The role of propeptide-mediated autoinhibition and intermolecular chaperone in the maturation of cognate catalytic domain in leucine aminopeptidase

The role of propeptide-mediated autoinhibition and intermolecular chaperone in the maturation of cognate catalytic domain in leucine aminopeptidase

G. Baltulionis[1,2], M. Blight[3], A. Robin[3], D. Charalampopoulos[2], K.A. Watson[1*]

School of Biological Sciences[1] and Department of Food and Nutritional Sciences[2], Whiteknights Campus, University of Reading, Reading, UK, RG6 6AS; Biocatalysts Limited[3], Cefn Coed Parc Nantgarw, Wales, UK, CF15 7QQ.

*Corresponding author

Abstract

Leucyl aminopeptidase A from *Aspergillus oryzae* RIB40 (AO-LapA) is an exo-acting peptidase, widely utilised in food debittering applications. AO-LapA is secreted as a zymogen by the host and requires enzymatic cleavage of the autoinhibitory propeptide to reveal its full activity. Scarcity of structural data of zymogen aminopeptidases hampers a better understanding of the details of their molecular action of autoinhibition and how this might be utilised to improve the properties of such enzymes by recombinant methods for more effective bioprocessing. To address this gap in the literature, herein we report high-resolution crystal structures of recombinantly expressed AO-LapA precursor (AO-*pro*LapA), mature LapA (AO-*m*LapA) and AO-*m*LapA complexed with reaction product l-leucine (AO-*m*LapA-Leu), all purified from *Pichia pastoris* culture supernatant. Our structures reveal a plausible molecular mechanism of LapA catalytic domain autoinhibition by propeptide and highlights the role of intramolecular chaperone (IMC). Our data suggest an absolute requirement for IMC in the maturation of cognate catalytic domain of AO-LapA. This observation is reinforced by our expression and refolding data of catalytic domain only (AO-*ref*LapA) from *Escherichia coli* inclusion bodies, revealing a limited active conformation. Our work supports the notion that known synthetic aminopeptidase inhibitors and substrates mimic key polar contacts between propeptide and corresponding catalytic domain, demonstrated in our AO-*pro*LapA zymogen crystal structure. Furthermore, understanding the atomic details of the autoinhibitory mechanism of cognate catalytic domains by native propeptides has wider reaching implications toward synthetic production of more effective inhibitors of bimetallic aminopeptidases and other dizinc enzymes that share an analogous reaction mechanism.

Keywords: leucine aminopeptidase, propeptide, autoinhibition, refolding, zymogen

Introduction

Aminopeptidases are utilised in the food industry for sensory applications and development of food flavour profile. Aminopeptidases are indispensable in protein hydrolysate debittering, since they remove hydrophobic amino acids from the N-terminus of peptide hydrolysates, which contribute to bitterness of the flavour profile [1-3]. Fungal aminopeptidases from *Aspergillus oryzae* are generally recognized as safe (GRAS) by the food industry, with a long history of utilisation in debittering applications. Efforts to characterise their role in this process, led to earlier work to over-express and biochemically characterize the extracellular leucine aminopeptidase A from *A. oryzae* RIB40 (AO-LapA) [4]. To date, no additional reports of exogenous over-expression of this high-utility aminopeptidase exist in the literature. Also lacking are attempts of successful production of such peptidases in *Pichia pastoris*, an important GRAS organism used in the food industry.

To our knowledge, there have been no published data for the crystal structures of the leucine aminopeptidases used in the food industry. The scarcity of precise structural data for these predominantly fungal peptidases inhibits our understanding of their reaction-inhibition mechanisms and substrate specificities. For example, there are only 12 amino acid differences between the full sequences of AO-LapA and AS-Lap1 (leucine aminopeptidase from *Aspergillus sojae*), which mostly occur outside the active site pocket, that is responsible for their quite distinct substrate preferences [5].

It is well established, that pro-domains of secretory endo- and exopeptidases play an important role in the inhibition of their cognate catalytic domains, which undergo activation by removal of the pro-domain either by autocatalytic processing or external endopeptidase activity. By secreting protease precursors in a zymogen form, undesirable cytosolic activation and proteolysis events are prevented [6]. Typically, the zymogen conversion does not involve conformational changes of catalytic residues and the active sites are sterically rendered

inaccessible to substrate by the unique residues of pro-domain or prosegment, thus preventing activity [7]. In addition, N-terminal propeptides are known to act as intramolecular chaperones (IMC), decreasing the energy barrier of transition from a molten globule intermediate to a kinetically trapped native state [8]. In this respect, a well-studied model enzyme, subtilisin, has been shown to be unable to escape the folding transition state between a molten-globule and a native conformation, without the presence of a pro-domain; a rate-limiting step in the folding process of subtilisin [9]. Remarkably, pro-domains of numerous proteinases have been reported to assist as chaperones in both inter- and intramolecular manner, following expression in trans- and cis- form, respectively [10-12]. An analogous mode of action for such pro-domains, with no requirement for a covalent link between propeptide and catalytic domain, has been discovered in other enzymes [13]. Indeed, both subtle and rather complex relationships between propeptide and catalytic domain were found in subtilisin E, where an unaltered, mature polypeptide chain of subtilisin E was shown to fold into a different conformation via a mutated intramolecular chaperone and to show altered substrate specificity; a phenomenon termed `protein memory` [14]. Normally, such intramolecular chaperones precede catalytic domain from the N-terminus, however, eukaryotic monozinc aminopeptidase A was found to recruit its C-terminal propeptide for this role [15]. Such findings demonstrate the diverse and complex relationships between the essential pro-domain and the cognate active conformation in the proteogenesis of proteases.

The aminopeptidase from *Vibrio proteolyticus* (AAP) is a model bimetallohydrolase enzyme, representing the M28 peptidase family (MH clan) whose extensive structural and functional studies have provided an in-depth understanding of the mechanism of action and the substrate preferences for these bi-metallic enzymes [16-25]. Furthermore, the role of the propeptide in AAP and other homologous leucine aminopeptidases has been extensively studied *in vitro*

[26-28]. A number of studies have provided structural insights into AAP inhibition via different synthetic compounds, acting as competitive inhibitors [18, 21, 22, 29]. However, until now, no report exists to provide a detailed mechanistic insight for leucine aminopeptidase inhibition by its natural competitive inhibitor – the N-terminal propeptide - at the molecular level.

In this article, we present the crystal structures of recombinantly expressed LapA precursor (AO-*pro*LapA), mature LapA (AO-*m*LapA) and mature form of the enzyme complexed with reaction product l-leucine (AO-*m*LapA-Leu), all purified from *P. pastoris* culture supernatant. To date, the exact mechanism of leucine aminopeptidase inhibition and the role of an intramolecular chaperone (IMC) by the pro-domain has not been described in detail. To the best of our knowledge, this is the first crystal structure of a leucine aminopeptidase precursor, revealing detailed structural aspects of catalytic domain inhibition by the cognate propeptide. This work shows that known synthetic aminopeptidase inhibitors and substrates mimic key polar contacts between propeptide and the catalytic domain. These findings could aid the future design of more effective inhibitors of bimetallic aminopeptidases and other dizinc enzymes, sharing a similar reaction mechanism.

Results & Discussion

Sequence comparison of Lap peptidases

Extracellular aminopeptidases from the M28 family share a very similar sequence composition of their open reading frames (ORF`s). In general, a short secretion signal sequence is preceded by a moderate length propeptide (removed in the maturation process), which defines a central catalytic domain. It is convenient to compare their primary sequences and their tertiary structure folds to establish conservation and evolutionary divergence at both the sequence and structure level. Evolutionary relationships between Lap1 from *A. sojae* and

eukaryotic, as well as microbial homologues, were previously investigated [30]. In our study, full ORF`s of several of these homologous fungal and microbial extracellular Lap enzymes were aligned together with LapA from *A. oryzae* (AO-*pro*LapA) and Lap1 from *A. sojae* (AS-proLapA) (**Figure S1**). Despite AO-*pro*LapA sharing 97% sequence identity to AS-proLapA, these enzymes display remarkably distinct preferences for substrates [31]. Compared to highly conserved catalytic core residues, a varying degree of homology was observed within the signal peptide and propeptide sequences. **Figure S1** shows that, of the 11 amino acids that differ in these two aminopeptidases, 5 reside within the propeptide region, which is removed in the maturation process, and the other 6 lie within the mature polypeptide. All of the different residues in AS-proLapA, except Gly192, are distributed far away from the catalytic centre, according to our model. Thus, the relationship between substrate specificities and differences in amino acid sequence, among these peptidases, is likely more subtle and translates either through a global network of intramolecular interactions via small modifications in the catalytic domain, or an even more complex relationship related to differences in the more variable propeptide sequence and its role as an intramolecular chaperone.

LapA from *A. oryzae*, Lap1 from *A. sojae*, AAP from *Aeromonas proteolytica* and more recently reported Lap enzymes from *Legionella pneumophila* [32] all belong to the M28 peptidase family; clan MH (MEROPS). AAP from *A. proteolytica* has been studied extensively and has been a model enzyme for this family [16, 17]. Lap enzymes are typical structural representatives of this family with two Zn(II) ions sequestered from the surrounding environment and coordinated in the active site by 5 conserved residues in sequence order His, Asp (coordinates both Zn(II) ions), Glu, Asp, Glu, His. Also, an activated water molecule, associated with both metal ions, acts as a nucleophile during catalysis. Additional Asp and Glu residues, considered to be essential for catalysis, reside in

close proximity within the His-Xaa-Asp and Glu-Glu motifs. Lap enzymes contain an archaic alpha/beta hydrolase fold with a five-stranded beta sheet (1 strand antiparallel to the rest) sandwiched between a layer of helices on each side. The active site is parallel to the end of these beta strands.

The coordinated divalent zinc ions form a functional charge/dipole complex along with conserved five amino acid ligands (AO-LapA numbering: H176, D195, E234, D261, H343). Another conserved E233 residue, within the Glu-Glu motif, is supposed to assist in deprotonation of the terminal catalytic water molecule and, subsequently, donating a proton to newly formed amino-terminus, described as a rate-limiting step in the catalysis [23]. Several mutant forms of this residue render homologous AAP enzymes inactive [33, 34]. Also, a complementary second-shell hydrogen bond network was identified in AAP and another homologous aminopeptidase from *Streptomyces griseus* (SGAP) [35]. We found that the Zn-D261-S315-Q259-D316-S240 network, identified in the AAP structure, also was conserved in our AO-LapA sequence and determined structure **(Figure 4,** a**)**. Six residues (M262, D338, Y312, C314, I342, F331) lining the hydrophobic specificity binding pocket of AO-LapA (**Figure 2**) were also mostly conserved in homologous extracellular aminopeptidases, including Lap1 from *A. sojae* (AS-proLap1) and AAP from *A. proteolytica* (**Figure S1**). This analysis clearly suggest that AO-LapA, AAP and AS-LapA share a conserved active site architecture, which in turn leads to highly conserved reaction mechanism in these M28 aminopeptidases.

