



**HAL**  
open science

# Etudes structurales et propriétés enzymatiques de deux nouvelles aminopeptidases TETs auto-compartmentées chez les archées

Hind Basbous

► **To cite this version:**

Hind Basbous. Etudes structurales et propriétés enzymatiques de deux nouvelles aminopeptidases TETs auto-compartmentées chez les archées. Biologie structurale [q-bio.BM]. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAV016 . tel-01842280

**HAL Id: tel-01842280**

**<https://theses.hal.science/tel-01842280>**

Submitted on 18 Jul 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **THÈSE**

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES**

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 25 mai 2016

Présentée par

**« Hind BASBOUS »**

Thèse dirigée par « **Bruno FRANZETTI** » et  
codirigée par « **Eric GIRARD** »

préparée au sein de l'**Institut de Biologie Structurale**  
dans l'**École Doctorale de Chimie et Sciences du Vivant**

## **Études structurales et propriétés enzymatiques de deux nouvelles aminopeptidases TETs auto- compartimentées chez les archées**

Thèse soutenue le « **19 Décembre 2016** »,  
devant le jury composé de :

**M. Lionel MOUREY**

Directeur de Recherche, CNRS, Toulouse (Rapporteur)

**M. Vincent DIVE**

Directeur de Recherche, CEA, Saclay (Rapporteur)

**Mme Vanessa DELFOSSE**

Chargé de Recherche, CNRS, Montpellier (Examineur)

**M. Mohamed JEBBAR**

Professeur des Universités, UBO, Brest (Examineur)

**M. Pascal FENDER**

Directeur de Recherche, CNRS, Grenoble (Président)

**M. Bruno FRANZETTI**

Directeur de Recherche, CNRS, Grenoble (Directeur de thèse)

**M. Eric GIRARD**

Chargé de Recherche, CEA, Grenoble (Co-Directeur de thèse)





## Remerciements

*Je tiens d'abord à exprimer toute ma reconnaissance et ma gratitude à mon directeur de thèse, Bruno Franzetti, pour m'avoir donné l'opportunité de réaliser ma thèse au sein de l'équipe ELMA. Je te remercie pour la confiance que tu m'as accordée, de m'avoir guidée, encouragée et soutenue durant ces trois années et surtout au cours de l'ultime phase de rédaction. Merci pour ton investissement, tes conseils avisés, ta disponibilité et ta grande gentillesse !*

*J'adresse aussi toute ma gratitude à mon co-directeur, Eric Girard, pour m'avoir appris avec patience la cristallographie. Merci pour tes explications, tes conseils, sans oublier les nouveaux termes français que tu m'as appris au cours de nos discussions ! « C'est en forgeant qu'on devient forgeron » !*

*J'aimerais remercier également tous les membres de l'équipe ELMA avec lesquels j'ai passé des moments sympathiques. Merci à Frank pour l'organisation des sorties raquettes chaque année ... j'ai souffert, mais j'ai bien aimé ! Merci à Pierre, le partisan du féminisme ! Merci à Jacques pour la lecture de mon introduction ... Merci à Dominique d'avoir partagé avec nous beaucoup d'aventures. Je n'oublierai jamais vos commentaires habituels (Dominique et Jacques) sur mon « looking » !! Merci à tous les thésards et les post-doc du laboratoire pour les discussions et les agréables moments partagés !*

*Mes remerciements sont ensuite adressés aux membres de mon jury Dr. Lionel Mourey, Dr. Vincent Dive, Dr. Vanessa Delfosse, Dr. Mohamed Jebbar et Dr. Pascal Fender qui m'ont fait l'honneur en acceptant d'évaluer mon travail de thèse.*

*Je n'oublierai pas les membres de mon comité de suivie de thèse, Jacques-Philippe Colletier et Jean-François Hernandez. Merci pour toutes vos propositions me permettant de surmonter les problèmes que j'ai affrontés durant ma thèse ! Grand merci à Jean-François, grâce à ta collection de peptides, j'ai pu identifier l'activité enzymatique de PhTET4. À Jacques-Philippe, je ne sais pas comment exprimer ma gratitude si particulière et sincère. Je te remercie de tout cœur pour ta grande confiance et ton soutien depuis mon master 1 jusqu'à présent ! J'espère en être digne.*

*Un très grand MERCI à toute la communauté libanaise à Grenoble ! Merci aux amis et amies, vous étiez et vous restez toujours ma deuxième famille ! Rida, Rim O. et Mariam mes compatriotes et mes collègues de laboratoire, merci pour toutes nos discussions, surtout de m'avoir supportée et entendue lorsque je me plaignais ... Merci à Maha, Mahmoud, Ali H. H., Wael, Rime K., Nour et Ali H. vous étiez toujours là quand j'avais besoin de vous ! Merci pour toutes les belles sorties et soirées ! Merci infiniment à la très fidèle Sandra, mon amie depuis très longtemps, de sorte que je n'arrive même pas à compter les années ! Nous avons débarqué ensemble dans cette nouvelle aventure qui s'appelait « La France ». Tu étais toujours présente lorsque j'avais besoin de toi, dans les mauvais jours tant comme dans les meilleurs !*

*Un merci exceptionnel est adressé au loyal et au catalan préféré, Martí Ninot, tu m'as prouvé que toute une « vraie » amitié peut résister à toute circonstance d'espace et de temps. Merci d'avoir supporté ma folie les derniers jours de rédaction !*

*Enfin, je voulais présenter toute ma gratitude et mon amour sincère à toute ma famille, particulièrement mon père Mohamad, ma mère Maha et mes deux sœurs Aya et Jamal. Je ne sais pas si ça va suffire de vous dire merci ! Mais, comme même je vais tenter ma chance ! Merci d'avoir cru en moi, de m'avoir toujours soutenue ! Sans vos efforts et vos sacrifices, je n'imaginerais pas que je serais là où je suis ! Je vous aime très, très fort ! Un dernier mot sera adressé à mes grand-mères et mon grand-père qui sont partis loin même sans pouvoir leur dire adieu ! Merci de m'avoir aimée comme si j'étais votre seule petite fille ! Je ne vous oublierai jamais !!!*

**Dieu, Merci !**

TABLE DES MATIÈRES

<b>1. INTRODUCTION.....</b>	<b>1</b>
<b>1.1. LA PROTÉOLYSE ET SES EFFECTEURS.....</b>	<b>1</b>
<b>1.2. LES MÉTALLOPEPTIDASES.....</b>	<b>2</b>
1.2.1. Les métallo-aminopeptidases.....	7
1.2.1.1. Caractérisation des sites métalliques.....	8
1.2.1.2. Rôles des aminopeptidases.....	9
<b>1.3. LES AMINOPEPTIDASES TÉTRAÉDRIQUES « TETs ».....</b>	<b>12</b>
1.3.1. Historique de la découverte des aminopeptidases TETs.....	13
1.3.2. Définition d'une aminopeptidase TET.....	13
1.3.3. Les aminopeptidases TETs existent dans les trois domaines du vivant.....	15
1.3.4. Caractéristiques communes à toutes les aminopeptidases TETs.....	20
1.3.4.1. Structure du monomère.....	20
1.3.4.2. Les aminopeptidases TETs : des particules dodécamériques uniques.....	24
1.3.4.3. Interfaces de dimérisation et d'oligomérisation d'une particule TET.....	25
1.3.4.4. Mécanisme d'assemblage et d'activation des particules TETs.....	28
1.3.4.5. Caractéristiques des poches catalytiques des aminopeptidases TETs.....	31
1.3.4.5.1. Sites catalytiques.....	31
1.3.4.5.2. Mécanisme catalytique.....	35
1.3.4.6. Reconnaissance et manipulation des peptides par une particule TET.....	36
1.3.5. Spécificités de différentes aminopeptidases TETs.....	39
1.3.5.1. Spécificités enzymatiques et poches de liaison de substrat « S1 ».....	39
1.3.5.2. Différences de propriétés de liaison aux métaux.....	42
1.3.5.2.1. Rôles des métaux dans les aminopeptidases TETs.....	43
1.3.6. Discussion du rôle physiologique des aminopeptidases TETs.....	45
<b>1.4. LES ARCHÉES : SOURCES DES MÉTALLO-AMINOPEPTIDASES TETs.....</b>	<b>48</b>
1.4.1. « <i>Woesian Revolution</i> » : Révélation d'un troisième domaine du vivant.....	49
1.4.2. Les archées : des organismes ubiquitaires ... et des hyperthermophiles.....	51
1.4.3. Les hyperthermophiles ... des vivants dans l'enfer.....	53
1.4.4. Adaptations moléculaires aux hautes températures.....	54
1.4.5. Phylogénétique des archées.....	55
1.4.5.1. <i>Methanocaldococcus jannaschii</i> .....	57
1.4.5.2. <i>Pyrococcus horikoshii</i> .....	59
1.4.6. Pourquoi des machineries archéennes sont-elles choisies comme modèles d'étude de la protéolyse ?.....	60
<b>2. PROBLÉMATIQUE ET OBJECTIFS DE LA THÈSE.....</b>	<b>62</b>
<b>3. MATÉRIELS ET MÉTHODES.....</b>	<b>65</b>
<b>3.1. MATÉRIELS ET MÉTHODES BIOCHIMIQUES.....</b>	<b>65</b>
3.1.1. Production de protéines.....	65
3.1.1.1. Synthèse de gènes et surexpression des protéines.....	65
3.1.1.2. Purification.....	66
3.1.2. Activité enzymatique.....	67
3.1.2.1. Substrats chromogènes et fluorogènes.....	67
3.1.2.2. <i>High performance liquid chromatography</i> (HPLC).....	68
<b>3.2. MATÉRIELS ET MÉTHODES BIOPHYSIQUES.....</b>	<b>69</b>
3.2.1. Microscopie électronique coloration négative.....	69
3.2.2. Cristallographie.....	69
3.2.2.1. Cristallogénèse.....	69
3.2.2.2. Enregistrement des données de diffraction.....	71
3.2.2.3. Traitement de données.....	72
3.2.2.4. Interfaces de dimérisation et d'oligomérisation.....	74
3.2.2.5. Création des figures.....	74
<b>4. REVUE SUR LES TETs M42 PROCARYOTES.....</b>	<b>75</b>

## TABLE DES MATIÈRES

---

4.1. TET peptidases: A family of tetrahedral complexes conserved in prokaryotes .....	76
<b>5. RÉSULTATS.....</b>	<b>86</b>
<b>5.1. CHAPITRE 1 : CARACTÉRISATIONS FONCTIONNELLES ET STRUCTURALES DE L'UNIQUE COMPLEXE TET ARCHÉEN DE <i>METHANOCALDOCOCBUS JANNASCHII</i> .....</b>	<b>86</b>
5.1.1. Présentation du travail .....	86
5.1.2. Crystallographic and cryo-EM studies of an archaeal TET peptidase complex from <i>Methanocaldococcus jannaschii</i> reveal internal network of loops responsible for peptide processing.....	89
5.1.3. Discussion et Perspectives.....	114
<b>5.2. CHAPITRE 2 : CARACTÉRISATIONS FONCTIONNELLES DE LA QUATRIÈME TET, PHTET4 DE <i>PYROCOCCUS HORIKOSHII</i>.....</b>	<b>116</b>
5.2.1. Présentation du travail .....	116
5.2.2. The hyperthermophilic archaeon <i>Pyrococcus horikoshii</i> possesses a fourth version of TET aminopeptidase complex with strict glyceryl-specificity and nickel-activation properties .....	119
5.2.3. État des connaissances sur les glyceryl aminopeptidases.....	139
5.2.3.1. États oligomériques .....	140
5.2.3.2. Spécificités enzymatiques.....	140
5.2.3.3. Propriétés catalytiques.....	140
5.2.3.4. Rôles supposés des glyceryl aminopeptidases.....	141
5.2.4. Discussion et Perspectives.....	143
<b>6. CONCLUSION GÉNÉRALE .....</b>	<b>145</b>
<b>7. RÉFÉRENCES BIBLIOGRAPHIQUES.....</b>	<b>148</b>

## TABLE DES ILLUSTRATIONS

<b>Figure 1</b> : Représentation schématique de la dégradation peptidique par les protéases ATP-dépendantes (Protéasome) suivie par les protéases ATP-indépendantes (Tricorne) et les aminopeptidases .....	2
<b>Figure 2</b> : Site catalytique de AAP, une leucyl aminopeptidase de <i>Vibrio proteolyticus</i> (anciennement nommée <i>Aeromonas proteolytica</i> ) (PDB ID : 1AMP) .....	4
<b>Figure 3</b> : Représentation schématique de différents types d'exopeptidases.....	5
<b>Figure 4</b> : Représentation schématique des résidus (P) du substrat peptidique et des sites actifs (S) d'une enzyme accueillant les acides aminés du peptide.....	8
<b>Figure 5</b> : Représentation schématique partielle du système RAS .....	11
<b>Figure 6</b> : Implication de certaines aminopeptidases de <i>Plasmodium falciparum</i> dans la gestion de l'hémoglobine humaine .....	12
<b>Figure 7</b> : A. Représentation schématique d'un tétraèdre TET. B. Dodécamère de PhTET1 (PDB ID : 2WYR) 14	
<b>Figure 8</b> : A. Localisation du <i>building block</i> (dimère) au niveau de l'arête du tétraèdre dodécamérique. B. Structure tertiaire du monomère de PhTET1 (PDB ID : 2WYR) comportant deux domaines.....	21
<b>Figure 9</b> : Structures tertiaires des monomères des familles M18, M20, M28 et M42 du clan MH .....	23
<b>Figure 10</b> : Représentation du dimère et du trimère dans le dodécamère de PhTET1 (PDB ID : 2WYR) .....	24
<b>Figure 11</b> : Interface de dimérisation entre deux monomères du <i>building block</i> d'une particule TET .....	25
<b>Figure 12</b> : Alignement des séquences des aminopeptidases TETs de la famille M42 archéennes et bactériennes (hormis ChFrvx).....	27
<b>Figure 13</b> : Interface d'oligomérisation de PhTET2 (PDB ID : 1XFO) .....	27
<b>Figure 14</b> : Implication d'un intermédiaire hexamérique dans le mécanisme d'assemblage de PhTET2 .....	29
<b>Figure 15</b> : Voie d'assemblage des aminopeptidases TETs révélée par l'hétéro-oligomère PhTET2-PhTET3 ...	30
<b>Figure 16</b> : Site actif binucléaire de PhTET2 de <i>P. horikoshii</i> .....	32
<b>Figure 17</b> : Alignement des séquences des protéines dodécamériques de deux familles M18 et M42 et mise en évidence des ligands liant les ions métalliques (étoiles vertes) et les résidus catalytiques (étoiles rouges).....	33
<b>Figure 18</b> : Site actif de hDNPEP (PDB ID : 4DYO) accueillant un résidu histidine d'une sous unité adjacente, du même <i>building block</i> .....	34
<b>Figure 19</b> : Alignement des séquences des aminopeptidases TETs M18 et M42 en utilisant la séquence de la TET humaine hDNPEP comme référence .....	35
<b>Figure 20</b> : Mécanisme catalytique proposé pour les aminopeptidases bimétaboliques (AAP de <i>A. proteolytica</i> ) .....	36
<b>Figure 21</b> : Potentiel électrostatique de l'intérieur de la particule PhTET1 (PDB ID : 2WYR) de <i>P. horikoshii</i> .....	38
<b>Figure 22</b> : Représentation des résidus de la poche spécifique de la fixation de substrat « S1 » de PhTET2 (PDB ID : 1XFO) de <i>P. horikoshii</i> .....	41
<b>Figure 23</b> : Implication de MHJ_0125 de <i>M. hyopneumoniae</i> dans l'infection et la colonisation de l'épithélium respiratoire porcin .....	47
<b>Figure 24</b> : Arbre universel du vivant adapté depuis Woese et al., 1990 et mise à jour par Forterre, 2015 .....	49
<b>Figure 25</b> : Structures cristallographiques de trois ARN polymérasas archéenne, eucaryote et bactérienne .....	50
<b>Figure 26</b> : Fumeur noir sur la dorsale Est Pacifique à 2600 mètres de profondeur (PHARE2002 Ifremer).....	53
<b>Figure 27</b> : Distribution phylogénétique de la thermophilie.....	54
<b>Figure 28</b> : Arbre phylogénétique des Archées comme proposé par Brochier-Armanet et ses collègues.....	57
<b>Figure 29</b> : <i>Methanocaldococcus jannaschii</i> observé sous microscope électronique.....	58
<b>Figure 30</b> : Observation d'une fine section de <i>Pyrococcus horikoshii</i> par microscopie électronique (González et al., 1998) .....	59
<b>Figure 31</b> : Représentation schématique de la phylogénie des TETs et de TRI (Tricorne peptidase) chez les archées.....	64
<b>Figure 32</b> : Représentation schématique de la réaction enzymatique exécutée par une TET (MjTET ou PhTET4) contre un substrat chromogénique (AA-pNA) et fluorogénique (AA-AMC).....	68
<b>Figure 33</b> : Représentation schématique du diagramme de phase de la cristallisation d'une protéine.....	70
<b>Figure 34</b> : Schéma décrivant la diffraction des rayons X selon la loi de Bragg.....	71
<b>Tableau 1</b> : Classification des métallopeptidases binucléaires en clans et familles, proposée par MEROPS ( <a href="http://merops.sanger.ac.uk">http://merops.sanger.ac.uk</a> ).....	6
<b>Tableau 2</b> : Aperçu de différentes TETs décrites dans la littérature chez les archées, les bactéries et les eucaryotes .....	16
<b>Tableau 3</b> : Nombre de structures secondaires composant les domaines catalytiques et les domaines de dimérisation (PDZ-like domain (domaine I) et petit feuillet $\beta$ (domaine II)) des TETs archéennes, bactériennes et eucaryotes.....	22



## LISTE DES ABBRÉVIATIONS

---

AAA-ATPase : ATPases Associated with diverse cellular Activities

ADN : Acide DésoxyriboNucléique

AMC : 7-Amino-4-MethylCoumarin

ARN : Acide RiboNucléique

ATP : Adénosine Triphosphate

CAPS : *N*-Cyclohexyl-3-AminoPropaneSulfonic acid

CCD : Charge-Coupled Device

CHES : *N*-CycloHexyl-2-aminoEthaneSulfonic acid

EDTA : EthyleneDiamineTetraacetic Acid

EM : Electron Microscopy

ESRF : European Synchrotron Radiation Facility

Ifremer : Institut Français de Recherche pour l'Exploitation de la MER

IPTG : IsoPropyl  $\beta$ -D-1-ThioGalactopyranoside

HPLC : High Performance Liquid Chromatography

LB : Lysogeny Broth

MPD : 2-Methyl-2,4-PentaneDiol

N-ter : N-terminal

PAN : Proteasome Activating Nucleotidase

pNA : p-NitroAniline

PDB : Protein Data Bank

Tris : Tris(hydroxymethyl)aminomethane

VIH : Virus de l'Immunodéficience Humaine

# 1. INTRODUCTION

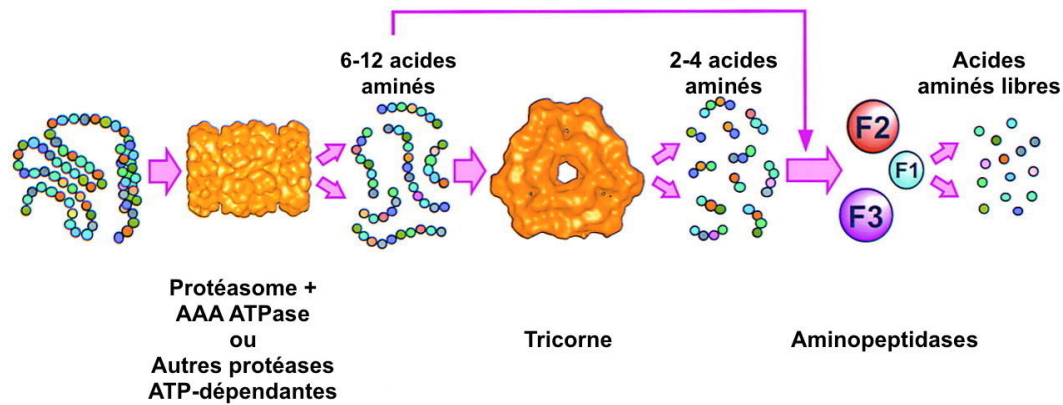
## 1.1. LA PROTÉOLYSE ET SES EFFECTEURS

La protéolyse est un processus régulant diverses fonctions cellulaires telles que le maintien de l'intégrité du protéome et la prévention des conséquences délétères de l'accumulation de protéines défectueuses dans le cytoplasme (Ward et al., 2002). Elle intervient dans le métabolisme énergétique et assure un renouvellement continu du « *pool* » d'acides aminés intracellulaires. Elle maintient aussi un équilibre dynamique entre la production de nouvelles protéines et le catabolisme des protéines « anciennes » (Goldberg, 2003; Russo and Baumann, 2004). Cette dégradation est une fonction de la demi-vie de la protéine et est une réponse aux modifications physiologiques et environnementales que subit la cellule. De plus, la protéolyse est impliquée dans la maturation post-traductionnelle et l'activation des protéines ainsi que dans la présentation des antigènes (Stoltze et al., 2000; Rock et al., 2002). Cette fonction est située au cœur de la réponse aux stress oxydatif et thermique (Berlett and Stadtman, 1997; Sherman and Goldberg, 2001) et est associée au processus de vieillissement. Toute dérégulation dans ce système protéolytique est à l'origine de certaines pathologies, comme le cancer (Yang et al., 2000; Gavioli et al., 2001) et les maladies neurodégénératives (Kitada et al., 1998; Sherman and Goldberg, 2001).

Selon leur dépendance à l'énergie, les effecteurs de la protéolyse sont classés comme ATP-dépendants ou ATP-indépendants. Les premiers sont gouvernés par le protéasome 20S et ses homologues tels que ClpP, HslV et Lon (Rohrwild et al., 1996; Baumeister et al., 1998; Maurizi et al., 1998; Stahlberg et al., 1999). Il s'agit d'endopeptidases nécessitant un apport d'énergie grâce à des sous-unités ou des domaines protéiques portant une activité ATPasique. Quant aux peptidases ATP-indépendantes, celles-ci interviennent en aval des peptidases ATP-dépendantes (Yao and Cohen, 1999), comme il a été proposé pour la protéase Tricorn et des protéines qui lui sont associées, transformant les peptides générés par le protéasome en acides aminés libres (**Figure 1**) (Tamura et al., 1998; Kim et al., 2002). On peut aussi citer la Thimet oligopeptidase qui intervient dans la dégradation des oligopeptides générés par la dégradation protéique par le protéasome. En hydrolysant les peptides précurseurs pour la présentation antigénique, cette endopeptidase non régulée par l'interféron- $\gamma$  limite, *in vivo*, la présentation antigénique par les complexes majeurs d'histocompatibilité de classe I (Saric et al., 2001). Non seulement les peptidases ATP-indépendantes dégradent les produits libérés par le

## INTRODUCTION

protéasome, mais elles semblent aussi impliquées dans plusieurs mécanismes physiologiques et pathologiques. C'est le cas, par exemple, de la Rénine qui joue un rôle important dans la régulation de la pression sanguine et la balance des électrolytes en transformant l'angiotensinogène en angiotensine I (Soffer, 1981).



