UNIVERSITY OF LEEDS

This is a repository copy of *How metal cofactors drive dimer–dodecamer transition of the M42 aminopeptidase TmPep1050 of Thermotoga maritima*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/159066/

Version: Accepted Version

Article:
Dutoit, R, Van Gompel, T, Brandt, N et al. (4 more authors) (2019) How metal cofactors drive dimer–dodecamer transition of the M42 aminopeptidase TmPep1050 of Thermotoga maritima. Journal of Biological Chemistry, 294 (47). pp. 17777-17789. ISSN 0021-9258

https://doi.org/10.1074/jbc.ra119.009281

© 2019 Dutoit et al. Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc. This is an author produced version of a paper published in Journal of Biological Chemistry. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

How metal cofactors drive dimer-dodecamer transition of the M42 aminopeptidase TmPep1050 of Thermotoga maritima

Raphaël Dutoit[1,2] *, Tom Van Gompel[3], Nathalie Brandt[2], Dany Van Elder[1], Jeroen Van Dyck[3], Frank Sobott[3,4], Louis Droogmans[1]

[1] Laboratory of Microbiology, Department of Molecular Biology, Université Libre de Bruxelles, rue des Professeurs Jeener et Brachet 12, B6041 Charleroi, Belgium
[2] Labiris Institut de Recherche, avenue Emile Gryzon 1, B1070 Brussels, Belgium
[3] Biomolecular & Analytical Mass Spectrometry, Department of Chemistry, Universiteit van Antwerpen, Groenenborgerlaan 171, B2020 Antwerpen, Belgium
[4] Astbury Centre for Structural and Molecular Biology, University of Leeds, LS2 9JT, Leeds, United Kingdom

Running title: Cobalt-driven dimer-dodecamer transition of TmPep1050

* To whom correspondence should be addressed: Raphaël Dutoit, rdutoit@ulb.ac.be, Tel. +32 2 526 72 83

Keywords: M42 aminopeptidase, TET aminopeptidase, MH clan, metallopeptidase, metal cofactors, oligomerization shift, thermostability, Thermotoga maritima.

Abstract

The M42 aminopeptidases are dinuclear aminopeptidases displaying a peculiar tetrahedral-shaped structure with twelve subunits. Their quaternary structure results from the self-assembly of six dimers controlled by their divalent metal ion cofactors. The oligomeric state transition remains debated despite the structural characterization of several archaeal M42 aminopeptidases. The main bottleneck is the lack of dimer structures, hindering the understanding of structural changes occurring during the oligomerization process. We present the first dimer structure of an M42 aminopeptidase, TmPep1050 of Thermotoga maritima, along with the dodecamer structure. The comparison of both structures allows to describe how the metal ion cofactors modulate the active site fold and, subsequently, affect the interaction interface between dimers. A mutational study shows that the M1 site strictly controls dodecamer formation. The dodecamer structure of TmPep1050 also reveals that a part of the dimerization domain delimits the catalytic pocket and could participate in substrate binding.

Introduction

Cells possess an arsenal of proteolytic enzymes to ensure specific and non-specific hydrolysis of proteins and peptides (for instance, up to 2% of human genome encodes proteases) (1). Proteolysis plays an important role in a wide array of cellular functions as diverse as homeostasis, stress responses, cell-cycle regulation, quorum sensing, stalled ribosome rescue, immune response, virulence, etc. (2–6). About 80% of protein degradation is achieved by the proteasome 26S in eukaryotes (3). Protein turnover is carried out by either the proteasome 20S in archaea and actinomycetes or several proteolytic complexes – notably HslUV, ClpAP, ClpXP, Lon, and FtsH – in other prokaryotes (4, 7). The proteasome and its related complexes generate peptides 6-9 amino acids in length (8) which are further processed by a set of endopeptidases and exopeptidases. Several peptidases have been proposed to act downstream the proteasome: the tricorn protease (9), the thimet oligopeptidase (10), the tripeptidyl peptidase II (11), the aminopeptidase N (12), and the TET-aminopeptidases (13). The latter peptidases are widely distributed in all kingdoms of life (14). They adopt a peculiar tetrahedron-

shaped structure compartmenting the active sites in a buried catalytic chamber (15). In Pyrococcus horikoshii, four TET-aminopeptidases - PhTET1, PhTET2, PhTET3, and PhTET4 - have been described and each of them displays a different substrate specificity – aspartyl-, leucyl-, lysyl-, and glycyl-aminopeptidase activity, respectively (15–18). Remarkably, heterocomplexes, made of PhTET2 and PhTET3, have been reported, leading to the assumption of the peptidasome particle existence (19, 20).

According to MEROPS classification, the TET-aminopeptidases are found in the M18 and M42 families (14, 21). Both families belong to the MH clan, encompassing metallopeptidases sharing a common α/β catalytic domain. The archetypal MH clan enzyme is the Vibrio proteolyticus aminopeptidase 1 (22). The M18 family is widely distributed in all domains of life (14) while the M42 family is unique to prokaryotes (23). Several structures of archaeal M42 aminopeptidases have been studied (13, 15, 17, 18, 24–26) but only one structure has been reported for bacteria (27). Franzetti et al. (2002) described the first TET-aminopeptidase structure of the Haloarcula marismortui M42 aminopeptidase, consisting of twelve subunits adopting a tetrahedron-shaped quaternary structure (25). High resolution structures of P. horikoshii PhTET2 have revealed that the M42 aminopeptidase subunit is composed of an α/β catalytic domain and a PDZ-like dimerization domain (13, 15). A similar quaternary structure has been reported for several M18 aminopeptidases from prokaryotes and eukaryotes (28–31) although their dimerization domain adopts a butterfly fold instead of a PDZ-like fold. The quaternary structure is often seen as the assembly of six dimers with a dimer positioned on each edge of the tetrahedron (13). At the center of the tetrahedron faces are found four gates leading to a wide inner cavity. Four exit pores are located at the tetrahedron vertices through which amino acids, generated during peptide hydrolysis, are released.

The catalytic site is characterized by (i) two catalytic residues, (ii) five residues forming the metal ion binding sites M1 and M2, and (iii) two divalent metal ions (usually Zn^{2+}, Co^{2+}, and Mg^{2+}). The metal ions have been described to bind the substrate, to facilitate nucleophile generation (water molecule deprotonation), and to stabilize the transition-state of peptide hydrolysis (32). In addition to their catalytic roles, the metal ions could also control the TET-aminopeptidase oligomerization. Such a structural role has been reported for three M42 aminopeptidases: PhTET2 and PhTET3 from P. horikoshii (13, 33, 34), and PfTET3 from P. furiosus (24). Under chelating conditions, dodecamers disassemble into either dimers (PhTET3) or monomers (PhTET2 and PfTET3). Colombo et al. (2016) inferred the role of the M1 and M2 sites in the oligomerization: the presence of a metal ion in the M2 site is required for the oligomerization while the M1 site controls protein flexibility (24). According to Macek et al. (2017), the dissociation is reversible as dodecamers are formed through random association of dimers (33). The understanding of dodecamer formation mechanism, however, suffers from the lack of dimer (or monomer) structure. A low-resolution structure of PhTET2 dimer has been reported but no significant structural change has been pinpointed (35). The current knowledge on M42 aminopeptidases relies on a set of archaeal enzymes while their bacterial counterparts remain ill-described.

In this work we focus on the bacterial M42 aminopeptidase model, TmPep1050 from Thermotoga maritima, previously characterized as a cobalt-activated leucylaminopeptidase (23). We present the first high resolution dimer structure of an M42 aminopeptidase, alongside the dodecamer structure. The dimer structure clearly showed the cofactor role in the active site fold which reverberates to the interaction interfaces between dimers. In addition, the thermostability of TmPep1050 is dependent on the oligomeric state. The dodecamer assembly/disassembly was further studied by native MS revealing the intermediate oligomeric states as well as the flexible parts in the dimer structure. Finally, we inferred the role of the M1 and M2 sites through the structural study of TmPep1050 variants. Our results showed that the M1 site strictly controls the dodecamer assembly. Based on the current knowledge, we should avoid drawing general dogma on the TET-aminopeptidases due to the lack of studies on metal ion binding behaviour and structures in complex with substrates and inhibitors.

Results and Discussions

TmPep1050 has a genuine TET-aminopeptidase structure.

Tm_1050 ORF was overexpressed in E. coli and the recombinant enzyme was purified to homogeneity through three chromatographic steps. Size-exclusion chromatography showed that the purified enzyme, subsequently named TmPep1050$_{12\text{-mer}}$, had an apparent molecular weight of 330 ± 15 kDa (SEM with n=12, molecular weight of a monomer: 36.0 kDa). Native MS analysis demonstrated that TmPep1050$_{12\text{-mer}}$ is a dodecamer with a measured mass of 435 323 ± 64 Da (theorical mass: 432 783 Da). Crystals were obtained in two crystallization conditions (see Table S1). X-ray fluorescence scanning detected Co^{2+} and Zn^{2+} in crystals grown in 0.18 M tri-ammonium citrate, 40% PEG3350 pH 7.5 while no trace of metal ions was found in crystals grown in 2.1 M malic acid pH 6.75 (see Figure S1). The absence of metal ions could be due to the use of malic acid, a weak chelating agent of divalent metal ions (K_d 10^{-3} M for Co^{2+} binding) (36), in the crystallization buffer. Under such conditions, TmPep1050$_{12\text{-mer}}$ was barely active although dodecameric oligomerization remained unaffected at room temperature (see below). A similar statement has been reported for PhTET2 (13): when co-crystallized with o-phenanthroline, PhTET2 remained dodecameric even if the two Zn^{2+} were chelated. The apo enzyme structure (apo-TmPep1050$_{12\text{-mer}}$) was determined by molecular replacement using the coordinates of YpdE, an uncharacterized M42 aminopeptidase of Shigella flexneri. The metal bound complex structure was determined by MR-SAD using apo-TmPep1050$_{12\text{-mer}}$ coordinates.

The TmPep1050$_{12\text{-mer}}$ quaternary structure consists of twelve subunits adopting a tetrahedron-shaped architecture (see Figure 1.A) like other available structures of M42 aminopeptidases (13, 17, 18, 24, 26, 27). The tetrahedron-shaped architecture is often described as the self-assembly of six dimers in such a manner that a dimer lies on each tetrahedron edge. Four entrances are located at the centre of the faces and four exit channels are located at the vertices. The entrances and exits lead to an inner cavity with the twelve catalytic sites oriented inwards, compartmenting the active sites. The interaction between dimers is maintained through a polar and hydrophobic interaction network at the vertices (see Figure S2.A) and nine salt-bridges between adjacent subunits (see Figure S2.B). The residues forming these salt-bridges are, however, not conserved among all M42 aminopeptidases. The apo-TmPep1050$_{12\text{-mer}}$ structure appears to be nearly identical to the TmPep1050$_{12\text{-mer}}$ structure obtained with bound metal ions (structural alignment RMSD = 0.235). There are, however, several differences observed in the catalytic site which are described below.

The α4 helix of the PDZ-like domain delimits the TmPep1050$_{12\text{-mer}}$ catalytic pocket.