LapA expression in recombinant hosts

Coding regions of LapA precursor (AO-*pro*LapA) and prodomain-deleted version (AO-*Δpro*LapA) were codon-optimized for use by a *P. pastoris* host. Multiple gene dosage transformants were obtained for both (pJ-npro-LapA and pJ-*Δpro*-LapA) constructs. However, no real-time qPCR was attempted to quantify the number of integrated gene copies,

in each case. Preliminary expression trials of the wild-type recombinant LapA with alpha mating factor signal sequence and native LapA signal peptide resulted in slightly higher secretion levels with the latter. Thus, it was decided to study further the expression of recombinant LapA harbouring native signal peptide.

Determined molecular weight of recombinantly produced proenzyme was 39 kDa. Whilst the mature enzyme showed the molecular weight of 32.5 kDa. The final molecular weight is dependant on the extent of N-linked glycosylation at Asn87. The enzyme was not capable of autoprocessing (autocatalytic cleavage of propeptide) following 72h incubation at ambient temperatures in the purified form. The expression of precursor proteins is well established in other metalloenzymes, such as subtilisin, thermolysin and homologous aminopeptidases from *Aeromonas caviae* T-64 and *Vibrio proteolyticus* [27, 28, 36, 37]. Usually, these are extracellularly secreted enzymes, where propeptide acts as a competitive inhibitor, including subtilisin E [38], carboxypeptidase A [39], a metalloprotease from *Brevibacillus brevis* [40]. This could be an evolutionary adaptation to prevent non-specific intracellular proteolysis. The catalytic nature of LapA can induce toxic effects to the host cell, therefore, most probably the activation and processing of the LapA precursor in native *A. oryzae* occur following secretion into the extracellular environment.

In the case of AO-mLapA and AO-*pro*LapA, expression levels reached approximately 0.3 g per litre. SDS-PAGE analysis revealed significantly lower expression levels for AO-*Δpro*LapA in *P. pastoris* supernatant (0.01 g/L) (**Figure 1**, a). We speculate that low expression levels were observed, due to the absence of the prosequence in the pJ-*Δpro*-LapA construct. Alternatively, it may be due to the integration of a lower copy number of LapA genes. However, pJ-*Δpro*-LapA transformants were found to grow on YPD agar containing 2000 ug/ml Zeocin, which is very unlikely for the *P. pastoris* strain harbouring only a single copy of the recombinant gene. The lack of propeptide in the AO-*Δpro*LapA expression

construct could have yielded misfolded protein with exposed hydrophobic patches, associating with molecular chaperones, the mechanism by which unfolded proteins are targeted for degradation within the ER [13, 41]. As a result, the observed expression levels (**Figure 1**, a, i) were significantly lowered compared to AO-*pro*LapA, which contained propeptide in the expression cassette (**Figure 1**, a, i).

AO-*pro*LapA and AO-*m*LapA recombinant proteins were obtained from the same expression cassette (pJ-npro-LapA) and were a result of clonal variation within *P. pastoris* transformants. Surprisingly, in one of the clones, unresolved processing of expressed recombinant LapA precursor (AO-*pro*LapA) occurred, yielding mature active AO-*m*LapA. Most likely, cleavage of propeptide was a post-translational event since genomic PCR-based sequencing within the coding region showed no deletion of propeptide (results not shown). We propose that co-secreted non-specific proteolytic activity by an endogenous *P. pastoris* peptidase was responsible for the processing of the LapA (AO-*pro*LapA) precursor protein. As a consequence, N-terminal sequencing of AO-*m*LapA and the observed electron density (**Figure 2**, a) revealed three additional residues (A77, V78, T79) preceding the N-terminal tyrosine, originally reported for AO-LapA [4]. The unobserved electron density for the C-terminal prodomain residues, a result of mobility, likely is required for better accessibility for activation by an endopeptidase involved in the maturation of AO-LapA precursor. Therefore, an unexplained endopeptidase activity, likely present in the *P. pastoris* supernatant, was different from that of native *A.oryzae*, which degrades the propeptide to N-terminal tyrosine (Y80). This would seem to be the only viable explanation since, in the other clones, the vast majority of recombinant protein was found in the form of the proenzyme (uncleaved propeptide) in the culture supernatant.

As expected, *P. pastoris* supernatants with secreted AO-*m*LapA and AO-*pro*LapA contained no other contaminant host proteins (**Figure 1**, a). This facilitated the subsequent purification of these recombinant LapA (AO-*pro*LapA, AO-*Δpro*LapA, AO-*m*LapA) enzymes in the absence of affinity tags. Importantly, repetitive AEX purification cycles were required to liberate AO-*pro*LapA and AO-*m*LapA from a brown pigment existing in *P. pastoris* supernatant. Elimination of these brown pigments proved very challenging, which was deemed mandatory for crystallisation trials. Two brown pigments (M_r < 2kDa) were reported previously to associate with other Lap peptidases and were briefly characterised by EPR to be spectroscopically active and contain a Fe(III) ligand [42].

Over-expression of AO-*ref*LapA in *E.coli* was performed to test the feasibility of soluble expression of the active catalytic domain (propeptide sequence omitted) of this eukaryotic enzyme. AO-*ref*LapA protein was only recovered from inclusion bodies by urea renaturation and subsequent 2-step purification (**Figure 1**, b). A dialysis step was required to liberate AO-*ref*LapA from 200 mM imidazole, present in the IMAC-eluted fraction. It is worth noting that, only a small portion of refolded AO-*ref*LapA was significantly active, following a size exclusion chromatography (SEC) step (fractions 7 and 8, **Figure 1**, b, ii). The specific activity of non-aggregated and refolded AO-*ref*LapA (fractions 7 - 8) reached 59.5 U/mg (**Table 1**). These fractions were pooled and utilised in subsequent CD analysis (**Figure S2**). A large population of AO-*ref*LapA separated as inactive, potentially aggregated protein (fractions 1 - 3, **Figure 1**, b, ii) using SEC. These high-MW aggregates were even visible under denaturing SDS-PAGE and reducing buffer conditions. Attempts to isolate AO-LapA in soluble form in *E. coli*, required an additional *in vitro* refolding step to restore some activity. However, the catalytic activity was not fully recoverable, likely due to absence of propeptide and its intramolecular chaperone effect (discussed later).

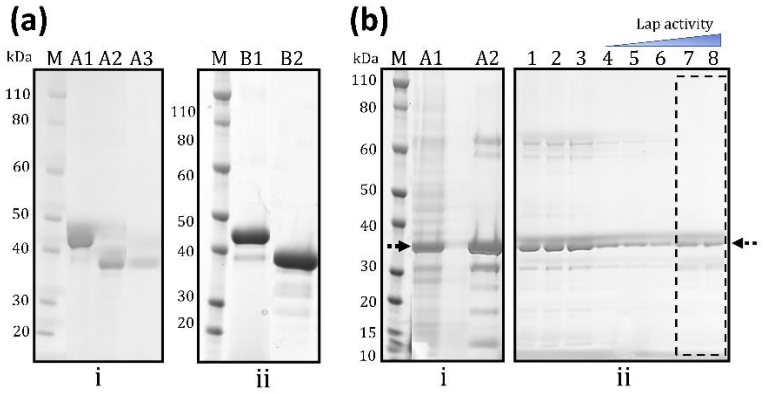


Figure 1 Purification of Lap proteins. (**a**) – purification of AO-*m*LapA, AO-*pro*LapA and AO-*Δpro*LapA: i) *P. pastoris* supernatants sampled (10 µl) at 120 h, containing recombinant LapA proteins: A1 – AO-*pro*LapA, A2 – AO-*m*LapA, A3 – AO-*Δpro*LapA; ii) purified from *P.pastoris* supernatant: B1 – AO-*pro*LapA, B2 – AO-*m*LapA; (**b**) – purification of AO-*ref*LapA (indicated by dotted arrows): i) initial IMAC purification step: A1 – resolubilized AO-*ref*LapA inclusion bodies in 8M urea, A2 – imidazole eluted fraction from IMAC column; ii) subsequent SEC purification of IMAC-eluted fraction and dialysis-refolded protein, each lane corresponds to individual fractions. Active (under blue triangular activity bar) and the most active (dashed rectangle) fractions are indicated. M is molecular weight marker (Novex Sharp, Invitrogen, US).

The hydrolytic activity profiles of all the recombinant LapA forms were obtained against LPNA substrate under similar conditions (pH 7.2, 37 ºC; **Table 1**), as previously reported in studies of homologous LapA proteins [30, 43]. The reaction and kinetic parameters of AO-*m*LapA were found to closely resemble wild-type VpAP (120 U/mg; k_{cat} (s^{-1}) = 82 ± 2; K_m (µM) = 18 ± 1) [43]. Also, AO-*m*LapA appeared significantly more efficient than another close homologue LP-proLapB protein: k_{cat}/K_m (s^{-1}/mM^{-1}) = 637.75 of AO-*m*LapA versus k_{cat}/K_m (s^{-1}/mM^{-1}) = 0.345 of LP-proLapB for its substrate Lys-*p*NA [44]. AO-*m*LapA

appeared to be two orders of magnitude more efficient at hydrolysing its substrate, LPNA. The observed catalytic efficiency for AO-*m*LapA represents a fully functional, activated recombinant protein, possessing an appropriate three-dimensional fold (discussed in the subsequent section).

(i)	Enzyme	Specific activity (U/mg)	(ii)	Reaction kinetics of AO-*m*LapA	
	AO-*pro*LapA	0.8 ± 0.1		k_{cat} (s^{-1})	76.53 ± 0.5
	AO-*m*LapA	140.15 ± 4.8		K_m (μM)	120 ± 1.1
	AO-*Δpro*LapA	0.34 ± 0.1		k_{cat}/K_m (s^{-1}/mM^{-1})	637.75 ± 2.5
	AO-*ref*LapA	59.5 ± 1.1			

Table 1 Specific activities of purified LapA proteins (i) and steady-state reaction kinetic measurements (ii) for AO-*m*LapA protein. Activity is averaged from 3 independent measurements with standard deviation determined.

Circular Dichroism (CD) of LapA proteins

The isolated LapA proteins were analysed by far-UV CD to estimate their secondary structure content and to study their folded states. The characteristic far-UV spectra of AO-*ref*LapA, AO-*m*LapA and AO-*Δpro*LapA proteins are shown in **Figure S2**.

The specific activity of AO-*ref*LapA was 2-fold reduced compared to AO-*m*LapA and it was predicted to exist in a molten globule state, according to CD analysis. This implies that absence of propeptide, in the expression construct of AO-*ref*LapA, gives rise to an improperly folded protein, which was unable to regain its full native conformation under 20 h

of refolding conditions. The lower content of secondary structure also was found in recombinant refolded aminopeptidase from *V. proteolyticus* (AAP) compared to natively expressed wild-type enzyme [45]. Finally, a spectral minimum at 200 nm and close-to-zero ellipticity near 222 nm suggested a lack of secondary structure in the AO-*Δpro*LapA protein and a predominant pre-molten globule state, which explains its completely diminished hydrolytic activity against Leu-*p*NA for this prodomain-truncated version of LapA protein (**Table 1**). The differences in hydrolytic activity observed between recombinant LapA in *P. pastoris* (AO-*Δpro*LapA) and recombinant LapA in *E. coli* (AO-*ref*LapA) likely arise from their different expression and isolation. AO-*ref*LapA was refolded *in vitro* from *E.coli* inclusion bodies, using standard refolding conditions to favour folded, functional protein [46]. However, AO-*Δpro*LapA was naturally folded in the cytoplasm of *P.pastoris* and secreted into the growth media, which could have led to misfolding (in the absence of its prodomain), thus yielding less active protein. These observations demonstrate the negative effect of the absence of the propeptide iand signifies a potential role for the LapA propeptide in assisting the mature domain to gain a fully folded, functional conformation essential for maximal activity.