**Figure 1 : Représentation schématique de la dégradation peptidique par les protéases ATP-dépendantes (Protéasome) suivie par les protéases ATP-indépendantes (Tricorne) et les aminopeptidases.** Cette cascade d'hydrolyse est proposée chez *Thermoplasma acidophilum*, une archée thermophile (Tamura et al., 1998).

Certains de ces effecteurs protéolytiques possèdent la capacité de former des complexes multimériques auto-compartmentés générés par l'assemblage de plusieurs sous unités homogènes ou hétérogènes (Stoll et al., 1973a; Appolaire et al., 2014a). La caractérisation structurale et fonctionnelle de ces grands assemblages représente une des principales thématiques du groupe ELMA « *Extremophiles and Large Molecular Assemblies* » au sein duquel j'ai réalisé mon projet de thèse.

## 1.2. LES MÉTALLOPEPTIDASES

Durant ma thèse, je me suis intéressée à la caractérisation enzymatique et structurale de nouvelles peptidases oligomériques ATP-indépendantes appartenant à la famille des métallo-aminopeptidases TETs. Ces deux macromolécules sont MjTET et PhTET4 exprimées par deux archées hyperthermophiles : *Methanocaldococcus jannaschii* et *Pyrococcus horikoshii*, respectivement. Ce chapitre est consacré à une présentation de l'état des connaissances sur les fonctions et le mode d'action des métallo-aminopeptidases en général. Le **chapitre 1.3.** sera dédié aux métallo-aminopeptidases TETs, sujet de cette thèse.

## INTRODUCTION

---

Les liaisons peptidiques sont extrêmement stables et possèdent une demi-vie de l'ordre de 350 à 600 ans à température ambiante et à pH compris entre 4 et 8 (Radzicka and Wolfenden, 1996). Ainsi pour hydrolyser de telles liaisons, il existe une classe d'enzymes spécifiques : les protéases et les peptidases. Les protéases sont considérées comme des enzymes spécialisées dans la dégradation des protéines et des longs peptides, alors que les peptidases agissent sur les peptides plus courts (Zhang et al., 2015). Cependant, la NC-IUBMB (*Nomenclature Committee of the International Union of Biochemistry and Molecular Biology*) recommande l'utilisation du terme « peptidase » pour désigner toute enzyme impliquée dans l'hydrolyse d'une liaison peptidique.

En se référant aux éléments clés impliqués dans le mécanisme catalytique, la banque de données des peptidases, MEROPS, répertorie les peptidases en 8 clans : le clan des peptidases à asparagine (N), à cystéine (C), à sérine (S), à thréonine (T), des peptidases aspartiques (A), glutamiques (G), mixtes (P) et les métallopeptidases (M). Un neuvième clan regroupe les peptidases non classées (U) (<http://merops.sanger.ac.uk>). Les sept premiers clans agencent des groupements organiques fonctionnels lors de la catalyse, tandis que les métallopeptidases disposent, au niveau de leurs sites catalytiques, d'ions métalliques coordonnant une molécule d'eau requise pour l'attaque nucléophile de la liaison peptidique (Rawlings et al., 2012; Zhang et al., 2015).

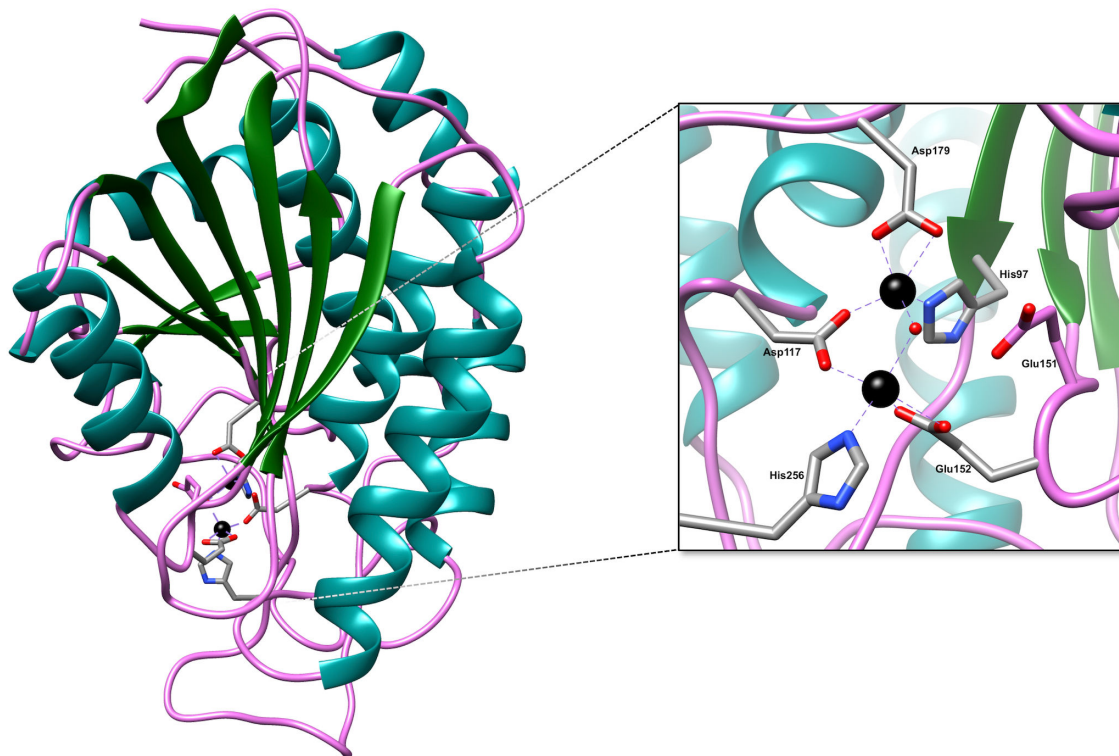
En 2000, Nelson et Cox ont défini le terme « métalloprotéome » comme étant l'ensemble des protéines ayant la capacité de fixer des métaux (Nelson and Cox, 2000). Les ions métalliques peuvent contribuer à la stabilité structurale et/ou à la régulation des activités catalytiques. Le rôle catalytique agencé par les ions métalliques permet de désigner les protéines qui les hébergent comme métallo-enzymes et dans le cas des peptidases : des métallopeptidases. Ces dernières représentent de l'ordre du tiers de toutes les enzymes connues (Mitić et al., 2014).

Les métallopeptidases appartiennent aux hydrolases pour lesquelles l'attaque nucléophile de la liaison peptidique implique une molécule d'eau. Des ions métalliques divalents, généralement des ions zinc et parfois des ions cobalt, magnésium, nickel ou cuivre sont hébergés dans le site catalytique des métallopeptidases et interviennent dans l'activation de la molécule d'eau (Barrett et al., 2004). En fonction du nombre d'ions métalliques requis pour la

## INTRODUCTION

catalyse, on distingue les métallopeptidases mononucléaires et les binucléaires (Zhang et al., 2015).

Les deux enzymes étudiées dans cette thèse sont des métallopeptidases binucléaires appelées aussi métallopeptidases co-catalytiques du fait de l'implication de deux ions divalents dans la catalyse enzymatique. Ces derniers sont coordonnés par cinq ligands dont l'un d'entre eux coordonne simultanément les deux cations métalliques (**Figure 2**). En plus des ligands, un résidu dit catalytique est présent au niveau du site actif. Celui-ci est communément un acide glutamique, mais d'autres résidus comme l'arginine, l'histidine, la lysine et la tyrosine peuvent aussi jouer le rôle de résidu catalytique (Barrett et al., 2004) (<http://merops.sanger.ac.uk>).

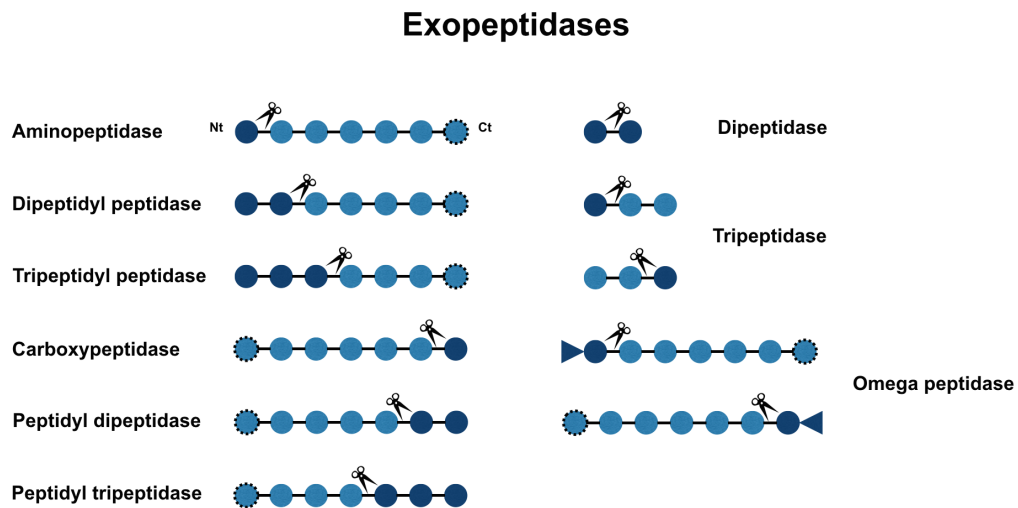


**Figure 2 : Site catalytique de AAP, une leucyl aminopeptidase de *Vibrio proteolyticus* (anciennement nommée *Aeromonas proteolytica*) (PDB ID : 1AMP).** AAP est une aminopeptidase monomérique appartenant à la famille M28 des métallopeptidases binucléaires. Le site catalytique abrite deux cations zinc penta-coordonnés (sphères noires) :  $Zn^{2+}_1$  coordonné par Glu152 et His256, et  $Zn^{2+}_2$  coordonné par His97 et Asp179. Un résidu Asp117 et une molécule d'eau (sphère rouge) assurent une connexion indirecte entre les deux ions métalliques. Le résidu catalytique de cette aminopeptidase est l'acide glutamique Glu151 (Chevrier et al., 1994a, 1996; Desmarais et al., 2006).

Dans l'ouvrage de référence « *Handbook of proteolytic enzymes* » (Barrett et al., 2004), toutes les métallopeptidases co-catalytiques sont considérées comme des exopeptidases aptes

## INTRODUCTION

à hydrolyser la liaison amide située aux extrémités N ou C-terminales (Nt ou Ct) de leurs substrats (Barrett et al., 2004). On y distingue les aminopeptidases, les carboxypeptidases et une diversité de peptidases, comme par exemple, les dipeptidyl peptidases et les peptidyl dipeptidases capables de cliver, respectivement, deux acides aminés de l'extrémité Nt et Ct de leurs substrats (**Figure 3**). Ces enzymes binucléaires sont répertoriées dans dix familles appartenant à six clans différents faisant partie des métallopeptidases où chaque famille est représentée par un holotype ou une protéine référence (Barrett and Rawlings, 2007) (**Tableau 1**) (<http://merops.sanger.ac.uk>). Notamment, les aminopeptidases TETs qui sont étudiées dans cette thèse sont classées comme étant des enzymes de la famille M42 du clan MH des métallopeptidases.



**Figure 3 : Représentation schématique de différents types d'exopeptidases.** Chaque cercle représente un acide aminé. Ceux qui sont colorés en bleu foncé sont cibles d'une hydrolyse enzymatique. Les cercles ayant une bordure pointillée signifient que la taille du peptide n'est pas limitée à sept acides aminés (comme représentée dans le schéma). Les cercles associés aux triangles désignent des acides aminés ayant subi des modifications (comme par exemple la cyclisation).

## INTRODUCTION

Métallopeptidases binucléaires								
Clan	Famille	Type de peptidase	Holotype	Ligands	Résidu(s) catalytique(s)	État oligomérique	PDB ID	Résidu(s) clivable(s)
<b>MF</b>	M17	Aminopeptidase	bILAP ou LAP3 de <i>Bos taurus</i>	Lys250, Asp255, Asp273, Asp332, Glu334	Lys262 / Arg336	Hexamère	1BLL	Cys/Leu
<b>MG</b>	M24	Aminopeptidase	MetAP1 d' <i>Escherichia coli</i>	Asp97, Asp108, His171, Glu204, Glu235	His79	Monomère	1MAT	Met
<b>MH</b>	M18	Aminopeptidase	Aminopeptidase I de <i>Saccharomyces cerevisiae</i>	His87, Asp235, Glu381, Asp435, His535*	Glu380*	Dodécamère	4EME	Leu
	M20	Peptidase / Amino / Carboxypeptidase	Glutamyl carboxypeptidase de <i>Pseudomonas sp.</i>	His112, Asp141, Glu176, Asp200, His385	Asp114 / Glu175	Dimère	1CG2	Glu
	M28	Amino / Carboxypeptidase	Aminopeptidase S de <i>Streptomyces griseus</i>	His85, Asp97, Glu132, Asp160, His247	Asp87 / Glu131	Monomère	1XJO	Leu/Phe
	<b>M42</b>	Aminopeptidase / Deblocking aminopeptidase	Glutamyl aminopeptidase de <i>Lactococcus lactis</i>	His61, Asp172, Glu205, Asp227, His306*	Asp63 / Glu204*	Dodécamère	2WYR	Asp/Glu
<b>MJ</b>	M19	Dipeptidase	Dipeptidase membranaire de <i>Homo sapiens</i>	His20, Asp22, Glu125, His198, His219	Asp288	Dimère	1ITQ	Cys/Gly
	M38	Peptidase	Isoaspartyl dipeptidase d' <i>E. coli</i>	His68, His70, Lys162, His201, His230	Asp285	Octamère	1ONW	Asp/β-Asp
<b>MN</b>	M55	Aminopeptidase /Peptidase putative	D-aminopeptidase de <i>Bacillus subtilis</i>	Asp8, Glu10, His60, His104, Glu133	His115	Décamère	1HI9	D-Ala/D-Leu/D-Phe
<b>MQ</b>	M29	Aminopeptidase	Aminopeptidase T de <i>Thermus aquaticus</i>	Glu250, Glu316, His345, His376, Asp378	Tyr352	Dimère	2AYI	Ala/Leu/Phe/Val

Tableau 1 : Classification des métallopeptidases binucléaires en clans et familles, proposée par MEROPS (<http://merops.sanger.ac.uk>).

## INTRODUCTION

---

Ce tableau détaille les principales caractéristiques décrites pour l'holotype de chaque famille (ligands, résidus catalytiques, états oligomériques, résidus clivables ainsi que les codes PDB correspondants). En absence de structure pour l'holotype (comme c'est le cas de l'aminopeptidase I de *Saccharomyces cerevisiae* et de la glutamyl aminopeptidase de *Lactococcus lactis*), les ligands et les résidus catalytiques présentés sont ceux d'une protéine appartenant à la même famille et/ou montrant une activité enzymatique similaire (aspartyl aminopeptidase de *Plasmodium falciparum* (PDB ID : 4EME) et glutamyl aminopeptidase de *Pyrococcus horikoshii* (PhTET1, PDB ID : 2WYR), respectivement).

### 1.2.1. Les métallo-aminopeptidases

Les aminopeptidases sont des exopeptidases qui libèrent sélectivement les résidus Nt à partir des polypeptides et des protéines. Elles regroupent 19 familles dont la plupart sont des métallo-aminopeptidases (Rawlings and Barrett, 1993). Ces enzymes sont très répandues chez les eucaryotes et les procaryotes. Elles peuvent être localisées dans le cytosol, dans certains compartiments cellulaires, ancrées à la membrane ou même être excrétées dans le milieu extracellulaire (Gonzales and Robert-Baudouy, 1996).

Divers critères de classification des aminopeptidases existent. Les plus fréquemment utilisés reposent sur la spécificité des substrats, la localisation cellulaire, le mécanisme catalytique et le pH optimal d'activité (Taylor, 1993; Rawlings et al., 2004). En se référant particulièrement à leurs spécificités de substrat, les aminopeptidases peuvent être classées en aminopeptidases spécifiques des résidus neutres, positifs ou négatifs. En outre, certaines sont dites aminopeptidases à large spectre d'activité du fait qu'elles clivent une diversité d'acides aminés N-terminaux. Ce groupe d'aminopeptidases à large spectre est dominé par les leucyl aminopeptidases (LAP) connues pour leur capacité à hydrolyser une grande diversité d'acides aminés. Par exemple, LAP de *Staphylococcus aureus* et MHJ\_0461 de *Mycoplasma hyopneumoniae*, en plus de leur capacité à cliver la leucine, agissent sur les acides aminés hydrophobes, basiques, polaires et acides (Carroll et al., 2013; Jarocki et al., 2015). En revanche, d'autres sont spécialisées dans le clivage d'un petit nombre de résidus, il s'agit des aminopeptidases à spectre d'activité restreint. Tel est le cas des méthionine aminopeptidases (Narayanan and Nampoothiri, 2012) et des prolyl iminopeptidases capables de dégrader les méthionines et les prolines situées en extrémité Nt, respectivement. De surcroît, l'efficacité catalytique de certaines métallo-aminopeptidases semble être affectée par la nature du résidu P1' situé juste après le résidu P1 cible de l'hydrolyse, en se référant à la nomenclature de Schechter et Berger (Schechter and Berger, 1967) (**Figure 4**). La prolyl aminopeptidase, nommée encore aminopeptidase P (AP-P) et clivant exclusivement les acides aminés suivis



## INTRODUCTION

---

par une proline en position P1' (Turner et al., 1997), constitue un bon exemple illustrant l'impact de P1' sur la dégradation.

**Substrat :** — P2 — P1 † P1' — P2' — P3' —

**Enzyme :** — S2 — S1 ★ S1' — S2' — S3' —

**Figure 4 :** Représentation schématique des résidus (P) du substrat peptidique et des sites actifs (S) d'une enzyme accueillant les acides aminés du peptide. Le résidu P2 est situé du côté N-terminal et est suivi par le résidu P1 (Schechter and Berger, 1967). † désigne la liaison peptidique hydrolysée et ★ désigne le site actif de l'enzyme.

### 1.2.1.1. Caractérisation des sites métalliques

Dans le chapitre des « Aminopeptidases Tétraédriques TETs », je présenterai les sites métalliques actifs des TETs (**paragraphe 1.3.4.5.1**). Ainsi, dans le but de comprendre l'origine de la nomenclature « M1 » et « M2 » des sites métalliques des peptidases binucléaires, j'expose dans ce paragraphe les études permettant de les caractériser.

Une description approfondie des sites métalliques des aminopeptidases binucléaires a été réalisée chez bLAP (*bovine lens leucyl aminopeptidase*), une aminopeptidase à zinc appartenant à la famille M17 des métallo-amino-peptidases. Ces travaux ont révélé la présence de deux sites de fixation des métaux ayant des affinités non équivalentes (Thompson and Carpenter, 1976). Cette propriété repose sur un mode d'échange différent des cations entre la protéine et le milieu réactionnel auquel sont ajoutés les ions divalents. Lors de l'addition du cobalt à bLAP, un site échange le  $Zn^{2+}$  natif rapidement par un ion  $Co^{2+}$ , contrairement à l'autre site où le  $Co^{2+}$  prend beaucoup plus de temps pour chasser l'ion zinc (environ une semaine) (Thompson and Carpenter, 1976). Conséquemment, le premier site désigné par M1 et caractérisé par une faible affinité pour les cations métalliques, fut dénommé « site régulateur » ou « *readily exchanging site* ». Quant au deuxième site M2, pourvu d'une affinité plus forte, plusieurs qualifications lui sont attribuées telles que « site catalytique », « site de spécificité » ou « *tightly exchanging site* » (Carpenter and Vahl, 1973; Thompson and Carpenter, 1976). Cette propriété a ensuite été validée grâce aux travaux de Kim et Lipscomb, qui décrivent la résolution des structures de bLAP natives (bLAP( $Zn^{2+}$ \_ $Zn^{2+}$ )) et activées par les ions  $Mg^{2+}$  (bLAP( $Mg^{2+}$ \_ $Zn^{2+}$ )). Deux densités électroniques différentes sont identifiées dans la chambre catalytique de bLAP( $Mg^{2+}$ \_ $Zn^{2+}$ ) contre deux autres similaires occupant le site catalytique de bLAP( $Zn^{2+}$ \_ $Zn^{2+}$ ). Les deux formes de bLAP possèdent en commun une

## INTRODUCTION

---

même densité électronique correspondant au « *tightly binding site* », M2. Tandis que la faible densité électronique observée pour bILAP(Mg<sup>2+</sup>\_Zn<sup>2+</sup>) correspond au site occupé par Mg<sup>2+</sup>, le « *readily exchanging site* », M1 (en raison du faible nombre d'électrons de Mg<sup>2+</sup> en comparaison avec ceux de Zn<sup>2+</sup>) (Kim and Lipscomb, 1993).

Le mécanisme catalytique adopté par une aminopeptidase co-catalytique et agencé par les ions M1 et M2, sera décrit dans le **paragraphe 1.3.4.5.2**.

### 1.2.1.2. Rôles des aminopeptidases

Les aminopeptidases ont longtemps été considérées comme des enzymes ubiquitaires jouant un rôle de « ménage », dans l'homéostasie cellulaire et le métabolisme de l'azote et du carbone dans des conditions de carence associée par exemple à la phase stationnaire de croissance (Lazdunski, 1989). Leur principale fonction était supposée être de participer au « *turnover* » des protéines endogènes en dégradant les peptides générés par l'activité des endopeptidases permettant ainsi de renouveler le réservoir d'acides aminés pour de nouvelles synthèses protéiques (Matsui et al., 2006). Les aminopeptidases sont également responsables de la digestion des peptides exogènes souvent utilisés comme source de nutriments par les microorganismes. Le cas le plus simple peut être attribué aux bactéries lactiques qui se développent en exploitant les peptides et les acides aminés issus de la dégradation des protéines du lait engendrée par une grande diversité de peptidases parmi lesquelles des aminopeptidases (Pritchard and Coolbear, 1993; Gonzales and Robert-Baudouy, 1996).

