White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/168466/

Version: Accepted Version

Article:
Dutoit, R, Brandt, N, Van Gompel, T et al. (2020) M42 aminopeptidase catalytic site: the structural and functional role of a strictly conserved aspartate residue. Proteins: Structure, Function, and Bioinformatics, 88 (12). prot.25982. pp. 1639-1647. ISSN: 0887-3585

https://doi.org/10.1002/prot.25982

© 2020 Wiley Periodicals LLC. This is the peer reviewed version of the following article: Dutoit, R, Brandt, N, Van Gompel, T et al. (4 more authors) (2020) M42 aminopeptidase catalytic site: the structural and functional role of a strictly conserved aspartate residue. Proteins: Structure, Function, and Bioinformatics, 88 (12). prot.25982. pp. 1639-1647. ISSN 0887-3585, which has been published in final form at https://doi.org/10.1002/prot.25982. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

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.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Exploring the structural and functional role of a strictly conserved aspartate residue of M42 aminopeptidase catalytic site

Raphaël Dutoit[1, 2], Nathalie Brandt[2], Tom Van Gompel[3], 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, 1 avenue Emile Gryzon, B1070 Brussels, Belgium

[2] Labiris Institut de Recherche, 1 avenue Emile Gryzon, B1070 Brussels, Belgium

[3] Biomolecular & Analytical Mass Spectrometry, Department of Chemistry, Universiteit van Antwerpen, 171 Groenenborgerlaan, B2020 Antwerpen, Belgium

[4] Astbury Centre for Structural and Molecular Biology, University of Leeds, LS2 9JT, Leeds, United Kingdom

Abstract

The M42 aminopeptidases are a family of dinuclear aminopeptidases widely distributed in Prokaryotes. They have been proposed to be associated to the proteasome, achieving the complete destruction of peptides. Their most peculiar characteristic is their quaternary structure, a tetrahedron-shaped particle made of twelve subunits. The catalytic site of M42 aminopeptidases is defined by seven conserved residues. Five of them are involved in metal ion binding which is important to maintain both the activity and the oligomeric state. The sixth conserved residue is the catalytic glutamate acting as a general base deprotonating the water molecule during peptide bond hydrolysis. The seventh residue is an aspartate whose function remains poorly understood. This aspartate residue, however, must have a critical role as it is strictly conserved in all MH clan enzymes. In addition, it takes part to some kind of catalytic triad with the histidine residue and the metal ion of the M2 binding site. We assess its role in TmPep1050, an M42 aminopeptidase of *Thermotoga maritima*, through a mutational approach. Asp-62 was substituted with alanine, asparagine, or glutamate residue. The three Asp-62 substitutions completely abolished TmPep1050 activity and impeded dodecamer formation. They also interfered with metal ion binding as only one cobalt ion is bound per subunit instead of two. The structural data showed that the Asp62Ala substitution has an impact on the active site folds becoming similar to TmPep1050 dimer. We propose a structural role for Asp-62, helping to stabilize a crucial loop in the active site. In its absence, the loop will

remain unstructured failing to position correctly the catalytic base and a metal ion ligand of the M1 site.

1. Introduction

The M42 aminopeptidases are dinuclear enzymes proposed to achieve peptide degradation downstream the proteasome in Bacteria and Archaea [1, 2]. In *Pyrococcus horikoshii*, four M42 aminopeptidases (PhTET1, PhTET2, PhTET3, and PhTET4) have been described, each having a different, but complementary, substrate specificity[3–6]. PhTET2 and PhTET3 have even been shown to form heterocomplexes, suggesting that peptidasome particles may exist[7, 8]. The M42 aminopeptidases are characterized by a genuine quaternary structure made of twelve subunits organized spatially as a tetrahedron[1, 3–6, 9–13]. The association of the twelve subunits is often described as the assembly of six dimers, each dimer being located at a tetrahedron edge[1]. Four gates are found at the middle of the tetrahedron faces, leading to a wide inner chamber. The gate size (between 12 Å and 20 Å) probably restricts the access to the inner chamber to unfolded peptides only[1, 10, 13, 14]. The twelve catalytic sites are oriented inward the chamber, compartmentalizing the activity. The amino acids, generated after peptide hydrolysis, exit the catalytic sites through four channels located at the tetrahedron vertexes. The oligomerization of M42 aminopeptidases is controlled by their metal ion cofactors, as shown for PhTET2, PhTET3, PfTET3 of *Pyrococcus furiosus*, and TmPep1050 of *Thermotoga maritima*[1, 12, 13, 15, 16]. Indeed, active dodecamers disassemble into inactive dimers upon the depletion of metal ions. The disassembly is although reversible as adding metal ions to dimers restores the dodecamer formation and activity.

The M42 family, ubiquitous to Bacteria and Archaea[17], belongs to the MH clan alongside the M18, M20, and M28[18]. All MH clan enzymes share a common α/β globular fold for their catalytic domain, the archetypal model being the aminopeptidase Ap1 of *Vibrio proteolyticus*[19]. The typical subunit of M42 aminopeptidases possesses a PDZ-like dimerization domain in addition to the catalytic domain[1]. Seven residues define the catalytic site of MH clan enzymes[18, 20]. Five of them are involved in the binding of two metal ion cofactors. The first metal ion binding site, M1, consists of three conserved residues, a histidine and two aspartates. One of the aspartates is shared with the second metal ion binding site, M2. Two other residues define the M2 site: a conserved histidine and a glutamate/aspartate. These five residues impose the tetrahedral coordination geometry to the two bound metal ions. The implication of both metal ions in the catalytic mechanism has been thoroughly studied in *V. proteolyticus* aminopeptidase Ap1[21–23]. Indeed, they interact with the free N-terminal amine and the first carbonyl of the peptide substrate. The M2 has been proposed to assist the

water molecule deprotonation. The hydroxide is then transferred to the M1 for peptide bond hydrolysis. In M42 aminopeptidases, the two metal ions are also involved in the dimer-dodecamer transition. We previously showed that the M1 strictly controls the oligomerization of TmPep1050 while M2 could have a stabilizer role[13]. Intriguingly, their roles were swapped in PfTET3[12], suggesting that other factors may influence oligomerization.