The TmPep1050$_{12\text{-mer}}$ subunit is composed of a catalytic domain and a dimerization domain (Figure 1.B). The catalytic domain adopts an α/β globular structure similar to the one of V. proteolyticus aminopeptidase Ap1 (see Figure S3) except that the β sheet displays an extension of three anti-parallel strands (β10, β11, and β17) connected to the β16 strand. The β sheet extension is conserved among the M42 family (13, 15, 21) and interacts with the neighbour subunit dimerization domain (13, 27). The TmPep1050$_{12\text{-mer}}$ dimerization domain adopts the typical fold of a PDZ-like domain (13). In all structurally characterized M42 aminopeptidases, the α4 helix of the PDZ-like domain is highly flexible (13, 15, 17, 18, 24, 26, 27). The α4 helix, however, was modelled in TmPep1050$_{12\text{-mer}}$ structure, probably due to its stabilisation in the crystal. In the quaternary structure, the α4 helices are positioned in such a manner that the entrances are restricted to 13 Å (see Figure S4), which is smaller than those reported for PhTET1, PhTET2 and SpPepA (13, 18, 27). The carboxylate functions of three glutamate residues (E110, E114 and E117) are oriented towards the entrance. In addition, the TmPep1050$_{12\text{-mer}}$ active sites appear to be more buried and delimited than those of other characterized M42 aminopeptidases as the α4 helix of one monomer delimits the active site pocket of the other monomer at the dimer level (see Figures 1.C and S5.A). Hence, the α4 helix flexibility could indicate a possible open/closed conformation of the enzyme. It is worth noticing that an α helix of the butterfly-fold dimerization domain of the

M18 aminopeptidases could also be highly flexible (30, 31) or, to the contrary, more stable (28, 29). Such an α helix, flexible or not, seems to be a conserved structural feature among TET aminopeptidases. In the human M18 aminopeptidase structure, this α helix is directly involved in substrate binding (28). A similar function could be expected for the α4 helix in M42 aminopeptidases.

Seven residues conserved in the MH clan define the TmPep1050$_{12\text{-mer}}$ active site.

The M42 aminopeptidase catalytic site consists of two divalent metal ions (M1 and M2) bound by five strictly conserved residues, and two catalytic residues. In the TmPep1050$_{12\text{-mer}}$ catalytic site, the M1 site consists of D168, E198, and H307 residues bound to a Co^{2+} ion while the M2 site consists of H60, D168, and D220 residues bound to a Zn^{2+} ion (see Figure 1.D). In the apo-TmPep1050$_{12\text{-mer}}$ structure, these residues are correctly positioned despite the absence of metal ions. The distance between the imidazole rings of H60 and H307 is, however, increased by 1.5 Å while the carboxylate of E198 is displaced by 0.4 Å away from the M1 site center in comparison to the TmPep1050$_{12\text{-mer}}$ structure with its metallic cofactors. Thus, the presence of metal ions seems to pull the side chains of metal binding residues closer, especially those of the M1 site (see Figure 1.D). The two metal ions also coordinate a water molecule involved in peptide bond hydrolysis. The water molecule is asymmetrically positioned, being closer to the M2 site rather than the M1 site (see Figure S5.B). In the monomer D, the distance between Zn^{2+} and the oxygen atom is so short that only a hydroxide ion could be modelled instead of a water molecule – suggesting the M2 site stabilizes the hydroxide ion prior to the nucleophilic attack. The TmPep1050$_{12\text{-mer}}$ structure contrasts with V. proteolyticus Ap1, PhTET1, and PhTET2 where the water molecule is positioned symmetrically (see Figure S5.B). In the TmPep1050$_{12\text{-mer}}$ catalytic site, the acid/base catalyst is E197 which is conserved in all MH clan members. Indeed, the mutation of E197 in glutamine completely abolished the activity (specific activity of L-Leu-pNA of x s^{-1}) while the oligomeric state remained unaffected. The seventh conserved residue is D62, which has been described as modulator of Lewis acid strength of M2 in V. proteolyticus Ap1 (37). The TmPep1050$_{12\text{-mer}}$ S1 binding pocket is similar to the one of PhTET2 since six out of seven residues are conserved (see Figure S5.C), which is in accordance with their substrate specificities as both enzymes are leucyl-aminopeptidases.

TmPep1050$_{12\text{-mer}}$ is a cobalt activated leucyl-aminopeptidase.

TmPep1050$_{12\text{-mer}}$ aminopeptidase activity was assayed with various L-aminoacyl-p-nitroanilide (pNA) derivatives. The substrate specificity is mainly towards nonpolar aliphatic L-aminoacyl-pNA, with a clear preference for L-Leu-pNA (Table 1). TmPep1050$_{12\text{-mer}}$ activity is maximal at pH between 7.0 and 7.8, and up to 90°C (see Figures S6 and S7). Its kinetic parameters k_{cat} and K_m were determined for L-Leu-pNA, L-Ile-pNA and L-Met-pNA (Table 2). These values differ greatly from our previous results obtained with His-tagged TmPep1050 (23), indicating that the polyhistidine tag could interfere in the binding of divalent metal ions to the M1 and M2 sites. Indeed, we reported a k_{cat} of 0.25 s^{-1} for His-tagged TmPep1050, even with a Co^{2+} to enzyme ratio of 500.

Dialysis of TmPep1050$_{12\text{-mer}}$ against 2.1 M malic acid pH 7.0 (one of the crystallization conditions), resulted in a loss of activity (dropping to 1 s^{-1}) while its oligomeric state remained unaffected as determined by gel filtration with an apparent molecular weight of 300 ± 10 kDa (SEM with n=3). The addition of Co^{2+} partly restored TmPep1050$_{12\text{-mer}}$ LAP activity while Zn^{2+} had no effect on LAP activity although it was found in the active site (see Table 3). Similar observations have been reported for other M42 aminopeptidases: while their structures contain Zn^{2+}, PhTET2, PhTET3, PfTET3 and SpPepA are activated by Co^{2+} (13, 17, 24, 27, 38). Other metal ions had no significant effect on TmPep1050$_{12\text{-mer}}$ LAP activity (see Table 3).

To quantify its affinity for Co^{2+}, metal ion-depleted TmPep1050$_{12\text{-mer}}$ was incubated with an increasing amount of Co^{2+}, ranging from 0 to 2560 µM. After 24 hour-incubation at 75°C, Co^{2+} binding was quantified with an Amplex UltraRed fluorescent probe. In parallel, LAP activity was measured to follow the reactivation of metal-depleted enzyme (see Figure 2). Co^{2+} content of metal ion-depleted enzyme was less than one atom per ten monomers (1.5 µM of Co^{2+}

for 20 µM of proteins). TmPep1050$_{12\text{-mer}}$ can bind Co^{2+} with an apparent association constant of 50 ± 5 µM.

The thermostability and oligomerization state of TmPep1050 are Co^{2+}-dependent.

As TmPep1050$_{12\text{-mer}}$ displays a LAP activity up to 90°, we expected TmPep1050$_{12\text{-mer}}$ to be a highly thermostable enzyme. Its thermostability was determined at 75°C and 95°C: at 95°C the half-life of TmPep1050$_{12\text{-mer}}$ is about 24 hours while at 75°C it is about 20 days. TmPep1050$_{12\text{-mer}}$ is highly thermostable, probably due to many ionic and hydrophobic interactions between dimers maintaining the whole quaternary structure in dodecamers. Shorter half-lives (few hours) have been reported for various M42 aminopeptidases (17, 38–41).

Remarkably, in the absence of its metal ion cofactors, the half-life at 75°C is dramatically reduced to merely one hour. Moreover, under these conditions, TmPep1050 dodecamers dissociated to dimers (namely TmPep1050$_{2\text{-mer}}$) after two-hour incubation at 75°C as an apparent molecular weight of 52 ± 3 kDa (SEM, n=12) was measured by gel filtration. Thermal shift assays (TSA) were conducted to further characterize the role of cobalt ions in TmPep1050$_{12\text{-mer}}$ thermostability. Cobalt-loaded TmPep1050$_{12\text{-mer}}$ had a T_m of about 97°C, confirming its remarkable thermostability. After dialysis against 2.1 M malic acid pH 7.0 (as mentioned above), the T_m dropped to 91°C meanwhile Co^{2+} addition restored its thermostability (see Figure 3).

The link between the metal cofactors and thermostability has been reported for several M42 aminopeptidases: Geobacillus stearothermophilus aminopeptidase I (42), Thermococcus onnurineus deblocking aminopeptidase (43) and PhTET3 (34). The dodecamer dissociation has often been achieved in harsh conditions such as EDTA treatment (24, 34, 42) or in acidic buffer (pH < 4) (33). Dimer formation has been reported for PhTET3 (34) and PhTET2 (13) while a further breakdown of dimers into monomers has been achieved for PfTET3 (24) and PhTET2 (33). Neither of these studies have, however, described the dodecamer dissociation in physiological compatible conditions. Even though our results focused on TmPep1050$_{12\text{-mer}}$ treated with malic acid, the same dissociation of dodecamers into dimers was achieved with an extensive dialysis against 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2 followed by a heat treatment at 75°C (see Table S2). Therefore, our data tend to support a dimer-dodecamer equilibrium depending on metal cofactor availability.

The oligomeric state transition is reversible and Co^{2+}-driven.

TSA suggested that dodecamer dissociation could be reversible but the different oligomeric states occurring during the association/dissociation process had to be identified. For that purpose, we opted for native mass spectrometry. After a buffer exchange in ammonium acetate, TmPep1050$_{12\text{-mer}}$ retained its oligomeric state (molecular weight = 300 kDa according to gel filtration) but had a LAP activity reduced by two-thirds, suggesting that metal cofactors were partly lost. Native mass spectrometry showed that, under these conditions, TmPep1050 was mainly a mixture of dimers and dodecamers (see Figure 4.A). MS/MS analysis of dodecamer +46 peak revealed two masses for dissociated monomer of 36076.2 ± 1.7 Da and 36136.9 ± 4.4 Da (see Figure S8). The first mass is almost equal to the theoretical molecular weight of 36065 Da while the second mass could correspond to a monomer with one cobalt ion bound.

When samples were loaded with Co^{2+} and heated at 75°C for 30 min. prior to MS analysis, peaks corresponding to dimers tended to disappear in favour of dodecamer peaks (see Figure 4.B). MS analysis of Co^{2+}-reloaded TmPep1050$_{12\text{-mer}}$ showed that monomer had a mass of about 176 Da higher than the theoretical molecular weight. Such difference could be explained by the presence of at least two cobalt ions per subunit. In addition, MS experiments showed the existence of intermediate oligomeric states: tetramers, hexamers and octamers (see Figure 4.C). The self-assembly pathway of M42 aminopeptidases is still poorly understood. Appolaire et al. (2013) proposed that dimers self-assemble into dodecamers via intermediate hexamers for PhTET2 (35). Macek et al. (2017) debated this theory arguing that dimers self-assemble randomly to form tetra-, hexa-, octa-, and decamer intermediates for PhTET2 (33). Our native MS data suggest that a dodecamer could result from the association of either two hexamers or a tetramer and an octamer. Such a transition of

oligomeric states could occur and be controlled in vivo due to the low availability of divalent metal ions.