En plus de ces rôles, les aminopeptidases possèdent des fonctions cellulaires plus spécifiques liées à des spécificités de substrats plus marquées. Ainsi, certaines aminopeptidases activent et régulent la demi-vie des protéines ou des peptides bioactifs. À titre d'exemple, certains peptides ou protéines sont synthétisés sous forme de précurseurs et nécessitent une maturation par des aminopeptidases pour exécuter leur fonction biologique. La plupart des protéines néosynthétisées sont modifiées par l'action des méthionines aminopeptidases. Cette action permet parfois une activation et le plus souvent est responsable de la régulation de la demi-vie intracellulaire de la protéine en modulant son adressage aux endopeptidases. Effectivement, d'après « *The N-end rule* », le clivage du résidu Nt déstabilise la protéine dont la demi-vie sera régulée par la nature du nouvel acide aminé en position P1 de la nouvelle terminaison Nt (Tobias et al., 1991). Les aminopeptidases interviennent ainsi

## INTRODUCTION

---

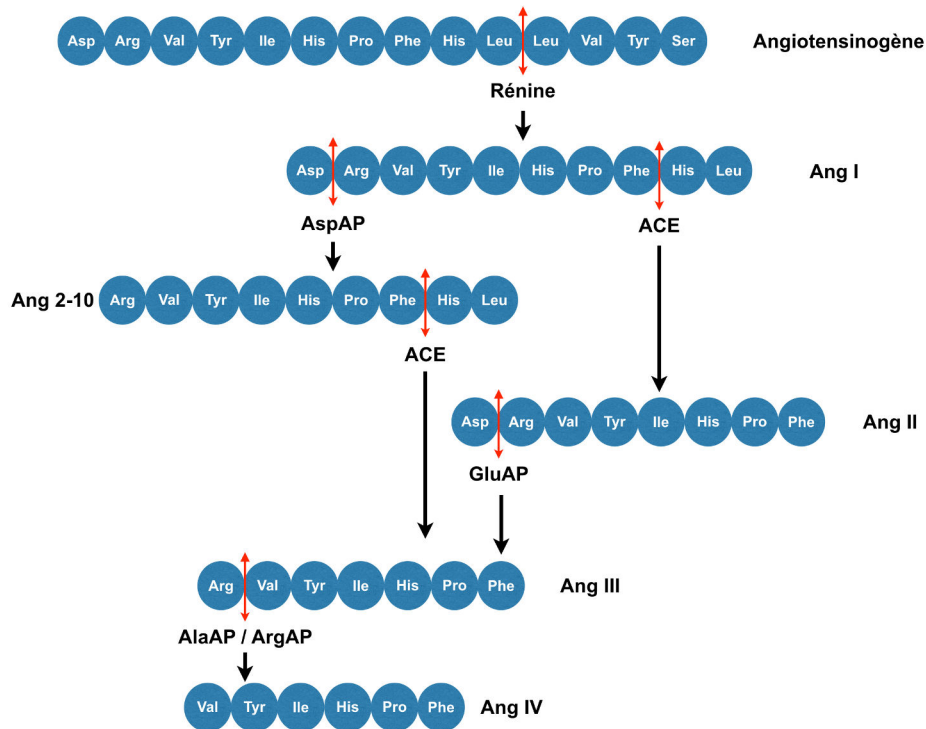
dans la première et la dernière étapes de la dégradation protéique (Gonzales and Robert-Baudouy, 1996).

Au-delà de la simple dégradation des protéines et des peptides, les aminopeptidases sont aussi des actrices importantes dans d'autres processus physiologiques tels que la présentation antigénique par les complexes majeurs d'histocompatibilité de classe I : CMH-I. En effet, la protéolyse cytosolique est une source majeure des ligands (antigènes) liés aux CMH-I (Rock and Goldberg, 1999; York et al., 1999). L'effecteur principal de cette tâche est le protéasome (Niedermann et al., 1999) qui, avec d'autres protéases et aminopeptidases, est impliqué dans une production efficace des antigènes et une augmentation de la variété des produits antigéniques (Kessler et al., 2002). C'est par exemple le cas de la leucine aminopeptidase cytosolique LAP ubiquitaire, décrite comme étant activée par l'interféron IFN- $\gamma$  (Beninga et al., 1998), de la bléomycine hydrolase, une aminopeptidase à cystéine, et d'une aminopeptidase du réticulum endoplasmique, qui coopèrent avec le protéasome pour la synthèse antigénique (Del-Val and López, 2002).

Une autre fonction a été rapportée pour les leucyl aminopeptidases LAP humaine et bLAP bovine qui sont censées hydrolyser le dépôt des protéines au niveau du cristallin de l'œil (Taylor et al., 1982). Ainsi, une altération de l'activité de ces aminopeptidases induit une accumulation des protéines cristallines, originaire de la cataracte. De plus, une diversité d'aminopeptidases (alanyl, arginyl, aspartyl, cysteinyl et glutamyl aminopeptidases) sont aussi décrites pour leur participation dans la cascade enzymatique du système RAS (*renin-angiotensin system*) et pour leur rôle crucial dans le contrôle de la pression sanguine (**Figure 5**) (Banegas et al., 2006).

Non seulement les aminopeptidases interviennent dans les mécanismes physiologiques, mais elles sont aussi des actrices majeures durant les infections virales, bactériennes et parasitaires. À titre d'exemple, le travail effectué par Pulido-Cejudo et ses collègues dévoile l'implication d'une LAP (leucyl aminopeptidase) humaine dans l'infection virale par le VIH (virus d'immunodéficience humaine). En effet, une activité « LAPase » élevée a été identifiée dans le sérum des patients infectés en comparaison aux individus sains. En outre, le traitement de certaines souches cellulaires avec des LAPs suivi de leur infection par le VIH, rend ces cellules beaucoup plus susceptibles à l'infection (Pulido-Cejudo et al., 1997).

## INTRODUCTION



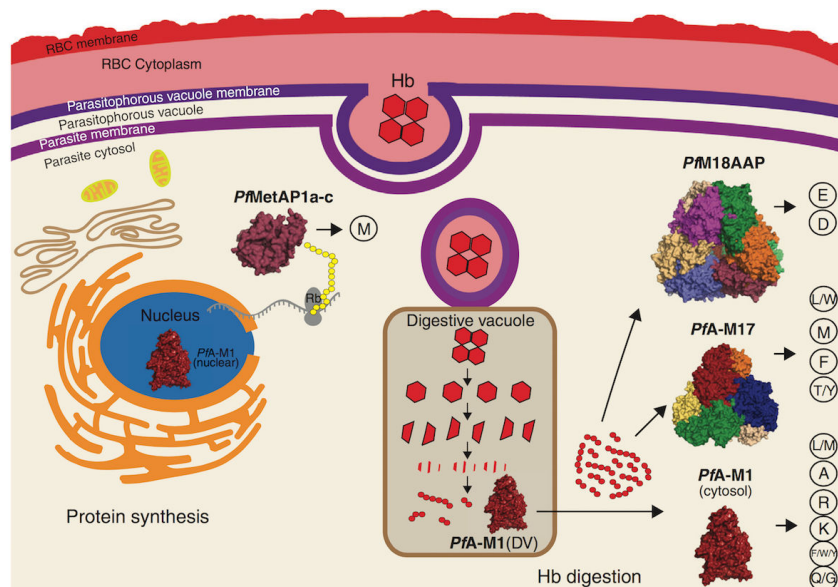
**Figure 5 : Représentation schématique partielle du système RAS.** Les aminopeptidases interviennent dans la cascade enzymatique de l'angiotensinogène et de différentes formes de l'angiotensine (Ang). AspAP, GluAP, AlaAP et ArgAP sont respectivement aspartyl, glutamyl, alanyl et arginyl aminopeptidases. ACE est *angiotensin-converting enzyme* (Banegas et al., 2006).

Un autre exemple illustrant le rôle des aminopeptidases dans les processus infectieux est celui de deux aminopeptidases de *Staphylococcus aureus* et de *Treponema denticola* qui sont qualifiées de cysteinylglycinase en raison de leur habilité à hydrolyser le dipeptide Cys-Gly, généré après hydrolyse du résidu glutamate de l'extrémité Nt du glutathion. Ainsi, ces deux aminopeptidases pourraient intervenir dans le métabolisme du glutathion, en produisant du sulfure d'hydrogène, H<sub>2</sub>S, toxique (Chu et al., 2008; Carroll et al., 2013). Ce dernier est connu pour son implication dans la dégradation des tissus touchés par une infection parodontale (Rizzo, 1967). L'aminopeptidase de *Staphylococcus aureus* fut la première connue pour son implication dans la virulence d'une bactérie à Gram positif (Carroll et al., 2012). Les études montrent que la mutation du gène *pepZ* codant pour cette aminopeptidase est accompagnée d'une atténuation de la virulence chez les modèles murins et les macrophages infectés par *S. aureus* (Carroll et al., 2012). Toutefois, les cibles potentielles de cette LAP restent encore inconnues (Carroll et al., 2013).

Pour finir, les aminopeptidases de certaines espèces de protozoaires représentent une cible pour le développement de drogues antiparasitaires (Howarth and Lloyd, 2000). C'est le cas

## INTRODUCTION

par exemple de *Plasmodium falciparum* qui exploite une diversité d'aminopeptidases pour la dégradation de l'hémoglobine humaine telles que l'alanyl (PfA-M1), la prolyl, la leucyl (PfA-M17), et l'aspartyl (PfM18AAP) aminopeptidases (Dalal and Klemba, 2007; Stack et al., 2007; Maric et al., 2009; McGowan, 2013) (**Figure 6**). Notamment, la leucine aminopeptidase PfA-M17, clivant la leucine Nt des peptides générés par la digestion de l'hémoglobine par les protéases parasitaires (Stack et al., 2007), joue un rôle vital pour le développement parasitaire. En effet, la leucine (constituant 24 % des acides aminés de l'hémoglobine humaine) est requise pour être échangée contre l'isoleucine absente dans l'hémoglobine humaine (Martin and Kirk, 2007) et que le parasite est incapable de synthétiser (Sherman, 1977). L'inhibition de PfM17LAP, hautement active durant le stade intra-érythrocytaire, par la bestatine (inhibiteur des leucine aminopeptidases) empêche la croissance du parasite (Skinner-Adams et al., 2007; Maric et al., 2009).



**Figure 6 : Implication de certaines aminopeptidases de *Plasmodium falciparum* dans la digestion de l'hémoglobine humaine.** Les aminopeptidases parasitaires jouent un rôle essentiel dans la dégradation des peptides générés par la digestion de l'hémoglobine humaine. PfA-M1 (alanyl aminopeptidase) et PfA-M17 (leucyl aminopeptidase) sont préférentielles des résidus neutres et chacune présente un large spectre d'activité. PfM18AAP est une aspartyl aminopeptidase spécifique des résidus acides (McGowan, 2013).

### 1.3. LES AMINOPEPTIDASES TÉTRAÉDRIQUES « TETs »

Comme nous venons de voir l'importance de la protéolyse dans les fonctions physiologiques et pathologiques, celle-ci constitue encore un aspect important d'adaptation aux environnements extrêmes (Ward et al., 2002). Jouissant de telles conditions, les archées

## INTRODUCTION

---

représentent une source riche en peptidases, vu leur capacité à puiser des protéines le carbone dont elles ont besoin (Adams, 1998). Ainsi, le chapitre suivant sera consacré aux métallo-aminopeptidases ATP-indépendantes dodécamériques tétraédriques, nommées TETs et constituant les objets principaux de mon projet de thèse.

### 1.3.1. Historique de la découverte des aminopeptidases TETs

La première description structurale d'une nouvelle peptidase ATP-indépendante remonte à 2002, grâce aux travaux réalisés par B. Franzetti et son équipe dans le cadre de la recherche des complexes chaperonnes et protéases associés au stress salin chez *Haloarcula marismortui*. C'est au cours de la purification du protéasome de l'archée halophile *H. marismortui* qu'un grand complexe de 500 kDa possédant une forme singulière est observé par microscopie électronique en coloration négative. Contrairement aux différentes structures de protéases et de peptidases déjà publiées à l'époque, ce nouveau complexe adoptait une configuration spéciale en tétraèdre creux, généré par l'auto-association de douze sous-unités (Franzetti et al., 2002). Celui-ci fut nommé « TET » en référence à la nomenclature du complexe peptidase « TRI » qui tire son nom de sa forme en chapeau tricorne (Tamura et al., 1996).

Depuis la découverte de cette nouvelle structure tétraédrique parmi les peptidases auto-compartimentées ATP-indépendantes, plusieurs laboratoires se sont intéressés à la biologie des TETs qui sont des machines cellulaires ubiquitaires présentes dans les trois domaines du vivant. Les études initiales ont montré que les TETs d'archées halophiles sont des aminopeptidases à large spectre d'action. Elles peuvent hydrolyser d'une façon non processive des polypeptides dépliés allant jusqu'à 39 acides aminés en composition (Franzetti et al., 2002). Actuellement, aucune activité endopeptidase, carboxypeptidase, ni di- ou tri-peptidase n'est relevée pour les TETs. La recherche de la fonction de ce nouveau type de peptidase géante à travers des études enzymatiques et structurales a été ensuite conduite chez les archées, les bactéries et les eucaryotes.

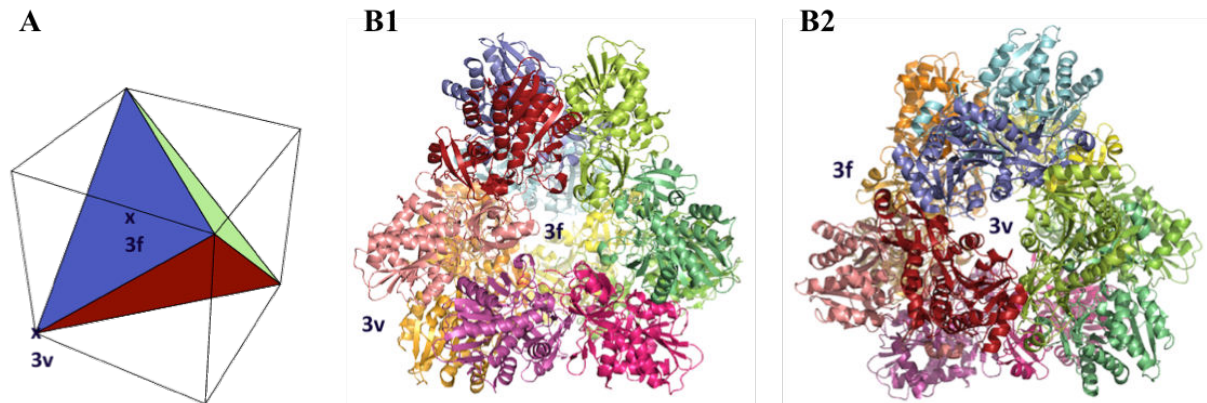
### 1.3.2. Définition d'une aminopeptidase TET

Comme souligné précédemment, le complexe TET tire son nom de sa structure quaternaire tétraédrique. Identifiées dans les trois domaines du vivant, les TETs sont des métallo-aminopeptidases appartenant aux familles M18 et M42 du clan MH (Métallo-protéase H) selon la classification MEROPS (Rawlings et al., 2012). Les caractérisations enzymatiques

## INTRODUCTION

des TETs sont entamées en utilisant une variété de substrats chromogéniques, fluorogéniques et de peptides de différentes tailles et compositions. Des études plus exhaustives ont été menées dans notre laboratoire en étudiant le profil de digestion de certains peptides par HPLC (*High Performance Liquid Chromatography*) (Franzetti et al., 2002; Durá et al., 2005; Appolaire et al., 2014a). Ces investigations démontrent que les TETs sont des métallo-exopeptidases exclusives agissant sur le côté N-terminal libre des substrats. Les cofacteurs métalliques activateurs de ces enzymes sont divers, mais chez la plupart des TETs archéennes, le cobalt est le principal activateur métallique (Ando et al., 1999; Durá et al., 2005; Schoehn et al., 2006; Durá et al., 2009; Durá and Franzetti, 2013; Colombo et al., 2016).

Le tétraèdre dodécamérique creux est édifié par l'auto-association de douze sous-unités formant quatre faces et six arêtes. L'unité de base (le *building block*) à l'origine de la formation du complexe TET dodécamérique est un dimère élaboré par l'association de deux monomères antiparallèlement orientés l'un par rapport à l'autre le long d'une symétrie d'ordre 2 (Borissenko and Groll, 2005; Schoehn et al., 2006; Rosenbaum et al., 2011; Appolaire et al., 2013). La particule TET est percée de quatre pores situés au niveau des faces, et de quatre orifices de diamètre plus réduit localisés au niveau des sommets du tétraèdre (**Figure 7**).



**Figure 7 : A. Représentation schématique d'un tétraèdre TET.** Les deux types de pores 3f et 3v sont localisés au niveau des faces (f) et des sommets (v), respectivement. Le chiffre 3 correspond à l'ordre de l'axe de symétrie. **B. Dodécamère de PhTET1 (PDB ID : 2WYR).** Chacune des douze sous-unités est représentée par une couleur. **B1.** Vue depuis le pore d'une face (3f), **B2.** Vue depuis le pore du sommet (3v) (Schoehn et al., 2006).

Les pores des faces ( $\sim 20$  Å) sont engendrés par la jonction de six sous-unités reliées par une symétrie d'ordre 3. Alors que les pores des apex ( $\sim 9$  Å) sont générés par l'association en mode circulaire de trois monomères liés par une symétrie d'ordre 3 (Franzetti et al., 2002; Schoehn et al., 2006; Durá et al., 2009). Il a été proposé que les pores des faces constituent la

## INTRODUCTION

---

porte d'entrée des polypeptides substrats dépliés vers l'intérieur de la particule où les sites actifs sont hébergés dans les quatre chambres catalytiques, chacune située au niveau du sommet du tétraèdre. Après l'hydrolyse du résidu Nt du substrat par l'un des trois sites actifs de la chambre catalytique, l'acide aminé est expulsé à l'extérieur de cette machine par un mécanisme qui reste à déterminer.

### 1.3.3. Les aminopeptidases TETs existent dans les trois domaines du vivant

Dans les génomes, la famille des peptidases M42 reste mal définie, puisque certains de ses membres sont annotés comme des cellulases partageant une similarité de séquence avec une endoglucanase CelM de *Clostridium thermocellum* (Dutoit et al., 2012). En effet, NCBI CDD (*Conserved Domain Database*) caractérise la famille M42 comme une famille regroupant des aminopeptidases, des endoglucanases et des protéines de l'opéron Fry, impliquée dans la biosynthèse et la dégradation des polysaccharides (Marchler-Bauer et al., 2011). Tel est le cas de TET de *Haloarcula marismortui* (HmTET) et TmPep1050 de *Thermotoga maritima*, désignées comme des endoglucanases putatives en raison de leur homologie de séquence avec la protéine CelM de *Halobacterium* sp. (Ng et al., 2000) et de *C. thermocellum*, respectivement. Ainsi A. Appolaire, durant sa thèse, a déterminé des motifs structuraux conservés facilitant l'identification des peptidases TETs dans les génomes (**paragraphe 1.3.4.3.**). Un travail (non publié) réalisé par S. Gribaldo à l'Institut Pasteur de Paris, en utilisant ces motifs conservés, a permis de montrer que les TETs sont présentes dans de nombreux de génomes archéens et bactériens, souvent en multiple copies voire allant à quatre (voir partie **Problématique et objectifs** de ce manuscrit).

Depuis la découverte de HmTET chez l'archée halophile *H. marismortui*, plusieurs types de peptidases TETs ont été identifiées et caractérisées au plan structural et/ou fonctionnel chez les archées, les bactéries et les eucaryotes (**Tableau 2**). Lorsqu'on les produit sous forme recombinante dans *Escherichia coli*, ces complexes ont tous en commun la faculté de s'auto-assembler en dodécamère stable possédant la forme de tétraèdre typique d'une TET.



## INTRODUCTION

		Espèce	Uniport ID	Famille	Activité aminopeptidase	Ion métallique activateur
TETs archéennes	HmTET	<i>Haloarcula marismortui</i>	Q5V3W6	M42	Spécifique des résidus neutres et basiques, à spectre large	Cobalt
	PhTET1	<i>Pyrococcus horikoshii</i>	O58255	M42	Glutamyl/ Aspartyl aminopeptidase	Cobalt
	PhTET2	<i>P. horikoshii</i>	O59196	M42	Leucyl aminopeptidase à spectre large	Cobalt
	PhTET3	<i>P. horikoshii</i>	O59485	M42	Lysyl aminopeptidase	Cobalt
	PfTET3	<i>P. furiosus</i>	Q8TZW4	M42	Lysyl aminopeptidase	Cobalt
	DkamTET	<i>Desulfurococcus kamchatkensis</i>	B8D484	M42	Spécifique des résidus neutres	Manganèse > Magnésium
	TkDAP1	<i>Thermococcus kodakarensis</i>	Q5JHA5	M42	Spécifique des résidus hydrophobes	Cobalt
	TkDAP2	<i>T. kodakarensis</i>	Q5JH87	M42	Spécifique des résidus hydrophobes	Cobalt
TETs bactériennes	SpPepA	<i>Streptococcus pneumoniae</i>	Q8DNJ7	M42	Glutamyl aminopeptidase	Cobalt ≥ Manganèse
	CelM	<i>Clostridium thermocellum</i>	P55742	M42	Leucyl aminopeptidase	Cobalt
	MHJ_0125	<i>Mycoplasma hyopneumoniae</i>	Q4AAK4	M42	Glutamyl aminopeptidase	Cobalt
	PaAP	<i>Pseudomonas aeruginosa</i>	Q9HYZ3	M18	Aspartyl aminopeptidase	Manganèse > Cobalt
TETs eucaryotes	PfM18AAP	<i>Plasmodium falciparum</i>	Q8I2J3	M18	Aspartyl aminopeptidase	Cobalt >> Manganèse
	hDNPEP	<i>Homo sapiens</i>	Q9ULA0	M18	Aspartyl aminopeptidase	Cobalt > Manganèse
	bDNPEP	<i>Bos taurus</i>	Q2HJH1	M18	Aspartyl aminopeptidase	Manganèse > Cobalt

**Tableau 2 : Aperçu de différentes TETs décrites dans la littérature chez les archées, les bactéries et les eucaryotes.** Leurs activités aminopeptidases et les cofacteurs métalliques requis pour leur activation sont mentionnés.

## INTRODUCTION

---

Chez les archées, le complexe TET de *H. marismortui* s'étant révélé difficile à produire sous une forme de protéine recombinante soluble, le laboratoire s'est tourné vers la caractérisation fonctionnelle et structurale de 3 protéines homologues présentes chez l'archée hyperthermophile anaérobie *Pyrococcus horikoshii*. Ces protéines sont codées par les trois gènes : *PH0519*, *PH1527* et *PH1821*. La chromatographie d'exclusion de taille montre que ces trois protéines sont des complexes dodécamériques (12-mères) (Porciero et al., 2005). Ces complexes enzymatiques ont été nommés PhTET1, PhTET2 et PhTET3, respectivement. L'objectif était d'obtenir les structures cristallographiques de ces TETs et de déterminer leurs spécificités catalytiques respectives afin de comprendre la fonction et le mode d'action de ce nouveau type de peptidase géante ATP-indépendante. Antérieurement, une première étude sur PhTET1 codée par le gène *PH0519* avait été publiée en 1999 par Ando et ses collaborateurs. Les auteurs avaient nommé cette protéine DAP car ils ont mis en évidence une activité *acylamino acid releasing (deblocking) aminopeptidase* capable de cliver les extrémités Nt libres et bloquées des peptides substrats (Ando et al., 1999). Cependant, le travail réalisé sur PhTET1, dans notre laboratoire, n'a révélé aucune activité *deblocking* aminopeptidase (Schoehn et al., 2006). En revanche, l'enzyme est bien une aminopeptidase possédant une spécificité de substrat très étroite dirigée contre les résidus acides : acides glutamique et aspartique (Durá and Franzetti, 2013). Sa structure cristallographique a été déterminée à 2,24 Å de résolution (PDB ID : 2WYR). L'utilisation de la microscopie électronique et le phasage avec des complexes de lanthanides développé par E. Girard dans le groupe ont permis d'identifier un réseau élaboré de canaux et de poches catalytiques situés à l'intérieur de la particule TET (Schoehn et al., 2006; Talon et al., 2011).