In addition to the five metal ion ligands, two other residues, a glutamate and an aspartate, define MH clan enzymes. These two residues are not involved directly in metal ion binding but are strictly conserved in the whole clan. The glutamate residue acts as a general base deprotonating the water molecule during the peptide bond hydrolysis[24]. It may also play a role in decreasing the Lewis acidity of the M2. As proposed for *V. proteolyticus* aminopeptidase Ap1, the catalytic base Glu-151 is within hydrogen bond distance to interact with His-97[21, 25]. The aspartate residue is not directly involved in hydrolysis but may have a role in decreasing the Lewis acidity of the M2[26]. In *V. proteolyticus* aminopeptidase Ap1, Asp-99 has been proposed to form some kind of catalytic triad (hereafter mentioned as Asp-His-Me) with His-97 and Zn^{2+} in the M2 binding site[21, 23, 25, 27]. Indeed, a strong hydrogen bond links the $O^{\delta 1}$ atom of Asp-99 with the N^{δ} atom of His-97, forcing the imidazole ring to be deprotonated. As a result, the Lewis acidity of Zn^{2+} would be decreased sufficiently to allow the transfer of the hydroxide to the M1 site, which is closer to the peptide bond to be hydrolyzed.

Still, how the aspartate residue participates in peptide bond hydrolysis remains intricate. Its role has only been hypothesized for *V. proteolyticus* aminopeptidase Ap1. To our knowledge, nothing has been well established about its implication in metal ion binding or active site stabilization for MH clan enzymes. In this work, we propose to study the Asp-His-Me triad of TmPep1050, an M42 aminopeptidase from *T. maritima*. TmPep1050 is a leucyl-aminopeptidase whose activity depends on its oligomeric states: inactive dimer and active dodecamer[13]. The transition between dimers and dodecamers is regulated by its metal ion cofactor, Co^{2+}. We present the characterization of three TmPep1050 variants for which Asp-62, equivalent to Asp-99 of *V. proteolyticus* aminopeptidase Ap1, was substituted with alanine, asparagine, or glutamate residue. All substitutions completely abolished the activity and dodecamer formation of TmPep1050. The Asp-62 variants were still able to bind one cobalt ion per subunit. As dodecamer formation relies on the M1 site, we propose that the M2 site remained functional despite the substitution. The structure of $TmPep1050_{D62A}$ was solved and, by comparing with the dimer and dodecamer structures, a structural role was inferred for Asp-62.

2. Materials and Methods

2.1. Mutagenesis of *Tm_1050*

Site-directed mutagenesis was performed following the single-primer reactions in parallel protocol (SPRINP)[28] using the pCEC43 as template vector. This vector allows the production of TmPep1050 under the regulation of P_{BAD} promoter in *E. coli*[13]. The primers used to generate the Asp-62 variants of TmPep1050 are described in Table S1. Phusion DNA polymerase (ThermoFisher Scientific) was used for DNA polymerization reactions. Cloning was carried out in the *E. coli* MC1061 strain[29]. The genetic constructs were checked by sequencing (Genetic Service Facility, University of Antwerp).

2.2. Production and purification

The Asp-62 variants of TmPep1050 were produced and purified to homogeneity according to the protocol described elsewhere[13]. It consists of four steps: (i) a heat treatment of crude cell extract at 70°C for 15 min., (ii) an anion-exchange chromatography using Source 15Q resin (GE Healthcare Life Sciences), (iii) an hydrophobic interaction chromatography using Source 15Phe resin (GE Healthcare Life Sciences), and (iv) a size-exclusion chromatography using a Superdex200 (GE Healthcare Life Sciences, XK16/20 column of 120-ml volume). During the last step of purification, an elution peak was observed at 95 ml for the three Asp-62 variants. The purified Asp-62 variants were concentrated to 250 µM using an Amicon Ultra-15 ultrafiltration unit with 30-kDa cutoff (Merck Millipore). Protein concentration was estimated by measuring the absorbance at 280 nm and applying the mass extinction coefficient of 18,910 M^{-1} cm^{-1}. For molecular weight determination, 50 µM of purified protein was loaded on a Superdex200 (GE Healthcare Life Sciences, XK16/20 column of 120-ml volume) using 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$, pH 7.2 as running buffer. The size exclusion column was calibrated using Gel-filtration standard (Biorad) and High-molecular-weight gel filtration calibration kit (GE Healthcare Life Sciences) under the same running conditions.

2.3. Mass spectrometry

Sample preparation and native MS analysis were performed as previously described[13]. $TmPep1050_{D62A}$ was conditioned in 20 mM Ammonium acetate, pH 7.2 using Zeba 7-kDa desalting columns (Thermo Fisher Scientific). Prior to MS analysis, the sample was diluted to 5 µM in 20 mM Ammonium acetate, pH 7.2. The spectrum was recorded in positive ion mode on a traveling-wave ion mobility Q-TOF instrument (Synapt G2 HDMS, Waters).