Crystal structure of mature LapA (AO-mLapA)

We determined a high-resolution crystal structure of the catalytic LapA domain to 1.97 Å. We observed an additional three propeptide-derived amino acids (A77, A78, A79) at the N terminus. Furthermore, additional electron density was observed for yeast-derived glycosylation at residue N87 (**Figure 2**, a). The determined monomeric LapA structure (AO-*m*LapA) forms a classical α/β globular domain. The hydrophobic core is comprised of a twisted 8-stranded β-sheet, sandwiched between α-helices, representing a typical archaic hydrolase fold. Secondary structure assignment, using the DSSP [47] server, showed that

45%, 16% and 39% of the residues are involved in α-helices, β-strands and coil structures, respectively. An accessible active site pocket is located at the surface of the protein. The catalytic binding pocket is formed by 6 conserved residues (H176, D195, D261, E233, E234, H343), coordinating two Zn ions, embedded in the loop region close to the core parallel β3 and β5 sheets, and hydrophobic specificity pocket residues, mostly arranged at the C-terminal end of the protein (**Figure 2**, b). The two Zn atoms are coordinated in a distorted tetrahedral geometry with an average interatomic distance of 3.6 Å. We also identified a nucleophilic water molecule bridging the two Zn atoms. Since both Zn ions significantly decrease the pK_a of this water molecule [48], at the pH for optimal activity (8.0 – 9.0), it may exist as an OH^- ion. However, under our crystallisation conditions for this structure (pH 5.0), it most likely adopts an HOH molecule.

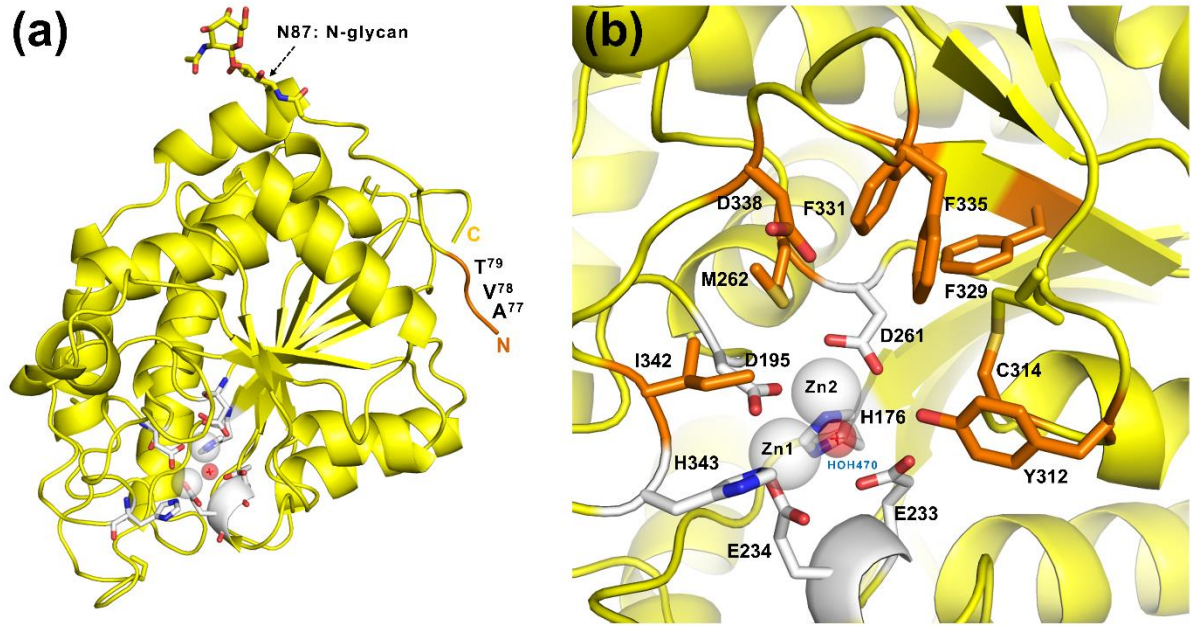


Figure 2 Crystal structure of LapA catalytic domain (AO-*m*LapA) determined to 1.97 Å resolution (**a**) and a close-up view of the active site and hydrophobic binding pockets (**b**). Catalytic residues and coordinated Zn atoms are coloured in white; nucleophilic water,

HOH470, coordinated between the two Zn ions, indicated by red sphere; conserved hydrophobic binding pocket residues are shown in orange sticks;

	AO-*pro*LapA	AO-*m*LapA	AO-*m*LapA-Leu
Data collection			
Beamline	DLS I02	DLS I02	DLS I04
Wavelength (Å)	0.9763	0.9763	0.9795
Space group	P 21 21 21	P 63 2 2	P 63 2 2
Unit-cell parameters			
a, b, c (Å)	43.43 94.39 99.11	153.92, 153.92, 88.64	154.1, 154.1, 87.83
α, β, γ (°)	90, 90, 90	90, 90, 120	90, 90, 120
Resolution range (Å)	42.61 - 1.61 (1.668 - 1.61)	73.81 - 1.97 (2.041 - 1.97)	73.37 - 2.48 (2.569 - 2.48)
No. of unique reflections	53338 (5269)	43992 (4297)	22302 (2195)
Redundancy	4.3 (4.2)	2.0 (2.0)	2.0 (2.0)
Completeness (%)	99.48 (99.55)	99.73 (99.33)	99.95 (99.95)
R_{sym}[a] or R_{merge} (%)	7.303 (89.15)	5.5 (61.81)	9.65 (48.29)
R_{meas}	0.08292	0.07729	0.1372
I/σ(I)	11.36 (1.59)	10.86 (1.21)	7.08 (1.55)
Refinement			
Resolution (Å)	42.61 - 1.61 (1.668 - 1.61)	73.81 - 1.97 (2.041 - 1.97)	73.37 - 2.48 (2.569 - 2.48)
No. reflections	53338 (5269)	43992 (4297)	22302 (2195)
R_{work}/R_{free} (%)[b]	15.27/17.39	16.45/20.36	15.69/21.50
No. atoms			
Protein	2698	2360	2390
Zn(II)	2	2	2
Chlorine	1	1	1
Water	482	462	360
Protein residues	341	301	302
Average B-factors (Å^2)			
Protein	23.1	34.58	39.08
Zn(II)	18.9	30.38	44.45
l-leucine ligand	-	-	59.04
RMSD[c]			
bond lengths (Å)	0.007	0.008	0.007
bond angles (°)	1.250	0.95	0.94
Ramachandran Plot[d]			
Favored, Allowed, Disallowed (%)	98, 2.00, 0.00	98.33, 1.67, 0.00	97.66, 2.01, 0.33

The values in parentheses are for the highest resolution shell.

[a] $R_{sym} = \sum_{hkl} \sum_i | I_i(hkl) - \langle I(hkl) \rangle | / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl) \rangle$ are the observed intensity and average intensity of multiple observations from symmetry-related reflections, respectively.

[b] R_{work} was calculated with 95% of the data used for refinement, and R_{free} was calculated with 5% of the data excluded from refinement.

[c] The RMSD stereochemistry value represents the deviation from ideal values.

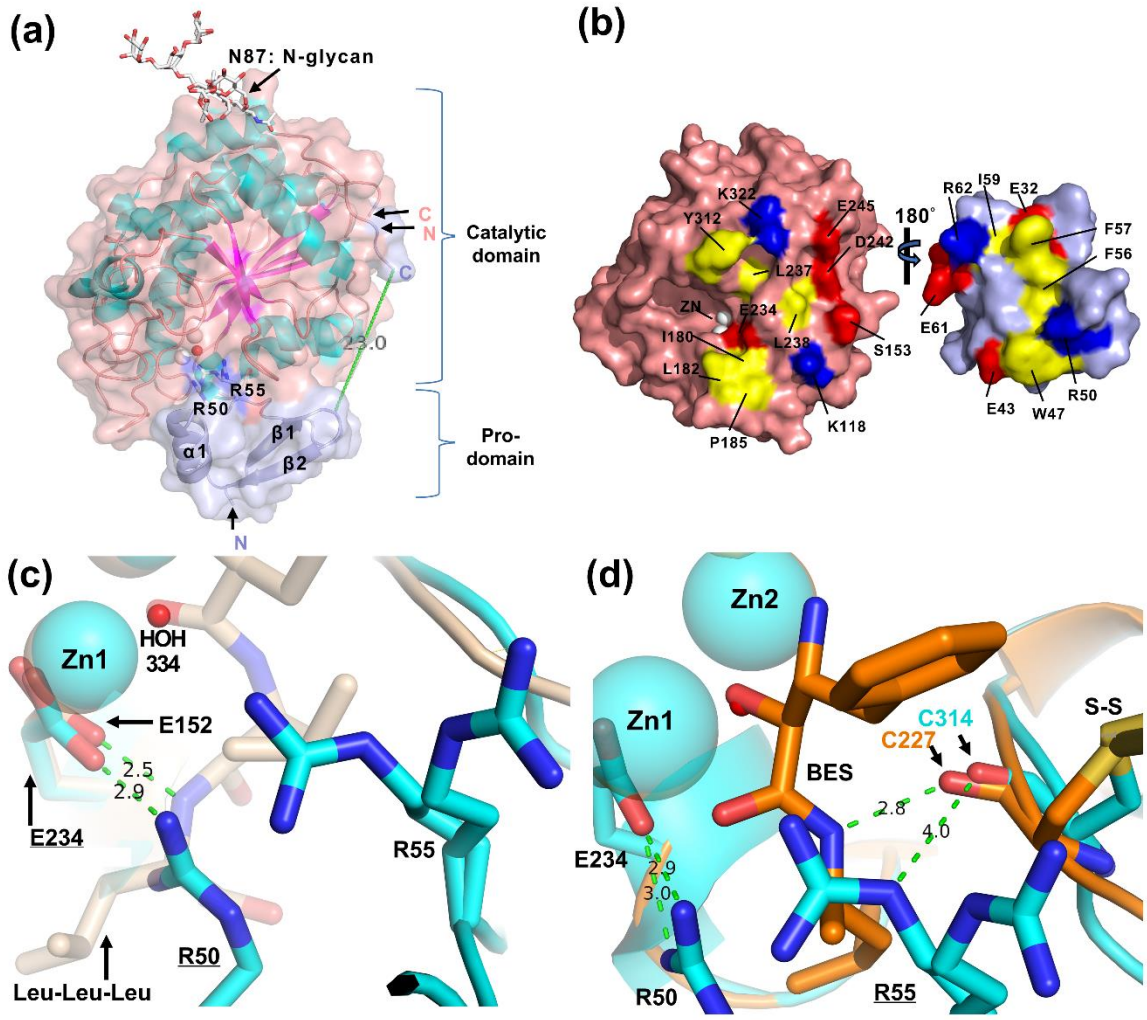
[d] Ramachandran analysis was carried out using Molprobity.

