intriguingly, this function of XPNPEP3 is mediated only by its structural fold, independent of its enzymatic activity (3).

In Icp55, XPNPEP3 homologs from yeast and plant have been shown to process a subset of nuclearly encoded polypeptides during import into the mitochondrial matrix. Icp55 acts on the N terminus of proteins subsequent to removal of the signal peptide by mitochondrial processing peptidase (MPP). In

The authors declare that they have no conflicts of interest with the contents of this article.

This article contains supplemental Fig. S1–S5.

The atomic coordinates and structure factors (code SX49) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 To whom correspondence may be addressed: Food Technology Division, Bhabha Atomic Research Centre, Mumbai 400085, India. Tel.: 91-22-25595375; E-mail: snjam@barc.gov.in.

2 To whom correspondence may be addressed: High Pressure and Synchrotron Radiation Physics Division, Bhabha Atomic Research Centre, Mumbai 400085, India. Tel.: 91-731-2442175; Fax: 91-731-2442140; E-mail: ravimakde@rrcat.gov.in.

3 The abbreviations used are: XPNPEP3, Xaa-Pro aminopeptidase 3; Icp55ag, intermediate-cleavage peptidase from A. gossypii; Icp55fg, intermediate-cleavage peptidase from F. graminearum; AMPPxc, aminopeptidase P from X. campestris, OPA, o-phthalaldehyde; MPP, mitochondrial processing peptidase; TEV, tobacco etch virus.
vivo substrate specificity suggests that Icp55 specifically cleaves off a destabilizing amino acid (Tyr, Phe, or Leu), leaving a stabilizing amino acid (Ser, Ala, or Thr) at the modified N terminus (4, 5, 7). This processing of mitochondrial proteins by Icp55 in yeast improves the life span of the proteins. A similar study on a plant, Arabidopsis thaliana, showed that absence of Icp55 processing compromises half-lives of not only Icp55 substrates but of the whole mitochondrial proteome.

Despite their notable biological significance, neither XPNPEP3 nor Icp55 enzyme has ever been purified and enzymatically characterized. Incidentally, humans have two other isoforms, XPNPEP1 (70 kDa), a cysteolytic enzyme, and XPNPEP2 (76 kDa), a glycosylphosphatidylinositol-linked glycoprotein located on the non-cytoplasmic face of plasma membrane. Both XPNPEP1 and XPNPEP2 have been identified as key activators of bradykinin peptide, a potent vasodilator and cardio-protective hormone (8–11). The crystal structure of XPNPEP3 has been reported earlier (9). Because XPNPEP2 has 42% sequence identity with XPNPEP1, it has been suggested that its structure would be similar. However, XPNPEP3 is an evolutionarily distinct enzyme as it shares only ~22% sequence identity with these isoforms and has a comparatively smaller polypeptide size (57 kDa). Interestingly, XPNPEP3 shares 33% sequence identity with a well characterized aminopeptidase P from Escherichia coli (AMPPec) (supplemental Fig. S2).

The biological role of XPNPEP3 was proposed on the basis of the Xaa-Pro aminopeptidase activity inferred from its classification as a member of the M24B subfamily (1). However, its homologs (Icp55) process peptides other than those typical for Xaa-Pro aminopeptidases during mitochondrial processing shown in vivo in yeast and plant. Thus, in the absence of in vitro enzyme activity data, it is difficult to infer the true substrate specificity for these enzymes. The earlier studies were also unable to explain the mitochondrial processing of cysteine desulfurase Nfs1, where Icp55 is reported to remove three amino acid) of the inhibitor apstatin. This analysis provides the molecular basis and structural framework for rationalizing the cellular roles of XPNPEP3 and Icp55 proteins.

**Results**

**Crystal structure of human XPNPEP3**

The structure of human XPNPEP3 was solved at 1.65 Å resolution with free R-factor (R<sub>free</sub>) of 17.75% (Table 1 and Fig. 1). An asymmetric unit consists of two protein chains, each composed of residues 54–507. The N-terminal 54-HPH<sup>506</sup>, C-terminal 504-SQAS<sup>507</sup>, and internal 87-AQGQSG<sup>92</sup> residues were not included in either chain of the final model due to poor electron densities. The two subunits in an asymmetric unit are related by 2-fold non-crystallographic symmetry (Fig. 1C). Each subunit shows a typical structural fold of aminopeptidase P (PDB code 5X49). It consists of the following two domains: N-terminal (N-domain, residues 54–245) and C-terminal (C-domain, residues 246–507). The N-domain is composed of a central β-sheet formed from five β-strands surrounded by seven α-helices. The C-domain adopts a well known pita-bread fold of α-β type.