2.4. Enzymatic and cobalt binding assays

The specific activity of Asp-62 variants was assayed using L-leucine-*p*-nitroanilide (L-Leu-*p*NA, Bachem AG) as substrate. 10 µl of 1 µM enzyme was added to 990 µl of reaction mix

containing 2.5 mM L-Leu-*p*NA in 50 mM MOPS, 10 % methanol, 1 mM $CoCl_2$, pH 7.2. The reaction was performed at 75°C during 1 h and stopped by adding 1 ml of 20% acetic acid. Released *p*-nitroaniline was quantified by measuring the absorbance at 410 nm. Prior to cobalt binding assays, apo-enzyme was prepared as followed for the three Asp-62 variants. 200 µl of 250 µM enzyme was diluted in 1.8 ml of 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, 10 mM 1,10-phenanthroline, pH 7.2. The sample was then dialyzed four times against 200 ml of 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, pH 7.2. After the dialysis, the sample was concentrated to 100 µl using an Amicon Ultra-15 ultrafiltration unit with 30-kDa cutoff (Merck Millipore). Protein concentration was determined as described above. To study the metal binding, apo-enzyme was diluted to 20 µM in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, pH 7.2 and $CoCl_2$ at a concentration ranging from 0 to 1280 µM. The reaction volume was 140 µl and the samples were incubated at 50°C for 24 h. The thermostability of the samples was determined by Thermal Shift Assay using SyproOrange (Thermofisher Scientific) as described previously[13]. Bound Co^{2+} was quantified using Amplex UltraRed fluorescent probe (ThermoFisher Scientific) as described previously[13].

2.5. X-ray crystallography

$TmPep1050_{D62A}$ was crystallized in 0.1 M citric acid, 5% PEG3350, pH 5.2 using the hanging drop diffusion method. Drops containing 2 µl of 250 µM $TmPep1050_{D62A}$ and 2 µl of crystallization reagent were set up in EasyXtal Tool plates (Qiagen). Fully grown crystals were obtained after a week at 292 K. Crystal shape and size were improved by microseeding. Diffraction data were collected on the FIP-BM30a beamline at the European Synchrotron Research Facility (Grenoble, France)[30, 31]. Data collection and refinement statistics are presented in Table 1. XDS program package[32] was used for indexing diffraction data. The phase was determined by molecular replacement using Phaser[33] with the monomer A coordinates of TmPep1050 dodecamer structure (PDB code 4P6Y). The model was build using *phenix.autobuild* in PHENIX software package[34]. The model was iteratively refined manually with Coot[35] and automatically with *phenix.refine*. Model stereochemical quality was assessed using MolProbity[36]. Protein structures were analyzed with PDBe Pisa[37], Arpeggio[38], and PyMOL Molecular Graphics System version 2.2 (Schrödinger,LLC).

3. Results and discussion

3.1. The role of Asp-62 in the activity and oligomerization of TmPep1050

Asp-62, equivalent to Asp-99 of *V. proteolyticus* aminopeptidase Ap1, was substituted with alanine, asparagine, or glutamate residue. The resulting variants were named $TmPep1050_{D62A}$, $TmPep1050_{D62N}$, and $TmPep1050_{D62E}$, respectively. The variants were recombinantly produced in *E. coli* and purified to homogeneity. Their molecular weights were

determined by size-exclusion chromatography (Figure 1). The substitution of Asp-62 had a dramatic impact on the TmPep1050 oligomerization state as only dimers were observed even in the presence of Co^{2+}. The oligomerization state of $TmPep1050_{D62A}$ was confirmed by native MS showing dimers mainly (Figure 2). MS experiments also highlighted a minor tetrameric form. Minor intermediate oligomers (tetramer, hexamer, and octamer) were reported previously for the wild-type enzyme[13]. These oligomers were inferred as intermediate forms occurring during the dimer-dodecamer transition. For $TmPep1050_{D62A}$, only tetramers were observed as intermediates, supporting the impaired oligomerization of the Asp-62 substitution. The three variants displayed less than 0.04% of relative activity on L-Leu-*p*NA compared to the dodecameric wild-type enzyme.

The loss of activity subsequent to Asp-62 substitution was expected according to the putative function of Asp-99 in *V. proteolyticus* Ap1. The negative charge and the side chain length are critical for the activity. Intriguingly, Asp-62 substitution prevents the self-assembly of dimers into dodecamers. Such an observation contrasts with the known role of the two metal ion binding sites in TmPep1050 oligomerization. Indeed, we previously demonstrated that the M1 binding site strictly controls the oligomerization as the His307Ala substitution prevented the self-assembly of dimers[13]. The substitution of His-60, a residue of the M2 binding site, perturbed the oligomerization to a lesser extent as both dimers and dodecamers were observed. As Asp-62 interacts with His-60, we could have expected that the Asp-62 substitution would have a rather small impact on oligomerization like His-60 substitution. Unexpectedly, the Asp-62 variants behaved as if the M1 binding site had been impaired as only dimers were observed. Therefore, Asp-62 substitution could somehow interfere with the metal ion binding in both M1 and M2 binding sites.