318 **Table 2** Data collection and refinement statistics

LapA precursor: propeptide-mediated inhibition of aminopeptidase domain

The determined three-dimensional structure of LapA precursor (AO-*pro*LapA) was refined to 1.61 Å resolution. All residues were well located in the electron density, with the exception of residues 1 – 25 and 66 – 76 at N- and C-terminal region of the pro-domain, respectively, due to inherent flexibility in these regions. The final model consisted of 341 residues, 482 water molecules, including the catalytic bridging water, a single chloride ion and an ethylene glycol molecule (arising from cryo-protection).

From **Figure 3** (a), it is evident that the pro-domain extends from the junction with the catalytic domain, follows alongside the catalytic core (residues 66 – 76 are missing, hence 23 Å gap) until it enters and occludes the catalytic cleft, using its 3-turn α-helix motif. The overall structure of this separate domain is highly disordered and flexible: only 23% of total residue content within the propeptide forms secondary structure. The middle part of the propeptide sequence forms a 3-turn α-helix (residues 43-52) and a short 2-stranded β-sheet (residues 29 - 31 and 39 - 41).



334

335 **Figure 3** Analysis of the crystal structure of LapA precursor (AO-*pro*LapA), determined to
336 1.6 Å resolution (PDBid: 6ZEP). (**a**) cartoon and surface representation of the final protein
337 model: pro-domain is coloured in light blue, key arginines (R50, R55) are displayed as white
338 sticks, catalytic Zn ions and coordinated water molecule are depicted as white and red
339 spheres, respectively; a single glycosylated asparagine (N87) is indicated by stick
340 representation of the attached sugar moieties. A missing, flexible region (residues 66 – 76),
341 connecting the pro-domain and the catalytic domain, is indicated by a dashed green line along
342 with a direct line measured distance in angstroms. (**b**) Contacts between the catalytic (left)
343 and pro-domain (right). The pro-domain is rotated 180 degrees to expose the corresponding
344 surface contacts. The domain interface is stabilised mainly by salt bridges (positively charged

residues – blue, negatively charged residues – red) and tight hydrophobic interactions (yellow). Di-nuclear Zn site is shown as white spheres. Mechanism of propeptide-mediated inhibition by the substrate (**c**) and inhibitor (**d**) mimicry. AO-*pro*LapA (in cyan) is superimposed with AAP in complex with Leu-Leu-Leu substrate (PDBid: 2IQ6) (in wheat) and inhibitor - bestatin (BES) (in orange) (PDBid: 1TXR). Key propeptide arginines R50 and R55, catalytic glutamates E234 (AO-*pro*LapA) and E152 (AAP), equivalent disulphide forming C314 (AO-*pro*LapA) and C227 (AAP) are labelled. Key electrostatic interactions are measured in angstroms.

The propeptide buries 8.8% of mature domain solvent-accessible surface area (1,005 Å of 11,474 Å). This does not take into account the missing 66 – 76 residue region of the propeptide. By corollary, the propeptide uses 29.7% of its surface area for the interface with the mature domain (**Table S1**). The inter-domain network of interactions involves 37 and 21 (almost a half) of the residues for the catalytic and pro-domains, respectively. More specifically, the inter-domain interface consists of 12 short-range (2.2 - 3.3 Å) H-bonds, 3 long-range (3.3 - 3.8 Å) H-bonds, 6 side chain-side chain salt bridges (< 3.3 Å), and at least 10 aliphatic side-chain residues within the propeptide that make tight hydrophobic contacts with the mature domain (**Table S2**) (**Figure 3**, b).

Lastly, we identified Leu49 as a key propeptide fold stabilising residue. In the crystal structure of LapA precursor, Leu49 was found to favourably interact with at least six adjacent residues (**Figure S3**). These multiple hydrophobic interfaces, between the aliphatic side chain of Leu49 and side chains of interconnecting residues, could be essential for stabilising the overall fold of the pro-domain. Congruently, Leu70 in the LP-proLapB structure was found to be involved in multiple adjacent hydrophobic contacts alongside the Cys37 – Cys77 disulphide link [44]. We conclude that the strategically placed leucine residues significantly contribute to stabilising the tertiary fold of the respective pro-domains, ensuring the correct

orientation of the α1 helix (**Figure 3**, a). The strategic orientation of the pro-domain helix could be instrumental for steric occlusion of the active site cleft, thereby assisting in the correct positioning of key inhibitory arginines (Arg50 and Arg55), discussed in the following section.

Mechanism of catalytic domain inhibition by LapA propeptide

The crystal structure of LapA proenzyme enabled elucidation of the core inhibitory interactions exerted by the propeptide domain, namely, two principal inhibitory arginines, Arg50 and Arg55, penetrating the catalytic cleft to form specific interactions with highly conserved residues, involved in peptide hydrolysis mechanism. Arg50, forms a tight 2.9 Å hydrogen bond with OE1 of the catalytic Glu234 in the main chain (**Figure 3**, c). This interaction directly mimics the analogous H-bond (2.5 Å) of Glu152 OE1 and the amide group of P3` LLL ligand observed in the AAP-LLL structure (PDBid: 2iq6); one of three essential H-bonds in forming the AAP-LLL pre-transition state complex [25]. The H-bond angle between donor and receptor atoms of Glu152 and leucine substrate is 108 ^{o} in the AAP-LLL complex, while the analogous H-bond contact between Arg50 of the AO-*pro*LapA and E234 is 147^{o}, indicating a more favourable electrostatic interaction, suggesting a strong inhibition mechanism by AO-*pro*LapA propeptide. The Arg50 – Glu234 interaction has a significant effect on the coordination of this residue to the Zn1 atom. The Zn1 – O distance of Glu234 OE1 (dangling oxygen) atom is 2.64 Å in the mature LapA structure, while in the propeptide-inhibited structure this increases to 2.90 Å due to a strong electrostatic interaction with both NH1 and NH2 nitrogens delivered by Arg50. Arg55 also contributes to the inhibitory mechanism of LapA propeptide by providing a flexible terminal guanidinium moiety, which (based on the electron density maps) we found to occupy three possible orientations, supporting its inherent flexibility. Each of the three nitrogens in the guanidinium

moiety can potentially form favourable van der Waals contact with the carbonyl oxygen of the adjacent Cys314 (**Figure 3**, d). This oxygen is involved in one of three key H-bond contacts and the only H-bond present in the AAP complex with the LLL substrate (PDBid: 2iq6) and aminopeptidase bestatin (BES) inhibitor (PDBid: 1txr), respectively. We propose, in a similar fashion, that the NE atom of the mobile Arg55 side chain is mimicking an essential hydrogen bond contact between the backbone carbonyl oxygen of Cys227 and the N1 atom of BES (2.8 Å). Also, the observed multiple orientations of Arg55 captured in our electron density maps illustrates the dynamic behaviour of the side chain of this residue, possibly distorting the di-nuclear Zn site by transiently interacting with the terminal nitrogens on the guanidinium moiety. In this respect Arg50, likely together with Arg55, prevents binding and initiation of substrate hydrolysis in AO-*pro*LapA. This autoinhibitory mechanism was not previously demonstrated in otherwise well studied M28 family of dizinc aminopeptidases.

Importantly, the mechanism of Glu234 inhibition by recruitment of the Arg50 side chain in the LapA pro-domain provides additional support to the proposed carboxylate-mediated nucleophile delivery hypothesis presented by Kumar et. al (2007) [25]. While our data does not directly indicate a role for Glu234 in catalysis, it does show that inhibition by the propeptide is via blocking of one of the principal H-bonding sites for substrate binding. In addition to sterically occluding the active site cleft, it is likely that the propeptide acts by neutralising key residues, Glu234 and Tyr312, thereby preventing peptide hydrolysis. Until more experimental data becomes available in the field, it can be confirmed that Zn1 ligand and its role in the catalytic process is due to tuning the Lewis acidity of Zn1. The hydrogen bond by Arg50 decreases the effective negative charge on the carboxylate group of Glu234 modulating the Zn1 Lewis acidity. The combination of steric hindrance of the active site along with change in Lewis acidity of Zn1 ligand results in the loss of AO-*pro*LapA activity.

This is well-established based on inhibited structures of AAP and other aminopeptidases within this class.

Pro-domain induced conformational changes in LapA precursor

Given the high conservation in active site geometries for dimetallic zinc M28 peptidases, the crystal structure of AAP in complex with the *bona fide* LLL (l-leucyl-l-leucyl-l-leucine) substrate (PDBid: 2iq6) and synthetic inhibitor bestatin (PDBid: 1txr) provides direct evaluation of propeptide-induced conformational changes within the substrate binding pocket and corresponding HBP (hydrophobic binding pocket) residues (**Figure 3**, c and d; **Figure 4**, c). LapA precursor was found to deviate from mature LapA only by 0.211 RMSD over an equivalent range of residues (77 - 377). The presence of propeptide elicits several structural rearrangements in the catalytic cleft of LapA precursor (**Figure 4**, c). Apart from the conformational changes in the hydrophobic binding pocket (HBP) residues (discussed later), the architecture of the active site is not altered significantly. However, very subtle changes in the interatomic distances of the active site residues were detected. The distance between the two catalytic Zn atoms is longer (3.6 Å) in the mature enzyme compared to LapA precursor (3.4 Å). This stems from a minor change in Zn coordination distances with catalytic residue ligands in the LapA precursor. Distinct loop movements in the residing regions of Cys310-314 and Phe335 residues causes lengthening of the bond distance to the carbonyl oxygen of Cys314, which forms direct contact with the amide nitrogen of the leucine substrate (P_1') (**Figure 4**, c). This subtle (2.9 Å to 3.3 Å) H-bond lengthening is sufficient to weaken this polar contact. Also, it is possible that the H-bond angle of this contact becomes less favourable in the LapA precursor. Alteration of this contact might contribute to an impaired ability to bind the substrate. Indeed, Cys314 has been identified as one of the key residues in the hydrophobic substrate binding pocket in Lap1 from *A. sojae* [5]. Similarly, residue Phe335 is directly involved in substrate binding in the mature enzyme, making a hydrophobic

444 contact (~ 4.5 Å) with the sidechain of P_1 moiety (**Figure 4**, c). A significant movement of
445 the Phe335 side chain out of the active site (increasing the contact distance to 10.4 Å), elicits
446 a less catalytically competent conformation in the LapA precursor.