Native MS analysis of TmPep1050$_{2\text{-mer}}$ (depleted with Co^{2+} and heat-treated) revealed peaks corresponding to dimers only (see Figure 4.D). The addition of Co^{2+} combined with a heat-treatment at 75°C completely changed the mass spectra. Indeed, dimer peaks became almost undetectable while dodecamers were the most abundant oligomer in presence of a six-times excess of cobalt (see Figure 4.D). In comparison only partial reassociation was reported for PhTET3 (34) and PhTET2 (35). A second experiment was conducted to quantify the different oligomers occurring during the reassociation. TmPep1050$_{2\text{-mer}}$ was incubated with an increasing concentration of Co^{2+} at 75°C for 30 min. and the oligomers were identified by size exclusion chromatography (see Figure 5.A). The intermediate oligomers were not detected due to their low abundance (as seen in MS experiments). As expected, the ratio of dodecamers to dimers increased according to the Co^{2+} concentration (see Figure 5.A and 5.B).

TmPep1050$_{2\text{-mer}}$ structure highlights the structural changes triggered by metal ion loss.

We showed that cobalt ions are important for activity, thermostability and oligomerization of TmPep1050. Native MS experiments demonstrated the reversibility of dodecamer dissociation and how it is dependent on metal cofactors. As dimers are unable to form dodecamers in absence of cobalt, structural changes must occur at oligomerization interfaces. To understand these changes, the structure of TmPep1050$_{2\text{-mer}}$ was solved by X-ray crystallography.

The overall shape of TmPep1050$_{2\text{-mer}}$ does not differ from a dimer in the TmPep1050$_{12\text{-mer}}$ structure (see Figure 6.A). However, several structural dissimilarities having a dramatic impact on the oligomerization capability are observed. Two segments are too flexible to be modelled in the TmPep1050$_{2\text{-mer}}$ structure: G203-G208 and F279-E292 corresponding to the α8 and α10 helices, respectively. To rule out potential crystallographic artefact linked to protein stacking in the crystal, TmPep1050$_{2\text{-mer}}$ backbone flexibility was probed by collision-induced dissociation. Mass spectra showed that proteins were fragmented at two preferential regions, G203-P212 and N283-T290, indicating the soundness of the crystallographic data (see Figure 6.B). At each vertex of the tetrahedron-shaped dodecamer, the α8 and α10 helices of three adjacent subunits form the exit tunnel via an interaction network between Y209 and R289 (see Figure S2.A). This interaction network has been described to be highly important for oligomerization (21, 35). In the TmPep1050$_{2\text{-mer}}$ structure, interaction between Y209 and R289 is prevented as Y209 is completely buried.

Structural alignment between dimer subunit and dodecamer subunit shows that the Q196-V202, K229-A235, and K247-S254 segments diverge greatly (see Figure 6.A). High B-factors were observed for these segments indicating that they are highly flexible (see Figure S9). The Q196-V202 segment contains two conserved residues of the catalytic site, E197 and E198, which are the catalytic general base and a metal ion binding residue of the M1 site, respectively. In the TmPep1050$_{2\text{-mer}}$ structure, the E197-G200 loop is so disordered that E197 and E198 side chains point outwards the catalytic site – which explains why the dimers are far less active than the dodecamers (see Figure 6.C). Thus, metal ion binding in the M1 site is strongly impaired in TmPep1050$_{2\text{-mer}}$. The displacement of E197 and E198 has probably an impact on H60 and D62 as their predicted pK_a increase from 8.5 to 10.1 and from 2.6 to 5.0, respectively. H60 plays a pivotal role in a H-bond network in the dodecamer subunit, interacting with D62, D168, D169, E197, E198, and D220. In the TmPep1050$_{2\text{-mer}}$ structure, the whole H-bond network is disrupted. In the dodecamer structure, the Q196-202 loop probably imposes the fold of the subsequent α8 helix via an entangled network of H-bonds and polar interactions. In addition, the Q196-V202 segment is closely connected to the α10 helix.

The K229-A235 and K247-S254 segments are important for oligomerization as K232, R233, and R249 are involved in the formation of salt-bridges in the dodecamer (see Figures S1.B). Systematic mutagenesis of these residues was set up to support the role of salt-bridges in the dodecamer structure. Single point mutations of K232, R233, and R249 in alanine residues did not impact the oligomerization (data

not shown). However, the triple mutation in either alanine or glutamate residues ($TmPep1050_{K232A/R233A/R249A}$ and $TmPep1050_{K232E/R233E/R249E}$, respectively) greatly disturbed the formation of dodecamers. Indeed, the oligomeric state of $TmPep1050_{K232A/R233A/R249A}$ is mainly dimeric as shown by gel filtration (see Figure 7). It also forms dodecamers and tetramers representing about 32% and 20 % of the purified sample, respectively. For $TmPep1050_{K232E/R233E/R249E}$, the mutations have a dramatic impact on the oligomerization as the dimeric fraction represents about 88% of the purified sample (see Figure 7). $TmPep1050_{K232A/R233A/R249A}$ has a reduced specific activity on L-Leu-pNA of 37.2 ± 0.9 s^{-1} while $TmPep1050_{K232E/R233E/R249E}$ is barely active with a specific activity of 0.15 ± 0.01 s^{-1}. Our results strongly support the role of K232, R233 and R249 to stabilize the dodecameric structure. Nevertheless, one should avoid generalizing such a conclusion to the whole M42 family as these residues are not conserved. Appolaire et al. (2013) reported the destabilization of PhTET2 by mutating five residues involved in inter-dimer interactions. Yet, in their study, the destabilization was only transitory as PhTET2 dimers slowly reassembled into dodecamers over time (35).

The metal ion cofactors are directly involved in structural changes during dodecamer dissociation.

The $TmPep1050_{2\text{-}mer}$ structure shows how the loss of metal ions induces dodecamer dissociation. The structural role of either M1 or M2 sites, however, were not so obvious as E198 (M1 site) interacts with H60 (M2 site). Therefore, systematic mutagenesis of H60 and H307 (M1 site) in alanine residues was undertaken - the resulting variants were named $TmPep1050_{H60A}$ and $TmPep1050_{H307A}$, respectively. The variants were barely active using L-Leu-pNA as substrate (specific activity of less than 0.1 s^{-1}). Their molecular states were determined by native MS: $TmPep1050_{H60A}$ was a mixture of dimers and dodecamers while $TmPep1050_{H307A}$ formed only dimers (see Figure 8.A). Size exclusion chromatography of $TmPep1050_{H60A}$ showed that the dodecamer to dimer ratio was about 0.75.

As $TmPep1050_{H307A}$ only exists as a dimer, the M1 site is probably the most important feature leading to $TmPep1050_{12\text{-}mer}$ formation. The structures of $TmPep1050_{H60A}$ (dimer only) and $TmPep1050_{H307A}$ were solved by X-ray crystallography. The H307A mutation provokes the same structural dissimilarities as observed in $TmPep1050_{2\text{-}mer}$ (see Figure 8.B and Figure S9). The α8 and α10 helices could not be modelled in the $TmPep1050_{H307A}$ structure while the α9 helix is elongated by three residues. E197 and E198 residues are also oriented outward the catalytic site. The M2 site presents the same spatial arrangement as in the dimer, indicating that only the apo-form was crystallized. Our data strongly support the structural role of the metal cofactors in addition to their catalytic role. Colombo et al. (2016) drew the same conclusion with PfTET3 and hypothesized that the metal ion bound to the M1 site could stabilize the oligomerization (24).

The presence of Co^{2+} in the M2 site is not strictly required for oligomerization as $TmPep1050_{H60A}$ is able to form a dodecamer. Still the H60A mutation impeded cobalt-binding in the M1 site, probably due to H-bond network disruption. The $TmPep1050_{H60A}$ dimer structure afforded a better understanding of conformational changes in the Q196-Y209 region. E197-I201 forms a wide loop instead of a tight U-turn. The α8 helix is completely unwound in the $TmPep1050_{H60A}$ dimer (see Figure 8.B).

Mutagenesis of H60 and H307 – involved in metal ion binding in the M2 and M1 sites, respectively – led to an interesting observation. In M42 aminopeptidases, the M1 site has been described to have a lower affinity to metal ion than the M2 site (15, 24). According to Colombo et al. (2016), the removal of Co^{2+} from the M1 site led to a partial dissociation of PfTET3 dodecamer into monomer. The dodecamers were fully dissociated when the second metal ion in the M2 site was removed (24). Surprisingly, the H307A mutation abolished the dodecamer formation for $TmPep1050_{H307A}$ while a partial dissociation was observed for $TmPep1050_{H60A}$. Our data strongly suggest that the M1 site strictly controls the TmPep1050 oligomerization. Further studies would be required to understand how M1 and M2 sites finely tune the oligomerization according to their affinity and selectivity.

Experimental procedures

Cloning and mutagenesis of Tm_1050

The Tm_1050 open reading frame (ORF) was amplified from BspHI-digested TmCD00089984 plasmid (Joint Center for Structural Genomics) using Pfu DNA polymerase (ThermoFisher Scientific) and primers ocej419 and ocej420 (see Table S3). The PCR product was inserted into the pBAD vector (ThermoFisher Scientific) by homologous recombination in E. coli according to the SLiCE protocol (44) giving rise to pCEC43. Briefly, the insertion of the PCR product is allowed via two 30 bp sequences, homologous to the insertion site of the pBAD vector, flanking the gene of interest. The homologous recombination is mediated by using a cell extract of PPY strain, expressing the λ prophage Red recombination system. Site-directed mutagenesis was carried out following the SPRINP protocol (45) except for the two vectors used for TmPep1050$_{K232A/R233A/R249A}$ and TmPep1050$_{K232E/R233E/R249E}$ production. In that case, two synthetic genes harbouring the desired mutations (GeneArt – ThermoFisher Scientific) were introduced into the pBAD vector by homologous recombination. The primers used to generate TmPep1050 variants are listed in Table S1. All genetic constructs were verified by sequencing (Genetic Service Facility, University of Antwerp) and are listed in Table S4. The E. coli MC1061 strain (46) was used for cloning and expression. Cells were grown on LB broth in presence of 100 µg/mL ampicillin for positive selection.