3.2. Asp-62 substitution partly impaired the Co^{2+} binding

To show the potential impact of Asp-62 substitution on Co^{2+} binding, thermal stability of Asp-62 variants were measured by thermal shift assay in presence of various concentrations of Co^{2+}. Using such a method, we previously showed that TmPep1050 dimers depleted of cobalt were more prone to thermal denaturation than TmPep1050 dodecamers, with melting temperatures of 91°C and 97°C, respectively. The addition of Co^{2+} restored the thermal stability of TmPep1050 dimers[13]. Prior to the binding assays, the Asp-62 variants were treated with 1,10-phenanthroline to remove any bound Co^{2+}. Then 20 µM of each apo-enzyme was incubated in presence of Co^{2+} at a concentration ranging from 0 to 1280 µM for 24 h. The melting temperatures were subsequently determined to monitor an eventual cobalt binding. Apo-$TmPep1050_{D62N}$ had a melting temperature of 89.1 ± 0.4°C, close to that of the wild-type dimer. Apo-$TmPep1050_{D62A}$ and apo-$TmPep1050_{D62E}$ had a melting temperature of 86 ± 0.1°C and 83.2 ± 0.2°C, respectively, indicating that the Asp62Ala and Asp62Glu substitutions

destabilized partly the structure. As the thermal stability increased with the cobalt concentration (Figure 3), the three variants were found to bind cobalt ions. Of note, none of them exhibited a thermostability comparable to the dodecamer upon Co^{2+} binding. Indeed, the maximal melting temperature of $TmPep1050_{D62N}$, $TmPep1050_{D62E}$, and $TmPep1050_{D62A}$ was 95.1°C, 91.1°C, and 90.7°C, respectively, for a Co^{2+}-to-protein molar ratio of 50.

Bound Co^{2+} was quantified to define how many metal ion binding sites are occupied in each variant. 20 µM of apo-enzyme was incubated with varying Co^{2+} concentration ranging from 0 to 1280 µM. The concentration of bound Co^{2+} was determined at the equilibrium using Amplex Ultra Red fluorescent probe for quantification. The results are showed in Figure 4. $TmPep1050_{D62A}$, $TmPep1050_{D62N}$, and $TmPep1050_{D62E}$ were able to bind Co^{2+} with an apparent association constant of 70 ± 12, 86 ± 2 µM, and 87 ± 13 µM, respectively. Only one cobalt ion was bound per monomer (maximal bound Co^{2+} being 1.3), suggesting that one of the two metal binding sites is impaired in the Asp-62 variants. In comparison, wild-type TmPep1050 could bind two cobalt ions with an apparent association constant of 50 µM[13]. For the Asp-62 variants, the cobalt quantification did not allow to determine whether both M1 and M2 sites were partially occupied or one of them remained vacant. The second hypothesis is the likeliest based on (i) the dimeric state of Aps-62 variants and (ii) the M1 site controlling the dimer-dodecamer transition. Thus, the M2 site was inferred to be functional even when Asp-62 was substituted. Our results contrast with what has been reported for the human copper/zinc superoxide dismutase. This enzyme displays an Asp-His-Me triad analogous to that of the MH clan enzymes. The substitution of the aspartate residue with an asparagine or alanine residue completely abolished both the activity and the metal ion binding[39].

3.3. The structure of $TmPep1050_{D62A}$ reveals how Asp-62 maintains the structural fold of the active site

To understand how Asp-62 could influence Co^{2+} binding in the M1 site, we solved the structure of $TmPep1050_{D62A}$ at 1.5 Å. The crystallization condition was 0.1 M citric acid, 5% PEG3350, pH 5.2. Unfortunately, X-ray fluorescence scanning of crystals did not detect any trace of metal ion, probably due to cobalt chelation by citric acid[40]. According to PDBe Pisa, the crystal structure confirmed the dimeric state of $TmPep1050_{D62A}$. When compared to TmPep1050 dodecamer, the subunit structure of $TmPep1050_{D62A}$ presents the same structural dissimilarities as TmPep1050 dimer (Figure 5A)[13]. Two segments, Lys-229 – Ala-235 and Lys-247 – Ser-254, diverge greatly between both structures. For instance, the backbone is displaced by 3 Å and 11 Å for Lys-232 and Arg-249, respectively. Both segments were described as critical in the oligomerization as they are involved in the interaction between dimers. The Asp-62 substitution also affects the catalytic pocket fold as the α8 and α10 helices

are too flexible to be modelled. Furthermore, the Gln-196 – Val-202 loop is displaced outwards the catalytic site, explaining the lack of activity of TmPep1050$_{D62A}$.

Because of Gln-196 – Val-202 loop displacement, Glu-197 and Glu-198 are not correctly positioned to fulfill their function (Figure 5B). The former is the catalytic base, assisting the deprotonation of a water molecule during the first step of the peptide bond hydrolysis. The latter is involved in the binding of a cobalt ion at the M1 site, along with His-307 and Asp-168. In the TmPep1050$_{D62A}$ structure, the His-307 side chain adopts a **m**-70° rotamer, compared to the **m**170° rotamer observed in the dodecamer structure. Consequently, the imidazole ring is tilted by about 28° between both structures. Asp-168 belongs to both M1 and M2 sites as it bridges the two cobalt ions. The Asp-62 substitution did not have any effect on the Asp-168 position. Regarding the M2 site, the side chains of Asp-220 and His-60 are oriented differently in TmPep1050$_{D62A}$ compared to dodecameric TmPep1050. The Asp-220 side chain adopts an **m**-20° rotamer instead of a **t**0° rotamer found in the dodecamer structure. Due to side chain orientation change, the carboxylate center is displaced by 2.4 Å and interacts with the carbonyl of Ile-221. In the monomer A of TmPep1050$_{D62A}$, the imidazole ring of His-60 is correctly positioned with a **p**-80° rotamer but is rotated by 180°. Consequently, $N^{\delta 1}$ and $N^{\varepsilon 2}$ of His-60 are found interacting with two water molecules which are not observed in the dodecamer structure. In the monomer B, His-60 side chain adopts two alternate rotamers, **m**90° and **p**-80°, with an occupation of 0.57 and 0.43, respectively.