447

448 Additionally, at least two more residues were significantly affected by the movement of the
449 Cys310-314s and Phe335s loops. An increase of > 2 Å from the leucyl hydrophobic side
450 chain (P_1) of the bound substrate was observed for the aromatic ring of Phe331. Remarkably,
451 the side chain of Y312 is almost completely flipped by 180 degrees. A conserved residue,
452 Tyr225, in an equivalent location is found in the AAP protein (PDBid: 2iq6). In the initial
453 model of peptide hydrolysis, Tyr225 was proposed to interact with the N-terminus of the
454 incoming substrate [20]. However, this concept was later challenged by the AAP-bestatin and
455 AAP-LPA complexes, where it was concluded that Tyr225 may not be involved
456 mechanistically, but it provides yet another stabilising hydrophobic contact [22].
457 Additionally, the phenolic oxygen of Tyr225 shares a H-bond with BES, Tris, LeuP and IDH
458 ligands. In contrast, we demonstrate that the N-terminus of the P_1' leucyl moiety is too distant
459 to form a H-bond with the hydroxide of the Tyr312 side chain, but it could make a favourable
460 hydrophobic contact with the side chain of P_1' (**Figure 4**, c). In any scenario, this highlights
461 the importance of Tyr312 in stabilising the hydrophobic substrate, which is impaired in the
462 flipped orientation of Tyr312 observed in the LapA precursor and other orthologous members
463 of the M28 peptidase family (**Figure 4**, b). In fact, the flipped orientation of this residue is a

single common trait identified in the propeptide-inhibited structures of these homologues (**Figure 4**, b). Therefore, the catalytic importance of this residue cannot be overlooked.

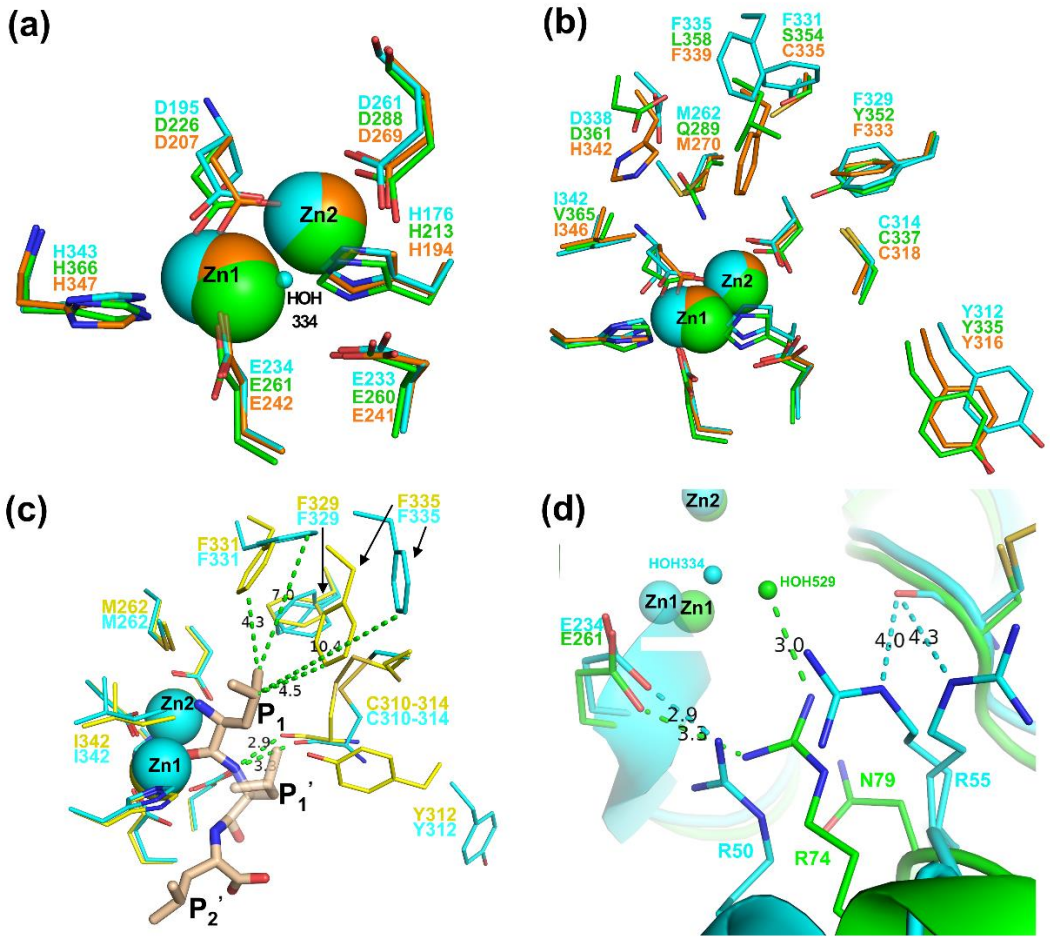


Figure 4 Structural comparison of LapA from *A. oryzae* to select orthologs from the M28 family. LapA precursor is shown in cyan, mature LapA – in yellow, AAP-LLL complex – in wheat (PDBid: 2iq6), LapA from *L. pneumphila* (PDBid: 6esl) – in orange, LapB from *L. pneumphila* (PDBid: 5gne) – in green. (**a**) Superimposition of active site residues; (**b**) Superimposition of active site and hydrophobic binding pocket residues; (**c**) Structural alterations in LapA precursor compared to mature protein, involving the leucine tripeptide taken from the AAP-LLL crystal structure; (**d**) Direct comparison of the mechanisms of propeptide-mediated inhibition in LapA from *A. oryzae* versus LapB from *L. pneumophila*. Residue labels are coloured corresponding to their respective molecules.

Structural comparison of LapA from A. oryzae and known orthologs of the M28 family

A number of crystal structures relating to family M28 aminopeptidases have been reported in the literature [32, 44]. Pairwise sequence alignment between aminopeptidase LP-*pro*LapB from *L. pneumophila* (PDBid: 5gne), a homologous M28 peptidase, and full-length AO-*pro*LapA revealed only 31% identity and 48.5% similarity, and another *L. pneumophila* aminopeptidase LP-*pro*LapA showed only 26% identity and 41.5% similarity. However, the tertiary fold similarity was strikingly high with AO-*pro*LapA deviating by only 0.793 Å and 0.924 Å RMSD from LP-*pro*LapA and LP-*pro*LapB, respectively, over an equivalent range of residues (26 – 377, AO-*pro*LapA numbering). The active site geometry was found almost identical in all three crystal structures, with highly conserved catalytic residues and dizinc metal cofactors superimposing, almost exactly (**Figure 4**, a). Despite only low sequence identity, these aminopeptidases displayed an identical conserved fold, which is in agreement with proteins having greater conservation in tertiary structures and confirming their classification as M28 family members.

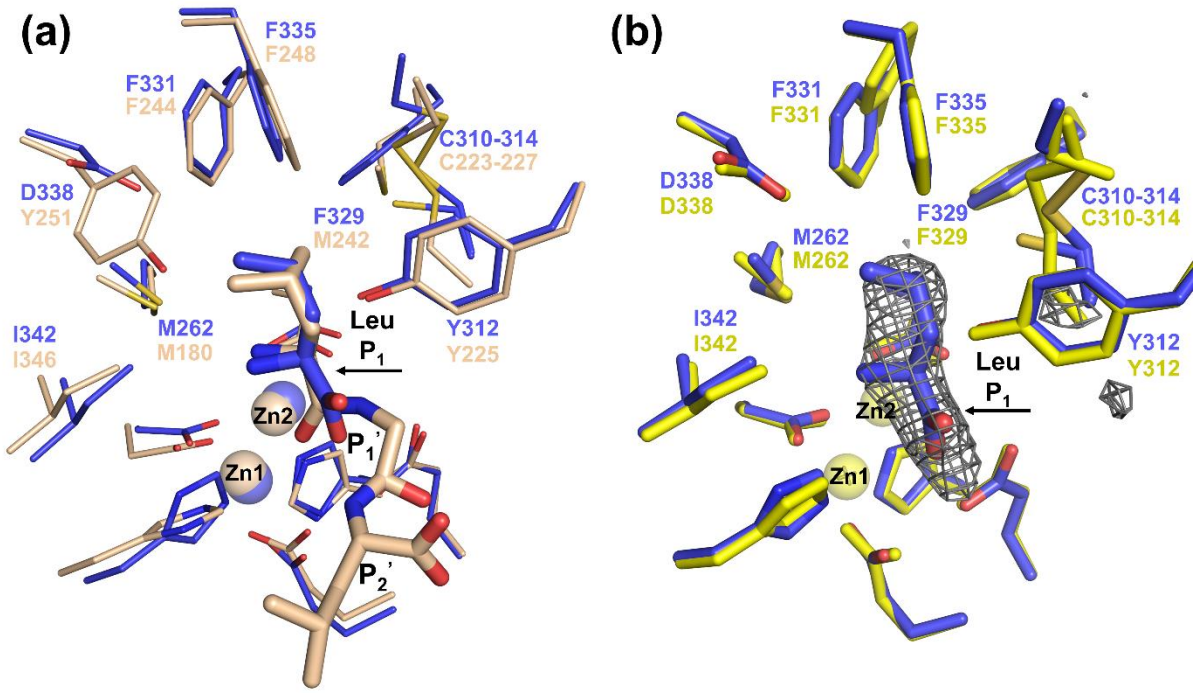
Notable differences in the substrate binding pockets of these aminopeptidases are identified (**Figure 4**, b). Conserved Phe329 and Phe331 are replaced by respective Tyr352 and Ser354 in LP-*pro*LapB, which may be responsible for the preferential hydrolysis of positively charged, hydrophilic substrates by this enzyme [49]. His342 and C335 substitutions in LP-*pro*LapA might be responsible for reported very broad specificity toward N-terminal hydrolysis targets [32]. Interestingly, the mode of autoinhibition in LP-*pro*LapB was quite similar to AO-*pro*LapA, albeit not identical. LP-*pro*LapB only recruits a single Arg74 to represent a substrate like-contact with Glu261 in a similar fashion to our observed mechanism in AO-*pro*LapA (**Figure 4**, d). Arg74 in LP-*pro*LapB also makes significant water-mediated contact with the active site Zn, not observed in our model of LapA precursor. Quite the opposite role for propeptide was observed in the study of LP-*pro*LapA

aminopeptidase, where no specific inhibitory-like interactions could be identified in the crystal structure of its precursor. A weaker catalytic domain – propeptide interface was exemplified by domain-swapping event captured in the dimer interface of the asymmetric unit [32]. Also, enzymatic removal of pro-domain significantly diminished catalytic activity in LP-*pro*LapA. A directly opposite effect was observed in autoinhibitory AO-*pro*LapA and LP-*pro*LapB peptidases, where the freed catalytic domain showed a substantial increase in a specific activity.

AO-LapA binding the reaction product – Leucine (*AO-*m*LapA-*Leu)

Crystal structure of the recombinant mature LapA (AO-*m*LapA) in complex with a single leucyl moiety was determined to 2.5 Å resolution. The crystallisation process involved briefly soaking the previously formed AO-*m*LapA crystals in leucine tripeptide (Leu-Leu-Leu). However, only a single leucine (P_1) could be reliably traced in the resulting electron density map (**Figures 5**). Most likely, the enzyme exhibited hydrolytic activity in the crystalline state, cleaving the scissile bond between P_1 and P_1' residues of the Leu-Leu-Leu substrate.

Several AAP crystal structures in complex with a natural substrate Leu-Leu-Leu tripeptide [25] and substrate analogue 1-Butaneboronic acid [50] have been reported. The coordination geometry of the leucyl moiety, in our AO-*m*LapA-Leu structure was almost identical to that observed in the crystal structure of AAP-LLL for the P_1 moiety (PDBid: 2iq6) (**Figure 5**, a). This observation reflects the conservation of the active site residues, substrate binding pocket geometries and reaction mechanism for these homologous leucine aminopeptidases, which display only 34% sequence identity.