Parallèlement à cette étude, la structure de PhTET2 (PDB ID : 1XFO/1Y0R) a été déterminée par deux équipes concurrentes, celle de Baumann et celle de Groll (Russo and Baumann, 2004; Borissenko and Groll, 2005). Ces structures représentent les premières structures cristallographiques des particules TETs. Enfin en 2009, notre laboratoire a publié la structure cristallographique de PhTET3 (PDB ID : 2WZN) (Durá et al., 2009). Toutes ces structures révèlent de grande similarité dans l'organisation et le mode d'action des particules TETs. Dans le cadre de la caractérisation des TETs, les spécificités et le mode d'action des trois aminopeptidases PhTET1, PhTET2 et PhTET3 de *P. horikoshii*, ainsi que leur dépendance aux métaux ont été décrits dans notre équipe (Durá et al., 2005, 2009; Durá and Franzetti, 2013). Ces travaux révèlent de grandes disparités dans leur capacité à hydrolyser les différents types d'acides aminés. Comme déjà dit, PhTET1 est une aspartyl/glutamyl

## INTRODUCTION

---

aminopeptidase, PhTET2 apparaît comme possédant un champ d'action plus large, surtout envers les acides aminés neutres et aliphatiques (Durá et al., 2005). Tandis que PhTET3 est une lysyl aminopeptidase qui coupe très efficacement les résidus basiques (Durá et al., 2009). En dépit de ces spécificités marquées, les trois complexes possèdent des similitudes prononcées au plan structural. Ainsi, l'étude comparative de leurs structures a révélé des caractéristiques communes aux complexes TETs et des spécificités pouvant expliquer leurs différentes activités. Chez les thermococcales, trois gènes sont identifiés chez *P. furiosus* (PF0369, PF1547 et PF1861) et *P. abyssi* (PAB0437, PAB2096, PAB2375) codant, respectivement, pour des analogues de PhTET1, 2 et 3 de *P. horikoshii* (Mori and Ishikawa, 2005). Parmi les TETs de *P. furiosus*, seule PFTET3 codée par le gène PF1861 a été étudiée et son identité de lysyl aminopeptidase a été confirmée (Story et al., 2005; Colombo et al., 2016).

De surcroît, chez *Desulfurococcus kamchatkensis*, une archée hyperthermophile, une autre TET : APDKam589 (nommée ici DkamTET) a été identifiée (Slutskaya et al., 2012; Petrova et al., 2015). Chez *Thermococcus kodakarensis*, deux protéines TkDAP1 et TkDAP2 (toujours nommées DAP par les auteurs) sont identifiées comme des complexes dodécamériques (Jia et al., 2011). Enfin chez *T. onnurineus*, TNA1\_DAP1 forme un édifice dodécamérique (Lee et al., 2007) dont la structure reste à déterminer. Bien que les caractérisations enzymatiques de ces enzymes aient été réalisées avec un jeu réduit de substrats synthétiques commerciaux, on peut néanmoins conclure qu'elles sont des *deblocking* aminopeptidases spécifiques des résidus hydrophobes et neutres.

Certaines aminopeptidases bactériennes adoptent également une structure tétraédrique. La première structure d'une TET bactérienne a été publiée pour SpPepA exprimée par *Streptococcus pneumoniae* (Kim et al., 2010). Les études structurales et enzymatiques montrent que SpPepA adopte une structure de tétraèdre dodécamérique et exhibe une activité glutamyl aminopeptidase. Ce complexe appartenant à la famille M42 est capable d'hydrolyser les résidus d'acides glutamiques et moins efficacement d'acides aspartiques. En 2012, deux autres protéines de la même famille M42 : CelM et TmPep1050 ont fait l'objet d'une étude fonctionnelle. CelM et TmPep1050 exprimées par *Clostridium thermocellum* et *Thermotoga maritima*, respectivement, ont été définies comme des endoglucanases. En revanche, aucune activité cellulase n'a été détectée pour ces deux enzymes. Au contraire, CelM et TmPep1050 présentent une activité aminopeptidase contre les résidus neutres avec une préférence pour la leucine (Dutoit et al., 2012). Grâce à ce travail, Dutoit et ses collègues ont souligné le

## INTRODUCTION

---

problème de la mauvaise annotation des aminopeptidases de la famille M42, évoqué précédemment. En solution, seule la protéine recombinante CelM génère des complexes dodécamériques contrairement à TmPep1050 qui existe sous forme dimérique en solution (Dutoit et al., 2012) et dodécamérique dans le cristal (PDB ID : 4P6Y). Une étude structurale de CelM est requise pour vérifier si le dodécamère généré adopte la conformation propre aux TETs. De plus, l'activité enzymatique et la prédiction de la structure d'une troisième TET bactérienne de la famille M42 : MHJ\_0125 de *Mycoplasma hyopneumoniae* ont été publiées en 2013. En utilisant des substrats chromogéniques commerciaux couplés à l'AMC (*7-Amino-4-methylcoumarin*), les auteurs ont pu révéler une activité glutamyl aminopeptidase de MHJ\_0125 avec une capacité beaucoup plus réduite à hydrolyser les acides aspartiques. Cette aminopeptidase est exprimée en solution sous forme dodécamérique dont la prédiction de structure confirme une structure quaternaire en tétraèdre creux (Robinson et al., 2013).

Les TETs bactériennes sont des aminopeptidases M42, hormis PaAP de *Pseudomonas aeruginosa*, qui appartient à la famille M18 du même clan MH que les M42. De manière identique à ses homologues bactériens, la caractérisation enzymatique et structurale de PaAP dévoile une activité aminopeptidase spécifique des résidus acides et une organisation en un tétraèdre dodécamérique. Une propriété d'aspartyl aminopeptidase est attribuée à PaAP qui peut dégrader les deux résidus acides avec une efficacité comparable (Nguyen et al., 2014).

Quant aux aminopeptidases eucaryotes, jusqu'à présent trois ont été décrites : une chez *Plasmodium falciparum* PfM18AAP (Sivaraman et al., 2012) et deux autres chez les mammifères : hDNPEP humaine (Chaikuad et al., 2012) et bDNPEP de *Bos taurus* (Chen et al., 2012). Les investigations réalisées sur les plans fonctionnel et structural montrent que les trois enzymes sont des aspartyl aminopeptidases adoptant la conformation caractéristique d'une TET. De plus chez *Saccharomyces cerevisiae*, Lap4 (ou API) une leucyl aminopeptidase (Chang and Smith, 1989) et Yhr113wp (ou Ape4), une aspartyl aminopeptidase (Yokoyama et al., 2006), sont décrites comme des complexes dodécamériques. Une détermination de la structure de ces deux aminopeptidases de la famille M18 sera requise pour déterminer la structure quaternaire de ces deux aminopeptidases. Contrairement aux TETs archéennes, les TETs eucaryotes sont les moins étudiées dans les trois domaines du vivant.

### 1.3.4. Caractéristiques communes à toutes les aminopeptidases TETs

Comme nous venons de voir, les TETs sont des enzymes ubiquitaires exprimées par les archées, les bactéries et les eucaryotes. Les caractéristiques communes partagées par ces aminopeptidases sont décrites dans la partie suivante.

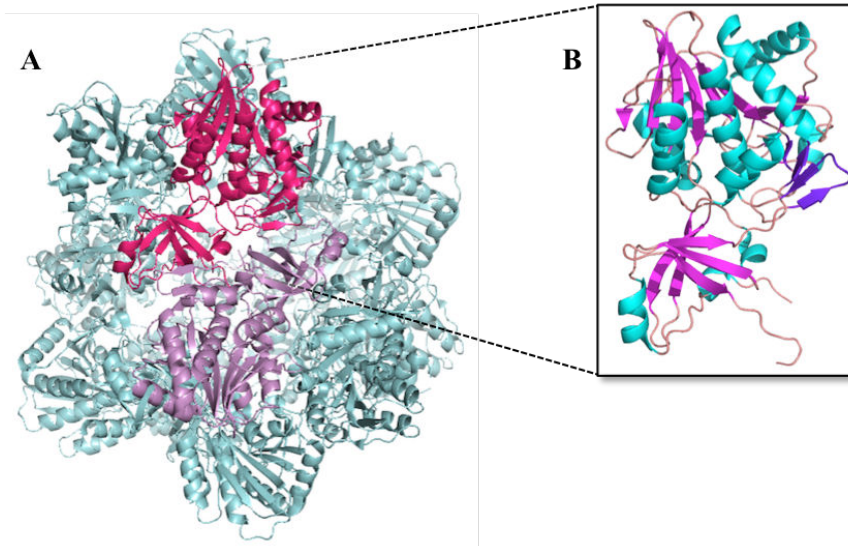
#### 1.3.4.1. Structure du monomère

En solution, les TETs adoptent une structure de tétraèdre creux qui les distingue des protéases et des peptidases ATP-dépendantes et indépendantes. Cet édifice est élaboré par l'assemblage de six *building blocks*. Chacun de ces *building blocks* est engendré par l'assemblage antiparallèle de deux monomères (**Figure 8A**). La structure tertiaire du monomère est conservée à travers les trois domaines du vivant (Russo and Baumann, 2004; Schoehn et al., 2006; Durá et al., 2009; Kim et al., 2010; Sivaraman et al., 2012; Chaikuad et al., 2012; Nguyen et al., 2014). Les études cristallographiques montrent qu'il est formé de deux domaines distincts : un domaine catalytique et un domaine de dimérisation dénommé encore domaine régulateur (Sivaraman et al., 2012).

Le domaine catalytique est un domaine globulaire aplati arrangé en trois couches, où un feuillet  $\beta$  central est pris en « sandwich » de part et d'autre par deux groupes d'hélices  $\alpha$ . À la périphérie du domaine catalytique, se trouve un petit feuillet  $\beta$  qui interagit avec le domaine de dimérisation d'une autre sous-unité adjacente pour former le dimère. Dans certaines publications, ce petit feuillet  $\beta$  est défini comme un deuxième domaine de dimérisation (II) indépendant du domaine catalytique (Borissenko and Groll, 2005; Kim et al., 2010).

Le domaine de dimérisation (I) adopte un repliement typique des *PDZ-like domains* en forme de « croissant » (Russo and Baumann, 2004; Schoehn et al., 2006). Il est composé d'hélices  $\alpha$  ainsi que d'un petit feuillet  $\beta$  organisé en une sorte de tonneau (**Figure 8B**).

La conservation des domaines catalytiques et de dimérisations entre les TETs est associée à une variation du nombre des structures secondaires : feuillets  $\beta$  et hélices, qui les composent. En dépit de ces variations, les structures des monomères des TETs sont globalement similaires. Le **Tableau 3** présente les structures secondaires édifiant les domaines catalytiques et de dimérisations TETs archéennes, bactériennes et eucaryotes.



**Figure 8 : A. Localisation du *building block* (dimère) au niveau de l'arête du tétraèdre dodécamérique. B. Structure tertiaire du monomère de PhTET1 (PDB ID : 2WYR) comportant deux domaines.** Le domaine catalytique (partie supérieure) est formé d'un feuillet  $\beta$  englobé par sept hélices  $\alpha$ . Le domaine de dimérisation I (*PDZ-like domain*) (partie inférieure) adopte une conformation en tonneau  $\beta$ . Un petit feuillet de quatre brins  $\beta$  (domaine de dimérisation II) (violet) est localisé en surface du domaine catalytique (Schoehn et al., 2006).

## INTRODUCTION

		Famille	PDB ID	Domaine catalytique	<i>PDZ-like domain</i> (domaine de dimérisation I)	Domaine de dimérisation II
TETs archéennes	PhTET1	M42	2WYR	10 brins $\beta$ , 7 hélices $\alpha$ et 1 hélice $3_{10}$	5 brins $\beta$ , 1 hélice $\alpha$ et 2 hélices $3_{10}$	4 brins $\beta$
	PhTET2	M42	1XFO	8 brins $\beta$ , 7 hélices $\alpha$ et 4 hélices $3_{10}$	6 brins $\beta$ , 1 hélice $\alpha$ et 2 hélices $3_{10}$	3 brins $\beta$
	PhTET3	M42	2WZN	8 brins $\beta$ , 7 hélices $\alpha$ et 2 hélices $3_{10}$	6 brins $\beta$ , 1 hélice $\alpha$ et 3 hélices $3_{10}$	3 brins $\beta$
	PfTET3	M42	4X8I	9 brins $\beta$ , 7 hélices $\alpha$ et 1 hélice $3_{10}$	6 brins $\beta$ , 1 hélice $\alpha$ et 2 hélices $3_{10}$	3 brins $\beta$
	DkamTET	M42	4WWV	9 brins $\beta$ , 7 hélices $\alpha$ et 1 hélice $3_{10}$	6 brins $\beta$ , 1 hélice $\alpha$ et 2 hélices $3_{10}$	3 brins $\beta$
TETs bactériennes	SpPepA	M42	3KL9	11 brins $\beta$ , 8 hélices $\alpha$ et 2 hélices $3_{10}$	6 brins $\beta$ , 1 hélice $\alpha$ et 2 hélices $3_{10}$	3 brins $\beta$
	MHJ_0125	M42	n.d.			
	PaAP	M18	4NJQ	11 brins $\beta$ , 8 hélices $\alpha$ et 1 hélice $3_{10}$	6 brins $\beta$ , 2 hélices $\alpha$ et 3 hélices $3_{10}$	3 brins $\beta$
TETs eucaryotes	PfM18AAP	M18	4EME	10 brins $\beta$ , 9 hélices $\alpha$ et 4 hélices $3_{10}$	6 brins $\beta$ , 2 hélices $\alpha$ et 3 hélices $3_{10}$	4 brins $\beta$
	hDNPEP	M18	4DYO	9 brins $\beta$ , 8 hélices $\alpha$ et 3 hélices $3_{10}$	6 brins $\beta$ , 2 hélices $\alpha$ et 3 hélices $3_{10}$	5 brins $\beta$
	bDNPEP	M18	3VAT	11 brins $\beta$ , 9 hélices $\alpha$ et 3 hélices $3_{10}$	6 brins $\beta$ , 2 hélices $\alpha$ et 3 hélices $3_{10}$	3 brins $\beta$

**Tableau 3 : Nombre de structures secondaires composant les domaines catalytiques et les domaines de dimérisation (*PDZ-like domain* (domaine I) et petit feuillet  $\beta$  (domaine II)) des TETs archéennes, bactériennes et eucaryotes.** Les structures secondaires ont été déterminées grâce aux programmes DSSP (Kabsch and Sander, 1983) et PYMOL.

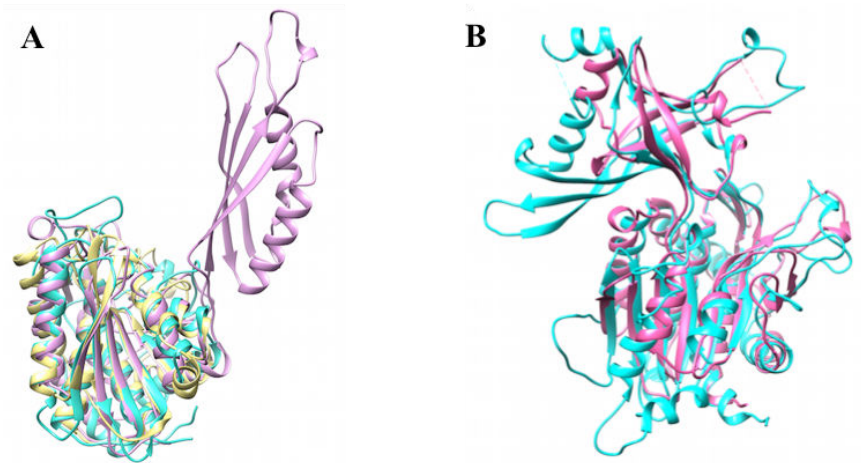
Comme nous l'avons vu précédemment, le clan MH comporte quatre familles de métallopeptidases : M18, M20, M28 et M42 (**Tableau 1, Chapitre 1.2.**). Les monomères des aminopeptidases tétraédriques des familles M18 et M42 partagent des structures tertiaires voisines, en comparaison aux familles M20 et M28. Toutefois, les domaines protéolytiques, particulièrement le cœur en feuillet  $\beta$  de ces quatre familles sont superposables (rmsd  $\sim$  2,3 Å) (Chaikuad et al., 2012). Cela suggère donc une stratégie évolutive conservée pour la coordination métallique et la catalyse enzymatique (Lowther and Matthews, 2002).

## INTRODUCTION

---

Or, la différence majeure entre ces familles et qui caractérise les monomères des peptidases TETs M42 et M18, est la présence d'un domaine de dimérisation particulier. Les membres de la famille M28 sont dépourvus du domaine de dimérisation ce qui affecte leur état oligomérique, comme le montrent les deux aminopeptidases monomériques : AAP (PDB ID : 1RTQ/1IGB) de *Aeromonas proteolytica* et SgApS (PDB ID : 1F2O) de *Streptomyces griseus* (Chevrier et al., 1994b; Desmarais et al., 2006; Gilboa et al., 2001). En revanche, les membres de la famille M20 présentent un domaine de dimérisation dont le repliement et l'arrangement tertiaires sont différents de ceux de M42 et M18, comme l'illustre la carboxypeptidase G2 de *Pseudomonas sp.* PsCPG2 (PDB ID : 1CG2) (Rowsell et al., 1997) (**Figure 9A**).

Quant aux aminopeptidases de la famille M18, elles présentent une topologie plus proche de celle des M42, bien que les chaînes polypeptidiques des aminopeptidases TETs M18 soient plus longues que celles des M42. Cela se traduit par la présence de structures secondaires supplémentaires chez ces premières, comme le montre le **Tableau 3**. De plus, une différence remarquable est observée au niveau du domaine de dimérisation I qui adopte une structure en tonneau  $\beta$  cylindrique chez M42, tandis que celui-ci est brisé par les hélices  $\alpha$  et  $3_{10}$  chez les M18 et exhibe une conformation de *butterfly-shaped domain* (Sivaraman et al., 2012; Chaikuad et al., 2012; Chen et al., 2012) (**Figure 9B**).

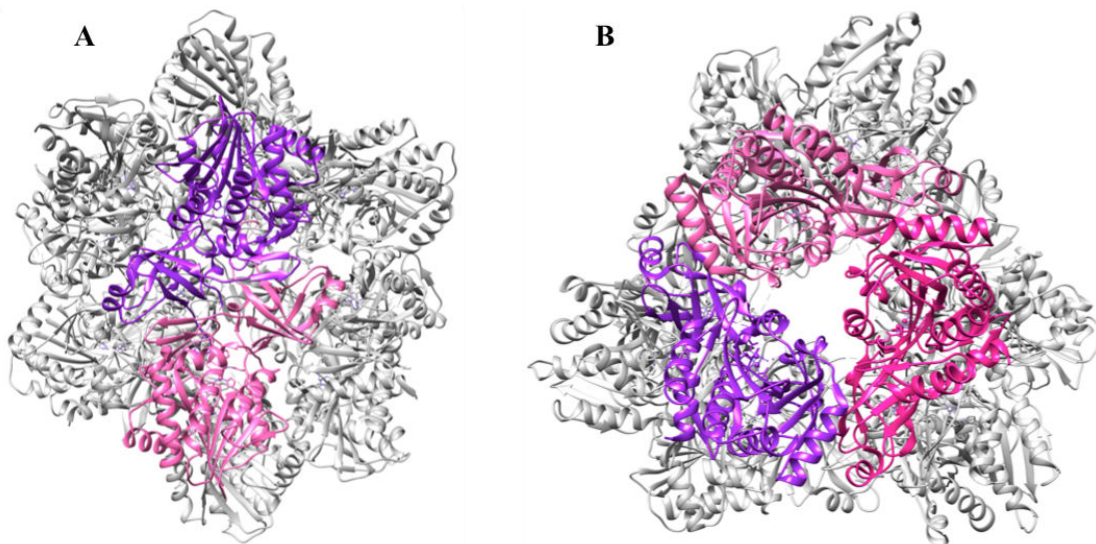


**Figure 9 : Structures tertiaires des monomères des familles M18, M20, M28 et M42 du clan MH. A.** Superposition des monomères de AAP (PDB ID : 1RTQ) (turquoise) et SgApS (PDB ID : 1F2O) (jaune) appartenant à M28 avec celui de PsCPG2 (PDB ID : 1CG2) (rose clair) de la famille M20. Les cœurs catalytiques des trois protéines sont superposables. PsCPG2 présente un domaine de dimérisation supplémentaire par rapport à AAP et SgApS. Celui-ci adopte une conformation complètement différente de celle des M18 et M42. **B.** Superposition des monomères de PhTET2 (PDB ID : 1XFO) (rose foncé) et bDNPEP (PDB ID : 3VAT) (cyan). La structure des cœurs catalytiques de deux aminopeptidases est conservée en dépit du domaine de dimérisation (partie supérieure), qui chez PhTET2 est en forme d'un tonneau cylindrique, alors qu'il est interrompu par deux hélices  $\alpha$  et une hélice  $3_{10}$  chez bDNPEP.



### 1.3.4.2. Les aminopeptidases TETs : des particules dodécamériques uniques

La configuration dodécamérique d'une TET résulte de la particularité de l'angle existant entre le domaine catalytique et le domaine de dimérisation. Il s'agit d'un angle tétraédrique équivalent à  $109^\circ$  chez PhTET2 à l'origine de l'auto-assemblage en 12-mères (Borissenko and Groll, 2005). Deux monomères s'orientent antiparallèlement l'un à l'autre et engendrent un dimère grâce à l'interaction entre le domaine de dimérisation I de l'un avec le petit feuillet  $\beta$  (domaine de dimérisation II) de l'autre. Le dimère ainsi formé constitue l'arête du tétraèdre. L'interface d'oligomérisation élaborée par les domaines catalytiques et les *PDZ-like domains* de trois monomères au niveau de chacune des quatre crêtes, édifie une particule tétraédrique entière (**Figure 10**). Cette structure singulière est traversée par quatre larges pores au niveau des faces du tétraèdre et quatre orifices plus étroits au niveau des sommets. Les premiers trous sont considérés comme les orifices d'entrée des substrats dans la macromolécule et les seconds sont proposés comme les trous de sortie des produits de l'hydrolyse enzymatique (Schoehn et al., 2006). En outre, des ions  $Mg^{2+}$  et des molécules de glycérol ont été observés au niveau des canaux de sortie de hDNPEP, suggérant ainsi un rôle supplémentaire de ces pores apicaux comme une voie de trafic de petites molécules (Chaikuad et al., 2012).