Excepting the alternate rotamers of His-60, the TmPep1050 dimer structure shows the same dissimilarities in the active site as TmPep1050$_{D62A}$. Despite being structurally similar, TmPep1050 dimers are still able to bind two cobalt ions and revert to dodecamers while TmPep1050$_{D62A}$ cannot. Hence the Asp-62 substitution must have a structural role for stabilizing the metal-bound form. In the dodecamer structure, $O^{\delta 1}$ and $O^{\delta 2}$ of Asp-62 are in tight H-bond interaction with the backbone nitrogen of Glu-197 and Glu-198, respectively (Figure 5B). Thus, Asp62Ala substitution probably prevents this interaction and, consequently, could not stabilize the Gln-196 – Val-202 loop. It has a dramatic impact on the M1 site as Glu-198 cannot be correctly positioned to allow the binding of a cobalt ion. The side chain length and charge of residue 62 are critical since TmPep1050$_{D62N}$ and TmPep1050$_{D62E}$ are also inactive dimers. The position of the Asp-62 carboxylate is probably imposed by H-bond interactions with Ser-15 and His-60. In the TmPep1050$_{D62A}$ structure, Ser-15 hydroxyl is rotated by 107.2° compared to wild-type dimer and dodecamer. In addition, charge repulsion may occur between Glu-18 and Asp-62. Placing a glutamate residue at position 62 would not allow the correct alignment of the carboxylate with Ser-15 and His-60. In the case of an asparagine residue at position 62, the amide group would not sustain the charge repulsion

from Glu-18. Furthermore, the amine group would clash with either the Ser-15 or His-60 due to short H-H distance.

4. Conclusion

The carboxylate-histidine-metal ion triad is commonly encountered in metalloenzymes. Christianson and Alexander reported that at least 35% of histidine residues bound to a metal ion interacts with either a glutamate or an aspartate residue[26]. The supposed role of the carboxylate is to modulate the basicity of the neighbor histidine residue, affecting the Lewis acidity of bound metal ion. It may also have an impact of metal ion binding affinity and nucleophilicity of the water molecule bound to the metal ion[41, 42]. The MH clan enzymes display such a carboxylate-histidine-metal ion triad at the M2 site. Both the aspartate and histidine residues of the triad are strictly conserved in the whole clan. A function of the aspartate residue has been postulated for the MH clan archetypal enzyme, the *V. proteolyticus* aminopeptidase Ap1[21–23, 25, 27]. Asp-99 has been described to decrease Lewis acidity of the M2 site Zn^{2+} via a strong H-bond between Asp-99 and His-97. Consequently, the interaction between the nucleophilic hydroxide and the M2 site Zn^{2+} is disfavored and may transitorily alter the coordination geometry of the metal ion.

In this work, we showed that the aspartate residue of the Asp-His-Me triad could have another function for the M42 aminopeptidases. Using TmPep1050 as a case of study, Asp-62, equivalent to Asp-99 of *V. proteolyticus* aminopeptidase Ap1, was substituted with alanine, asparagine, or glutamate residue. This substitution completely abolished the activity, as expected, but also interfered in dodecamer formation. We suppose that the Asp-62 substitution affects the M1 binding site as the oligomerization of TmPep1050 strictly relies on the M1. The structure of $TmPep1050_{D62A}$ presented the same structural dissimilarities as observed for the wild-type dimer. Notably, the Gln-196 – Val-202 loop is disordered, preventing the correct positioning of Glu-197 and Glu-198 (the catalytic base and a ligand of the M1 site). Asp-62 has probably a structural role for stabilizing this active site loop since its carboxylate function is in close interaction with the backbone nitrogen of Glu-197 and Glu-198. Nevertheless, we could not ascertain the structural role of Asp-62 to all MH clan enzymes, as no structure of equivalent variant has been reported in other families. It would be interesting to know whether substituting the aspartate residue would destabilize the active site of a monomeric enzyme like *V. proteolyticus* aminopeptidase Ap1. If the destabilization is also observed for such an enzyme, it will support the structural role of this strictly conserved aspartate residue thorough the whole MH clan. If not, it will indicate that the destabilization of the active site as observed for $TmPep1050_{D62A}$ could be an adaptation to oligomerization.

Nevertheless, both hypotheses agree with the dependency of M42 aminopeptidase oligomerization on the active site fold.

5. Reference

1. Borissenko L, Groll M. Crystal Structure of TET Protease Reveals Complementary Protein Degradation Pathways in Prokaryotes. Journal of Molecular Biology. 2005;346:1207–19.

2. Appolaire A, Colombo M, Basbous H, Gabel F, Girard E, Franzetti B. TET peptidases: A family of tetrahedral complexes conserved in prokaryotes. Biochimie. 2016;122:188–96.

3. Russo S, Baumann U. Crystal Structure of a Dodecameric Tetrahedral-shaped Aminopeptidase. Journal of Biological Chemistry. 2004;279:51275–81.