Production and purification of recombinant enzymes

Cultures and protein extracts were prepared following previously published procedures (23) with two modifications: (i) cells from 1L culture were disrupted in 40 mL of 50 mM MOPS, 1 mM $CoCl_2$ pH 7.2 and (ii) protein extracts were heated at 70°C for 10 min. The purification consisted in thr chromatographic steps. The first step was an anion exchange chromatography on Source 15Q resin (GE Healthcare Life Sciences, Tricorn 10/150 column) in 50 mM MOPS, 1 mM $CoCl_2$ pH 7.2. Elution was performed with a gradient step from 0 to 0.5 M NaCl for five column volumes (CV). Fractions containing the protein of interest (2 CV) were pooled and $(NH_4)_2SO_4$ powder was added to a concentration of 1.5 M $(NH_4)_2SO_4$. The second chromatographic step was a hydrophobic interaction chromatography on Source 15Phe resin (GE Healthcare Life Sciences, XK 16/20 column) in 50 mM MOPS, 1.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$ pH 7.2. Elution was performed with a gradient step from 1.5 to 0 M $(NH_4)_2SO_4$ for five CV. Fractions (1.5 CV) containing the protein of interest were pooled and concentrated using an Amicon Ultra-15 ultrafiltration unit with 30 kDa cut-off (Merck Millipore). The third step consisted of a gel filtration on Superdex 200 resin (GE Healthcare Life Sciences, XK 16/70 column) in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$ pH 7.2. The concentration of $(NH_4)_2SO_4$ had to be maintained at 0.5 M to avoid protein precipitation. Purified proteins were finally concentrated using an Amicon Ultra-15 ultrafiltration unit with 30 kDa cut-off. The presence and purity of the recombinant enzymes were checked throughout the purification procedure by SDS-PAGE and enzymatic assays with L-Leu-pNA as substrate (see below). Proteins were quantified by measuring the absorbance at 280 nm and applying the extinction coefficient of 18 910 M^{-1} cm^{-1}. This purification protocol allowed the purification of about 10 mg of TmPep1050 from 1 L of culture. The concentrations of purified TmPep1050 are given in Table S1). Molecular weights were determined by gel filtration on Superdex 200 resin (GE Healthcare Life Sciences, XK 16/70 column) and Superpose 6 10/300 GL (GE Healthcare Life Sciences) using 50 mM MOPS 0.5 M $(NH_4)_2SO_4$ pH 7.2 as running buffer. The gel filtration columns were calibrated using both HMW gel filtration calibration kit (GE Healthcare Life Sciences) and gel filtration standard (Bio-Rad).

Enzymatic assays

Aminoacyl-p-nitroanilide (pNA) substrates were purchased from Bachem AG. Aminopeptidase (EC 3.4.11.1) activity was assayed as previously described (23) except that enzymatic reactions were stopped by adding 1 mL of 20% acetic acid to 1 mL of reaction mix (200 µL of substrate in 10% methanol, 790 µL of 50 mM MOPS pH 7.2, and 10 µL of enzyme at a concentration ranging from 10 to 50 nM, depending on the substrate). For the determination of substrate specificity of TmPep1050$_{12\text{-mer}}$, aminopeptidase assays were

carried out at 75°C with an enzyme concentration of 30 nM and the appropriate amino acid-pNA substrate. All substrates were used at 2.5 mM, except L-Gly-pNA (1.25 mM), L-Phe-pNA (0.75 mM), L-His-pNA and L-Glu-pNA (0.5 mM). The effect of metal ions, pH and temperature on the activity was studied as previously described (23). For the determination of kinetic parameters, assays were performed at 75°C with an enzyme concentration of 10 nM for L-Leu-pNA and L-Ile-pNA and an enzyme concentration of 50 nM for L-Met-pNA. Reaction mixes were supplemented with 250 µM $CoCl_2$. Kinetic parameters (k_{cat}, K_m and k_{cat}/K_m) were determined from the initial reaction rates, using Lineweaver-Burk linearization of the Michaelis-Menten equation. Activation energies were calculated from the slope of the trend line obtained by plotting the logarithm of the specific activity vs. the inverse of the temperature. For thermostability assays, $TmPep1050_{12\text{-}mer}$ was diluted to 1 µM in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$ pH 7.2, and it was incubated at either 75°C or 95°C. At various time-intervals, the activity was measured at the incubation temperature (75°C or 95°C) by diluting the enzyme to 10 nM in 1 mL of reaction mix containing 2.5 mM L-Leu-pNA as substrate.

Cobalt binding assays

A 1 mM $TmPep1050_{12\text{-}mer}$ sample was diluted in ten volumes of 2.1 M malic acid pH 7.0 and concentrated back to one volume using an Amicon Ultra-15 ultrafiltration unit with a 30 kDa cut-off. The malic acid treated sample was dialysed four times against 100 volumes of 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2 using SnakeSkin™ dialysis tubing with a 3.5 kDa cut-off (ThermoFisher Scientific). To monitor metal ion removal, the specific activity was measured with L-Leu-pNA as substrate. Cobalt binding assays were performed by incubating 100 µL of 20 µM cobalt-depleted $TmPep1050_{12\text{-}mer}$ with $CoCl_2$ at a concentration ranging from 0 to 2560 µM in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2 for 24 H at 75°C. After incubation, the specific activity was measured using 10 nM of enzyme and 2.5 mM L-Leu-pNA as substrate without added cobalt in the reaction mix. Other metal ions were tested following this procedure to identify the metal cofactor of $TmPep1050_{12\text{-}mer}$.

In parallel, cobalt concentration was determined using AmplexUltraRed (ThermoFisher Scientific), a fluorescent probe that binds specifically cobalt at high pH. The protocol established by Tsai & Lin (2013) was adapted to fit a 384-well black microplate (Corning) (47). Fluorescence was measured on a SpectraMax 5 (Molecular Device) with excitation wavelength set on 495 nm and emission wavelength set on 570 nm. Prior to cobalt ion quantification, samples were diluted twice in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2. In addition to total cobalt concentration, unbound cobalt was quantified after filtering samples using an Amicon Ultra-0.5 ultrafiltration unit with a 30 kDa cut-off. The affinity constant K_d was determined from Scatchard plot data representation.

To study the reassocation of dimers into dodecamers, 100 µL of 50 µM $TmPep1050_{2\text{-}mer}$ was incubated with Co^{2+} at a concentration ranging from 0 to 5 mM for 30 min. at 75°C. The oligomers were detected and quantified by size exclusion chromatography using Superdex200 resin (GE Healthcare Life Sciences, XK16/20 column). 50 mM MOPS 0.5 M $(NH_4)_2SO_4$ pH 7.2 buffer was used for this assay.

Thermal shift assay

SyproOrange™ (ThermoFisher Scientific) was diluted 1:125 in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2. The fluorescence probe was mixed with protein samples conditioned in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2 and Co^{2+} at a concentration ranging from 0 to 2560 µM. The working SyproOrange™ dilution was 1:1000 and the protein concentration was 20 µM for a reaction volume of 20 µL. Thermal shift assays were performed in 96-well plate on a StepOnePlus™ Real-Time PCR System (ThermoFisher Scientific). Fluorescence curves were treated with StepOne™ Software.

Native mass spectrometry

Samples for native mass spectrometry were transferred into 20 mM ammonium acetate (AmAc) pH 7.2. This was done by Zeba 7 kDa desalting columns (ThermoFisher Scientific). If further desalting was needed, Bio-spin P-6 gel columns (Bio-Rad) were used. The protein samples were diluted in 100 mM AmAc pH 7.2 to working concentration of 5 µM unless stated otherwise. In house prepared borosilicate gold

coated needles filed with 2-3 µl sample were used to introduce the protein into the gas phase using nano electrospray ionization. The spectra were recorded in positive ion mode on a traveling-wave ion mobility Q-TOF instrument (Synapt G2 HDMS, Waters). Different settings were tuned to optimize sample measurement. The most important settings applied during the measurements, unless stated differently at the figure legend, were 20 Volts sampling cone, 10 Volts trap collision energy, and pressures 8.30 e^{-3} mbar and 5.31 e^{-2} mbar of source and trap respectively.

Crystallization

$TmPep1050_{12\text{-}mer}$, $TmPep1050_{2\text{-}mer}$, $TmPep1050_{H60A}$ and $TmPep1050_{H307A}$ were crystallized using the hanging-drop vapour diffusion method, at 292 K in EasyXtal Tool plates (Qiagen). Drops contained 2 µL of recombinant enzyme mixed with 2 µL of well buffer. Crystallization conditions are described in Table S1. One cycle of seeding was necessary to get monocrystals of each species.

Structure determination and analysis

For apo-$TmPep1050_{12\text{-}mer}$, diffraction data were collected on the FIP-BM30a beamline at ESRF (Grenoble, France) (48, 49). For $TmPep1050_{12\text{-}mer}$, $TmPep1050_{2\text{-}mer}$, $TmPep1050_{H60A}$, and $TmPep1050_{H307A}$, diffraction data were collected on Proxima 2 beamline at SOLEIL (Saint-Aubin, France). The data collection and refinement statistics are presented in Table 4. Diffraction data were processed using the XDS program package (50). Molecular replacement and model building were performed using PHENIX software package v.1.10.1-2155 (51). The initial solution of $TmPep1050_{12\text{-}mer}$ was determined by molecular replacement with MR-Rosetta using the coordinates of YpdE of Shigella flexneri (PDB ID: 1YLO) as the search model (TFZ-value = 23.9, log-likelihood gain = 5150). For $TmPep1050_{2\text{-}mer}$, $TmPep1050_{H60A}$ and $TmPep1050_{H307A}$, molecular replacement was achieved with Phaser-MR using the coordinates of $TmPep1050_{12\text{-}mer}$ as search model. The models were built using phenix.autobuild. Iterative manual building was done in COOT (52). Multiple rounds of refinement were performed using phenix.refine. Model stereochemical quality was assessed using MolProbity (53). Protein structures were analysed with PDBe Pisa (54), Arpeggio (55), Rosetta pKa protocol (56, 57), APBS (58), and PyMOL Molecular Graphics System version 2.2 (Schrödinger, LLC).

Acknowledgements

The authors thank Martine Roovers for proofreading this paper and giving constructive comments. Access to BM30A beamline (ESRF synchrotron) was supported by Fonds National de la Recherche Scientifique under the contract IISN 4.4503.11F. Access to Proxima 2 beamline (SOLEIL synchrotron) was within Block Allocation Groups 20151139 and 20171555.

Declaration of interests

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

Author contributions

Conceptualization: R.D., F.S., L.D.; Formal analysis: R.D., N.B.; Investigation: R.D., T.V.G., N.B., J.V.D.; Methodology: R.D., N.B.; Project administration: F.S., L.D.; Resources: D.V.; Software: D.V.; Supervision: F.S., L.D.; Validation: R.D., N.B., Visualization: R.D., T.V.G., Writing – original draft: R.D., J.V.D, L.D., Writing – review and editing: R.D., L.D.