The active site with a binuclear metal center is located in an oval depression at the center of the C-domain (Fig. 1D). The identity and positions of manganese ions have been determined by difference Fourier anomalous map computed from data collected at the manganese K-edge (6.53 keV/1.8961 Å) (Fig. 1B).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Data collection, refinement, and model statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P 2&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>Unit cell parameters (Å, °)</td>
<td>a = 62.90, b = 135.10, c = 67.21, β = 99.87</td>
</tr>
<tr>
<td>Resolution limit (Å)</td>
<td>47.68–1.65 (1.68–1.65)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>No. of observed reflections</td>
<td>489,774 (24,401)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.7 (3.7)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.8 (99.8)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.2 (63.6)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>12.9 (22.3)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>R&lt;sub&gt;f&lt;/sub&gt; (</td>
<td>Fo−Fc</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>17.75</td>
</tr>
<tr>
<td>R&lt;sub&gt;sim&lt;/sub&gt; (%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.7 (38.3)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>CC&lt;sub&gt;1/2&lt;/sub&gt; (%)</td>
<td>99.8 (74.2)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of atoms for asymmetric unit</td>
<td>6714</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>673</td>
</tr>
<tr>
<td>Bond length r.m.s.d. (Å)</td>
<td>0.027</td>
</tr>
<tr>
<td>Bond angle r.m.s.d. (°)</td>
<td>2.408</td>
</tr>
<tr>
<td>Average B-factor (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>24.30</td>
</tr>
<tr>
<td>All atoms</td>
<td>22.83</td>
</tr>
<tr>
<td>Protein</td>
<td>14.12</td>
</tr>
<tr>
<td>Manganese</td>
<td>11.73</td>
</tr>
<tr>
<td>Water</td>
<td>34.58</td>
</tr>
<tr>
<td>Other molecules</td>
<td>35.20</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
</tr>
<tr>
<td>Favored</td>
<td>97.64%</td>
</tr>
<tr>
<td>Allowed</td>
<td>2.36%</td>
</tr>
<tr>
<td>Outliers</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

| PDB code | 5X49 |

<sup>a</sup> Values in parentheses are for the highest resolution shell.

<sup>b</sup> R<sub>merge</sub> = ΣₙΣₗΣᵢΣᵢ|Iₗ(hkl) − Σᵢ|Iₗ,hkl|/ΣₙΣₗΣᵢ|Iₗ(hkl)|, where |Iₗ,hkl| and |Iₗ,khl| represent the diffraction intensity values of the individual measurement and the corresponding mean value, respectively.

<sup>c</sup> R<sub>free</sub> was calculated using 5% of the data that were excluded from refinement.
water (observed in AMPpec structure (14, 15)) at 2.1- and 2.2-Å distances from MnA and MnB, respectively. The O₁ of Glu-451 forms a hydrogen bond with this oxygen of O₁B. The amino group of O₁B is coordinated to MnB at a distance of 2.3 Å and forms a cation–π interaction with the conserved Tyr-300. One of the oxygens of the free carboxylate of O₁B is bound to N⁶¹ of His-431, whereas the other oxygen of the carboxylate binds to water. The phenyl ring of O₁B is bound into a hydrophobic pocket lined by Tyr-300, Pro-301, Val-316, His-314, and Val-303 of the C-domain, as well as Ile-111 and Pro-112 residues of a loop protruding from the other subunit (Fig. 1B). The other hydrolyzed product of apstatin, Pro-Pro-Ala peptide, is not traceable in the crystal structure. Two molecules of dimethyl sulfoxide (DMSO) are refined well into the void created due to absence of this product (Fig. 1D).

XPNPEP3 and Icp55 proteins exist as dimer in solution as shown by gel filtration elution profiles on Superdex200 (10/300GL) columns (supplemental Fig. S3). A dimer observed in an asymmetric unit of the XPNPEP3 structure seems to be equivalent to the biological dimer (Fig. 1C). This view is corrob-
Structural and enzymatic characterization of XPNPEP3 and Icp55

Structural comparison with bacterial AMPPec

Dali (17) search revealed that AMPPec (PDB code 1WL9 (15)) is the closest structural homolog of XPNPEP3. Their crystal structures superpose very well with each other with an overall r.m.s.d. of 1.58 Å between 440 aligned Ca atoms (Fig. 3, A and B). Whereas C-domains show r.m.s.d. of 1.14 Å, the N-domains deviate more with r.m.s.d. of 1.94 Å. The highest deviation is seen in the region 216YMQ...NKV233 of human XPNPEP3 (Fig. 4A). There is a deletion of five residues with respect to AMPPec, which results in a kink in the helix and shortening of the neighboring loop. The corresponding region in AMPPec is involved in further oligomerization of dimers to tetramer (Fig. 4A). The structural differences explain why the human protein does not form the higher order oligomer equivalent to that of AMPPec.