**Figure 10 : Représentation du dimère et du trimère dans le dodécamère de PhTET1 (PDB ID : 2WYR). A.** Observation d'un dimère formant une des arêtes du dodécamère de PhTET1. **B.** Observation d'un trimère situé au niveau du sommet de la particule PhTET1.

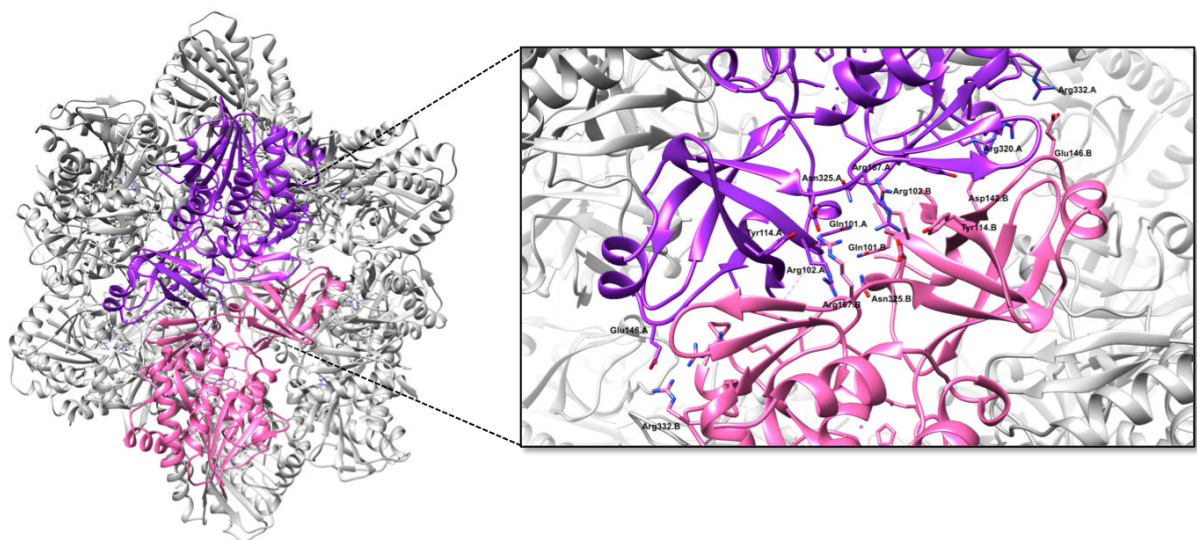
L'importance du domaine de dimérisation dans la formation des complexes oligomériques

## INTRODUCTION

est abordée précédemment (**paragraphe 1.3.4.1.**) pour les peptidases de la famille M28 du clan MH. Il a été proposé que, contrairement aux aminopeptidases et carboxypeptidases monomériques secrétées, l'acquisition du domaine de dimérisation semble avoir eu lieu au cours de l'évolution. Elle a permis la formation des complexes oligomériques et leur localisation cytosolique (Schoehn et al., 2006).

### 1.3.4.3. Interfaces de dimérisation et d'oligomérisation d'une particule TET

La genèse d'une particule dodécamérique tétraédrique TET est initiée par la formation du *building block* dimérique constituant l'une des six arêtes de l'oligomère. Les deux monomères mettent en place un réseau de liaisons hydrogène, de ponts salins et des interactions hydrophobes entre le domaine de dimérisation I d'une sous-unité et le petit feuillet  $\beta$  (domaine de dimérisation II) de l'autre (**Figure 11**).



**Figure 11 : Interface de dimérisation entre deux monomères du *building block* d'une particule TET.** Agrandissement de l'interface de dimérisation entre deux monomères (violet et rose) de PhTET2 (PDB ID : 1XFO). Les résidus qui sont impliqués dans la formation de cette interface sont représentés en *sticks*. Les interfaces de dimérisation ont été générées par PDBePISA.

Les travaux réalisés par A. Appolaire ont montré une conservation de certains résidus appartenant à l'interface de dimérisation chez les aminopeptidases TETs, notamment entre les TETs archéennes (**Figure 12A**). Une séquence consensus, IDIGAXXXE, a été identifiée à l'interface de dimérisation comme motif structural caractéristique d'une particule TET (Appolaire, 2014). En effet, la présence des domaines de dimérisation ne signifie pas que la

## INTRODUCTION

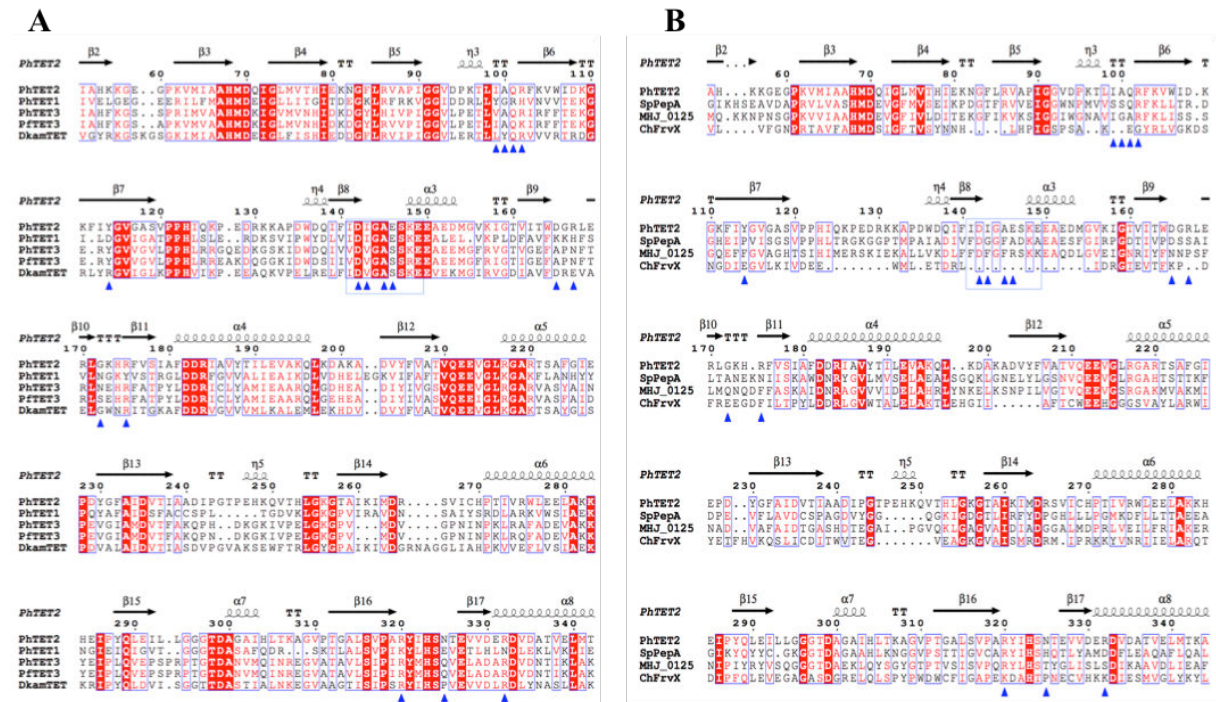
---

particule qui sera formée soit obligatoirement une TET. C'est le cas par exemple de la protéine codée par le gène *frvX* (PDB ID : 3CPX) de la bactérie *Cytophaga hutchinsonii* et appartenant à la famille M42. Cette protéine existe dans l'unité asymétrique du cristal sous forme de trimère à partir duquel un dodécamère ne peut être généré en appliquant les symétries cristallographiques. L'analyse de la séquence de la protéine de *C. hutchinsonii* montre l'absence de la séquence consensus, une cause pouvant expliquer alors son oligomérisation non tétraédrique (**Figure 12B**).

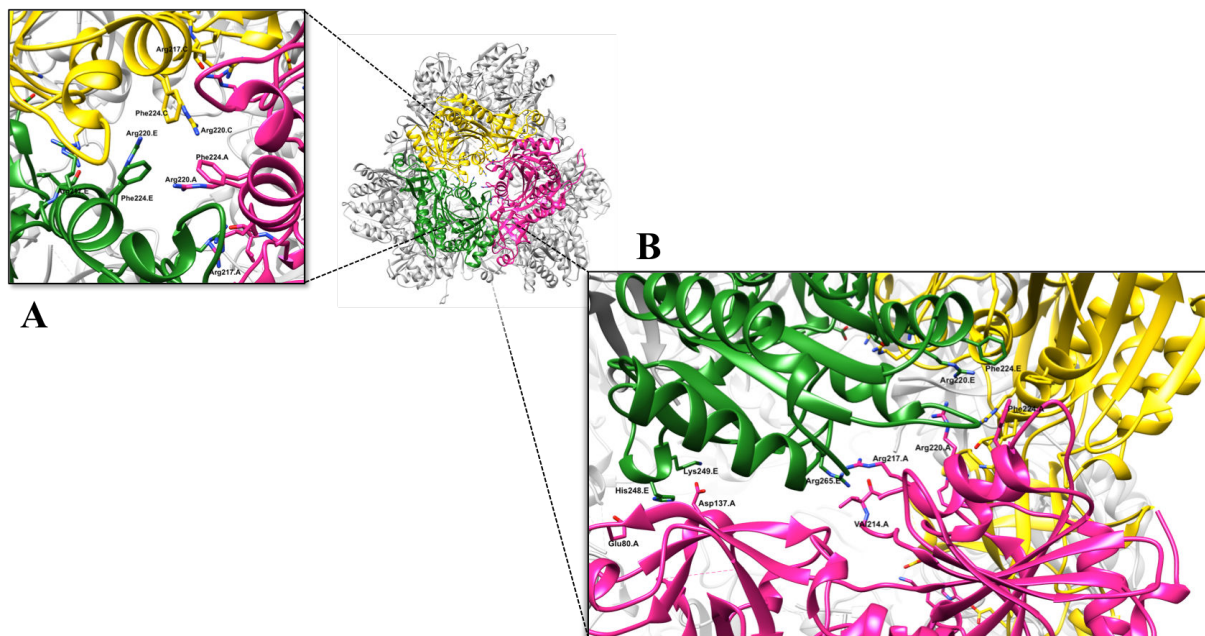
Une fois le *building block* formé, le complexe TET est édifié grâce aux interactions entre les dimères au niveau des sommets du tétraèdre. Celles-ci élaborent un second type d'interface dit interface d'oligomérisation. Ces zones dont la surface est identique à celles des interfaces de dimérisation ( $15125 \text{ \AA}^2$ ) (d'après le serveur d'analyse des surfaces PDBePISA), sont engendrées par un nombre plus restreint de liaisons hydrogène et de ponts salins.

Au niveau de cette interface d'oligomérisation, A. Appolaire a évoqué l'importance de l'hélice  $\alpha 5$  du domaine catalytique (en référence à la séquence de PhTET2) dans la formation du dodécamère (Appolaire, 2014). Cette hélice contient les résidus Val214, Arg217, Arg220 et Phe224 d'un monomère. Ceux-ci interagissent avec les résidus équivalents de deux monomères adjacents au niveau de l'apex de la particule (**Figure 13**). Ces acides aminés sont très similaires entre les TETs archéennes (PhTET1, 2, 3, PfTET3 et DkamTET).

# INTRODUCTION



**Figure 12 : Alignement des séquences des aminopeptidases TETs de la famille M42 archéennes et bactériennes (hormis ChFrvx).** La numérotation des résidus et les structures secondaires correspondent à PhTET2. Les triangles bleus représentent les résidus impliqués dans la formation du réseau de liaisons hydrogène et ponts salins à l'interface de dimérisation entre deux monomères de PhTET2. Le motif IDIGAXXE est marqué par un rectangle bleu. **A.** Alignement des séquences des TETs archéennes : PhTET1, 2 et 3 de *P. horikoshii*, PFTET3 de *P. furiosus* et DkamTET de *D. kamchatkensis*. **B.** Alignement de la séquence de PhTET2 avec celles des TETs bactériennes : SpPepA de *S. pneumoniae* et MHJ\_0125 de *M. hypopneumoniae*. ChFrvX est une protéine M42 codée par le gène *frvx* de *C. hutchinsonii*.

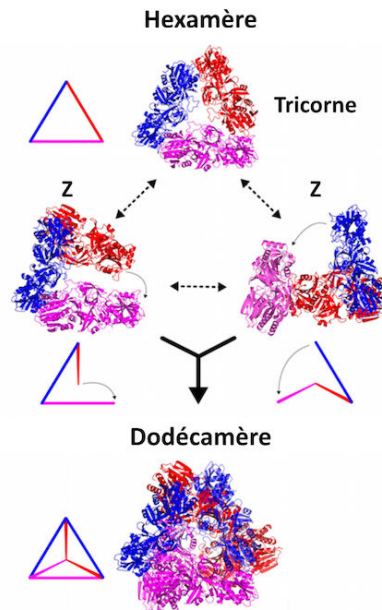


**Figure 13 : Interface d'oligomérisation de PhTET2 (PDB ID : 1XFO).** **A.** L'interface d'oligomérisation est observée du côté du sommet de la particule, généré par l'association de trois monomères (colorés en vert, jaune et rose) reliés par une symétrie d'ordre trois. **B.** Observation du profil de l'interface d'oligomérisation engendré par deux monomères. Les résidus impliqués dans les ponts salins et les liaisons hydrogènes sont représentés en stick. Les interfaces de dimérisation sont générées par PDBEPIISA.

### 1.3.4.4. Mécanisme d'assemblage et d'activation des particules TETs

Dans le but d'élucider le mécanisme d'oligomérisation des édifices dodécamériques et les intermédiaires qui interviennent dans ce processus, cinq mutations (R217S, R220S, F224S, H248S et I292A) sont élaborées au niveau de l'interface dimère-dimère de PhTET2 (Appolaire et al., 2013). Ces mutations ralentissent la genèse du dodécamère en permettant de stabiliser et de purifier une population dimérique capable de s'auto-associer en dodécamère. Les mêmes types de population, dimérique et dodécamérique, sont retrouvés *in vivo* dans *P. horikoshii* et dans *Halobacterium salinarum*, une archée halophile (Appolaire et al., 2013), fait intéressant qui pourrait traduire un mécanisme de régulation original. L'incubation des espèces dimériques résultant de la mutation de PhTET2 à pH 7,5 et à une concentration saline croissante révèle l'apparition successive d'intermédiaires structurés révélés par électrophorèse en conditions natives. Également lors de l'assemblage métal-dépendant des dimères de PhTET3 de *P. horikoshii* (Rosenbaum et al., 2011) et de la purification de PfM18AAP de *P. falciparum* (Teuscher et al., 2007; Sivaraman et al., 2012) et hormis les espèces dodécamériques, d'autres complexes oligomériques sont identifiés, tels que des dimères, des tétramères, des hexamères et des octamères. Ces observations, en combinaison avec la conservation des interfaces d'oligomérisation, permettent de proposer l'existence d'un mécanisme d'oligomérisation commun pour les TETs.

L'apparition chronologique des intermédiaires permet de conclure que l'oligomérisation de PhTET2 est un processus progressif et non pas aléatoire. La détermination de la structure du précurseur dimérique par SAXS ainsi que l'étude de différents intermédiaires par ultracentrifugation analytique et microscopie électronique ont permis de proposer que l'hexamère représente l'espèce intermédiaire de la voie d'assemblage du dodécamère PhTET2. Cet intermédiaire exhibe deux structures différentes : un complexe fermé tricorne et une chaîne ouverte en forme Z. Un modèle dynamique expliquant la séquence d'assemblage a ainsi pu être proposé (**Figure 14**).

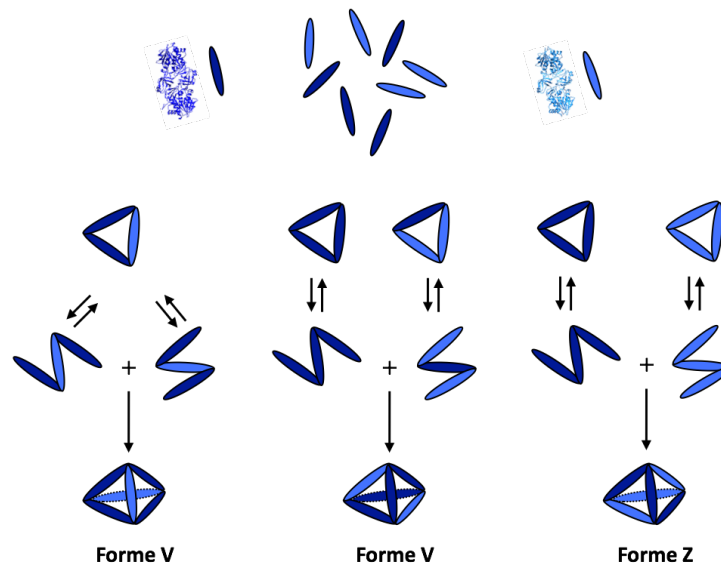


**Figure 14 : Implication d'un intermédiaire hexamérique dans le mécanisme d'assemblage de PhTET2.** Un équilibre dynamique existe entre les deux conformations, fermée (tricorne) et ouverte (Z), de l'hexamère. Deux formes ouvertes en Z s'auto-associent pour édifier une particule dodécamérique stable (Appolaire et al., 2013).

Ce modèle a ensuite été validé par une approche originale utilisant la diffusion neutronique et le marquage des sous-unités par perdeutération. Dans ces expériences, c'est l'hétéro-complexe TET formé par les sous-unités PhTET2 et PhTET3 de *P. horikoshii* qui a été étudié. En effet, Appolaire et ses collègues ont montré la capacité de PhTET2 et PhTET3 à former des hétéro-oligomères *in vitro* et *in vivo* (Appolaire et al., 2014a). Cette aptitude est supportée par une identité de séquence de 48 % entre les deux TETs et une conservation des résidus Arg217 et Arg220 décrits pour leur implication dans l'interface d'oligomérisation entre deux dimères (Appolaire et al., 2014a). Par la suite, cette hétéro-dodécamérisation a été exploitée en utilisant le SANS (*Small Angle Neutron Scattering*) et la variation du contraste entre les sous-unités deutérées de PhTET2 (dPhTET2) et celles hydrogénées de PhTET3 (hPhTET3) pour mieux comprendre le mécanisme de dodécamérisation (Appolaire et al., 2014b). Les résultats de cette étude ont montré que non seulement l'oligomérisation suit un mécanisme régulier, mais encore qu'elle favorise une meilleure topologie au niveau des sommets, combinant ainsi des sites actifs différents et complémentaires (PhTET2 et PhTET3) (**Figure 15**). Cette édification particulière des sites actifs promeut une hydrolyse enzymatique plus efficace des substrats peptidiques par l'hétéro-oligomère comparée à un mélange équimolaire d'homo-oligomères de PhTET2 et PhTET3 (Appolaire et al., 2014a). De plus, la genèse d'une particule dodécamérique pourrait contribuer à une meilleure thermostabilité du complexe oligomérique par rapport à la forme dimérique, comme cela a été décrit pour

## INTRODUCTION

PhTET3 (Rosenbaum et al., 2011).



**Figure 15 : Voie d'assemblage des aminopeptidases TETs révélée par l'hétéro-oligomère PhTET2-PhTET3.** Chaque *building block* est représenté par un bâtonnet en bleu foncé : dPhTET2 (d = deutéré) et bleu clair hPhTET3 (h = hydrogéné). Deux conformations hexamériques : une ouverte (Z) et une autre fermée (tricorne), représentent les intermédiaires de l'assemblage du dodécamère. Deux types de complexes sont générés selon le positionnement des sous-unités PhTET2 : forme V et forme Z conférant une hétérogénéité des trois sites actifs situés au niveau de chaque apex (Adapté depuis (Appolaire et al., 2014b)).

L'étude des propriétés peptidases du précurseur dimérique des TETs montre que l'auto-assemblage en un complexe dodécamérique déclenche l'activité catalytique de TET contre les substrats peptidiques en comparaison au *building block* seulement actif contre les petits substrats (Appolaire et al., 2013). En outre, il est possible que cette oligomérisation favorise encore une interaction physique et fonctionnelle avec des substrats protéiques ou d'autres complexes de dégradation. Alors, une réponse rapide en cas de stress pourrait être exécutée facilitant le fonctionnement cellulaire et la survie dans de telle condition (Borissenko and Groll, 2005). Également, il a été proposé pour DkamTET de *D. kamchatkensis* de pouvoir lier d'autres protéines ou des ligands ayant des structures spécifiques *via* les groupements cycliques des résidus tryptophanes situés à la surface du dodécamère (Samanta et al., 2000; Petrova et al., 2015).

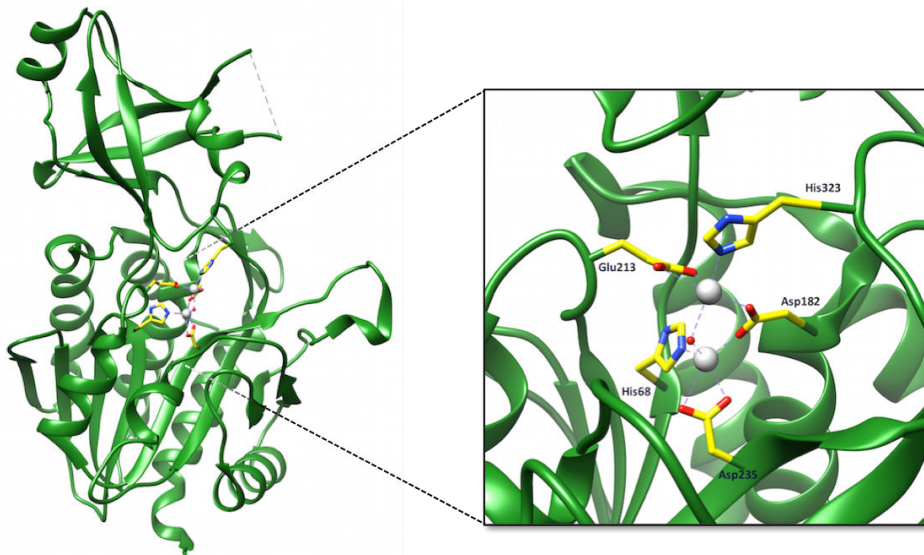
### 1.3.4.5. Caractéristiques des poches catalytiques des aminopeptidases TETs

#### 1.3.4.5.1. Sites catalytiques

Le site actif du domaine catalytique de chaque monomère de TET héberge deux ions métalliques formant un centre métallique binucléaire. Le site M1 est situé à l'entrée du site actif, alors que le site M2 est enfoui dans la structure. Ils sont adjacents à la poche de spécificité de fixation de substrat « S1 » (Lowther and Matthews, 2002). Généralement, les ions métalliques du site actif des aminopeptidases appartenant aux familles M18 et M42 sont des cations  $Zn^{2+}$  (Rawlings et al., 2012). À titre d'exemple, la spectrométrie d'émission plasma et la diffusion anormale ont révélé la présence de deux ions zinc par monomère de PfTET3 native (Story et al., 2005; Colombo et al., 2016). De même, deux ions  $Zn^{2+}$  ont été identifiés dans les sites actifs de PhTET2 et PfM18AAP en s'appuyant sur la diffusion anormale et dans le site actif de PhTET3 par fluorescence X opérée sur le cristal (Russo and Baumann, 2004; Sivaraman et al., 2012; Durá et al., 2009). Quant à PhTET1, deux ions cobalts ont été modélisés dans le site actif du fait que la solution de cristallisation en contient (Schoehn et al., 2006). Les ligands qui coordonnent ces ions métalliques sont hautement conservés, et exhibent HDEDH comme signature caractéristique. À noter que leur position dans la séquence d'acide aminé varie d'une TET à l'autre. Dans le site actif de PhTET2, la coordination des deux ions  $Zn^{2+}$  (distants de 3,8 Å) est tétraédrale complétée par une molécule d'eau.  $Zn^{2+}_1$  est coordonné par Glu213 et His323, alors que  $Zn^{2+}_2$  est coordonné par His68 et Asp235. Un résidu Asp182 assure le pontage entre les deux ions  $Zn^{2+}$  (Russo and Baumann, 2004) (**Figure 16**).