References

1. Puente, X. S., Sánchez, L. M., Overall, C. M., and López-Otín, C. (2003) Human and mouse proteases: a comparative genomic approach. Nature Reviews Genetics. **4**, 544–558
2. Butler, S. M., Festa, R. A., Pearce, M. J., and Darwin, K. H. (2006) Self-compartmentalized bacterial proteases and pathogenesis. Molecular Microbiology. **60**, 553–562
3. Collins, G. A., and Goldberg, A. L. (2017) The Logic of the 26S Proteasome. Cell. **169**, 792–806

4. Gur, E., Biran, D., and Ron, E. Z. (2011) Regulated proteolysis in Gram-negative bacteria — how and when? Nature Reviews Microbiology. **9**, 839–848
5. Janssen, B. D., and Hayes, C. S. (2012) The tmRNA ribosome-rescue system. in Advances in Protein Chemistry and Structural Biology, pp. 151–191, Elsevier, **86**, 151–191
6. Konovalova, A., Søgaard-Andersen, L., and Kroos, L. (2014) Regulated proteolysis in bacterial development. FEMS Microbiology Reviews. **38**, 493–522
7. Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998) The Proteasome: Paradigm of a Self-Compartmentalizing Protease. Cell. **92**, 367–380
8. Kisselev, A. F., Akopian, T. N., Woo, K. M., and Goldberg, A. L. (1999) The Sizes of Peptides Generated from Protein by Mammalian 26 and 20 S Proteasomes: Implications for Understanding the Degradative Mechanism and Antigen Presentation. Journal of Biological Chemistry. **274**, 3363–3371
9. Brandstetter, H., Kim, J.-S., Groll, M., and Huber, R. (2001) Crystal structure of the tricorn protease reveals a protein disassembly line. Nature. **414**, 466–470
10. Saric, T., Graef, C. I., and Goldberg, A. L. (2004) Pathway for Degradation of Peptides Generated by Proteasomes: a Key Role for Thimet Oligopeptidase and other Metallopeptidases. Journal of Biological Chemistry. **279**, 46723–46732
11. Chuang, C. K., Rockel, B., Seyit, G., Walian, P. J., Schönegge, A.-M., Peters, J., Zwart, P. H., Baumeister, W., and Jap, B. K. (2010) Hybrid molecular structure of the giant protease tripeptidyl peptidase II. Nature Structural & Molecular Biology. **17**, 990–996
12. Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., Ito, T., and Yoshimoto, T. (2006) Crystal Structure of Aminopeptidase N (Proteobacteria Alanyl Aminopeptidase) from Escherichia coli and Conformational Change of Methionine 260 Involved in Substrate Recognition. Journal of Biological Chemistry. **281**, 33664–33676
13. Borissenko, L., and Groll, M. (2005) Crystal Structure of TET Protease Reveals Complementary Protein Degradation Pathways in Prokaryotes. Journal of Molecular Biology. **346**, 1207–1219
14. Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., and Finn, R. D. (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Research. **46**, D624–D632
15. Russo, S., and Baumann, U. (2004) Crystal Structure of a Dodecameric Tetrahedral-shaped Aminopeptidase. Journal of Biological Chemistry. **279**, 51275–51281
16. Basbous, H., Appolaire, A., Girard, E., and Franzetti, B. (2018) Characterization of a Glycyl-Specific TET Aminopeptidase Complex from Pyrococcus horikoshii. Journal of Bacteriology. 10.1128/JB.00059-18
17. Durá, M. A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F. M. D., and Franzetti, B. (2009) The structural and biochemical characterizations of a novel TET peptidase complex from Pyrococcus horikoshii reveal an integrated peptide degradation system in hyperthermophilic Archaea: Characterization of P. horikoshii TET3 peptidase. Molecular Microbiology. **72**, 26–40
18. Schoehn, G., Vellieux, F. M. D., Asunción Durá, M., Receveur-Bréchot, V., Fabry, C. M. S., Ruigrok, R. W. H., Ebel, C., Roussel, A., and Franzetti, B. (2006) An Archaeal Peptidase Assembles into Two Different Quaternary Structures: a Tetrahedron and a Giant Octahedron. Journal of Biological Chemistry. **281**, 36327–36337
19. Appolaire, A., Girard, E., Colombo, M., Durá, M. A., Moulin, M., Härtlein, M., Franzetti, B., and Gabel, F. (2014) Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase. Acta Crystallographica Section D Biological Crystallography. **70**, 2983–2993
20. Appolaire, A., Durá, M. A., Ferruit, M., Andrieu, J.-P., Godfroy, A., Gribaldo, S., and Franzetti, B. (2014) The TET2 and TET3 aminopeptidases from P yrococcus horikoshii form a hetero-subunit peptidasome with enhanced peptide destruction properties: TET aminopeptidase multi-subunit complex. Molecular Microbiology. **94**, 803–814

21. Appolaire, A., Colombo, M., Basbous, H., Gabel, F., Girard, E., and Franzetti, B. (2016) TET peptidases: A family of tetrahedral complexes conserved in prokaryotes. Biochimie. **122**, 188–196
22. Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.-M., Moras, D., and Tarnus, C. (1994) Crystal structure of Aeromonas proteolytica aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. Structure. **2**, 283–291
23. Dutoit, R., Brandt, N., Legrain, C., and Bauvois, C. (2012) Functional Characterization of Two M42 Aminopeptidases Erroneously Annotated as Cellulases. PLoS ONE. **7**, e50639
24. Colombo, M., Girard, E., and Franzetti, B. (2016) Tuned by metals: the TET peptidase activity is controlled by 3 metal binding sites. Scientific Reports. **6**, 20876
25. Franzetti, B. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. The EMBO Journal. **21**, 2132–2138
26. Petrova, T. E., Slutskaya, E. S., Boyko, K. M., Sokolova, O. S., Rakitina, T. V., Korzhenevskiy, D. A., Gorbacheva, M. A., Bezsudnova, E. Y., and Popov, V. O. (2015) Structure of the dodecamer of the aminopeptidase APDkam598 from the archaeon Desulfurococcus kamchatkensis. Acta Crystallographica Section F Structural Biology Communications. **71**, 277–285
27. Kim, D., San, B. H., Moh, S. H., Park, H., Kim, D. Y., Lee, S., and Kim, K. K. (2010) Structural basis for the substrate specificity of PepA from Streptococcus pneumoniae, a dodecameric tetrahedral protease. Biochemical and Biophysical Research Communications. **391**, 431–436
28. Chaikuad, A., Pilka, E. S., De Riso, A., von Delft, F., Kavanagh, K. L., Vénien-Bryan, C., Oppermann, U., and Yue, W. W. (2012) Structure of human aspartyl aminopeptidase complexed with substrate analogue: insight into catalytic mechanism, substrate specificity and M18 peptidase family. BMC Structural Biology. **12**, 14
29. Nguyen, D. D., Pandian, R., Kim, D., Ha, S. C., Yoon, H.-J., Kim, K. S., Yun, K. H., Kim, J.-H., and Kim, K. K. (2014) Structural and kinetic bases for the metal preference of the M18 aminopeptidase from Pseudomonas aeruginosa. Biochemical and Biophysical Research Communications. **447**, 101–107
30. Sivaraman, K. K., Oellig, C. A., Huynh, K., Atkinson, S. C., Poreba, M., Perugini, M. A., Trenholme, K. R., Gardiner, D. L., Salvesen, G., Drag, M., Dalton, J. P., Whisstock, J. C., and McGowan, S. (2012) X-ray Crystal Structure and Specificity of the Plasmodium falciparum Malaria Aminopeptidase PfM18AAP. Journal of Molecular Biology. **422**, 495–507
31. Su, M.-Y., Peng, W.-H., Ho, M.-R., Su, S.-C., Chang, Y.-C., Chen, G.-C., and Chang, C.-I. (2015) Structure of yeast Ape1 and its role in autophagic vesicle formation. Autophagy. **11**, 1580–1593
32. Holz, R. C. (2002) The aminopeptidase from Aeromonas proteolytica: structure and mechanism of co-catalytic metal centers involved in peptide hydrolysis. Coordination Chemistry Reviews. **232**, 5–26
33. Macek, P., Kerfah, R., Erba, E. B., Crublet, E., Moriscot, C., Schoehn, G., Amero, C., and Boisbouvier, J. (2017) Unraveling self-assembly pathways of the 468-kDa proteolytic machine TET2. Science Advances. **3**, e1601601
34. Rosenbaum, E., Ferruit, M., Durá, M. A., and Franzetti, B. (2011) Studies on the parameters controlling the stability of the TET peptidase superstructure from Pyrococcus horikoshii revealed a crucial role of pH and catalytic metals in the oligomerization process. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. **1814**, 1289–1294
35. Appolaire, A., Rosenbaum, E., Durá, M. A., Colombo, M., Marty, V., Savoye, M. N., Godfroy, A., Schoehn, G., Girard, E., Gabel, F., and Franzetti, B. (2013) Pyrococcus horikoshii TET2 Peptidase Assembling Process and Associated Functional Regulation. Journal of Biological Chemistry. **288**, 22542–22554
36. Cannan, R. K., and Kibrick, A. (1938) Complex Formation between Carboxylic Acids and Divalent Metal Cations. Journal of the American Chemical Society. **60**, 2314–2320
37. Stamper, C., Bennett, B., Edwards, T., Holz, R. C., Ringe, D., and Petsko, G. (2001) Inhibition of the Aminopeptidase from Aeromonas proteolytica by L-Leucinephosphonic Acid. Spectroscopic

and Crystallographic Characterization of the Transition State of Peptide Hydrolysis. Biochemistry. **40**, 7035–7046
38. Durá, M. A., Receveur-Brechot, V., Andrieu, J.-P., Ebel, C., Schoehn, G., Roussel, A., and Franzetti, B. (2005) Characterization of a TET-like Aminopeptidase Complex from the Hyperthermophilic Archaeon Pyrococcus horikoshii. Biochemistry. **44**, 3477–3486
39. Ando, S., Ishikawa, K., Ishida, H., Kawarabayasi, Y., Kikuchi, H., and Kosugi, Y. (1999) Thermostable aminopeptidase from Pyrococcus horikoshii. FEBS Letters. **447**, 25–28
40. Mori, K., and Ishikawa, K. (2005) New Deblocking Aminopeptidases from Pyrococcus horikoshii. Bioscience, Biotechnology, and Biochemistry. **69**, 1854–1860
41. Roncari, G., Stoll, E., and Zuber, H. (1976) [43] Thermophilic aminopeptidase I. Methods in Enzymology, **45**, 522–530
42. Moser, P., Roncari, G., and Zuber, H. (1970) Thermophilic aminopeptidases from Bac.stearothermophilus: II. Aminopeptidasc I (AP I): physico-chemical properties; thermostability and activation; formation of the apoenzyme and subunits. International Journal of Protein Research. **2**, 191–207
43. Lee, H. S., Cho, Y., Kim, Y. J., Nam, K., Lee, J.-H., and Kang, S. G. (2007) Biochemical characterization of deblocking aminopeptidase from hyperthermophilic archaeon Thermococcus onnurineus NA1. Journal of Bioscience and Bioengineering. **104**, 188–194
44. Zhang, Y., Werling, U., and Edelmann, W. (2012) SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Research. **40**, e55–e55
45. Edelheit, O., Hanukoglu, A., and Hanukoglu, I. (2009) Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. BMC Biotechnology. **9**, 61
46. Casadaban, M. J., and Cohen, S. N. (1980) Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. Journal of Molecular Biology. **138**, 179–207
47. Tsai, C.-Y., and Lin, Y.-W. (2013) A highly selective and sensitive fluorescence assay for determination of copper(ii) and cobalt(ii) ions in environmental water and toner samples. The Analyst. **138**, 1232
48. Ferrer, J.-L. (2001) Automated data processing on beamline FIP (BM30A) at ESRF. Acta Crystallographica Section D Biological Crystallography. **57**, 1752–1753
49. Roth, M., Carpentier, P., Kaïkati, O., Joly, J., Charrault, P., Pirocchi, M., Kahn, R., Fanchon, E., Jacquamet, L., Borel, F., Bertoni, A., Israel-Gouy, P., and Ferrer, J.-L. (2002) FIP: a highly automated beamline for multiwavelength anomalous diffraction experiments. Acta Crystallographica Section D Biological Crystallography. **58**, 805–814
50. Kabsch, W. (2010) XDS. Acta Crystallographica Section D Biological Crystallography. **66**, 125–132
51. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX : a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D Biological Crystallography. **66**, 213–221
52. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallographica Section D Biological Crystallography. **66**, 486–501
53. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity : all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D Biological Crystallography. **66**, 12–21
54. Krissinel, E., and Henrick, K. (2007) Inference of Macromolecular Assemblies from Crystalline State. Journal of Molecular Biology. **372**, 774–797

55. Jubb, H. C., Higueruelo, A. P., Ochoa-Montaño, B., Pitt, W. R., Ascher, D. B., and Blundell, T. L. (2017) Arpeggio: A Web Server for Calculating and Visualising Interatomic Interactions in Protein Structures. Journal of Molecular Biology. **429**, 365–371
56. Kilambi, K. P., and Gray, J. J. (2012) Rapid Calculation of Protein pKa Values Using Rosetta. Biophysical Journal. **103**, 587–595
57. Lyskov, S., Chou, F.-C., Conchúir, S. Ó., Der, B. S., Drew, K., Kuroda, D., Xu, J., Weitzner, B. D., Renfrew, P. D., Sripakdeevong, P., Borgo, B., Havranek, J. J., Kuhlman, B., Kortemme, T., Bonneau, R., Gray, J. J., and Das, R. (2013) Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). PLoS ONE. **8**, e63906
58. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences. **98**, 10037–10041

Footnotes

Atomic coordinates of TmPep1050 structures are available in the Protein Data Bank under accession numbers: 4P6Y, 5NE6, 5NE7, 5NE8, and 6NW5.