Key structural differences are observed in dimeric organization (Fig. 3B). The loop containing Ile-111 and Pro-112 residues in XPNPEP3, which contributes toward the active site of the other monomer in a dimer, is found to have a shifted location in AMPPec (Fig. 3, D and F). Due to this shift, we call it the sliding loop (residues 102–116). Sliding loops from two subunits of a dimer interact extensively with each other in AMPPec (Fig. 3F). However, these loops do not interact in XPNPEP3 and are 4–5 Å apart with their tips capping the S1 substrate site (Fig. 3, C and E). This deviation is attributed to sliding of monomers in a dimer and also to the subtle rearrangement of two domains in XPNPEP3 monomer. The surface on each monomer buried upon dimerization in AMPPec is 2064 Å², which is significantly more than that of XPNPEP3. This is primarily due to interaction of sliding loops in AMPPec. The proline pocket at S1’ site in XPNPEP3 is wider than that of AMPPec (Fig. 3, C and D). This is mainly due to the 1.5–2-Å shift in position of a loop (312TLHY315). The S2’ site in XPNPEP3 is more pronounced than that of AMPPec (Fig. 3, C and D). This pocket is formed by the side chains of Arg-438 and His-431 residues. These distinguishing features of S1’ pocket size and sliding loop position are invariant in other bacterial AMP structures (PDB codes 4PV4 and 3IG4). Furthermore, surface electrostatic potentials of human XPNPEP3 are starkly distinct from those of the bacterial enzyme AMPPec (Fig. 4, B and C). Although the bacterial enzyme has uniform negative surface potentials all over the protein, human XPNPEP3 has patches of positive potentials, mainly located on the N-domain.

In vitro enzymatic activities

We carried out enzymatic assays on the purified recombinant human XPNPEP3 and two putative Icp55 proteins from fungi, A. gossypii and F. graminearum (supplemental Fig. S2). We also carried out enzymatic assays on its bacterial homolog from X. campestris. Based on M24B classification of these enzymes, we examined the peptidase activity for several Xaa↓Pro peptides (Xaa↓Pro dipetides and Xaa↓Pro-Ala tripeptides). Because isozymes of XPNPEP3 (XPNPEP1 and XPNPEP2) are known to process the bradykinin peptide (Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg), substrates based on this peptide sequence were also tested (Fig. 5). We assayed the enzymes for non-proline peptides based on earlier in vivo sub-
strate specificity of Icp55 enzymes (mitochondrial processing activity) (4, 7). In these peptides canonical proline at the P1 position is substituted with Ser, Ala, or Gly (Fig. 6B). Overall, we found that their Xaa-Pro peptidase (Xaa-Pro aminopeptidase and prolidase) activity is much higher than that for non-proline peptides. Comparison of specificity constants (kcat/Km) for different substrates also clearly indicates that the peptides with proline at the P1' position are much preferred substrates over non-proline peptides (Fig. 6A).

Icp55 and XPNPEP3 enzymes showed significant Xaa-Pro dipeptidase (prolidase, Xaa↓Pro) and Xaa-Pro aminopeptidase (Xaa↓Pro-Xaa(n)) activities similar to their bacterial counterpart, AMPPxC (Fig. 5, A and B). The aminopeptidase activity was severalfold higher than the prolidase activity. The specific activities of XPNPEP3 and bacterial enzyme for Xaa↓Pro peptides were significantly higher compared with that of fungal enzymes (Fig. 5). Also, XPNPEP3 and bacterial enzymes showed much broader specificity toward Xaa↓Pro-Ala tripeptides compared with their fungal counterparts. In contrast to the bacterial enzyme, XPNPEP3 and Icp55 enzymes showed insignificant activity for the tripeptides with Asp at the P1 position (Fig. 5A). Icp55fg enzyme showed a narrower sub-