523

524 **Figure 5** Comparison of the substrate binding pockets in AO-*m*LapA-Leu and AAP-LLL
525 (PDBid: 2iq6) crystal structures. (**a**) Superimposition of active site residues, bound substrates
526 and HBP residues. The isobutyl sidechain of the leucyl P_1' was omitted for clarity; (**b**)
527 Superimposition of active site and HBP residues in resting (AO-*m*LapA) and Leu-bound
528 (AO-*m*LapA-Leu) enzymes; Colour code: AO-*m*LapA-Leu – violet, AAP-LLL – wheat, AO-
529 *m*LapA - yellow. The calculated interatomic distances between ligand and protein in the AO-
530 *m*LapA-Leu structure are provided in **Table S3** and an Fo-Fc omit map for the Leu ligand is
531 shown in grey mesh.

532

533 The distance between the two Zn atoms in the AO-*m*LapA-Leu complex is 3.3 Å, which is
534 slightly less than that observed in the resting LapA molecule (3.6 Å). The isobutyl sidechain
535 of the P_1 ligand interacts with an expansive, hydrophobic S_1 subsite, formed by the
536 intervening sidechains of residues Tyr312, Cys314, Phe335, Phe329, Phe331, Met262 and
537 Ile342, which all are within 5 Å of the ligand sidechain (**Figure 5**, b). Accommodation of the

leucine ligand elicited no major conformational changes of the key HBP residues, indicated by an almost ideal superimposition with equivalent residues in resting AO-*m*LapA crystal structure (RMSD 0.175 Å). The HBP residue sidechains 'lock in' the leucine moiety in the active site pocket, which adopts a catalytically competent geometry that resembles the unbound state. By contrast, we observed significant movements of key HBP residues in the catalytically compromised state of autoinhibited enzyme (AO-*pro*LapA) (**Figure 4**, c). Due to the differences in ligand soaking times (into preformed crystals) compared to the published AAP-LLL complex (PDBid: 2iq6), it appears that we likely trapped a post-catalysis state showing only a single leucine bound to AO-*m*LapA.

Conclusion

In this work, the role of propeptide in the expression construct of LapA precursor was investigated, using a *P. pastoris* recombinant expression platform. Due to the propeptide deletion AO-*Δpro*LapA, recombinant LapA was secreted at significantly reduced levels compared to full LapA precursor (AO-*pro*LapA). Lack of the propeptide in AO-*Δpro*LapA expression cassette could have yielded inappropriately folded protein with exposed hydrophobic patches, associating with molecular chaperones, the mechanism by which unfolded proteins are targeted for degradation within the ER of yeast [13, 41]. Hence, the observed expression levels of AO-*Δpro*LapA protein were significantly reduced compared to AO-*pro*LapA, which contained propeptide in the expression construct, in turn escaping the degradation mechanism. Using CD spectroscopy, AO-*Δpro*LapA was shown to exhibit a lack of secondary structure, consequently, to exist in a predominantly pre-molten globule state, while LapA was able to acquire a native fold in the presence of propeptide. This experiment alone indicated a requirement of prosequence, acting as a chaperone in an intramolecular

562 manner. However, our crystal structures also reveal a plausible molecular mechanism of AO-
563 LapA catalytic domain inhibition by propeptide, further supporting a role for its N-terminal
564 domain as an intramolecular chaperone, in the process of maturation of AO-LapA zymogen.
565 We observed that the mature catalytic domain of AO-LapA is not able to achieve its native
566 conformation when produced in the absence of N-terminal prosegment. This observation is
567 reinforced by expression and refolding study of hexahistidine-tagged catalytic domain only
568 (AO-*ref*LapA) from *E. coli* inclusion bodies, revealing a limited active conformation.

569 This work also demonstrates that AO-LapA propeptide acts as an intramolecular chaperone
570 and is essential for the maturation of the catalytic domain when expressed in *cis*. We show
571 that AO-LapA prodomain may be able to function in an intermolecular manner and assist
572 with secretion and folding of the mature polypeptide when expressed in *trans*. This is in
573 agreement with other reported studies of prodomain expression in non-covalently linked state
574 (*trans*) for homologous and more distantly related peptidases. For example; expression of the
575 propeptide and mature domains, as separate polypeptides, resulted in the recovery of activity
576 for membrane Type 1-matrix metalloproteinase [51]; impeded secretion of the mature domain
577 of aspartic proteinase Sap1p from *C.albicans* in *P. pastoris* was restored upon co-expression
578 of the pro-polypeptide in *trans* [10]; and unlinked propeptide co-expression in *P. pastoris*
579 also was shown to rescue the appropriate secretion and activity levels of *Streptomyces*
580 *mobaraensis* Transglutaminase [13].
581 More generally, this work contributes to the knowledge of currently available crystal
582 structures of distantly related orthologous peptidases and offers an evolutionary and
583 biologically important model of propeptide-mediated inhibition for metalloenzymes of the
584 M28 family. Based on the available structural data from distant prokaryotic orthologs and our
585 high resolution Lap crystal structures presented herein, we also conclude that the catalytic
586 domain of AO-*pro*LapA is more specifically and strongly inhibited by its cognate propeptide,

recruiting long side chains of positively charged arginine residues. In the parent *A. oryzae* eukaryotic cell this most likely evolved in order to prevent non-specific proteolysis of essential amino peptides before secretion to the extracellular environment, whereas this may not be so important in a less complex prokaryotic organism, such as *L. pneumophila*. Only LP-*pro*LapB showed autoinhibitory control over its catalytic domain, while removal of LP-*pro*LapA propeptide even reduced the catalytic rates and slightly modulated substrate preference. Such detailed understanding of the autoinhibitory mechanism of cognate catalytic domains by native propeptides could aid synthetic production of more effective inhibitors of bimetallic aminopeptidases and other dizinc enzymes that share an analogous reaction mechanism.

Materials and Methods

Strains and plasmids

P. pastoris BG10 strain was obtained from Biogrammatics Inc. (Carlsbad, CA, USA). Secretory expression vector, pJ902, was provided by DNA 2.0 (USA). *E. coli* DH5α (*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) cloning strain was purchased from New England Biolabs (NEB,UK).

All primers used in this study were obtained from Sigma (UK). Zeocin™ was purchased from Invitrogen (USA). All other reagents were purchased from Sigma (UK) and were of analytical grade.

Construction of recombinant expression vectors and transformations in Pichia pastoris

Recombinant LapA open reading frames (Genebank accession: XP_001825745.1) were codon-optimised and cloned into the pJ902 vector, by DNA 2.0, using the *Eco*RI and *Not*I restriction sites. In this way, an expression construct harbouring the LapA native signal peptide and propeptide sequences (pJ-npro-LapA) was obtained. A second expression construct, containing a deletion of the LapA propeptide (pJ-*Δpro*-LapA) was produced, utilising the QuickChange™ II site-directed mutagenesis kit from Stratagene (USA). Deletion of the propeptide sequence was achieved by performing PCR with designed primers 1; 5'-GGACGCTATCTGGATAAGCTAATGCGCTTGCC-3' and 2; 5'-GGCAAGCGCATTAGCTTATCCAGATAGCGTCC-3', using the pJ-npro-LapA construct as a template.

Following *Sac*I-mediated linearization and clean-up, recombinant plasmids were transformed into an X33 expression strain by electroporation (a single 1500 V pulse in a 2 mm electroporation cuvette). Immediately following electroporation, the cells were resuspended

623 in 0.5ml 1M sorbitol and 0.5ml YPD (1% yeast extract Y, 2% soy peptone P, 2% dextrose D)
624 and incubated for 2h at 30ºC under gentle agitation. The revived cells were then plated onto
625 YPD agar (1% yeast extract, 2% soy peptone, 2% dextrose, 2% agar), containing either 100
626 µg/ml or 2000 µg/ml Zeocin™ and incubated at 30 ºC for 2-4 days. The resulting *P. pastoris*
627 colonies were verified by PCR for appropriate genomic integration, using the method
628 previously described [52]. In each case, 10 transformants were screened picking the largest,
629 single colonies. Each single colony was patch-plated onto a separate YPD agar plate and
630 residual cells were re-suspended in 100µl lysis solution (200 mM lithium acetate, 1% SDS),
631 followed by incubation at 70 ºC for 15 min. Following 96% ethanol precipitation, the DNA
632 pellet was re-suspended in 100 µl sterile water and 1µl (in each case) was used in each
633 subsequent genomic PCR. Clone ID (Lucigen, USA), PCR pre-*m*ix and a pair of AOX
634 primers (AOX1-F; 5'-GACTGGTTCCAATTGACAAGC-3' and AOX1-R; 5'-
635 GCAAATGGCATTCTGACATCC-3') were used, under the following conditions: 1 cycle of
636 98 ºC for 2 min, 30 cycles at 98 ºC for 30 s, 56 ºC for 30 s, 72 ºC for 1 min, finishing with 72
637 ºC for 5 min. The PCR products of predicted size were visualised on a 1% agarose gel.

638 ***Construction of recombinant expression vectors and transformation in E. coli***

639 Mature *A. oryzae* LapA polypeptide (Genebank accession: XP_001825745.1) was codon-
640 optimised and recombinantly cloned into a pET-21a expression vector (Eurofins, UK), using
641 *Nde*I and *Bam*HI restriction sites. In this way, a pET-6xHis*Δpro*LapA expression construct
642 resulted, comprised of the T7 bacteriophage promoter, an N-terminal hexahistidine tag,
643 mature LapA polypeptide (minus propeptide) and an ampicillin resistance marker. This LapA
644 expression construct was transformed into BL21 (DE3) competent cells, according to the
645 manufacturer's instructions (New England Biolabs, US).

Over-expression of recombinant LapA in P. pastoris

Recombinant LapA proteins were over-expressed in a *P. pastoris* X-33 strain (Mut^+), under control of the P_{AOX1} methanol inducible promoter. Shake flask expression trials were conducted in BMGY (1% yeast extract, 2% soy peptone, 1% glycerol, 1.34% YNB, 4E-5% biotin, 100mM potassium phosphate, pH 6.0) and BMMY (as for BMGY with 1% glycerol replaced by 0.5% methanol). Genomic PCR-verified *P. pastoris* clones were inoculated into 50 ml BMGY medium in 250 ml baffled conical flasks and incubated in a shaker (250 rpm) for 16 h at 28 ºC. Once the OD_{600} reached 2-8, an equivalent amount of cells (corresponding to OD_{600} of 1) were harvested and re-suspended in 100 ml of BMMY and shaken (250 rpm) in 500 ml conical baffled flasks at 28 ºC for 4-5 days. Every 24 h, the cultures were supplemented with 0.5% final concentration of methanol to maintain induction and OD_{600} cell densities were recorded. After 120 h of maintaining recombinant protein expression, the cell cultures were harvested by centrifugation at *6000 x g* for 15 min. The supernatant was sampled and recombinant LapA protein detected by SDS-PAGE (NuPAGE® Novex® 4-12% Bis-Tris gels).