The abbreviations used are: TET, tetrahedral shape; SEM, standard error of the mean; RMSD, root-mean-square deviation; pNA, p-nitroanilide; LAP, leucyl-aminopeptidase; k, specific activity; k_{cat}, turnover number; T_m, melting temperature; HMW, high molecular weight; n.d., not determined; mol. enz., molecule of enzyme; Abs, absorbance; V_e, elution volume; $[Co^{2+}]$, cobalt concentration; $R_{12/2}$, dodecamer to dimer ratio.

Tables

Substrates	k (s^{-1})
L-Leu-pNA	118.3 ± 3.6
L-Ile-pNA	42.9 ± 1.6
L-Val-pNA	7.2 ± 0.8
L-Met-pNA	6.3 ± 0.2
L-Phe-pNA	0.9 ± 0.02
L-Ala-pNA	0.1 ± 0.01
L-Glu-pNA	n.d.
Gly-pNA	n.d.
L-His-pNA	n.d.
L-Lys-pNA	n.d.
L-Pro-pNA	n.d.
Ac-Leu-pNA	n.d.

Table 1 $TmPep1050_{12\text{-}mer}$ aminopeptidase activity against L-aminoacyl-pNA derivatives. k = specific activity, n.d. = not detectable ($k < 0.05\ s^{-1}$).

Substrates	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\ M^{-1}$)
L-Leu-pNA	1750 ± 250	138 ± 12	7.9×10^4
L-Ile-pNA	1100 ± 100	53 ± 4	4.8×10^4
L-Met-pNA	1750 ± 100	16 ± 1	9.1×10^3

Table 2 $TmPep1050_{12\text{-}mer}$ kinetic parameters. The ratio k_{cat}/K_m represents the catalytic efficiency.

Metal ion	k (s^{-1})
Co^{2+}	40.1 ± 1.1
Zn^{2+}	1.9 ± 0.02
Mn^{2+}	1.3 ± 0.02
Mg^{2+}	1.4 ± 0.01
Ni^{2+}	1.1 ± 0.03
Cu^{2+}	1.4 ± 0.01
Ca^{2+}	1.6 ± 0.01

Table 3 Effect of divalent metal ions on $TmPep1050_{12\text{-}mer}$ LAP aminopeptidase activity. Metal-ion depleted $TmPep1050_{12\text{-}mer}$ was diluted at 20 µM in 50 mM MOPS, 0.5 m $(NH_4)_2SO_4$ pH 7.2 supplemented with 1.28 mM of various metal chloride. After 24-hour incubation at 75°C, specific activities (k) were measured with L-Leu-pNA as substrate.

	apo-$TmPep1050_{12\text{-}mer}$	**$TmPep1050_{12\text{-}mer}$**	**$TmPep1050_{2\text{-}mer}$**	**$TmPep1050_{H60A}$**	**$TmPep1050_{H307A}$**
DATA COLLECTION					
Temperature (K)	100	100	100	100	100
Radiation source	ESRF BM30a	SOLEIL Proxima 2	SOLEIL Proxima 2	SOLEIL Proxima 2	SOLEIL Proxima 2
Wavelength (Å)	0.9797	0.9801	0.9800	0.9801	0.9801
Detector	ADSC QUANTUM 315r	DECTRIS EIGER X 9M	DECTRIS EIGER X 9M	DECTRIS EIGER X 9M	DECTRIS EIGER X 9M
Rotation range (°)	0.37	0.10	0.20	0.10	0.10
Exposure time (s)	20	0.025	0.025	0.025	0.025
Space group	P 1	H 3	C 2 2 21	C 2 2 21	P 1 21 1
Unit cell parameters					
α, β, γ (°)	114.46, 91.71, 105.69	90.00, 90.00, 120.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 110.51, 90.00
a, b, c (Å)	114.26, 114.57, 114.04	131.15, 131.15, 285.61	42.55, 114.71, 267.69	42.63, 114.22, 267.96	42.79, 138.65, 61.25
Resolution (Å)	44.05-2.20 (2.28-2.20)	47.60-1.70 (1.74-1.70)	48.25-2.00 (2.05-2.00)	43.46-1.84 (1.89-1.84)	40.11-1.75 (1.79-1.75)
Unique reflections	237152	201316	45086	57222	67094
R_{merge} (%)	8.9 (39.2)	8.2 (67.0)	9.5 (69.1)	8.7 (60.6)	5.6 (56.3)
Redundancy	3.2 (2.2)	10.34 (10.41)	13.2 (13.0)	12.0 (9.8)	6.7 (6.7)
$\langle I/\sigma \rangle$	8.56 (2.22)	15.81 (2.81)	16.81 (3.46)	16.52 (2.63)	18.04 (2.38)
Completeness (%)	93.5 (84.8)	99.8 (97.6)	99.9 (99.2)	99.4 (91.7)	99.5 (97.2)
$CC_{1/2}$ (%)	99.4 (81.1)	99.9 (87.1)	99.8 (90.0)	99.9 (90.0)	99.9 (88.1)
REFINEMENT					
Resolution	44.05-2.20	47.60-1.70	48.25-2.00	43.46-1.84	40.11-1.75
Reflections	237090	201316	45076	57213	67094
R_{free} set test count	11854	9715	2254	2862	3363
R_{work}/R_{free}	0.212/0.247	0.143/0.164	0.166/0.203	0.167/0.195	0.165/0.185
Protein molecules per ASU	12	4	2	2	2
V_M ($Å^3$/Da)	2.98	3.27	2.26	2.26	2.36
Solvent content (%)	58.7	62.4	45.6	45.6	47.9
Protein/solvent atoms	29969/2223	10759/1474	4610/362	4730/500	4621/335
RMSD bond lengths (Å)	0.009	0.019	0.004	0.004	0.012
RMSD bond angles (°)	1.254	1.680	0.647	0.672	1.263
Average B-factors ($Å^2$)	37.0	24.0	37.8	33.2	28.9
Favoured/disallowed Ramachandran φ/ψ (%)	95.71/0.46	95.35/0.34	94.86/0.17	94.95/0.00	93.64/0.52
Twin law	-	-	-	-	h, -k, -h-l
PDB ID	4P6Y	6NW5	5NE6	5NE7	5NE8