Figure 3. Comparison of XPNPEP3 structure with AMPPec. A, superimposed monomers of XPNPEP3 and bacterial AMPPec (PDB code 1WL9). The manganese ions and 01B ligands are shown in violet spheres and magenta sticks, respectively. B, superimposed dimers of XPNPEP3 and bacterial AMPPec. C and D, comparison of active sites of XPNPEP3 (C) and AMPPec (D) with modeled tripeptide Tyr-Pro-Leu. The substrate was modeled based on the Val-Pro-Leu bound in the crystal structure of AMPPec (PDB code 2V3Z), and the tyrosine residue was modeled based on the phenyl ring of 01B ligand in XPNPEP3. The sliding loops from dimeric partner are shown in ribbon and stick. E and F, relative positions of sliding loops in XPNPEP3 (E) and AMPPec (F) (PDB code 1WL9) structures. The 01B ligands (magenta sticks) are shown for the reference. The 01B ligand in AMPPec structure is modeled based on the XPNPEP3 structure.
strate range than other enzymes as it cleaved substrate with only Leu, Met, Tyr, or Phe at the P1 position. In contrast to fungal enzymes, XPNPEP3 and bacterial enzymes showed excellent activity toward bradykinin-based peptides (Fig. 5C). However, replacing Arg at the P1 position by Leu or Met in these peptides significantly improved the activity in fungal enzymes. In contrast, replacement by Ser, Asp, Val, Lys, or Ala lowered the activity in both XPNPEP3 and fungal enzymes, while retaining enzymatic activity in the bacterial counterpart (Fig. 5C). The substitutions at P3’ position in the bradykinin-based peptides had insignificant effect on the activity of all enzymes (Fig. 5C).

We tested enzymatic activity toward several non-proline substrates as follows: dipeptides (Tyr/Leu/Met ↓ Ser/Ala), tripeptides (Xaa ↓ Ser/Ala/Gly-Ala/Ser) and longer peptides (Tyr/Leu ↓ Ser-Ala-Ala, Tyr ↓ Ser-Ser-Ala-Ala-Ala-Ala) (Fig. 6B). All the enzymes showed minuscule activity for the non-proline dipeptides. XPNPEP3 and Icp55 enzymes showed significant activity for tripeptides and longer peptides with Ala or Ser at the P1’ position. However, they did not show any activity for a tripeptide with Gly at the P1’ position. Substituting aromatic or bulky hydrophobic residues (Tyr, Phe, Leu, or Met) at the P1 position with charged or smaller residues (Ala, Arg, or Asp) completely obliterated the enzymatic activity. The bacterial enzyme did not show any activity against any non-proline peptide other than Met ↓ Ala-Ala. In general, human XPNPEP3 showed higher specific activity for non-proline peptides than its fungal counterpart. However, the $K_m$ values for non-proline peptides in XPNPEP3 were generally higher than those of fungal enzymes (Tables 2 and 3). The non-proline substrates with Tyr at P1 position were the most preferred substrates for the XPNPEP3 and Icp55 enzymes (Fig. 6B). XPNPEP3 and Icp55 enzymes showed improvement in enzymatic activity with an increase in the length of substrate.

**Nfs1 processing**

Previous *in vivo* studies show that the Nfs1 protein in *Saccharomyces cerevisiae* is processed by the Icp55 enzyme after MPP processing. To confirm this activity *in vitro*, we conducted activity assays on peptides mimicking the N-terminal sequence of Nfs1 that remains after MPP processing. As mentioned earlier, we were unable to purify yeast Icp55 in its active form. Hence, the peptide substrates (Tyr-Ser-Pro-Pro-Ala and Ser-Pro-Pro-Ala) were subjected to proteolysis by the highly active Icp55 homolog, XPNPEP3. The reaction mixture was analyzed at the different time intervals (Fig. 7). Peptides and amino acids in the reaction mixture with the free α-amino group were first derivatized by α-phthalaldehyde (OPA) reagent and detected at 320 nm by reverse phase-HPLC analysis (Fig. 7A). Amino acids were identified based on standard elution profiles (Fig. 7C). With an increase in reaction time interval, a steady decrease in Tyr-Ser-Pro-Pro-Ala substrate coupled with an increase in products Ser-Pro-Pro-Ala and Tyr were observed. At the reaction interval of 120 min, a significant quantity of Ser amino acid was also observed. This implies sequential cleavage of Tyr and Ser from the peptide Tyr-Ser-Pro-Pro-Ala. Absence of a peak corresponding to free Ala implies that the peptide was not completely digested. Independent enzymatic cleavage of Ser-Pro-Pro-Ala peptide also showed cleavage of Ser and absence of free Ala (Fig. 7B).