4. Schoehn G, Vellieux FMD, Asunción Durá M, Receveur-Bréchot V, Fabry CMS, Ruigrok RWH, et al. An Archaeal Peptidase Assembles into Two Different Quaternary Structures: *A TETRAHEDRON AND A GIANT OCTAHEDRON*. Journal of Biological Chemistry. 2006;281:36327–37.

5. Durá MA, Rosenbaum E, Larabi A, Gabel F, Vellieux FMD, Franzetti B. 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. 2009;72:26–40.

6. Basbous H, Appolaire A, Girard E, Franzetti B. Characterization of a Glycyl-Specific TET Aminopeptidase Complex from *Pyrococcus horikoshii*. Journal of Bacteriology. 2018;200:e00059-18.

7. Appolaire A, Girard E, Colombo M, Durá MA, Moulin M, Härtlein M, et al. Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase. Acta Crystallographica Section D Biological Crystallography. 2014;70:2983–93.

8. Appolaire A, Durá MA, Ferruit M, Andrieu J-P, Godfroy A, Gribaldo S, et al. 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. 2014;94:803–14.

9. Franzetti B, Schoehn G, Hernandez J-F, Jaquinod M, Ruigrok RWH, Zaccai G. Tetrahedral aminopeptidase: a novel large protease complex from archaea. The EMBO Journal. 2002;21:2132–8.

10. Kim D, San BH, Moh SH, Park H, Kim DY, Lee S, et al. Structural basis for the substrate specificity of PepA from *Streptococcus pneumoniae*, a dodecameric tetrahedral protease. Biochemical and Biophysical Research Communications. 2010;391:431–6.

11. Petrova TE, Slutskaya ES, Boyko KM, Sokolova OS, Rakitina TV, Korzhenevskiy DA, et al. Structure of the dodecamer of the aminopeptidase APDkam598 from the archaeon *Desulfurococcus kamchatkensis*. Acta Crystallographica Section F Structural Biology Communications. 2015;71:277–85.

12. Colombo M, Girard E, Franzetti B. Tuned by metals: the TET peptidase activity is controlled by 3 metal binding sites. Scientific Reports. 2016;6. doi:10.1038/srep20876.

13. Dutoit R, Van Gompel T, Brandt N, Van Elder D, Van Dyck J, Sobott F, et al. How metal cofactors drive dimer–dodecamer transition of the M42 aminopeptidase TmPep1050 of *Thermotoga maritima*. J Biol Chem. 2019;294:17777–89.

14. Gauto DF, Estrozi LF, Schwieters CD, Effantin G, Macek P, Sounier R, et al. Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. Nat Commun. 2019;10:2697.

15. Rosenbaum E, Ferruit M, Durá MA, Franzetti B. 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. 2011;1814:1289–94.

16. Macek P, Kerfah R, Erba EB, Crublet E, Moriscot C, Schoehn G, et al. Unraveling self-assembly pathways of the 468-kDa proteolytic machine TET2. Science Advances. 2017;3:e1601601.

17. Dutoit R, Brandt N, Legrain C, Bauvois C. Functional Characterization of Two M42 Aminopeptidases Erroneously Annotated as Cellulases. PLoS ONE. 2012;7:e50639.

18. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Research. 2018;46:D624–32.

19. Chevrier B, Schalk C, D'Orchymont H, Rondeau J-M, 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.

20. Neuwald AF, Liu JS, Lipman DJ, Lawrence CE. Extracting protein alignment models from the sequence database. Nucleic Acids Research. 1997;25:1665–77.

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. Holz RC. The aminopeptidase from *Aeromonas proteolytica*: structure and mechanism of co-catalytic metal centers involved in peptide hydrolysis. Coordination Chemistry Reviews. 2002;232:5–26.

23. Schürer 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. Bzymek KP, Holz RC. The Catalytic Role of Glutamate 151 in the Leucine Aminopeptidase from *Aeromonas proteolytica*. J Biol Chem. 2004;279:31018–25.

25. 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.

26. Christianson DW, Alexander RS. Carboxylate-histidine-zinc interactions in protein structure and function. J Am Chem Soc. 1989;111:6412–9.

27. 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.

28. Edelheit O, Hanukoglu A, Hanukoglu I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. BMC Biotechnology. 2009;9:61.

29. Casadaban MJ, Cohen SN. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. Journal of Molecular Biology. 1980;138:179–207.

30. Ferrer J-L. Automated data processing on beamline FIP (BM30A) at ESRF. Acta Crystallographica Section D Biological Crystallography. 2001;57:1752–3.

31. Roth M, Carpentier P, Kaïkati O, Joly J, Charrault P, Pirocchi M, et al. FIP: a highly automated beamline for multiwavelength anomalous diffraction experiments. Acta Crystallographica Section D Biological Crystallography. 2002;58:805–14.

32. Kabsch W. *XDS*. Acta Crystallographica Section D Biological Crystallography. 2010;66:125–32.

33. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *Phaser* crystallographic software. J Appl Crystallogr. 2007;40:658–74.

34. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D Biological Crystallography. 2010;66:213–21.

35. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of *Coot*. Acta Crystallographica Section D Biological Crystallography. 2010;66:486–501.

36. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. *MolProbity*: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D Biological Crystallography. 2010;66:12–21.