Table 4 Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

Figures

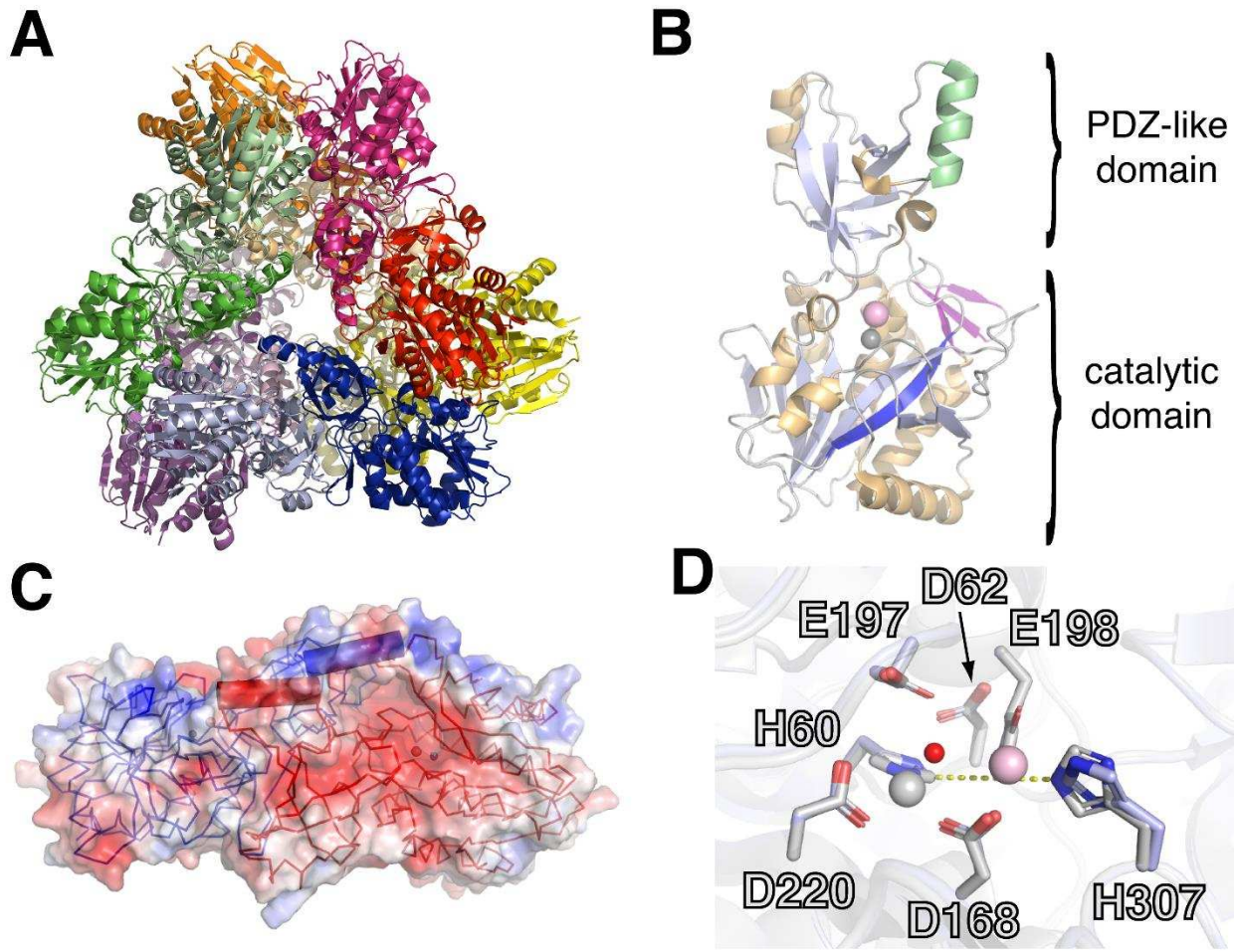

Figure 1 The TmPep1050$_{12\text{-mer}}$ structure. **A** Schematic representation of TmPep1050$_{12\text{-mer}}$ quaternary structure (PDB ID 6NW5) centred on one of the tetrahedron faces. As the asymmetric unit contains 4 monomers, the quaternary structure was reconstituted by PDBe Pisa. **B** Schematic representation of TmPep1050$_{12\text{-mer}}$ subunit. The α4 helix, β16 strand, and β sheet extension (β10, β11, β17) are coloured in green, dark blue, and purple, respectively. Zn^{2+} and Co^{2+} are displayed as grey and pink spheres, respectively. **C** Schematic representation of subunits A and B (red and blue, respectively) composing a dimer in the TmPep1050$_{12\text{-mer}}$ structure. The subunit A catalytic pocket is indicated by an arrow, the α4 helices are schematized as cylinders. **D** Structural alignment of the five metal ion binding residues and the two catalytic residues of TmPep1050$_{12\text{-mer}}$ (in light grey) and apo-TmPep1050$_{12\text{-mer}}$ (PDB ID 4P6Y, in light blue). Zn^{2+}, Co^{2+} and H_2O are represented as, respectively, grey, pink, and red spheres. The distance between H60 and H307 Nε2 is about 6.7 Å in the TmPep1050$_{12\text{-mer}}$ structure. In the apo-TmPep1050$_{12\text{-mer}}$ structure, this distance is increased to about 8.2 Å.

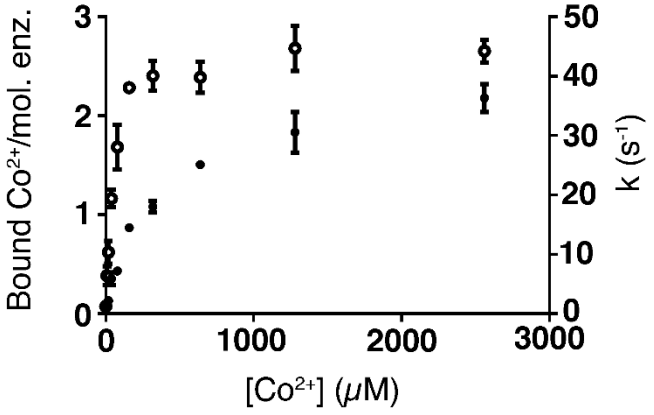

Figure 2 Evolution of bound Co^{2+} per molecule of enzyme (mol. enz.) (open circles) and LAP activity (k, closed circles) in response to an increasing Co^{2+} concentration ([Co^{2+}]) ranging from 0 to 2560 µM.

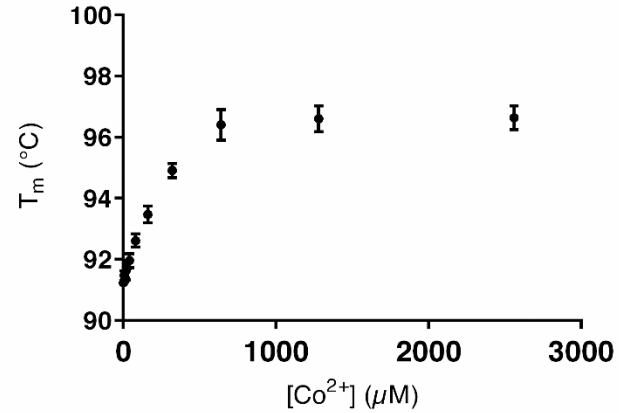

Figure 3 Thermal shift assays of Co^{2+} effect on $TmPep1050_{12\text{-}mer}$ thermostability. Metal ion-depleted $TmPep1050_{12\text{-}mer}$ was diluted at 20 µM in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$ pH 7.2 supplemented with a Co^{2+} concentration ([Co^{2+}]) ranging from 0 to 2560 µM. After 24-hour incubation at 75°C, the thermostability was determined through measuring T_m.

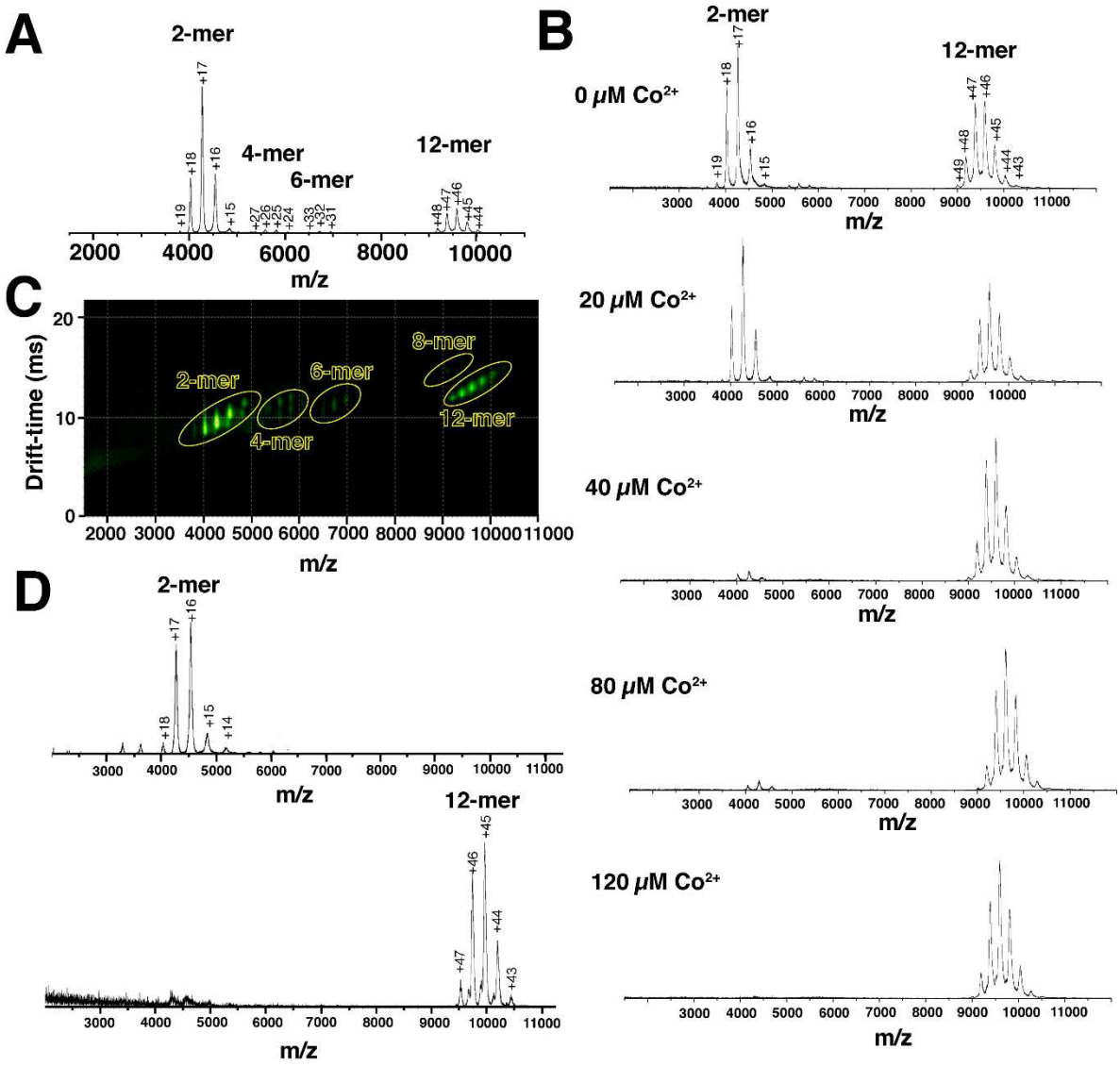

Figure 4 Oligomeric states determination by native MS. **A** Mass spectrum of TmPep1050$_{12\text{-mer}}$ at a concentration of 50 µM, trap collision energy set to 10 V. **B** Mass spectra of TmPep1050$_{12\text{-mer}}$ in presence of an increasing Co^{2+} concentration. 20 µM enzyme was incubated during 30 min. at 75°C with Co^{2+} at a concentration ranging from 0 to 120 µM. **C** Ion-mobility mass spectra of TmPep1050$_{12\text{-mer}}$ (50 µM). **D** Mass spectra of TmPep1050$_{2\text{-mer}}$ (upper) and its reassociation into dodecamers (lower). For the reassociation experiment, 20 µM enzyme was incubated with 120 µM Co^{2+} during 30 min. at 75°C.

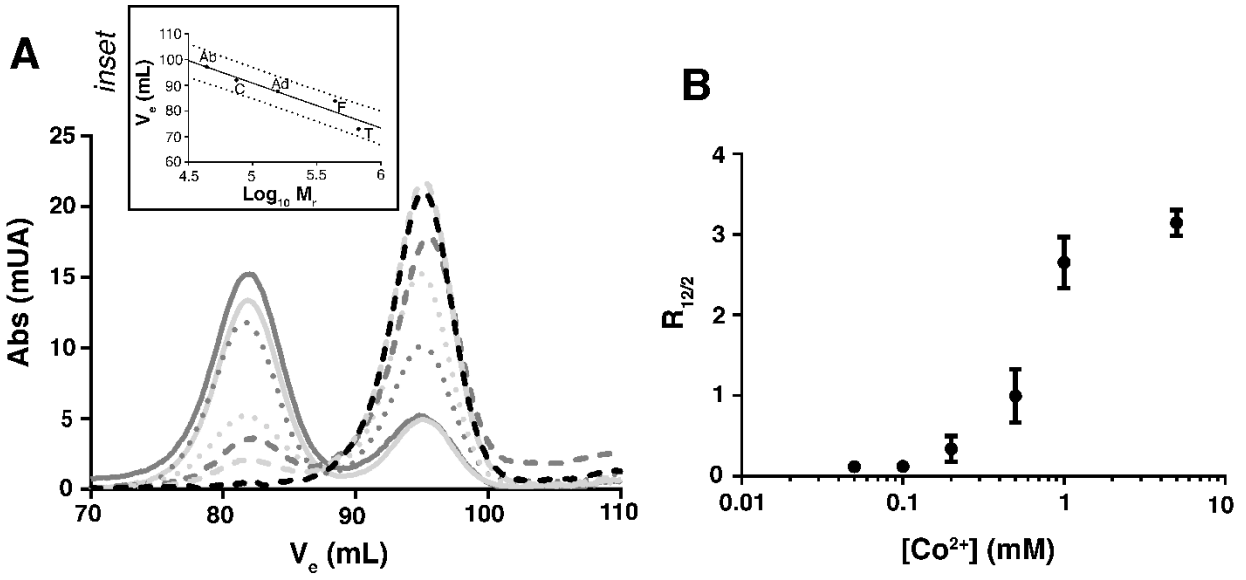

Figure 5 Reassociation of dimers into dodecamers in response to Co^{2+} concentration. $TmPep1050_{2\text{-}mer}$ was incubated with Co^{2+} concentration ranging from 0 to 5 mM for 30 min. at 75°C. Oligomers were detected and quantified by gel filtration on a Superdex 200 column (volume of 120 mL). **A** Gel filtration chromatogram of $TmPep1050_{2\text{-}mer}$ after incubation with increasing Co^{2+} concentrations: 0 mM (black dashes), 0.05 mM (light grey dashes), 0.1 mM (grey dashes), 0.2 mM (light grey dots), 0.5 mM (grey dots), 1 mM (light grey line), and 5 mM (grey line). The peak at elution volume (V_e) of 95.0 mL corresponds to dimers while the peak at V_e of 81.8 mL corresponds to dodecamers. Inset The calibration of the Superdex 200 column with albumin (Ab), conalbumin (C), aldolase (Ad), ferritin (F), and thyroglobulin (T) as standards. The correlation between V_e (mL) and the logarithm of the relative mass (M_r) is linear with R^2 of 0.91. The 95% confidence intervals of the linear regression are shown in dots. **B** Ratio of dodecamers to dimers ($R_{12/2}$) after incubating $TmPep1050_{2\text{-}mer}$ with Co^{2+}. Ratios were calculated based on the peak areas.