Une liaison peptidique *cis* est décrite entre Asp182 et Asp183 dans le site actif de PhTET2. Ce type de liaison peptidique ne renfermant pas de résidu proline est rare dans les protéines, mais une fois présente, cette conformation est attribuée aux molécules intimement impliquées dans la catalyse (Jabs et al., 1999; Borissenko and Groll, 2005). Le même type de liaison peptidique *cis* a été aussi décrit au niveau du site catalytique de bDNPEP bovine. Celle-ci est engendrée par Asp260 et Asn261, où Asp260 représente le ligand commun impliqué dans la coordination de deux ions  $Zn^{2+}$  de bDNPEP (Chen et al., 2012).



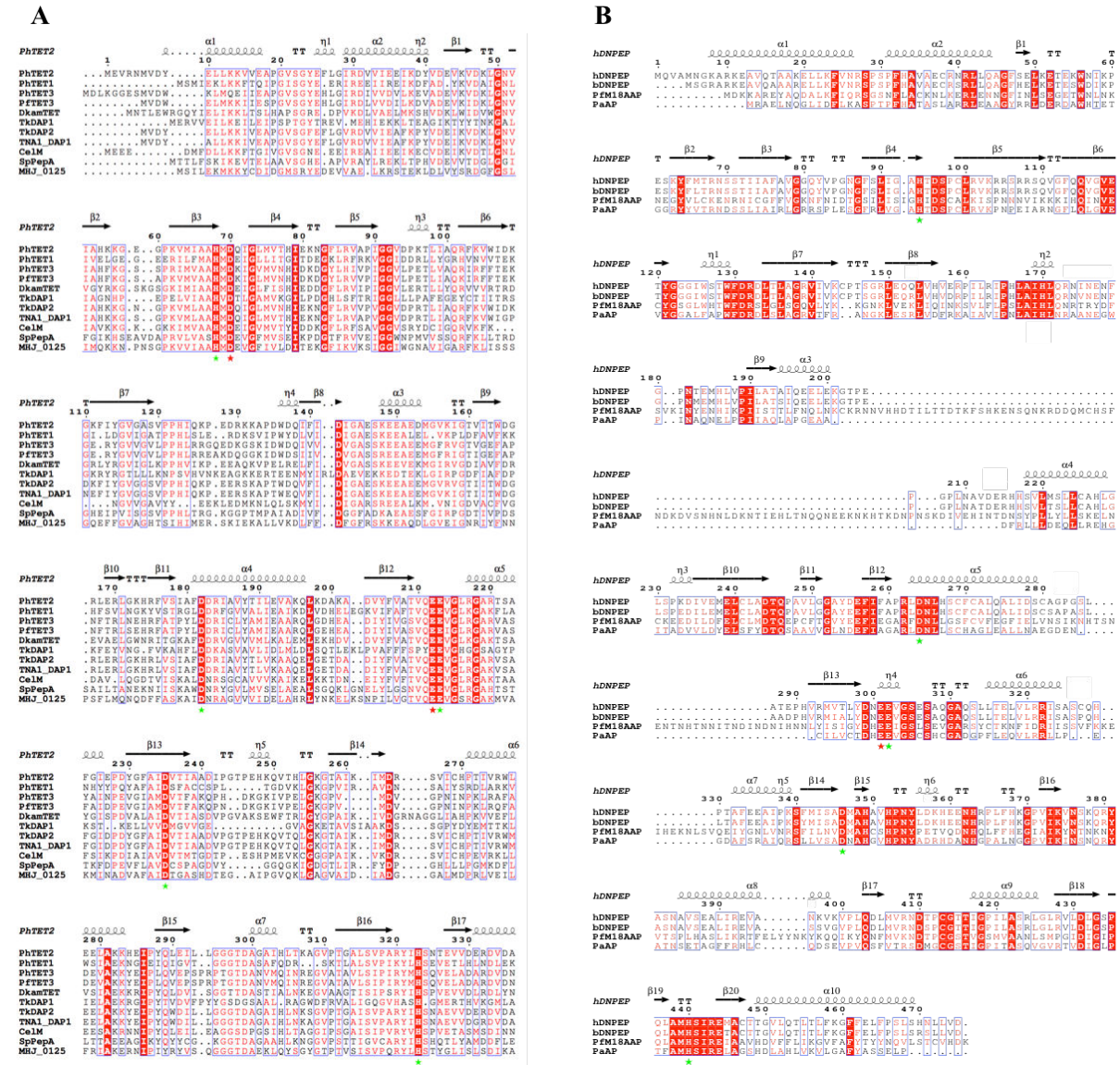


**Figure 16 : Site actif binucléaire de PhTET2 de *P. horikoshii*.** Le domaine catalytique de PhTET2 (PDB ID : 1XFO) héberge un site actif bimétallique renfermant deux ions  $Zn^{2+}$  (sphères grises).  $Zn^{2+}_1$  est coordonné par Glu213 et His323, tandis que  $Zn^{2+}_2$  est coordonné par His68 et Asp235. Simultanément Asp182 et une molécule  $H_2O$  (sphère rouge) coordonnent  $Zn^{2+}_1$  et  $Zn^{2+}_2$  (Russo and Baumann, 2004).

Les résidus Asp70 et Glu212 voisins des ligands His68 et Glu213 qui coordonnent  $Zn^{2+}_2$  et  $Zn^{2+}_1$  respectivement, sont proposés comme résidus actifs jouant le rôle d'une base générale dans la catalyse enzymatique de PhTET2 (Russo and Baumann, 2004; Borissenko and Groll, 2005; Durá et al., 2005). Ces deux résidus sont conservés au sein de toute la famille M42 des peptidases TETs archéennes et bactériennes (**Figure 17A**).

Également dans le site actif de hDNPEP appartenant à la famille M18, les deux ions  $Zn^{2+}$  (distants de 3,4 Å) sont coordonnés simultanément par Asp264.  $Zn^{2+}_1$  est encore coordonné par Glu302 et His440, alors que His94 et Asp346 sont les ligands de  $Zn^{2+}_2$ . Le résidu catalytique impliqué dans l'activation d'une molécule d'eau : Glu301 (Chaikuad et al., 2012) est conservé chez les TETs de la famille M18 eucaryotes et bactériennes (**Figure 17B**). Celui-ci est équivalent à Glu212 de PhTET2. Chez hDNPEP, le résidu équivalent au résidu catalytique Asp70 de PhTET2 est Asp96. Pourtant, son implication dans la catalyse n'est pas encore prouvée.

# INTRODUCTION



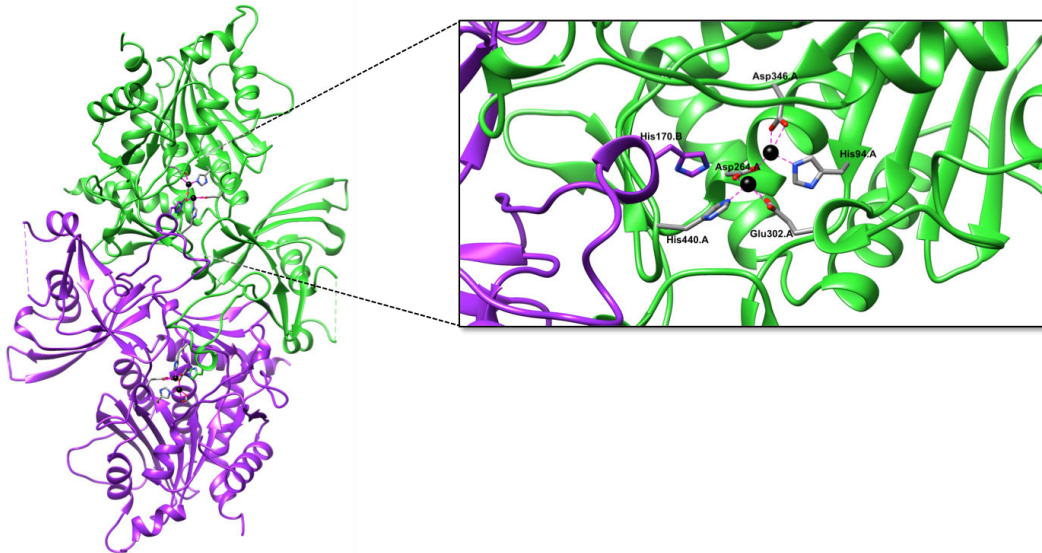
**Figure 17 :** Alignement des protéines dodécamériques de deux familles M18 et M42 et mise en évidence des ligands liant les ions métalliques (étoiles vertes) et les résidus catalytiques (étoiles rouges).

**A.** Alignement des protéines dodécamériques M42 archéennes : PhTET1, 2, 3 de *P. horikoshii*, PFTET3 de *P. furiosus*, DkamTET de *D. kamchatkensis* et bactériennes : SpPepA de *S. pneumoniae* et MHJ\_0125 de *M. hypopneumoniae*. PhTET2 est utilisée comme protéine de référence. **B.** Alignement des protéines dodécamériques M18 eucaryotes : hDNPEP humaine, bDNPEP de *Bos taurus*, PfM18AAP de *P. falciparum* et bactérienne : PaAP de *P. aeruginosa*. hDNPEP est utilisée comme protéine de référence.

D'ailleurs dans les sites actifs de hDNPEP humaine et bDNPEP bovine de la famille M18, une histidine singulière a été décrite. Appartenant au domaine de dimérisation d'une sous-unité « b », celle-ci se trouve enfouie dans le site bimétallique du domaine catalytique d'une autre sous-unité adjacente « a » (Chaikwad et al., 2012; Chen et al., 2012) (**Figure 18**). Ces résidus sont : His170 et His166 chez hDNPEP et bDNPEP, respectivement. Selon Chen, His166 de bDNPEP forme une liaison hydrogène avec une molécule d'eau catalytique. Quant

## INTRODUCTION

à His170 de hDNPEP, elle intervient dans l'interaction avec le substrat (oxygène du groupement amide de la liaison peptidique), notamment dans la translocation des résidus distaux vers le site actif (Chaikuad et al., 2012). La mutation de cette histidine par une phénylalanine abolit l'activité enzymatique de hDNPEP (Wilk et al., 2002).



**Figure 18 :** Site actif de hDNPEP (PDB ID : 4DYO) accueillant un résidu histidine d'une sous-unité adjacente, du même *building block*. Le résidu His170 du monomère B (en violet) de hDNPEP est ramené par la boucle  $\beta$ 8- $\beta$ 9 vers le site actif d'un autre monomère A voisin (en vert) (Chaikuad et al., 2012).

Cette histidine semble être conservée à travers les aminopeptidases M42 et M18. En effet, l'alignement de séquences de hDNPEP et bDNPEP avec les TETs de *P. aeruginosa*, *M. hyopneumoniae* et *P. horikoshii* montre une conservation de cette histidine au travers les différents homologues TETs (**Figure 19**). En revanche, les boucles où réside cette histidine sont complètement ou partiellement désordonnées (comme c'est le cas par exemple des TETs de *P. horikoshii*). Ce désordre pourrait être attribué à l'absence de substrat, suggérant alors qu'une modification conformationnelle puisse être induite par la fixation de substrat, et soit requise pour compléter le site catalytique (Chaikuad et al., 2012). En outre, l'enfouissement de la boucle maintenant His170 de hDNPEP est une caractéristique structurale commune aux aminopeptidases M18 et M42 pour générer le dimère (Chaikuad et al., 2012).

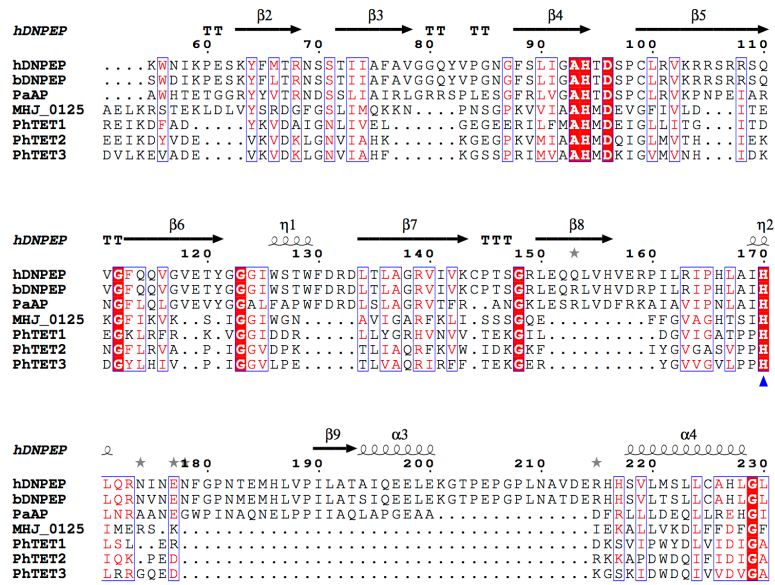


Figure 19 : Alignement des séquences des aminopeptidases TETs M18 et M42 en utilisant la séquence de la TET humaine hDNPEP comme référence. Histidine 170 (marquée par un triangle bleu) décrite chez hDNPEP pour son implication dans l'interaction avec le substrat est conservée à travers les familles M18 et M42.

### 1.3.4.5.2. Mécanisme catalytique

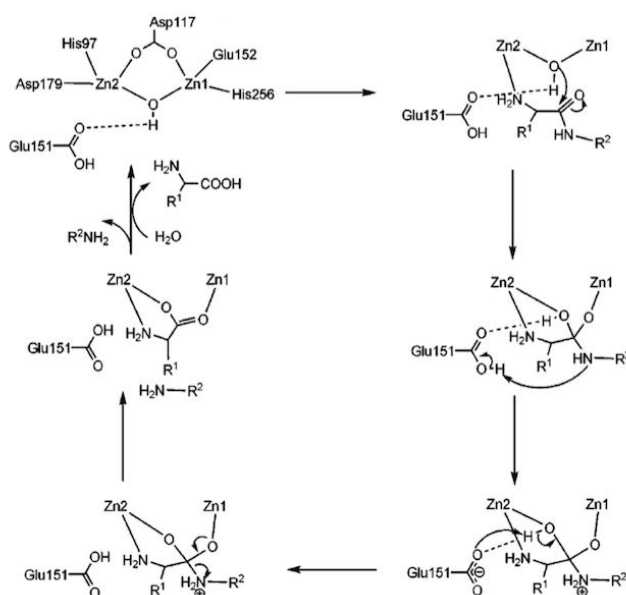
Le domaine catalytique des aminopeptidases TETs est très similaire à celui de l'aminopeptidase AAP de *A. proteolytica*, un prototype des enzymes co-catalytiques à zinc du clan MH (Chevrier et al., 1994b). La conservation des ligands qui coordonnent les cations métalliques (H.D.E.D.H) suggère que les TETs adoptent un mécanisme enzymatique hautement similaire, voire identique à celui de AAP (Schoehn et al., 2006; Durá et al., 2009).

Au niveau du site actif de AAP,  $Zn^{2+}_1$  est coordonné par Glu152, His256, alors que  $Zn^{2+}_2$  est coordonné par His97, Asp179. Asp117 et une molécule d'eau coordonnent simultanément les deux ions  $Zn^{2+}$ . Le processus d'hydrolyse exécuté par AAP et proposé par Chen et ses collaborateurs, est représenté par la Figure 20 (Chen et al., 2008). Glu151 joue le rôle d'une base activant la molécule d'eau en un ion hydroxyle en lui arrachant un proton.  $Zn^{2+}_2$  lie le groupement amine du substrat et oriente la liaison peptidique vers le groupement hydroxyle nucléophile.

Une fois le substrat lié, l'ion hydroxyle  $OH^-$  liant les deux ions métalliques devient terminal par rapport à  $Zn^{2+}_1$  et s'en dissocie. Ainsi,  $OH^-$  réalise l'attaque nucléophile du carbone du groupement carbonyle de la liaison peptidique. La double liaison  $C=O$  devient simple et l'oxygène développe une charge négative permettant son interaction avec  $Zn^{2+}_1$ .

## INTRODUCTION

Ensuite, pour qu'un proton puisse être transféré de Glu151 vers l'azote de la liaison amide, celui-ci (Glu151) s'éloigne de Asp179 et la liaison hydrogène entre eux sera rompue. L'azote devient alors protoné et la liaison amide devient plus longue. Glu151, en cédant son proton à l'azote, devient anionique et met en place une liaison hydrogène avec l'ion  $\text{OH}^-$ . La liaison peptidique est prête à être clivée mais nécessite un deuxième transfert de proton. Celui-ci est transféré à partir de  $\text{OH}^-$  à Glu151. Ainsi, la liaison entre le carbone et l'oxygène originaire du  $\text{OH}^-$  devient plus courte et la liaison entre l'azote et le carbone de la liaison amide devient de plus en plus longue. La liaison peptidique est alors clivée, libérant un groupement amine neutre et un groupement carboxyle déprotoné stabilisé par  $\text{Zn}^{2+}_1$  et  $\text{Zn}^{2+}_2$ . Ce dernier sera libéré du site catalytique grâce à une molécule d'eau permettant la régénération du site actif (Chen et al., 2008).



**Figure 20 : Mécanisme catalytique proposé pour les aminopeptidases bimétalliques (AAP de *A. proteolytica*).** Les résidus qui figurent dans le site actif sont ceux de AAP. Glu151 est décrit comme base activatrice de la molécule d'eau pour l'attaque nucléophile (Chen et al., 2008).

### 1.3.4.6. Reconnaissance et manipulation des peptides par une particule TET

La particule TET est ouverte *via* deux types de pores : quatre pores situés au niveau des faces et quatre autres localisés au niveau des sommets. Les pores des facettes sont plus larges ( $\sim 20$  Å de diamètre) que ceux des apex ( $\sim 9$  Å de diamètre) et permettent aux peptides dépliés d'accéder à la cavité interne. Grâce à la différence du potentiel électrostatique, ces peptides

## INTRODUCTION

---

naviguent à l'intérieur de l'édifice vers l'un des trois sites actifs de l'une des poches catalytiques présentes au niveau des sommets (Schoehn et al., 2006). Une fois que les acides aminés des extrémités Nt des substrats sont clivés, ils sont expulsés à l'extérieur *via* les pores apicaux. En revanche, un autre type de pore est révélé par Borissenko et Groll qui proposent que les pores d'évacuation des acides aminés puissent être situés aux alentours des pores d'entrée. En effet, une mutation des résidus qui bordent ces pores et leur obstruction conduisent à 70 % d'inhibition de l'activité enzymatique de PhTET2 (Borissenko and Groll, 2005). Le mécanisme authentique reflétant l'expulsion des acides aminés hydrolysés hors de l'édifice catalytique reste donc méconnu.

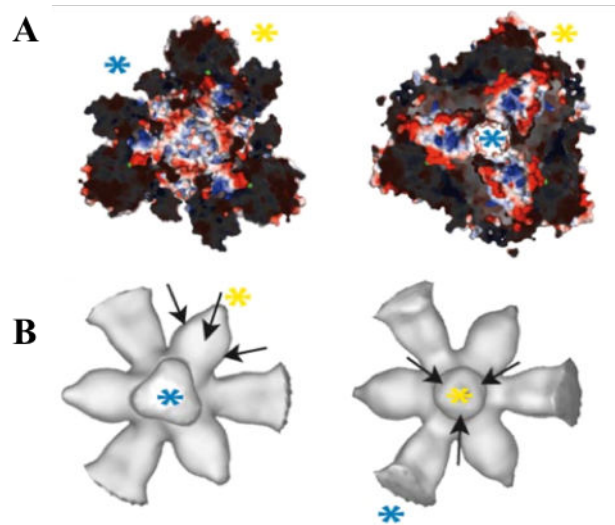
L'organisation de la cavité interne de PhTET1 révélée par microscopie électronique ainsi que l'analyse des propriétés de surface des particules archéennes : PhTET2 et PhTET3, bactériennes (MHJ\_0125) et eucaryotes (hDNPEP), ont permis de proposer un modèle expliquant le trafic des peptides au sein de la particule TET (Schoehn et al., 2006; Durá et al., 2009; Robinson et al., 2013; Chaikuad et al., 2012).

Le mécanisme le plus détaillé est décrit par Schoehn et ses collègues en 2006 pour PhTET1 (aspartyl / glutamyl aminopeptidase). Les pores d'entrée donnent accès à quatre larges canaux qui se croisent au centre de la particule. Ces canaux s'étendent sous forme d'entonnoir vers les sommets de la particule, où sont localisées les chambres catalytiques. Ces dernières débouchent par la suite vers l'extérieur *via* les trous de sortie (**Figure 21**). La variation du potentiel électrostatique le long des canaux assure une translocation en un mouvement continu des substrats vers les poches catalytiques. À l'entrée du canal, le potentiel électrostatique est négatif (Asp74, Glu75). Au fur et à mesure de la pénétration au sein de la particule, il devient positif puis négatif avant d'atteindre l'intersection des canaux ayant la forme d'une sphère bordée par les chaînes latérales des douze résidus Lys122. La Lysine 122 appartient à une boucle flexible intervenant dans l'orientation du substrat vers le site actif. À partir de cette intersection, la chambre catalytique (bordée par Arg120, His115, Asn252, Lys212) hautement basique est accessible. La répulsion de charges entre l'extrémité Nt des substrats et les résidus qui bordent la chambre catalytique dirige le peptide vers l'un des trois sites bimétalliques hautement acides, et oriente la chaîne latérale de l'acide aminé en position P1 du substrat vers la poche « S1 » de fixation de substrat. Dans la direction du pore de sortie, une surface de charge négative (Asp291) attire l'extrémité Nt de l'acide aminé clivé et précède une couche formée par Tyr219. Cette tyrosine permet d'éjecter le résidu hydrolysé

## INTRODUCTION

---

vers l'extérieur du dodécamère grâce à sa chaîne latérale (Schoehn et al., 2006).