37. Krissinel E, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. Journal of Molecular Biology. 2007;372:774–97.

38. Jubb HC, Higueruelo AP, Ochoa-Montaño B, Pitt WR, Ascher DB, Blundell TL. Arpeggio: A Web Server for Calculating and Visualising Interatomic Interactions in Protein Structures. Journal of Molecular Biology. 2017;429:365–71.

39. Banci L, Bertini I, Cabelli DE, Hallewell RA, Tung JW, Viezzoli MS. A characterization of copper/zinc superoxide dismutase mutants at position 124 Zinc-deficient proteins. Eur J Biochem. 1991;196:123–8.

40. Wyrzykowski D, Chmurzyński L. Thermodynamics of citrate complexation with Mn^{2+}, Co^{2+}, Ni^{2+} and Zn^{2+} ions. J Therm Anal Calorim. 2010;102:61–4.

41. El Yazal J, Pang Y-P. Proton Dissociation Energies of Zinc-Coordinated Hydroxamic Acids and Their Relative Affinities for Zinc: Insight into Design of Inhibitors of Zinc-Containing Proteinases. J Phys Chem B. 2000;104:6499–504.

42. Lin, Lim C. Factors Governing the Protonation State of Zn-Bound Histidine in Proteins: A DFT/CDM Study. J Am Chem Soc. 2004;126:2602–12.

	TmPep1050$_{D62A}$
Data collection	
Temperature (K)	100
Radiation source	ESRF BM30A
Wavelength	0.9797
Detector	ADSC QUANTUM 315r
Oscillation range (°)	0.5
Exposure time (s)	
Space group	C 2 21 1
Unit cell parameters	
α, β, γ (°)	90.00, 90.00, 90.00
a, b, c (Å)	42.18, 113.96, 267.23
Resolution	48.00 – 1.50 (1.58-1.50)
Unique reflections	103,939
R_{merge} (%)	0.09
Redundancy	5.8
$<I/\sigma>$	14.80 (3.34)
Completeness (%)	99.7 (89.4)
$CC_{1/2}$ (%)	99.8 (87.6)
Refinement	
Resolution	48.00 – 1.50
Reflections	103,177
R_{free} set test count	5,158
R_{work}/R_{free}	0.183/0.215
Protein molecules per ASU	2
V_M (Å^3/Da)	2.27
Solvent content (%)	45.79
Protein/solvent atoms	4,950/820
r.m.s.d. bond lengths (Å)	0.36
r.m.s.d. bond angles (°)	0.55
Average B-factors (Å^2)	20.0
Favored/disallowed Ramachandran φ/ψ (%)	97.71/0.00
PDB code	5L6Z

Table 1 Data collection and refinement statistics. Values in parentheses are for the highest-resolution shell.

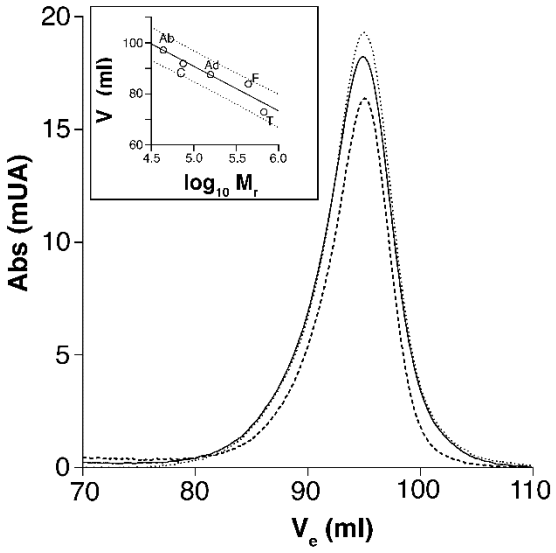


Figure 1 Size exclusion chromatography of TmPep1050$_{D62A}$ (plain line), TmPep1050$_{D62N}$ (dot line), and TmPep1050$_{D62E}$ (dashed line). An elution peak was observed at 95 ml for the three variants using a Superdex 200 column (V = 120 ml). *Inset:* The column was calibrated using albumin (Ab), conalbumin (C), aldolase (Ad), ferritin (F), and thyroglobulin (T) as standards. The correlation between the logarithm of the relative mass and the elution volume is linear, with a R^2 of 0.91. The 95% confidence intervals are represented as dots.

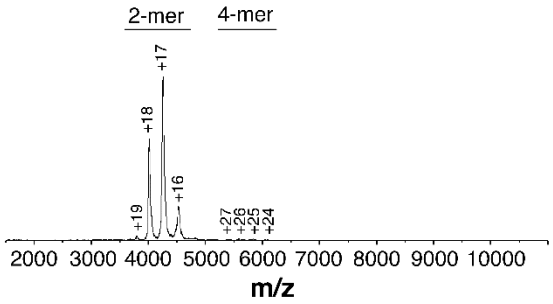


Figure 2 Native mass spectrum of TmPep1050$_{D62A}$. Dimers are the major observed species. The sample was conditioned in 20 mM Ammonium acetate and diluted to 5 µM prior to native MS analysis.