The peptides with N-terminal Pro could not be derivatized by OPA reagent and thus could not be observed at 320 nm. Hence, the proteolytic digestion of the Ser-Pro-Pro-Ala peptide was also analyzed at 220 nm (peptide bond $\gamma_{max}$) (Fig. 7, D and E). The elution profiles at different reaction time intervals showed a gradual decrease in Ser-Pro-Pro-Ala substrate coupled with an increase in the Pro-Pro-Ala product. Because of poor absorbance of a single peptide bond in Pro-Ala peptide, it was observed only after complete digestion of the Pro-Pro-Ala peptide (50 min of reaction). Hence, the enzyme cleaves the Tyr-Ser, Ser-Pro, and Pro-Pro bond and not the Pro-Ala bond of the peptide substrates. These results clearly demonstrate the capability of the enzymes of this family to process the N terminus of Nfs1.

**Discussion**

**Significance of in vitro enzymatic activity**

The mitochondrial role of Icp55 was earlier defined on the basis of *in vivo* substrate processing in *S. cerevisiae* (4). This mitochondrial processing activity involves the proteolysis of peptides with the N-terminal sequence motif Yaa ↓ Ser/Ala-Xaa$_{in}$ (Yaa denotes Tyr, Phe, or Leu) (supplemental Fig. S5). We were unable to purify the active form of this Icp55 protein. In lieu of that, we purified its two orthologs, one (Icp55ag) from a closely related fungus, *A. gossypii* (18) and the other (Icp55fg)
from distantly related fungus, *F. graminearum*. Mitochondrial signal sequence analyses in the putative substrates of Icp55ag and Icp55fg show the conservation of sequence motifs for the MPP and Icp55 processing (supplemental Fig. S5). Our *in vitro* results clearly show the presence of characteristic Yaa↓ Ser/Ala−Xaa(3) substrate activity in two putative fungal Icp55 proteins and also in human XPNPEP3. This activity is entirely absent in the bacterial AMPPxc. XPNPEP3 and Icp55 proteins.
show enhancement in the activity with increasing lengths of peptide substrates (Fig. 6B). The Yaa\n2 Ser/Ala-Xaa(n) substrate activity highlighted by earlier in vivo studies is a minor activity as compared with the major Xaa-Pro aminopeptidase activity in both XPNPEP3 and Icp55 enzymes. Moreover, substrate specificity profiles for both enzymes are similar. Based on these similarities in their activities and their reported common cellular localization, we suggest that XPNPEP3 could perform a similar function in human mitochondria as reported for Icp55 in yeast and plant (4, 5), i.e. processing of mitochondrial matrix proteins for their stabilization as per the N-end rule (19).

The major Xaa-Pro peptidase activity observed in vitro in Icp55 and XPNPEP3 proteins could have significant implication for mitochondria. Mitochondrial proteins get damaged
due to high oxidative stress in the organelle. This results in a large turnover of mitochondrial proteome, about 10% degradation per h in exponentially growing yeast cells (20). In the mitochondrial matrix, ATP-dependent proteases like Pim1/Lon, ClpP, and m-AAA process these damaged proteins and produce 5–30-amino acid-long peptides (21). Because 70% of degraded protein is effluxed out of mitochondria in the form of amino acids (22), peptides are most likely digested to amino acids in the mitochondrial matrix. This process could be limited by the presence of Xaa-Pro peptide bonds, which are unpro-

Figure 7. In vitro processing of Nfs1-derived substrates. A, elution profile of derivatized hydrolytic product of YSPPA substrate at different reaction time intervals. B, elution profile of derivatized hydrolytic product of SPPA substrate at different reaction time intervals. C, standard elution profile of derivatized amino acids and peptides. D, elution profile of underivatized hydrolytic product of SPPA substrate at different reaction time intervals. E, standard elution profile of underivatized peptides.
processed by most peptidases due to restricted rotation around the Ca-N bond. Xaa-Pro bond at the N terminus of peptide is cleaved exclusively by Xaa-Pro aminopeptidase or prolidase enzymes, members of M24B metallopeptidases (1). To the best of our knowledge, Icp55 or XPNPEP3 is the only known member of M24B peptidases found inside the mitochondrial matrix and likely facilitates the degradation of oligopeptides to individual amino acids. Effective degradation of mitochondrial proteins is crucial for mitochondrial function, integrity, and homeostasis (23, 24). Therefore, mitochondrial processing and Xaa-Pro aminopeptidase activities of XPNPEP3 and Icp55 could be important for the proper mitochondrial functionality.