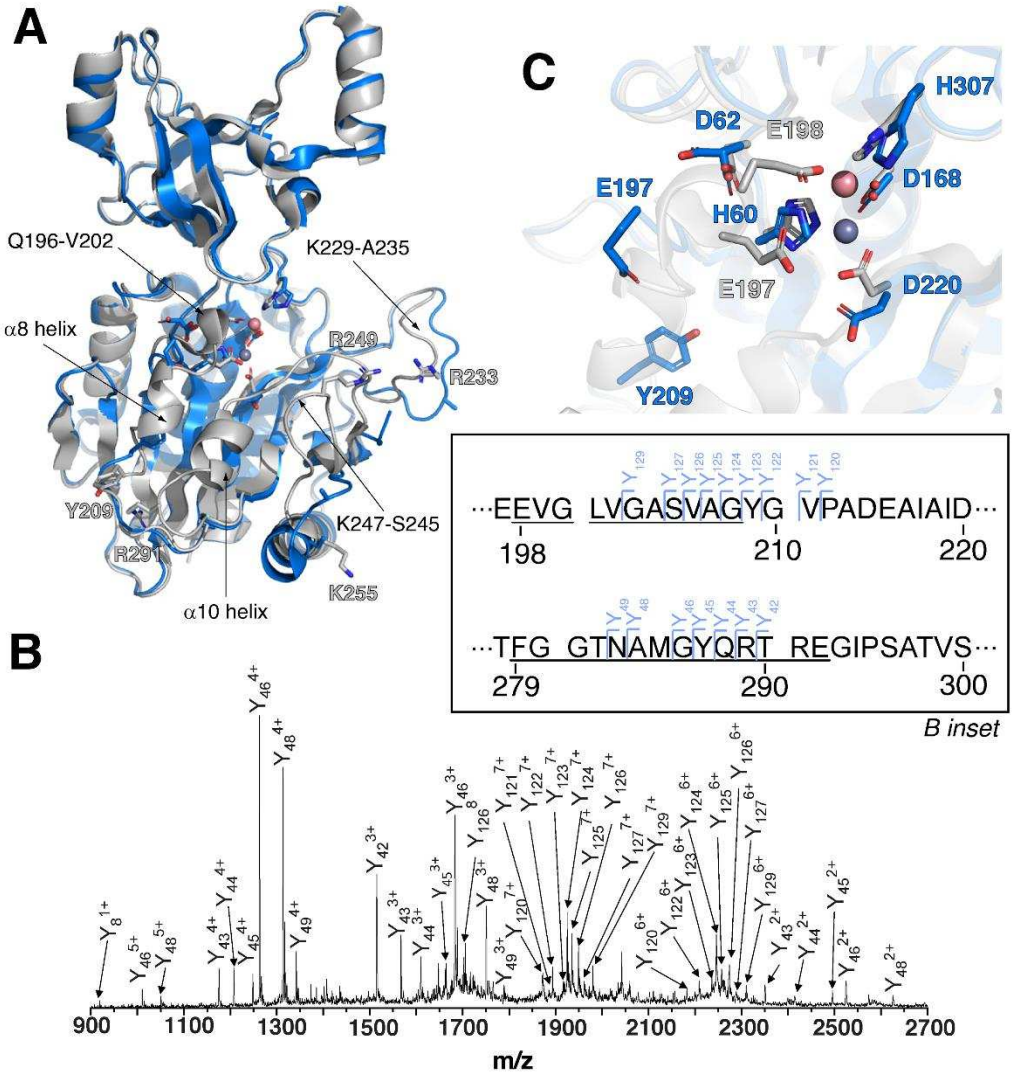

Figure 6 The TmPep1050$_{2\text{-mer}}$ structure. **A** Structural alignment of TmPep1050$_{2\text{-mer}}$ subunit (PDB ID 5NE6, blue) vs. TmPep1050$_{12\text{-mer}}$ subunit (PDB ID 6NW5, light grey). Arrows point the structural dissimilarities. The metal binding residues, catalytic residues, and residues involved in the oligomerization are shown in stick representation. **B** MS fragmentation pattern of TmPep1050$_{2\text{-mer}}$ under 125V collision trap energy. In the inset, preferential fragmentations are annotated in blue. The sequences underlined are the unmodelled flexible parts of the TmPep1050$_{2\text{-mer}}$ structure. **C** Close-up of the TmPep1050$_{2\text{-mer}}$ (blue) active site vs. the TmPep1050$_{12\text{-mer}}$ active site (light grey). The side chains of the metal binding residues and the catalytic residues are represented as sticks. The metal cofactors of TmPep1050$_{12\text{-mer}}$, Co^{2+} and Zn^{2+}, are shown as pink and grey spheres, respectively.

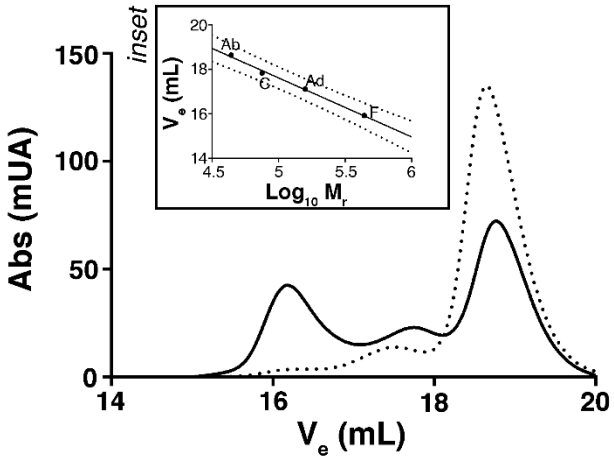

Figure 7 Oligomeric states of TmPep1050$_{K232A/R233A/R249A}$ (plain line) and TmPep1050$_{K232E/R233E/R249E}$ (dots) determined by gel filtration on Superose 6 column. The peaks at elution volume (V_e) of 16.2, 17.1 and 18.3 mL correspond to dodecamers (347 kDa), tetramers (151 kDa) and dimers (54 kDa), respectively. Inset The calibration of the Superose 6 column with albumin (Ab), conalbumin (C), aldolase (Ad), and ferritin (F) as standards. The correlation between V_e (mL) and the logarithm of the relative mass (M_r) is linear with R^2 of 0.99. The 95% confidence intervals of the linear regression are shown in dots.

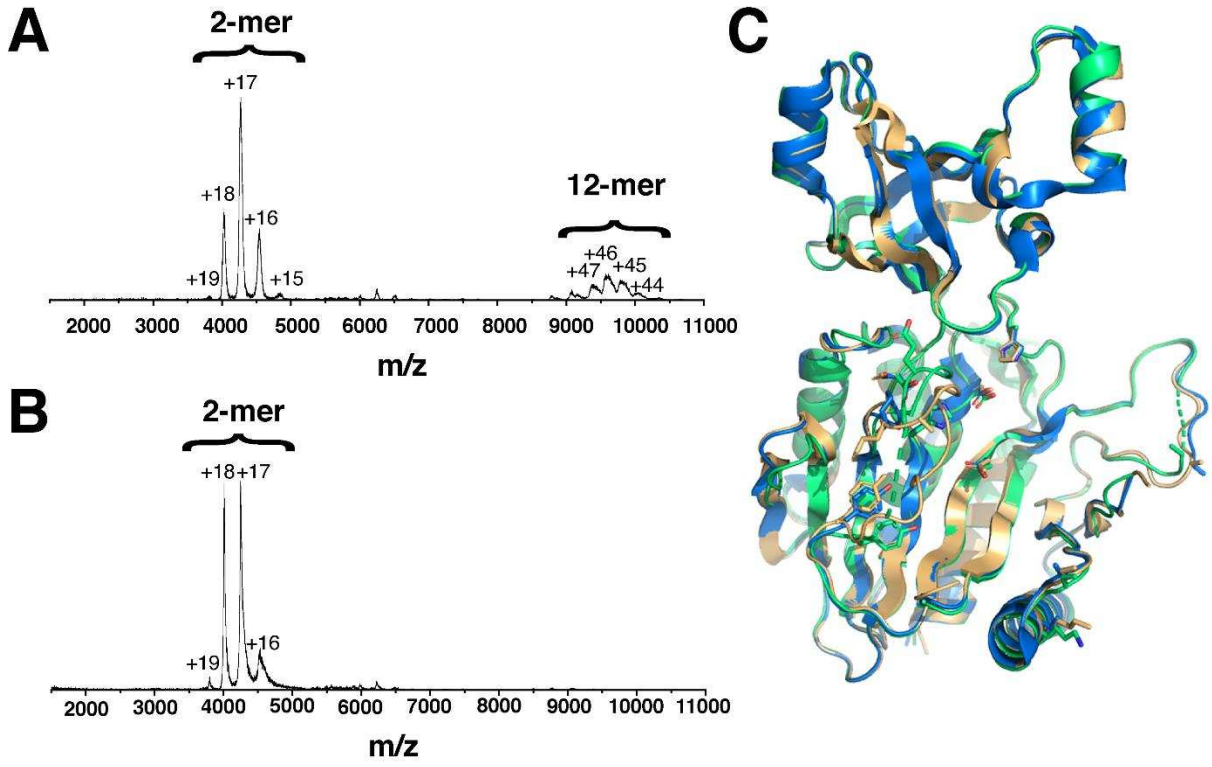

Figure 8 A Native mass spectrum of TmPep1050$_{H60A}$ in 100 mM ammonium acetate. Dimers and dodecamers were the two major species being observed. **B** Native mass spectrum of TmPep1050$_{H307A}$ in 100 mM ammonium acetate. Dimers were the main oligomeric form being observed. **C** Structural alignment of TmPep1050$_{2\text{-mer}}$ (blue), TmPep1050$_{H60A}$ structure (gold), and TmPep1050$_{H307A}$ (green). RMSD values are 0.122 and 0.468 for TmPep1050$_{H60A}$ and TmPep1050$_{H307A}$, respectively.