Purification of recombinant LapA from P. pastoris supernatant

Over-expressed AO-*pro*LapA, AO-*m*LapA, AO-*Δpro*LapA proteins were purified from *P. pastoris* supernatant employing two different protocols: 1) crude *P. pastoris* supernatant, containing recombinant LapA, was dialysed against 20 mM Tricine pH 8.0, 1 mM $ZnCl_2$ for 16 h at 4 ºC, then filtered through a 0.22 µm steri-filter prior to loading onto a 5 mL HiTrap™ Q HP anion exchange column (AEX), pre-equilibrated with 20 mM Tricine pH 8.0, 1 mM $ZnCl_2$. The column was washed (5 – 10 CV) with equilibration buffer until no trace of protein appeared in the eluate. Recombinant LapA was eluted using an increasing gradient (0 - 1M) of NaCl (20 mM Tricine pH 8.0, 1 mM $ZnCl_2$, 1M NaCl). Active fractions were pooled and concentrated, using ultrafiltration membranes (10 kDa MWCO). In each case, the

respective protein concentrate was loaded onto a HiLoad Superdex^{TM} 16/60 200pg (GE Healthcare, UK) column, pre-equilibrated with 20 mM Tricine pH 8.0, 1 mM $ZnCl_2$. Homogeneous recombinant LapA fractions were concentrated and stored at -20 ºC for further analysis; or 2) crude *P. pastoris* supernatant, containing the respective recombinant LapA constructs, was saturated to 90% ammonium sulphate by adding hard salt, followed by stirring for 1h at ambient temperature. LapA precipitant was pelleted by centrifugation at *16,000 x g* for 20 minutes. The collected pellet was re-suspended in 20 mM Tricine pH 8.0, 1 mM $ZnCl_2$, followed by dialysis against the same buffer overnight at 4 ºC. At this point, the purification procedure was equivalent to 1), starting at the anion exchange chromatography step. Recombinant proteins purified using protocols 1) and 2) were analysed on SDS-PAGE for purity.

Over-expression of recombinant LapA in E. coli

All preliminary recombinant LapA over-expression trials were carried out using 10 ml cultures. A single BL21 (DE3) colony harbouring pET-6xHis*Δpro*LapA expression construct was transferred to 2 mL LB broth and incubated at 37 ºC (shaking at 220 rpm) overnight. 200 µL of the overnight inoculum was transferred to 10 ml of fresh LB media. The optical density (OD) at 600 nm was monitored regularly to reach a value between 0.5 and 1.0 absorbance units, after which IPTG (1 mM final concentration) was added to induce protein expression culture. The pre-induction sample was taken for further analysis on SDS-PAGE. The induced cultures were then incubated for 16-20 h at either 37 ºC or 20 ºC. Following harvest at *16,000 x g*, cells were lysed with BugBuster® Master Mix (Novagen) and soluble versus insoluble protein fractions were analysed on SDS-PAGE.

Purification and refolding of* AO-*ref*LapA *protein from E. coli inclusion bodies

The pellet with inclusion bodies was re-solubilised in 8M urea buffer (50 mM Tricine pH 8.0, 500 mM NaCl, 8 M urea) gently agitating for 1 hour. The non-dissolved debris was separated

by centrifugation for 20 min at *16,000 x g*. The resulting supernatant was loaded onto a HisTrap™ FF 5 ml (GE Healthcare, UK) nickel immobilisation affinity column (IMAC), pre-equilibrated with wash buffer (50 mM Tricine pH 8.0, 500 mM NaCl, 8 M urea) and re-solubilised AO-*ref*LapA was eluted in 50 mM Tricine pH 8.0, 500 mM NaCl, 8 M urea and 200 mM imidazole. AO-*ref*LapA was allowed to slowly refold by reducing the concentration of urea in the buffer: IMAC-eluted fraction was dialysed against refolding buffer (10 mM Tricine pH 8.0, 1 mM $ZnCl_2$) for 20 h at 4 ºC. The AO-*ref*LapA was further purified using a HiLoad Superdex™ 16/60 200pg (GE Healthcare, UK) gel filtration column pre-equilibrated with 20 mM Tricine pH 8.0, 1 mM $ZnCl_2$. Sample collection (2 ml fractions) of fractions exhibiting the highest specific activity were pooled and concentrated on Vivaspin 15R – 10,000 Dalton Centrifugal Filter Units (Sartorius, UK) containing the HydroSart® membrane.

Aminopeptidase assay

LapA specific activity was determined spectrophotometrically, using a synthetic substrate l-leucine-p-nitroanilide (LPNA) and monitoring the formation of the *p*-nitroanilide moiety (*p*NA), as absorbance, at 405 nm ($\varepsilon = 9.9$ $mmol^{-1}cm^{-1}$) in a cuvette with a 1 cm path length. Peptidase assays were performed on a CE1020 1000 Series spectrophotometer (Cecil Instruments). One unit of LapA activity was defined to catalyse the hydrolysis of one micromole of LPNA to l-leucine and p-nitroaniline, per minute at pH 7.2, 37 ºC.

Quantification of recombinant LapA

LapA proteins were quantified either by Bradford assay (Bio-Rad, USA), using BSA as a standard, or spectrophotometrically, employing a NanoDrop® ND-1000 (Thermo Scientific), using the LapA absorbance extinction coefficient (23,505 $M^{-1}cm^{-1}$). The concentration of recombinant LapA in crude *P. pastoris* supernatant could only be determined by Bradford assay due to impurities interfering with spectrophotometric measurements.

CD spectroscopy of recombinant LapA

The circular dichroism spectra for all resultant proteins were recorded in the range 360 to 190 nm at 20 ºC, using a Chirascan series 800 spectrophotometer. Measurements were taken using a 1 nm step size in a 0.1 mm path length quartz cell. CD data were averaged from 5 independent scans for each sample. Samples of AO-*ref*LapA, AO-*m*LapA and AO-*Δpro*LapA proteins were analysed at 36 µM, 21 µM and 40 µM concentrations, respectively, buffered in 20 mM HEPES pH 8.0.

Crystallisation and structure determination of recombinant LapA

Crystallisation screening was carried out using HTP robotics at Oxford Protein Production Facility (OPPF-UK) (96 well commercial sparse matrix screens, sitting-drop vapour diffusion method) and crystallisation optimisation was carried out in the home laboratory using 24 well, hanging-drop vapour diffusion. AO-*pro*LapA, AO-mLapA recombinant proteins were crystallised at concentrations of 13 mg/ml and 11 mg/ml, respectively. Both proteins were buffer exchanged into 20 mM Tricine pH 8.0, 1 mM $ZnCl_2$ prior to crystallisation. AO-*pro*LapA crystals with approximate dimensions 0.5 × 0.04 × 0.005 mm formed under F10 condition (0.1 M Bis-Tris, pH 5.5, 0.2 M NaCl, 25% PEG 3,350) in the Index HT™ screen (Hampton Research, USA) over 3 days at 21 °C. AO-mLapA crystals were produced in 24-well plate format at room temperature under previously identified F2 condition (0.1M Citrate, pH 5.0, 3.2 M Ammonium sulphate) from the JCSGPlus HT^{TM} screen (Molecular Dimensions, UK).

Prior to data collection, crystals were cryo-protected by quick transfer through a mixture of mother liquor and 20% ethylene glycol, followed by flash cooling to 100K in liquid nitrogen. Diffraction data were collected at Diamond Light Source (UK), on beamlines I02 and I03. The diffraction data were recorded, by fine slicing (0.1° degree oscillation), on a Pilatus 6M detector. Indexing and integration were completed automatically by *XDS* [53] and *xia2* [54].

All LapA structures were solved by maximum likelihood molecular replacement with PHASER [55], using the coordinates of an established LapA homology model; LapA homology model was generated using the IntFold2 server [56], which identified aminopeptidase from *Vibrio sp.* (34% sequence identity) as a closely related structural homologue (PDBid: 1RTQ), as the template. The resulting LapA models were manually built and fitted to the electron density maps using COOT [57]. All structures were refined using PHENIX [58], polder omit maps [59] were used to validate the presence of any potential ligands. Structural comparisons and superimpositions were generated using the TM-align server [60]. Final structural refinement statistics are provided in **Table 2.** Structural images were prepared using PyMOL [The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC].

Accession numbers

Structural data were deposited in the PDB database with the following accession codes: 6ZEP (AO-*pro*LapA), 6ZEQ (AO-*m*LapA) and 6ZES (AO-*m*LapA-Leu).

CRediT authorship contribution statement

Gediminas Baltulionis: Conceptualization, Investigation, Methodology, Formal analysis, Visualisation, Writing - original draft, Writing - review & editing. **Kimberly A. Watson:** Funding acquisition, Conceptualization, Resources, Methodology, Formal analysis, Visualisation, Supervision, Writing - review & editing, . **Mark Blight:** Conceptualization, Resources, Methodology. **Aelig Robin:** Conceptualization, Resources, Methodology. **Dimitris Charalampopoulos:** Funding acquisition, Conceptualisation, Resources, Supervision, Writing - review & editing.

Acknowledgments

We thank Diamond Light Source, University of Reading and Biocatalysts Ltd. for access to excellent facilities. This work was funded by a BBSRC CASE award BB/K012053/1 (to GB).

Declarations of Interest: None.

References

[1] Stressler T, Eisele T, Schlayer M, Lutz-Wahl S, Fischer L. Characterization of the recombinant exopeptidases PepX and PepN from Lactobacillus helveticus ATCC 12046 important for food protein hydrolysis. PLoS One. 2013;8:e70055.

[2] Nampoothiri KM, Nagy V, Kovacs K, Szakacs G, Pandey A. l-leucine aminopeptidase production by filamentous Aspergillus fungi. Lett Appl Microbiol. 2005;41:498-504.

[3] Lin LL, Hsu WH, Wu CP, Chi MC, Chou WM, Hu HY. A thermostable leucine aminopeptidase from Bacillus kaustophilus CCRC 11223. Extremophiles. 2004;8:79-87.

[4] Matsushita-Morita M, Tada S, Suzuki S, Hattori R, Marui J, Furukawa I, et al. Overexpression and Characterization of an Extracellular Leucine Aminopeptidase from Aspergillus oryzae. Curr Microbiol. 2011;62:557-64.

[5] Huang WQ, Zhong LF, Meng ZZ, You ZJ, Li JZ, Luo XC. The Structure and Enzyme Characteristics of a Recombinant Leucine Aminopeptidase rLap1 from Aspergillus sojae and Its Application in Debittering. Appl Biochem Biotechnol. 2015;177:190-206.

[6] Lazure C. The peptidase zymogen proregions: nature's way of preventing undesired activation and proteolysis. Current pharmaceutical design. 2002;8:511-31.

[7] Khan AR, Khazanovich-Bernstein N, Bergmann EM, James MN. Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. Proc Natl Acad Sci U S A. 1999;96:10968-75.

[8] Subbian E, Williamson, D. and Shinde, U. Protein Folding Mediated by an Intramolecular Chaperone: Energy Landscape for Unimolecular Pro-Subtilisin E Maturation. Advances in Bioscience and Biotechnology. 2015;6:73 - 88.

[9] Shinde U, Inouye M. Intramolecular chaperones: polypeptide extensions that modulate protein folding. Semin Cell Dev Biol. 2000;11:35-44.