**XPNPEP3 mutation induced renal pathology, an insight**

Mutations in XPNPEP3 gene in human result in nephronophthisis, a fatal cystic kidney disease (2). XPNPEP3 crystal structure supports the view that both the mutations (F543 and A131) would lead to structural collapse and preclude the existence of active enzyme. F543 is a frameshift mutation leading to change of residues from position 311 to 314 and truncation of the protein beyond it (Fig. 8A). However, in A131, aberrant splicing leads to a change of residues from position 453 to 456 and truncation of the protein beyond it (Fig. 8B). Earlier, it has been suggested that these mutations impede normal ciliary functions during pathogenesis of the disease (2). Furthermore, the role of Xaa-Pro aminopeptidase activity of XPNPEP3 in the cytoplasm for N-terminal processing of ciliary cystogenic proteins has been proposed. Indeed, the in vivo processing of some ciliary proteins has been shown to be important for preventing the disease phenotype. Interestingly, XPNPEP1, a cytosolic Xaa-Pro aminopeptidase, is shown to express well in kidney cells (6, 25, 26). In the scenario of XPNPEP3 mutation, XPNPEP1 could process the suspected ciliary cystogenic proteins preventing the disease phenotype. Our studies suggest that XPNPEP3 could contribute to mitochondrial fitness either due to its Xaa-Pro peptidase activity and/or mitochondrial processing activity in the mitochondrial matrix. Individuals with XPNPEP3 mutation showed other symptoms like cardiomyopathy, hypertension, seizures, tremor, mental retardation, memory disorders, and hearing disorders (2), which were shown to be linked with mitochondrial defects (27). Individuals with frameshift mutation in the XPNPEP3-encoding gene, where no functional XPNPEP3 protein can be produced, have shown functional defect in mitochondrial respiratory chain complex 1 (RCCI) (2). Components of the respiratory chain complex and the proteins involved in their assembly have already been shown to be the substrates of Icp55 in yeast (4). Together, these observations support the mitochondrial connection to the cystic kidney disease, which was suspected earlier (2).

**Processing of Nfs1**

Nfs1 is a mitochondrial cysteine desulfurase in yeast and is reported as one of several substrates of Icp55 (13). Icp55 is shown to remove three N-terminal residues (Tyr-Ser-Pro→Pro-Ala-) from the polypeptide subsequent to MPP processing. Based on earlier reported in vivo substrate specificity, it is inconceivable as to how Icp55 might process Nfs1, and the role of an additional unknown peptidase was suspected (12). Our studies show that Icp55 processes both types of substrates: Yaa2Ser/Ala-Xaa(n) and Xaa2Pro-Xaa(n) (Figs. 6B and 5A). The enzyme can hydrolyze all three peptide bonds (Tyr→Ser, Ser→Pro, and Pro→Pro) but is unable to process the Pro→Ala peptide bond. Therefore, we propose Icp55 alone could remove three consecutive amino acid residues from the N terminus in three discrete steps for the complete processing of Nfs1 (Fig. 9).

To support this, we have shown the cleavage of Nfs1-derived peptides by the homologous protein XPNPEP3 (Fig. 7). Amino acids and peptide fragments observed in the reaction mixture clearly show sequential cleavage of Tyr→Ser and Ser→Pro bonds. The cleavage of the Pro→Pro bond was inferred from the presence of the Pro-Ala peptide in the final reaction mixture. Absence of free Ala in all the reaction mixtures confirms the inability of the enzyme to process the Pro→Ala bond.
Structural and enzymatic characterization of XPNPEP3 and Icp55

Figure 10. Structural adaptation of XPNPEP3. A and B, conformation of S1’ pocket and sliding loop in AMPPxc (PDB code 1WL9) in structure (A) and schematic (B). Active site pocket is shown in yellow; sliding loop in orange, and bound substrate in magenta. C and D, conformation of S1 pocket and sliding loop in XPNPEP3 in structure (C) and schematic (D). Active-site pocket is shown in cyan, sliding loop in marine blue, and bound substrate in magenta. The substrate Tyr-Ser-Leu was modeled into the active site as given in Fig. 3.

Structural basis for enzymatic activity

Human XPNPEP3 and Icp55 have significant activity for non-proline substrates like Yaa↓Ser/Ala-Xaa(ω) in contrast to their bacterial counterpart AMPPxc. The wider S1’ pocket in XPNPEP3 can easily accommodate substrates with Ser or Ala side chains but may not provide effective binding for catalysis (Fig. 10C). It is likely that this ineffective binding at S1’ pocket is compensated by improvement in S1 site binding. This improvement is attributed to perfect molecular complementarity of S1 site pocket for Tyr or Phe side chains (Figs. 3C and 10D). In addition, the S1 site is capped by a sliding loop for improving molecular binding by van der Waal contacts. In the absence of these structural adaptations (Figs. 3D and 10, A and B), the bacterial enzyme does not show significant activity for non-proline substrates, Yaa↓Ser/Ala-Xaa(ω).