**Figure 21 : Potentiel électrostatique de l'intérieur de la particule PhTET1 (PDB ID : 2WYR) de *P. horikoshii*.** **A.** À gauche, vue de l'espace interne de PhTET1 du côté du pore d'entrée, le pore de sortie apparaît au fond. À droite, vue opposée du côté du pore de sortie, le pore d'entrée apparaît en face. Les sphères vertes correspondent aux ions  $\text{Co}^{2+}$  des sites actifs. **B.** Représentation schématique des canaux internes traversant PhTET1 depuis les pores d'entrée (étoile bleue) jusqu'aux pores de sorties (étoile jaune). À gauche, trou d'entrée est situé en face. À droite, observation du côté opposé, depuis le trou de sortie. Les trois flèches indiquent les trois sites actifs voisins des pores de sortie (Schoehn et al., 2006).

Également, les peptides à l'intérieur de la cavité de PhTET2 (leucyl aminopeptidase) et PhTET3 (lysyl aminopeptidase) sont guidés par la variation du potentiel électrostatique. La charge négative du trou d'entrée (Asp135 et Asp137 de PhTET2 ; Asp83, Asp138 et Asp140 de PhTET3) attire les extrémités Nt positives des substrats vers la cavité centrale de la particule. À l'intérieur, l'extrémité Nt fait face à un ensemble de charges positives (His78, Lys81, Lys148, Arg220, His248 et His253, en référence à PhTET2) permettant son adressage au site actif où elle sera clivée. Les pores de sortie de PhTET2 et PhTET3 chargés positivement sont bloqués par les résidus phénylalanine (Phe224) et tyrosine (Tyr227), respectivement, de chacun des trois monomères. Une modification conformationnelle de ces résidus est donc requise pour ouvrir le pore de sortie et expulser les acides aminés clivés (Russo and Baumann, 2004; Durá et al., 2009). De même les pores de sortie de MHJ\_0125 (glutamyl aminopeptidase) sont positivement chargés par la présence de résidus lysines facilitant l'expulsion des résidus acides (Robinson et al., 2013).

Le pore d'entrée de hDNPEP (aspartyl aminopeptidase) est bordé par des résidus positivement chargés pouvant attirer les résidus acides en position Nt des substrats

## INTRODUCTION

---

peptidiques. Alors que le potentiel électrostatique au niveau des trous de sortie est mixte et varie entre positif et négatif permettant une translocation continue des acides aminés hydrolysés vers l'extérieur de la macromolécule. Le trou de sortie de hDNPEP exhibe un diamètre très étroit de 3 Å, nécessitant une modification conformationnelle des résidus qui l'édifient pour pouvoir expulser les produits de l'hydrolyse enzymatique vers l'extérieur (Chaikuad et al., 2012).

Chez PfM18AAP (aspartyl aminopeptidase), deux boucles flexibles s'étendent à la périphérie du pore d'entrée et jouent le rôle d'une barrière contrôlant l'accès à la cavité de cette aminopeptidase et régulant alors sa fonction enzymatique. De surcroît, ces boucles flexibles peuvent jouer le rôle d'une pince piégeant les substrats pour cette TET parasitaire (Sivaraman et al., 2012).

### 1.3.5. Spécificités de différentes aminopeptidases TETs

L'étude des propriétés enzymatiques de différentes TETs caractérisées à ce jour révèle que, en dépit d'une grande conservation de structure et de mode d'action, les particules TETs peuvent présenter de grandes disparités vis-à-vis de leurs spécificités de substrat et de leurs mécanismes d'activation *via* les métaux. Dans ce paragraphe, je présente les différences des propriétés structurales qui peuvent être à l'origine des comportements fonctionnels divers.

#### 1.3.5.1. Spécificités enzymatiques et poches de liaison de substrat « S1 »

En vue de caractériser l'activité amidolytique des TETs, une variété de substrats chromogéniques, fluorogéniques, et de peptides de tailles et de compositions variées ont été utilisés. Tenant compte de leur spécificité de substrat, les TETs peuvent être réparties en trois groupes distincts. Certaines sont spécifiques des résidus acides, d'autres sont à large spectre d'activité exprimant une spécificité pour les résidus neutres. Quelques unes sont décrites comme des TETs spécifiques des résidus basiques (**Tableau 2, paragraphe 1.3.3.**). Ces résidus sont situés en position Nt au niveau du site P1 de la liaison peptidique cible de l'hydrolyse. La spécificité enzymatique est gouvernée par la nature des acides aminés localisés dans la poche de liaison spécifique de substrat « S1 » adjacente au site actif de l'enzyme.

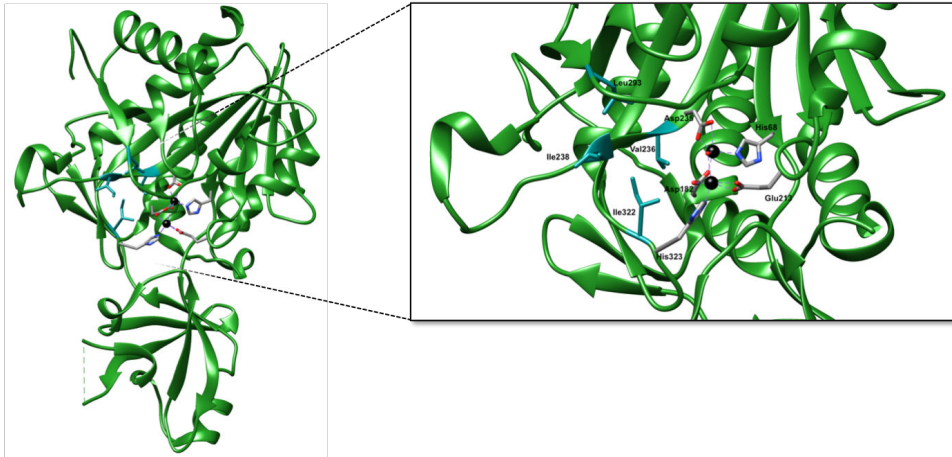


## INTRODUCTION

---

PhTET1 de *P. horikoshii*, SpPepA de *S. pneumoniae*, MHJ\_0125 de *M. hyopneumoniae*, PaAP de *P. aeruginosa* ainsi que les trois TETs décrites chez les eucaryotes : PfM18AAP de *P. falciparum*, hDNPEP humaine et bDNPEP bovine sont des métallo-aminopeptidases spécifiques des résidus acides. Grâce aux chaînes latérales positivement chargées de l'arginine, la lysine et l'histidine, présentes dans la poche S1, ces résidus interagissent avec les chaînes latérales des acides glutamiques et aspartiques des substrats. Selon leurs préférences pour l'acide glutamique ou l'acide aspartique, les TETs acides sont classées comme glutamyl aminopeptidases : PhTET1, SpPepA et MHJ\_0125 (Durá and Franzetti, 2013; Kim et al., 2010; Robinson et al., 2013), ou aspartyl aminopeptidases : PaAP, PfM18AAP, hDNPEP et bDNPEP (Nguyen et al., 2014; Teuscher et al., 2007; Chaikuad et al., 2012; Chen et al., 2012). Chaikuad et ses collègues ont proposé que la conservation de Lys374 de hDNPEP puisse être utilisée comme marqueur de prédiction d'une activité aspartyl aminopeptidase. Cette lysine est conservée chez hDNPEP, PaAP, PfM18AAP et bDNPEP. Cependant, cette lysine est substituée par une arginine chez PhTET1 et SpPepA et par un acide aspartique chez MHJ\_0125, qui sont décrites comme des glutamyl aminopeptidases (Kim et al., 2010; Robinson et al., 2013).

Les TETs à large spectre d'activité sont caractérisées par une poche S1 plus large que celle des TETs à spécificité restreinte, comme PhTET1. Elle est bordée par des résidus hydrophobes tels que Val236, Ile238, Leu293 et Ile322 chez PhTET2 (**Figure 22**). PhTET2, HmTET, CelM et DkamTET hydrolysent les résidus hydrophobes non chargés, notamment la leucine et moins efficacement la méthionine, l'isoleucine, l'alanine, la sérine et la valine (Durá et al., 2005; Franzetti et al., 2002; Dutoit et al., 2012; Slutskaya et al., 2012).



**Figure 22 : Représentation des résidus de la poche spécifique de la fixation de substrat « S1 » de PhTET2 (PDB ID : 1XFO) de *P. horikoshii*.** Val236, Ile238, Leu293 et Ile322 (en cyan foncé) représentés en *sticks*, sont localisés au voisinage du site bimétallique de PhTET2. Ils permettent l'accommodation des chaînes latérales des substrats hydrophobes clivables par cette leucyl aminopeptidase.

Jusqu'à présent, uniquement deux TETs spécifiques des résidus basiques ont été identifiées chez deux espèces d'archée : PhTET3 de *P. horikoshii* et PftET3 de *P. furiosus* (Durá et al., 2009; Story et al., 2005). Elles clivent préférentiellement la lysine Nt et moins efficacement l'arginine. Les acides aminés non polaires comme la leucine, la phénylalanine et l'alanine sont hydrolysables dans une moindre mesure (Durá et al., 2009). Cette fonction est en accord avec leur poche S1 négativement chargée permettant d'accueillir la chaîne latérale positive des lysines et des arginines. Cette poche est bordée par les mêmes résidus chez PhTET3 et PftET3 : Asp262, Thr240, Thr295 and Thr297 (en référence à PhTET3).

Récemment, la présence d'un hétéro-dodécamère PhTET2-PhTET3 a été mise en évidence *in vivo* (Appolaire et al., 2014a). Celui-ci combine les activités enzymatiques de PhTET2 et PhTET3 et présente une capacité à hydrolyser un peptide mixte (Val-Leu-Lys-pNA) plus efficacement qu'un mélange stœchiométrique d'homo-dodécamères de PhTET2 et PhTET3. Comme précédemment évoquée (voir **paragraphe 1.3.4.4.**), cette capacité est expliquée par la présence de deux types de poches S1 différentes, complémentaires, au niveau de chaque sommet de cet édifice hétéro-oligomérique (Appolaire et al., 2014a). Un autre hétéro-dodécamère de l'aminopeptidase API de *Bacillus (Geobacillus) stearothermophilus*, une bactérie à Gram positif, a été aussi publié dans les années 70. Cette particule est élaborée par une combinaison à plusieurs ratios de deux sous-unités  $\alpha$  et  $\beta$ . La sous-unité  $\alpha$  est spécifique des résidus neutres, et la sous-unité  $\beta$  clive préférentiellement les résidus acides.

## INTRODUCTION

---

Ainsi, l'hétéro-oligomère généré exprime une spécificité enzymatique plus large que les homo-oligomères (Roncari and Zuber, 1970; Stoll et al., 1972, 1973b).

Jusqu'à présent uniquement deux TETs de l'archée hyperthermophile *T. kodakarensis* : TkDAP1 et TkDAP2 sont caractérisées comme des *deblocking* aminopeptidases capables de cliver les extrémités Nt libres et bloquées, polaires et apolaires des substrats peptidiques (Jia et al., 2011). Toutefois, en absence de structure il est difficile de définir les caractéristiques à l'origine de cette spécificité.

En conclusion, il apparaît que la spécificité de substrat est essentiellement manœuvrée par les caractéristiques de dimensions et de charges des poches « S1 ».

### 1.3.5.2. Différences de propriétés de liaison aux métaux

Dans plusieurs études, les conditions optimales d'activation des TETs vis-à-vis des cofacteurs métalliques activateurs ont été déterminées après avoir identifié le substrat préférentiel de chacune (voir **Tableau 2, paragraphe 1.3.3.**).

La grande majorité des TETs archéennes (HmTET, PhTET1, PhTET2, PhTET3, PFTET3) sont activées par l'ajout des cations cobalt dans le milieu réactionnel (Franzetti et al., 2002; Schoehn et al., 2006; Durá et al., 2005, 2009; Colombo et al., 2016). La concentration optimale du cobalt pour l'activation de PFTET3 est 300  $\mu$ M, au-delà de cette valeur un excès de cobalt devient inhibiteur de l'activité amidolytique de cette TET (Colombo et al., 2016). En outre, le cobalt a été identifié comme cofacteur physiologique de PFTET3. En effet, des études du métallome intracellulaire ont montré une concentration élevée en cobalt dans *P. furiosus* (Cameron et al., 2012). Également, les TETs bactériennes : CelM, MHJ\_0125 et eucaryotes : PFM18AAP sont activées en présence du cobalt (Dutoit et al., 2012; Robinson et al., 2013; Sivaraman et al., 2012). SpPepA de *S. pneumoniae* et hDNPEP humaine sont aussi activées par le cobalt mais le manganèse induit aussi une activation significative bien que légèrement inférieure à celle du cobalt (Kim et al., 2010; Chaikuad et al., 2012).

Seules trois TETs présentent une exception notable à l'activation par le cobalt : DkamTET de *T. kodakarensis*, PaAP de *P. aeruginosa* et bDNPEP de *Bos taurus*. Pour ces TETs, l'activité est stimulée en présence des ions manganèse (Slutskaya et al., 2012; Nguyen et al., 2014; Chen et al., 2012).

### 1.3.5.2.1. Rôles des métaux dans les aminopeptidases TETs

Les sites M1 et M2 présentent des affinités différentes pour lier les ions métalliques : le site M1 est le *readily exchanging site*, alors que M2 est le *tightly binding site* (Allen et al., 1983; Colombo et al., 2016). Il a été proposé que la nucléophilicité ou la polarisabilité d'une molécule d'eau ou le pKa d'une base impliquée dans la déprotonation de la molécule d'eau peut être affectée par la nature des ions métalliques coordonnés dans le site actif. Ainsi, toute substitution des ions métalliques dans le site actif est à l'origine d'une modification de l'activité catalytique de la métallo-enzyme (Gomis-Rüth et al., 1994; Lowther and Matthews, 2002). De même, le changement de la géométrie de la coordination de ces cations peut influencer l'activité amidolytique (Chaikuad et al., 2012).

Des études ont été réalisées afin de comprendre le rôle joué par M1 et M2 dans le mécanisme catalytique des métallopeptidases bimétalliques.

Dans l'étude de Nguyen et ses collègues, trois résidus parmi ceux impliqués dans la coordination des ions métalliques de PaAP, une aspartyl aminopeptidase de *P. aeruginosa* ont été mutés en alanine. Ces ligands sont : Asp236 qui coordonne M1 et M2, His82 qui lie M2 et His401 ligand de M1. L'étude de l'activité catalytique du mutant D236A ne montre aucune fonction enzymatique en présence de divers ions métalliques. Ce qui est en accord avec la structure cristallographique de D236A qui montre une structure cristallographique dépourvue d'ions métalliques. Par contre, H82A dont le site actif héberge un ion métallique au niveau du site M1 exhibe une activité catalytique considérable en présence du cobalt, contrairement à H401A qui est moins actif que H82A. Nguyen et ses collègues ont donc conclu que M1 est le site impliqué dans le mécanisme catalytique de PaAP (Nguyen et al., 2014).

Les résultats obtenus par Nguyen et ses collègues sont à nuancer. En effet, le rôle catalytique de M1 semble être inversé chez PfTET3 de *P. furiosus*. Effectivement, le site métallique M2 semble être le site responsable de la catalyse enzymatique. Celui-ci est coordonné par un ion  $\text{Cl}^-$  qui peut, comme  $\text{F}^-$ , jouer le rôle d'un inhibiteur non compétitif et déplacer l'ion hydroxyle de la molécule d'eau active (Ercan et al., 2010; Colombo et al., 2016). En outre, en utilisant la diffusion anormale, Colombo et ses collègues ont montré que seuls les ions  $\text{Zn}^{2+}$  des sites M1 sont substitués par le  $\text{Co}^{2+}$ , engendrant une activation de

## INTRODUCTION

---

PfTET3. L'analyse des paramètres cinétiques  $K_m$  et  $k_{cat}$  en présence et en absence de cobalt révèle une diminution de  $K_m$  et une invariabilité de  $k_{cat}$ . Grâce à ces observations, les auteurs ont dévoilé le rôle catalytique du métal du site M2 et ont montré que M1 contrôle la flexibilité du site actif de PfTET3 à haute température. De plus, Colombo et ses collègues ont décelé la présence d'un troisième site métallique M3. Celui-ci est localisé dans la poche spécifique de fixation de substrat « S1 » et est coordonné par Glu281, Asp254 et Thr232. Il a été proposé que M3 joue un rôle dans la régulation de la préférence enzymatique de PfTET3 qui présente une spécificité pour les résidus basiques, en lui permettant d'hydrolyser des résidus acides, préférentiellement l'acide glutamique (Colombo et al., 2016).

L'incubation de bDNPEP bovine avec des ions  $Mg^{2+}$  ou  $Mn^{2+}$  entraîne une substitution des ions  $Zn^{2+}$  du site M2 par ces cations métalliques (Chen et al., 2012), ce qui montre une inversion des affinités métalliques entre M1 et M2 en comparaison avec PfTET3. Comme mentionné précédemment, l'ion  $Mn^{2+}$  est le cofacteur activateur de bDNPEP, induisant une augmentation de  $k_{cat}$  et une diminution de  $K_m$ . Alors, similairement à PfTET3, le site M2 peut être responsable de l'activité catalytique de cette métallo-aminopeptidase eucaryote.

Comme illustrée par les exemples ci-dessus, la disparité de la spécificité vis-à-vis des cofacteurs métalliques activateurs des TETs et le mécanisme d'activation lié à la liaison au métal restent mal compris. De même, on ne comprend pas les bases structurales qui déterminent le choix du métal activateur.

Non seulement les ions métalliques interviennent dans la catalyse enzymatique des particules TETs, mais ils sont aussi impliqués dans la régulation de leur état oligomérique. En effet, les études entreprises par Rosenbaum, Colombo et leurs collègues montrent que la chélation des cations du site métallique actif par de l'EDTA induit une dissociation du dodécamère PhTET3 en dimère et PfTET3 en monomère (Rosenbaum et al., 2011; Colombo et al., 2016). Une explication a été proposée chez PhTET3 où un des ligands coordonnant le métal M1, Glu216 est maintenu par une hélice où sont localisés deux résidus (Arg220 et Arg223) impliqués dans un réseau de liaisons hydrogènes au niveau de l'interface dimère-dimère. Ainsi, la chélation du métal du site M1 peut induire un changement conformationnel au niveau de cette boucle altérant ainsi la position de Arg220 et Arg223 et engendrant une déstabilisation des interfaces d'oligomérisation et une dissociation de la macromolécule (Rosenbaum et al., 2011).

## INTRODUCTION

---

En revanche, le cobalt peut avoir un effet opposé à celui décrit pour PhTET3 et PfTET3. Effectivement, l'addition du  $\text{Co}^{2+}$  à TNA1-DAP1 (partageant 82 % d'identité avec PhTET2), une enzyme dodécamérique de *T. onnurineus* NA1, est accompagnée par une activation de l'enzyme mais une dissociation du dodécamère en dimère (Lee et al., 2007). La cause de cette dissociation reste mal connue en raison du manque de données structurales disponibles pour cette protéine.

En plus du rôle joué par les ions métalliques dans la régulation de l'état oligomérique des particules TETs, ceux-ci, notamment le cobalt, interviennent dans la régulation de la demi-vie et la thermostabilité de l'aminopeptidase à haute température (Story et al., 2005; Colombo et al., 2016).

### 1.3.6. Discussion du rôle physiologique des aminopeptidases TETs

Aujourd'hui, le rôle physiologique des métallo-aminopeptidases TETs reste méconnu. On considère généralement que le rôle des aminopeptidases est d'intervenir après l'action endoprotéolytique des protéases ATP-dépendantes pour terminer le processus de dégradation des protéines cellulaires en acides aminés libres. Ainsi, les aminopeptidases TETs pourraient être impliquées dans le processus de protéolyse et le métabolisme énergétique des espèces qui les expriment, notamment en conditions de carence nutritionnelle. Il a alors été proposé que TET est un partenaire du protéasome 20S sur la base d'expériences démontrant que PhTET2 exerce une activité sur des peptides issus de l'activité du protéasome (Borissenko and Groll, 2005). L'activité aminopeptidase à large spectre de PhTET2 suffit à expliquer ce phénomène qui pourtant ne démontre en aucune manière l'existence d'une voie protéolytique associant spécifiquement TET au protéasome.

Il est également envisageable que TET joue un rôle important dans le métabolisme nutritionnel des archées et des bactéries. En effet, les peptides présents dans l'environnement représentent une source importante de nutriments. Le fait que les archées du genre Thermococcales possèdent plusieurs versions de la particule TET pourrait aller dans ce sens. L'archée hyperthermophile abyssale *P. horikoshii* exprime quatre aminopeptidases TETs différentes : PhTET1, PhTET2, PhTET3 et PhTET4. Les trois premières TETs ont été caractérisées antérieurement et la quatrième sera décrite dans mon manuscrit de thèse. PhTET1, 2 et 3 possèdent des spécificités enzymatiques différentes non redondantes, suggérant une meilleure efficacité dans la dégradation d'une diversité de peptides. Cette

## INTRODUCTION

---

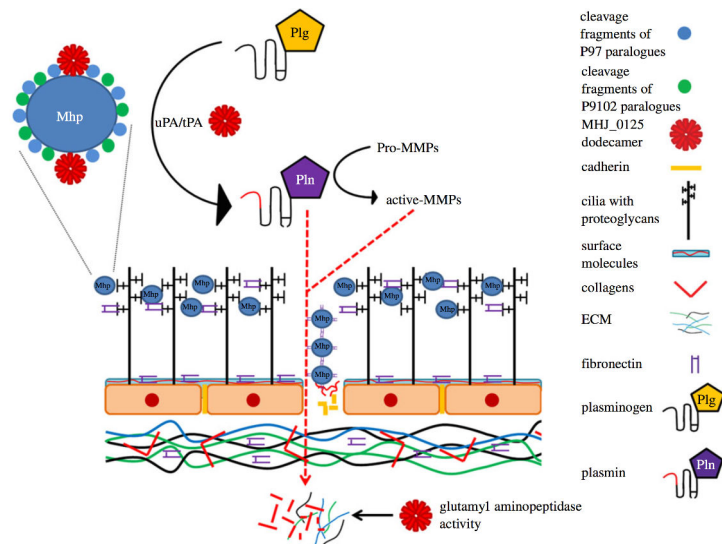
capacité à hydrolyser une variété d'acides aminés est en accord avec l'hétérotrophie de *P. horikoshii* qui puise son besoin en carbone à partir des peptides présents dans son écosystème (Karl, 2002). En outre, PhTET2 et PhTET3 peuvent s'assembler sous forme d'hétéro-dodécamère PhTET2-PhTET3, dénommé peptidasome en raison de sa capacité accrue à dégrader des chaînes polypeptidiques complexes et à attaquer des peptides que même un mélange de PhTET2 et PhTET3 ne peut dégrader (Appolaire et al., 2014a). Il est ainsi possible que cet hétéro-oligomère soit spécifiquement dédié à la destruction non spécifique des peptides et que son activité soit régulée pour répondre aux besoins du métabolisme. Un argument qui va dans ce sens est que l'assemblage des TETs sous leur forme dodécamérique fonctionnelle apparaît être modifié en fonction de l'état physiologique des cellules par un mécanisme qui demeure inconnu. Ceci a été montré dans notre laboratoire chez *H. salinarum* dans des conditions de carences (Marty, 2006) (résultats non publiés). Cette hypothèse est également soutenue par une évaluation du taux d'expression d'un gène codant pour une protéine homologue de PhTET1 chez *P. furiosus*. Ce gène est surexprimé dès lors que les cellules des *Pyrococcus* sont cultivées dans un milieu riche en peptides (Schut et al., 2003). De plus, la surexpression de TkDAP2, *deblocking* aminopeptidase de *Thermococcus kodakarensis*, mise en évidence une fois les cellules soumises à un stress oxydatif ou thermique, montre l'implication de TkDAP2 dans la réponse au stress (Jia et al., 2011).