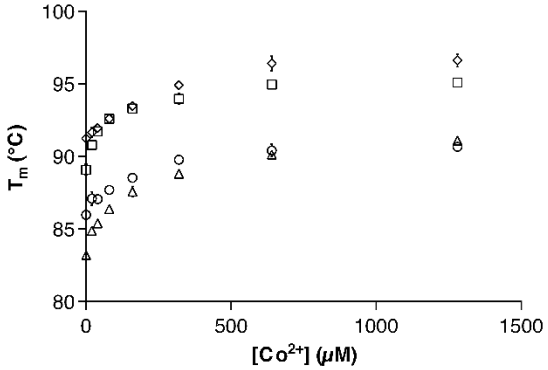


Figure 3 Thermal shift assay showing Co^{2+} binding effect on Asp-62 variant thermostability. The thermostability of $TmPep1050_{D62A}$ (circle), $TmPep1050_{D62N}$ (square), and $TmPep1050_{D62E}$ (triangle) was determined by measuring T_m at each Co^{2+} concentration. From our previous study, the thermostability of wild-type TmPep1050 dimer is presented as rhombi[13]. Error bars represent standard error with n = 3.

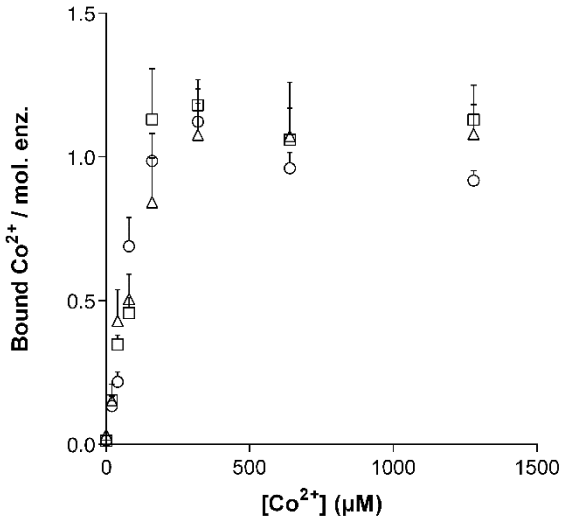
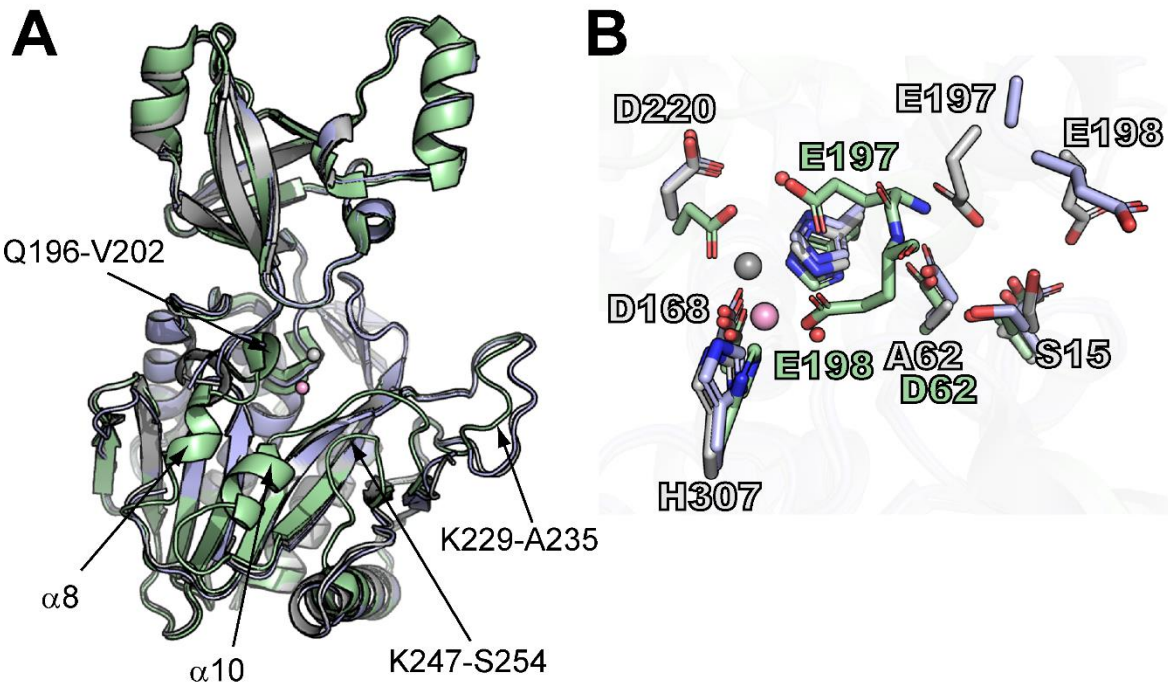


Figure 4 Co^{2+} binding of apo-$TmPep1050_{D62A}$ (circle), apo-$TmPep1050_{D62N}$ (square), and apo-$TmPep1050_{D62E}$ (triangle). The results are expressed as the number of bound Co^{2+} per molecule of enzyme (mol. enz.) in response to a varying concentration of Co^{2+} ranging from 0 to 1280 µM. Error bars represent standard error with n = 3.



457

458 **Figure 5 The structure of TmPep1050$_{D62A}$.** (A) Structural alignment of TmPep1050$_{D62A}$
459 subunit (PDB code 5L6Z, light grey) with TmPep1050 dodecamer subunit (PDB code 6NW5,
460 green), and TmPep1050 dimer subunit (PDB code 5NE6, light blue). The structural
461 dissimilarities are indicated with arrows. Zn^{2+} and Co^{2+} are represented as grey and pink
462 spheres, respectively. (B) Close-up view of the active sites of TmPep1050$_{D62A}$ (light grey),
463 TmPep1050 dodecamer (green) and dimer (light blue). The water molecules interacting with
464 His-60 in the TmPep1050$_{D62A}$ structure are showed as red spheres.