Cytosolic isofrom of XPNPEP3 and TNF-TNFR2 signaling

To the best of our knowledge, the isofrom corresponding to XPNPEP3c has not been identified in any mammals other than human, and still no function has been assigned to this isofrom. XPNPEP3c lacks the mitochondrial localization signal and a large part of a critical N-terminal helix (Fig. 1A). This helix guards the hydrophobic face of the central β-sheet in the N-domain. Hence, XPNPEP3c is likely to be structurally unstable and may require a partner protein to stabilize itself. This hypothesis is consistent with our in vitro protein expression studies. XPNPEP3c expresses well into soluble form, whereas XPNPEP3c forms insoluble inclusion bodies under same conditions (data not shown). Lack of the helix limits the ability of XPNPEP3c to act as an adaptor molecule in TNF-TNFR2 signaling (3). Furthermore, stark differences in surface electrostatic potentials (Fig. 4, B and C) in human and bacterial enzyme (AMPPec) could have implications in TNF-TNFR2 signaling. The positive surface charges in the N-domain of human enzyme may have evolved to facilitate protein-protein interactions.

In conclusion, this study characterizes XPNPEP3 and Icp55 as genuine Xaa-Pro aminopeptidases. It also provides in vitro backing to the previously observed mitochondrial role of Icp55 and explains its processing of the Nfs1 substrate. Similar activity profiles of XPNPEP3 and Icp55 suggest that they play similar functional roles in mitochondria. Structural and functional studies of XPNPEP3 provide valuable insights into TNF-TNFR2 signaling and mitochondrial connection of cystic kidney disease.

Experimental procedures

Protein expression and purification

The coding DNA sequence of human XPNPEP3 (UniProt: Q9NQH7) was codon-optimized, and the gene was procured from Genscript. The coding DNA of putative Icp55 proteins from two fungi, A. gossypii (UniProt: Q75CD7) and F. graminearum (UniProt: I1RY72), and aminopeptidase P from bacteria, X. campestris (UniProt: Q8PS57) were amplified by PCR from the respective genomic DNA. These were later cloned into pST501gr (28), a T7-promoter-based expression plasmid, to form in-frame translational fusion with streptavidin-His6-TEV tag (STRHISTEV, 3.8 kDa). The exact polypeptide sizes of the proteins used in this study and their sequence alignment are given in supplemental Fig. S2. Icp55ag, Icp55fg, and AMPPxc clones were expressed in E. coli Rosetta (DE3) pLysS strain. XPNPEP3 was coexpressed with GroEL/ES in E. coli BL21(DE3) strain. The cultures were grown in 2×TY containing ampicillin (50 μg ml⁻¹) and chloramphenicol (34 μg ml⁻¹) and induced by isopropyl β-d-1-thiogalactopyranoside for 24 h at 18 °C in the presence of 1 mM MnCl₂. In case of XPNPEP3 clone, the culture at 0.6 A₆₀₀nm was induced by 5 μg ml⁻¹ of L-arabinose 30 min prior to its induction by isopropyl β-d-1-thiogalactopyranoside. The proteins were purified using metal affinity chromatography followed by TEV cleavage to remove the affinity tag. The proteins were further purified by size-exclusion chromatography on a Superdex200 10/300 GL column.

Enzyme assay

The peptides were procured from Sigma, Genic Bio (China), Biojoure (India), or USV Ltd. (India). The amidohydrolase activity of the enzymes toward different Xaa-Pro (except for Pro-Pro), Xaa-Pro-Ala (except for Pro-Pro-Ala), bradykinin-based peptides, as well as other peptides was measured by detecting the exposed α-NH₂ group using a modified ninhydrin assay (29) involving use of cadmium-ninhydrin reagent. The activity toward Pro-Pro and Pro-Pro-Ala was measured by colorimetric estimation of released proline using the ninhydrin method, as suggested by Yaron and Mlynar (30) with some modifications (31). The detailed method for ninhydrin assays are given as supplementary information. One unit of enzyme
activity in the assays was defined as the amount of enzyme that released 1 μmol of glycine eq or proline per min under the assay condition. The kinetic constants \( K_m \) and \( k_{cat} \) were determined for different substrates by the Michaelis–Menten curve fitting using Graph Pad Prism (version 6.0). The temperature optima of all the four enzymes (XPNPEP3, Icp55ag, Icp55fg, and AMPPec) were determined using Met-Pro-Ala as a substrate and in the temperature range of 25 to 60 °C. The pH optima of the enzymes were determined using 50 mM of different buffers, viz. MES (pH 6.5) and Tris–HCl (pH 7.0, 7.5, 8.0, and 8.5). The optimal pH for the activity of the enzymes was 8.0, and the optimal temperature was 55 °C.