[10] Beggah S, Lechenne B, Reichard U, Foundling S, Monod M. Intra- and intermolecular events direct the propeptide-mediated maturation of the Candida albicans secreted aspartic proteinase Sap1p. Microbiol-Uk. 2000;146:2765-73.

[11] Tang B, Nirasawa S, Kitaoka M, Hayashi K. The role of the N-terminal propeptide of the pro-aminopeptidase processing protease: refolding, processing, and enzyme inhibition. Biochem Bioph Res Co. 2002;296:78-84.

[12] Safina D, Rafieva L, Demidyuk I, Gasanov E, Chestukhina G, Kostrov S. Involvement of Propeptides in Formation of Catalytically Active Metalloproteinase from Thermoactinomyces sp. Protein Peptide Lett. 2011;18:1119-25.

[13] Yurimoto H, Yamane M, Kikuchi Y, Matsui H, Kato N, Sakai Y. The pro-peptide of Streptomyces mobaraensis transglutaminase functions in cis and in trans to mediate efficient secretion of active enzyme from methylotrophic yeasts. Biosci Biotechnol Biochem. 2004;68:2058-69.

[14] Shinde UP, Liu JJ, Inouye M. Protein memory through altered folding mediated by intramolecular chaperones. Nature. 1997;389:520-2.

[15] Rozenfeld R, Muller L, El Messari S, Llorens-Cortes C. The C-terminal domain of aminopeptidase A is an intramolecular chaperone required for the correct folding, cell surface expression, and activity of this monozinc aminopeptidase. J Biol Chem. 2004;279:43285-95.

[16] Schalk C, Remy JM, Chevrier B, Moras D, Tarnus C. Rapid purification of the Aeromonas proteolytica aminopeptidase: crystallization and preliminary X-ray data. Archives of biochemistry and biophysics. 1992;294:91-7.

[17] Chevrier B, Schalk C, D'Orchymont H, Rondeau JM, Moras D, Tarnus C. Crystal structure of Aeromonas proteolytica aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. Structure. 1994;2:283-91.

[18] Chevrier B, D'Orchymont H, Schalk C, Tarnus C, Moras D. The structure of the Aeromonas proteolytica aminopeptidase complexed with a hydroxamate inhibitor. Involvement in catalysis of Glu151 and two zinc ions of the co-catalytic unit. European journal of biochemistry / FEBS. 1996;237:393-8.
[19] Bennett B, Holz RC. Spectroscopically distinct cobalt(II) sites in heterodimetallic forms of the aminopeptidase from Aeromonas proteolytica: characterization of substrate binding. Biochemistry. 1997;36:9837-46.
[20] Chen G, Edwards T, D'Souza V M, Holz RC. Mechanistic studies on the aminopeptidase from Aeromonas proteolytica: a two-metal ion mechanism for peptide hydrolysis. Biochemistry. 1997;36:4278-86.
[21] Stamper C, Bennett B, Edwards T, Holz RC, Ringe D, Petsko G. Inhibition of the aminopeptidase from Aeromonas proteolytica by L-leucinephosphonic acid. Spectroscopic and crystallographic characterization of the transition state of peptide hydrolysis. Biochemistry. 2001;40:7035-46.
[22] Stamper CC, Bienvenue DL, Bennett B, Ringe D, Petsko GA, Holz RC. Spectroscopic and X-ray crystallographic characterization of bestatin bound to the aminopeptidase from Aeromonas (Vibrio) proteolytica. Biochemistry. 2004;43:9620-8.
[23] Schurer G, Lanig H, Clark T. Aeromonas proteolytica aminopeptidase: an investigation of the mode of action using a quantum mechanical/molecular mechanical approach. Biochemistry. 2004;43:5414-27.
[24] Desmarais W, Bienvenue DL, Bzymek KP, Petsko GA, Ringe D, Holz RC. The high-resolution structures of the neutral and the low pH crystals of aminopeptidase from Aeromonas proteolytica. J Biol Inorg Chem. 2006;11:398-408.
[25] Kumar A, Periyannan GR, Narayanan B, Kittell AW, Kim J-J, Bennett B. Experimental evidence for a metallohydrolase mechanism in which the nucleophile is not delivered by a metal ion: EPR spectrokinetic and structural studies of aminopeptidase from Vibrio proteolyticus. Biochem J. 2007;403:527-36.
[26] Zhang ZZ, Nirasawa S, Nakajima Y, Yoshida M, Hayashi K. Function of the N-terminal propeptide of an aminopeptidase from Vibrio proteolyticus. Biochem J. 2000;350 Pt 3:671-6.
[27] Bzymek KP, D'Souza VM, Chen G, Campbell H, Mitchell A, Holz RC. Function of the signal peptide and N- and C-terminal propeptides in the leucine aminopeptidase from Aeromonas proteolytica. Protein expression and purification. 2004;37:294-305.
[28] Nirasawa S, Nakajima Y, Zhang ZZ, Yoshida M, Hayashi K. Intramolecular chaperone and inhibitor activities of a propeptide from a bacterial zinc aminopeptidase. Biochem J. 1999;341 (Pt 1):25-31.
[29] Hanaya K, Suetsugu M, Saijo S, Yamato I, Aoki S. Potent inhibition of dinuclear zinc(II) peptidase, an aminopeptidase from Aeromonas proteolytica, by 8-quinolinol derivatives: inhibitor design based on Zn2+ fluorophores, kinetic, and X-ray crystallographic study. J Biol Inorg Chem. 2012;17:517-29.
[30] Chien HC, Lin LL, Chao SH, Chen CC, Wang WC, Shaw CY, et al. Purification, characterization, and genetic analysis of a leucine aminopeptidase from Aspergillus sojae. Biochim Biophys Acta. 2002;1576:119-26.
[31] !!! INVALID CITATION !!! [4, 5].
[32] White RC, Gunderson FF, Tyson JY, Richardson KH, Portlock TJ, Garnett JA, et al. Type II Secretion-Dependent Aminopeptidase LapA and Acyltransferase PlaC Are Redundant for Nutrient Acquisition during Legionella pneumophila Intracellular Infection of Amoebas. MBio. 2018;9.
[33] Bzymek KP, Holz RC. The catalytic role of glutamate 151 in the leucine aminopeptidase from Aeromonas proteolytica. J Biol Chem. 2004;279:31018-25.

[34] Bzymek KP, Swierczek SI, Bennett B, Holz RC. Spectroscopic and thermodynamic characterization of the E151D and E151A altered leucine aminopeptidases from Aeromonas proteolytica. Inorg Chem. 2005;44:8574-80.
[35] Gilboa R, Greenblatt HM, Perach M, Spungin-Bialik A, Lessel U, Wohlfahrt G, et al. Interactions of Streptomyces griseus aminopeptidase with a methionine product analogue: a structural study at 1.53 A resolution. Acta Crystallogr D Biol Crystallogr. 2000;56:551-8.
[36] Ikemura H, Takagi H, Inouye M. REQUIREMENT OF PRO-SEQUENCE FOR THE PRODUCTION OF ACTIVE SUBTILISIN-E IN ESCHERICHIA-COLI. J Biol Chem. 1987;262:7859-64.
[37] Gao X, Wang J, Yu DQ, Bian F, Xie BB, Chen XL, et al. Structural basis for the autoprocessing of zinc metalloproteases in the thermolysin family. Proc Natl Acad Sci U S A. 2010;107:17569-74.
[38] Hu Z, Haghjoo K, Jordan F. Further evidence for the structure of the subtilisin propeptide and for its interactions with mature subtilisin. J Biol Chem. 1996;271:3375-84.
[39] Segundo BS, Martinez MC, Vilanova M, Cuchillo CM, Aviles FX. The severed activation segment of porcine pancreatic procarboxypeptidase A is a powerful inhibitor of the active enzyme. Isolation and characterisation of the activation peptide. Biochimica et biophysica acta. 1982;707:74-80.
[40] Serkina AV, Gorozhankina TF, Shevelev AB, Chestukhina GG. Propeptide of the metalloprotease of Brevibacillus brevis 7882 is a strong inhibitor of the mature enzyme. FEBS Lett. 1999;456:215-9.
[41] McIver KS, Kessler E, Olson JC, Ohman DE. The elastase propeptide functions as an intramolecular chaperone required for elastase activity and secretion in Pseudomonas aeruginosa. Mol Microbiol. 1995;18:877-89.
[42] !!! INVALID CITATION !!! [16, 19, 41, 42].
[43] Hartley M, Yong W, Bennett B. Heterologous expression and purification of Vibrio proteolyticus (Aeromonas proteolytica) aminopeptidase: a rapid protocol. Protein expression and purification. 2009;66:91-101.
[44] Zhang N, Yin S, Zhang W, Gong X, Zhang N, Fang K, et al. Crystal Structure and Biochemical Characterization of an Aminopeptidase LapB from Legionella pneumophila. J Agric Food Chem. 2017;65:7569-78.
[45] Hernandez-Moreno AV, Villasenor F, Medina-Rivero E, Perez NO, Flores-Ortiz LF, Saab-Rincon G, et al. Kinetics and conformational stability studies of recombinant leucine aminopeptidase. International journal of biological macromolecules. 2014;64:306-12.
[46] Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins. J Biosci Bioeng. 2005;99:303-10.
[47] Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983;22:2577-637.
[48] Desmarais WT, Bienvenue DL, Bzymek KP, Holz RC, Petsko GA, Ringe D. The 1.20 A resolution crystal structure of the aminopeptidase from Aeromonas proteolytica complexed with tris: a tale of buffer inhibition. Structure. 2002;10:1063-72.
[49] Rossier O, Dao J, Cianciotto NP. The type II secretion system of Legionella pneumophila elaborates two aminopeptidases, as well as a metalloprotease that contributes to differential infection among protozoan hosts. Appl Environ Microbiol. 2008;74:753-61.
[50] De Paola CC, Bennett B, Holz RC, Ringe D, Petsko GA. 1-Butaneboronic acid binding to Aeromonas proteolytica aminopeptidase: a case of arrested development. Biochemistry. 1999;38:9048-53.
[51] Cao J, Hymowitz M, Conner C, Bahou WF, Zucker S. The propeptide domain of membrane type 1-matrix metalloproteinase acts as an intramolecular chaperone when expressed in trans with the mature sequence in COS-1 cells. J Biol Chem. 2000;275:29648-53.

[52] Looke M, Kristjuhan K, Kristjuhan A. Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques. 2011;50:325-8.
[53] Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010;66:125-32.
[54] Winter G, Lobley CM, Prince SM. Decision making in xia2. Acta Crystallogr D Biol Crystallogr. 2013;69:1260-73.
[55] McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr. 2005;61:458-64.
[56] Roche DB, Buenavista MT, Tetchner SJ, McGuffin LJ. The IntFOLD server: an integrated web resource for protein fold recognition, 3D model quality assessment, intrinsic disorder prediction, domain prediction and ligand binding site prediction. Nucleic acids research. 2011;39:W171-6.
[57] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60:2126-32.
[58] Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66:213-21.
[59] Liebschner D, Afonine PV, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, et al. Polder maps: improving OMIT maps by excluding bulk solvent. Acta Crystallogr D Struct Biol. 2017;73:148-57.
[60] Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic acids research. 2005;33:2302-9.