**Proteolysis of Nfs1-based peptide**

Nfs1-based peptide substrates (YSSPA, SSPA) were subjected to proteolytic processing by XPNPEP3. The 5 mM substrate was incubated with 2 μg of XPNPEP3 enzyme at 50 °C. At different time intervals, an aliquot of the homogeneous reaction mixture (40 μl) was removed and allowed to react with 60 μl of OPA reagent (0.5% (w/v) OPA prepared in the solution containing 200 mM borate (pH 10.4), 80% methanol, 25 μM 3-mercaptopropionic acid) for 3 min to derivatize the product of the reaction mixture. The analysis was performed by reverse phase HPLC by injecting the derivatized products into a C18 column, equilibrated with 0.1% trifluoroacetic acid (TFA) in 100% acetonitrile. The separation was carried out at flow rate of 0.5 ml min\(^{-1}\) by gradient elution between solution A and solution B (12.5 mM phosphate buffer (pH 7) in 50% acetonitrile). The gradient elution was carried out as follows: 0–2 min, 100% solution A; 2–35 min, 50% solution A, 50% of solution B; 35–40 min, 100% solution B. The elution was monitored at 330 nm. To monitor the peptides with N-terminal Pro residue, an aliquot of reaction mixture was subjected to reverse phase HPLC by injecting the derivatized products into a C18 column, equilibrated with 12.5 mM phosphate buffer (pH 7) (solution A). Separation of the reaction mixture was performed at a flow rate of 1 ml min\(^{-1}\) by gradient elution between solution A and solution B (12.5 mM phosphate buffer (pH 7) in 50% acetonitrile). The gradient elution was carried out as follows: 0–2 min, 100% solution A; 2–35 min, 50% solution A and 50% of solution B; 35–40 min, 100% solution B. The elution was monitored at 220 nm.

**Crystallization, structure determination, and refinement**

The purified XPNPEP3 protein (12–15 mg ml\(^{-1}\)) was mixed with 5 mM apstatin (Santa Cruz Biotechnology) and incubated on ice for 30 min prior to crystallization trials. The crystallization was performed by under the oil microbatch method at 21 °C (32). The diffraction quality crystals were obtained by microseeding in a crystallization condition, 100 mM HEPES (pH 7.5) and 20% polyethylene glycol (PEG) 3350. The crystals were cryoprotected by Paratone N (Hampton Research HR-632). The diffraction data were collected on MAR225 CCD (Rayonix) detector at PX-BL21 beamline of Indus-2 Synchrotron, India (33). The high-resolution data were collected at a wavelength of 0.9794 Å at 100 K. The data were processed by XDS (34) and AIMLESS (35). The initial phases for the structure were obtained by molecular replacement (MR) by Phaser using the structure of AMPPec (PDB code 1WL9 (15)) as the MR search model. The model was built using phenix.autobuild (36). Further improvement in the model was carried out by iterative cycles of manual model building and refinement using phenix.refine (36), respectively. The structure was finally refined to \( R_{free} \) of 17.75% at a resolution of 1.65 Å. The quality of the model was analyzed with MolProbity (38). The data processing and refinement statistics are summarized in the Table 1. The structural analyses and illustrations were performed using PyMOL. The surface electrostatic calculations were performed using the ABPS server through the PDB2PQR web portal (39).

**Author contributions**—R. D. M. and S. N. J. designed the experiments. R. S. purified and crystallized biomolecules. S. N. J. and R. S. performed the enzymatic assays. R. S., A. K., and B. G. collected and processed the diffraction data. R. D. M. refined the crystal structure. R. S., V. D. G., S. N. J., and R. D. M. analyzed the data. R. D. M., R. S., and V. D. G. wrote the manuscript. All the authors have reviewed and commented on the manuscript.

**Acknowledgments**—We are grateful to Drs. S. M. Sharma, P. D. Gupta, and P. A. Naik for their support during this project. We thank Narisjeet Singh Sabharwal (RRCAT project trainee) for initial work on protein solubilization. We thank the staff at Indus-2 (RRCAT, India) for providing synchrotron beam. We acknowledge the use of PX-BL21 beamline and its associated biochemical facilities.

**References**