On ne peut toutefois pas exclure que les particules TETs possèdent également des rôles physiologiques spécifiques en dégradant ou en modifiant l'extrémité Nt de certains types de peptides. En effet, la capacité de cliver des résidus spéciaux en Nt suggère que ces aminopeptidases peuvent intervenir dans l'activation des peptides biactifs (Chen et al., 2012) et le contrôle de leur demi-vie (Bachmair et al., 1986). D'ailleurs, une étude de l'oligomérisation *in vivo* de PhTET1, 2 et 3 montre qu'au moment où PhTET2 et PhTET3 sont sous forme de 12-mères, PhTET1 existe sous forme dimérique. Ceci permet de proposer que PhTET1 possède une fonction différente de ses homologues PhTET2 et 3 et que son activation puisse être induite dans des conditions différentes de celles de PhTET2 et PhTET3 (Appolaire et al., 2014a).

Récemment, certains rôles plus spécifiques ont été accordés aux TETs bactériennes et eucaryotes. Chez les bactéries, MHJ\_0125, une TET de *M. hyopneumoniae*, est décrite comme étant localisée dans le cytosol, liée à la membrane externe et excrétée dans le milieu extracellulaire (Robinson et al., 2013). Cette bactérie est un pathogène obligatoire qui

## INTRODUCTION

nécessite un apport d'acides aminés de l'hôte pour pouvoir synthétiser ses protéines essentielles. MHJ\_0125 peut jouer un rôle important lors de l'infection et la colonisation des voies respiratoires porcines vu sa capacité à lier le plasminogène et les glycosaminoglycane des tissus épithéliaux. Il a été proposé que ces interactions soient agencées par les extrémités Ct et Nt supposées être exposées à la surface du dodécamère. Une fois MHJ\_0125 liée au plasminogène, celui-ci en présence de tPA (*Tissue Plasminogen Activator*) est activé en plasmine, une sérine protéase, qui active les métallo-protéases de la matrice extracellulaire. À leur tour, ces protéases interviennent dans la dégradation de cette matrice et des cadhérines des jonctions intercellulaires permettant ainsi à *M. hyopneumoniae* de gagner les tissus subépithéliaux. Également, elles génèrent des substrats peptidiques pour MHJ\_0125 (Robinson et al., 2013) (**Figure 23**).



**Figure 23 : Implication de MHJ\_0125 de *M. hyopneumoniae* dans l'infection et la colonisation de l'épithélium respiratoire porcine.** MHJ\_0125 (représentée en rouge) lie le plasminogène (Plg) et induit son activation en plasmine (Pln), en présence de tPA. La plasmine active les métallo-protéases de la matrice extracellulaire (MMPs) et toutes deux dégradent la matrice extracellulaire, les protéines de l'hôte ainsi que la cadhérine (en jaune) permettant probablement à la bactérie de gagner les tissus subépithéliaux (Robinson et al., 2013).

Chez les eucaryotes, le parasite intra-érythrocytaire de la malaria, *P. falciparum*, digère et exploite l'hémoglobine des érythrocytes pour synthétiser ses propres protéines. En effet, 16 % de l'hémoglobine digérée sont utilisés pour la production des protéines parasitaires (Krugliak et al., 2002). De nombreuses endopeptidases, dipeptidases, aminopeptidases et neuf métallo-aminopeptidases sont engagées pour transformer l'hémoglobine en acides aminés libres. Parmi ces neuf aminopeptidases, PfM18AAP est une aspartyl aminopeptidase spécifique des acides aspartiques et glutamiques constituant 8 % des acides aminés composant



## INTRODUCTION

---

l'hémoglobine (Stack et al., 2007). Le *knockdown* de PfM18AAP est accompagné par des altérations cellulaires du parasite comme la rupture de la vacuole digestive dans les stades schizontes et une accumulation de l'hémoglobine dans les trophozoïtes (Teuscher et al., 2007). En outre, l'altération du gène codant pour cette métallo-aminopeptidase révèle son implication dans la réplication du parasite au stage sanguin (mérozoïte) (Dalal and Klemba, 2007). La localisation de PfM18AAP dans la vacuole parasitophore et sa liaison avec des protéines membranaires de l'érythrocyte suggèrent que l'enzyme est impliquée dans la rupture cellulaire et la libération des mérozoïtes des globules rouges (Teuscher et al., 2007; Lauterbach and Coetzer, 2008).

Toujours chez les eucaryotes, la cascade enzymatique du système rénine-angiotensine (RAS) fait appel à plusieurs aminopeptidases (comme déjà présentée dans la **Figure 5, paragraphe 1.2.1.2.**) (Banegas et al., 2006), notamment la glutamyl aminopeptidase ENPEP (ou aminopeptidase A) et l'aspartyl aminopeptidase DNPEP. ENPEP et DNPEP clivent l'angiotensine II (Ang II) en angiotensine III (Ang III) et l'angiotensine I (Ang I) en angiotensine 2-10, respectivement (Wright and Harding, 1997; Wilk et al., 1998; Reaux et al., 1999; Min et al., 2000). Elles hydrolysent un résidu Asp du côté Nt des substrats pour générer une nouvelle espèce d'angiotensine. L'aspartyl aminopeptidase bovine bDNPEP hautement exprimée dans l'œil et la rétine, est proposée comme régulateur clef dans le système RAS oculaire (Chen et al., 2012). Les travaux réalisés par Chen et ses collaborateurs révèlent que bDNPEP, une TET spécifique des résidus acides, est capable d'agir sur les peptides bioactives : Ang I, Ang II, ainsi que CCK-8 (un neuropeptide), *in vitro*. Ces observations consolident la fonction d'angiotensinase de bDNPEP, régulatrice de la pression sanguine et des peptides neuronaux dans l'œil et le système nerveux central.

### 1.4. LES ARCHÉES : SOURCES DES MÉTALLO-AMINOPEPTIDASES TETs

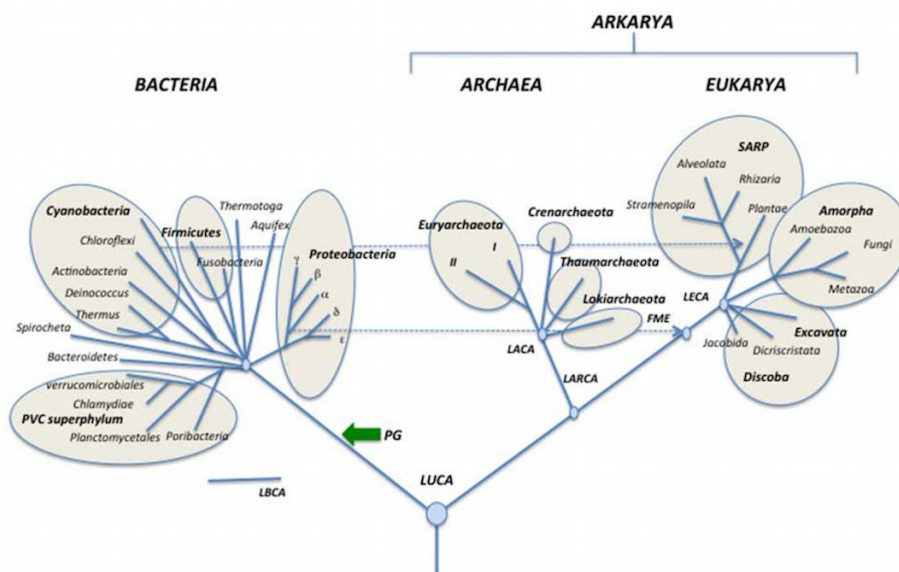
Mon projet de thèse porte sur la caractérisation fonctionnelle et structurale de deux nouvelles aminopeptidases tétraédriques issues de *Methanocaldococcus jannaschii* et *Pyrococcus horikoshii*, deux archées hyperthermophiles issues des sources hydrothermales abyssales. Beaucoup de peptidases TETs ont été caractérisées chez les archées aux plans structural et fonctionnel. L'étude des systèmes archéens est ainsi la plus à même de permettre de comprendre le rôle physiologique de ces machines cellulaires, leur lien avec l'adaptation

## INTRODUCTION

environnementale et les raisons de l'existence de diverses versions de particules TETs dans une même cellule. Dans ce chapitre, je présenterai donc le domaine des archées, son intérêt pour l'étude de la protéolyse ainsi que les principales caractéristiques des deux souches modèles.

### 1.4.1. « *Woesian Revolution* » : Révélation d'un troisième domaine du vivant

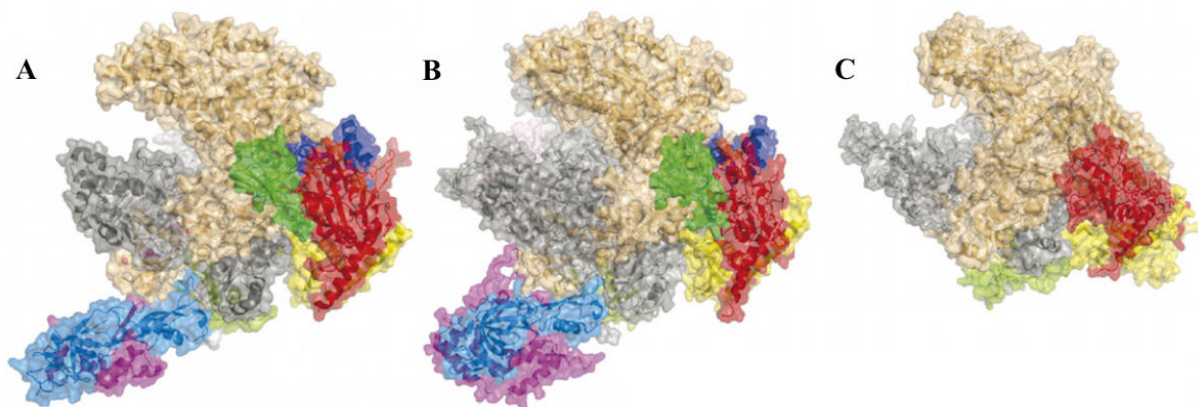
Dans les années 70, les analyses des séquences d'ARN ribosomal ont permis à Carl Woese et George Fox d'identifier, au sein des procaryotes, un nouvel ensemble d'organismes. Caractérisés par un métabolisme supposé primitif, ceux-ci furent classés parmi les « Archaeobactéries » du grec signifiant « ancien » (Woese and Fox, 1977). Au fil du temps, les études phylogénomiques ont clairement révélé que les « Archaeobactéries » forment un troisième domaine du vivant bien distinct de celui des eucaryotes et des bactéries. Conséquemment, un arbre phylogénétique à trois branches principales : *Archaea*, *Bacteria* et *Eukarya* a été édifié afin de décrire le monde du vivant (**Figure 24**) (Forterre, 2015).



**Figure 24 : Arbre universel du vivant adapté depuis Woese et al., 1990 et mis à jour par Forterre, 2015.** Bacteria, Archaea et Eukarya constituent les trois branches de l'arbre phylogénétique du vivant. Arkarya est une proposition pour regrouper le domaine Archaea avec le domaine Eukarya pour un arbre à deux domaines. Les termes LUCA, LBCA, LARCA, LACA et LECA désignent le dernier ancêtre commun de chaque embranchement (*Last Universal / Bacterial / Arkarya / Archaeal / Eukaryotic Common Ancestor*) (Forterre, 2015).

## INTRODUCTION

Mises à part leurs spécificités distinctives, les archées présentent un lien évolutif qui les rapprochent plus des eucaryotes que des bactéries, notamment pour ce qui est des processus de transmission de l'information génétique, tels que la réplication, la transcription, la traduction et la réparation de l'ADN (Gribaldo et al., 2011). À titre d'exemple, les ARN polymérases archéennes ressemblent plus à leurs homologues eucaryotes que bactériens du point de vue structural et dans la nature des sous-unités qui les composent (**Figure 25**) (Werner, 2008). De même les exosomes qui sont des complexes multimériques impliqués dans la dégradation et la maturation des molécules d'ARN, sont conservés entre les eucaryotes et la plupart des archées (Buttner et al., 2006; Hartung and Hopfner, 2009). De surcroît, de nombreuses enquêtes ont révélé le lien évolutif existant entre les archées et eucaryotes comme la présence des histones (Cubonova et al., 2005), les voies de sécrétion et le système ubiquitine-protéasome : Ubiquitine-19S-20S chez les eucaryotes et SAMP-PAN-20S chez les archées (Gribaldo and Brochier-Armanet, 2006; Maupin-Furlow, 2011). Non seulement les machineries des archées partagent une forte similitude avec leurs congénères eucaryotes, mais elles dévoilent encore une simplicité assez remarquable en comparaison avec ces dernières, comme le soulignent les exosomes (Buttner et al., 2006) et le protéasome 20S. En effet, la version archéenne de ce complexe protéolytique est formée généralement de sept copies identiques de chacune des sous-unités  $\alpha$  et  $\beta$  formant ses quatre anneaux heptamériques. Tandis que chacune des sous-unités  $\alpha$  et  $\beta$  du protéasome eucaryote ont divergé en sept différentes variantes (Groll et al., 1997) (Unno et al., 2002).



**Figure 25 : Structures cristallographiques de trois ARN polymérases archéenne, eucaryote et bactérienne.** **A.** ARN polymérase archéenne de *Sulfolobus solfataricus* (PDB ID : 2PMZ) et **B.** ARN polymérase eucaryote de *Saccharomyces cerevisiae* (PDB ID : 1NT9) présentent une forte homologie de structure et du nombre de sous-unités qui les composent. **C.** En revanche, l'ARN polymérase bactérienne de *Thermus aquaticus* (PDB ID : 1L6V) présente une structure plus simple du fait de l'absence de certaines sous-unités présentes chez ses homologues archéens et eucaryotes (Werner, 2008).

## INTRODUCTION

---

En dépit des caractéristiques communes entre les archées et les eucaryotes, ces premières disposent de certaines propriétés qui leur sont singulières. Avec l'existence de voies métaboliques qui leur sont propres, une des plus remarquables porte sur la composition de leur membrane phospholipidique. Chez les eucaryotes et les bactéries, les chaînes latérales des acides gras sont liées par une liaison ester au glycérol-3-phosphate. Chez les archées, les chaînes isoprénoïdes des membranes sont liées *via* une liaison éther au glycérol-1-phosphate (Peretó et al., 2004). De plus, leur membrane phospholipidique est entourée par une paroi externe de protéines ou de glycoprotéines, dénommée « *S-layer* » (*surface layer*) (Sleytr et al., 2007).

### 1.4.2. Les archées : des organismes ubiquitaires ... et des hyperthermophiles

Il est possible de définir pour chaque espèce une valeur minimale, maximale et optimale de facteurs biotiques et abiotiques sous lesquels elle mène tranquillement son cycle vital. L'approche habituelle en biochimie suggère que les conditions physiques standards représentées par une température modérée (entre 20 et 40 °C), une pression atmosphérique et une disponibilité de l'oxygène sont les lois qui dominent toute la biosphère (Jaenicke, 1981). En franchissant ces frontières, de nouvelles conditions pouvant atteindre des valeurs dites extrêmes, règnent. Dans ces nouvelles conditions, aucun signe de vie ne peut être détecté pour un individu considéré comme non extrémophile (Prieur, 2012). Cependant dans ces mêmes circonstances, se développent d'autres organismes : les extrémophiles.

Pendant longtemps, on avait pensé que les archées étaient uniquement des organismes que l'on pouvait qualifier de « héros de l'extrême ». Pourtant, les études de microbiologie environnementale, en particulier les récentes avancées dues à la métagénomique, dévoilent que ces créatures peuvent coloniser une grande diversité d'écosystèmes terrestres et marins où les conditions de vie sont dites « normales » (Reeve and Schleper, 2011). Ainsi, les archées peuvent se développer dans le sol, les zones marécageuses, les océans où elles constituent 20 % des microorganismes (DeLong and Pace, 2001), les eaux d'égout, le tractus intestinal de certains animaux ou encore font partie du microbiome humain (DeLong, 1998; Lepage et al., 2013). D'autres vivent dans des communautés symbiotiques avec des organismes marins (comme *Crenarchaeum symbiosum* vivant en symbiose avec une éponge marine) (Preston et al., 1996). Toutefois, aucune espèce archéenne pathogène n'a été découverte jusqu'à présent

## INTRODUCTION

---

(Reeve and Schleper, 2011).

Quant aux archées extrémophiles, celles-ci prospèrent dans des biotopes où règnent des pH extrêmement acides ou alcalins, des températures infernales ou glaciales, des pressions colossales et même des concentrations salines proches de la saturation. Ainsi, on distingue divers comportements au sein de l'extrémophilie : l'acidophilie, l'alcalophilie, la thermophilie, la psychrophilie, la piézophilie et l'halophilie, respectivement (Stetter, 1999). Certaines espèces sont même décrites comme des polyextrémophiles, combinant une exigence simultanée de plusieurs caractères extrêmes.

Dans cette étude, nous nous sommes intéressés à l'étude des archées hyperthermophiles isolées à partir des cheminées hydrothermales situées dans les abysses océaniques où les températures peuvent atteindre les 400 °C. A certains endroits, le plancher océanique présente d'énormes accidents de reliefs qui sont le résultat d'immenses mouvements affectant la croûte terrestre, nommés tectonique des plaques. C'est ici que surgissent des sources hydrothermales (**Figure 26**). Où les plaques tectoniques s'écartent, le magma en fusion monte, puis s'étale et se solidifie en se craquelant. L'eau de mer s'y infiltre alors, se réchauffe au contact du basalte, puis remonte par d'autres fissures. L'eau qui jaillit est très chaude, 300 à 400 °C, acide (le pH est compris entre 2,0 et 5,9), totalement dépourvue d'oxygène et enrichie en sels minéraux, en ions polymétalliques (Fe(II), Mn(II), Zn(II), Cu (II) et Cd(II), Pb(II), Co(II), Ni(II)) et en gaz dissous (H<sub>2</sub>S, CO<sub>2</sub>, CH<sub>4</sub> et H<sub>2</sub>). La rencontre de ce fluide hydrothermal avec l'eau de mer provoque une précipitation des sels polymétalliques qu'il contient et engendre de hautes cheminées hydrothermales dépassant parfois 20 m : les fumeurs noirs. C'est à l'interface entre le fluide hydrothermal et l'eau de mer, là où on imaginait trouver un paysage désertique, qu'apparut une biomasse luxuriante abyssale (Le Bris, 2013) formée par des communautés archéennes, bactériennes et eucaryotes originales qui se sont développées au cours des temps géologiques (Querellou et al., 2001). Cette biomasse puise son énergie des composés chimiques réduits émis par ces fumeurs par le biais des micro-organismes chimio-synthétiques (Jannasch, 1984). Ainsi, les sources hydrothermales sont considérées comme des « oasis » du fait de la biodiversité qu'elles abritent en contraste avec les océans profonds considérés comme milieu hostile abyssal, froid et obscur.

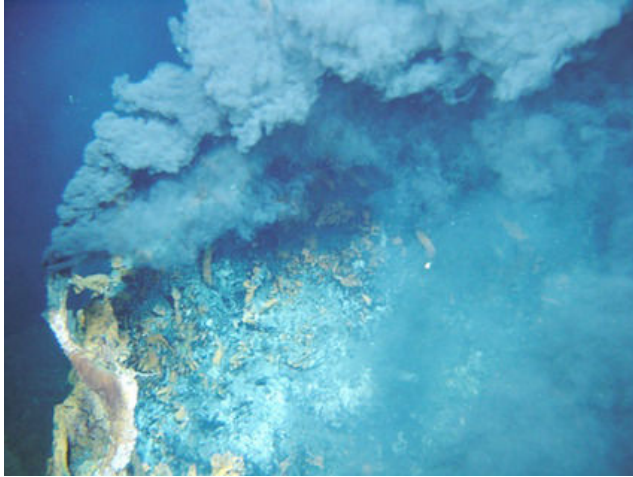


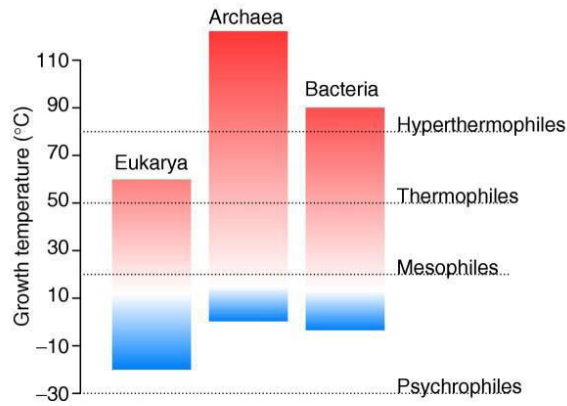
Figure 26 : Fumeur noir sur la dorsale Est Pacifique à 2600 mètres de profondeur (PHARE2002 Ifremer).

### 1.4.3. Les hyperthermophiles ... des vivants dans l'enfer

*« Une vie à haute température ? Longtemps, les scientifiques sont restés sceptiques quant à cette éventualité. Pourtant dans l'enfer (paradisique) de Vulcano et les sources brûlantes des abysses, la vie fleurit... Les microbes hyperthermophiles présentent un avantage certain pour le chercheur : ils n'habitent pas n'importe où. Ils ont le bon goût de peupler bien souvent des endroits parmi les plus magiques, les plus reculés et plus fascinants de notre planète ... »*

*Patrick Forterre, 2007, Microbes de l'enfer.*

Les organismes vivants sont classés en 4 groupes conformément à leur comportement thermophile : les psychrophiles qui se développent à des températures inférieures à 20 °C, les mésophiles qui vivent entre 20-50 °C, les thermophiles qui croissent entre 50-80 °C et les hyperthermophiles qui colonisent des biotopes où les températures sont supérieures à 80 °C (**Figure 27**) (Vieille and Zeikus, 2001).



**Figure 27 : Distribution phylogénétique de la thermophilie.** L'habilité de croître à haute température est commune chez les archées, peu répandue chez les bactéries et relativement absente chez les eucaryotes (Hickey and Singer, 2004).

Plus particulièrement, les hyperthermophiles sont isolés à partir des environnements aquatiques naturels ou industriels (Stetter, 1982, 1996) : champs de fumerolles, geysers, îles volcaniques, solfatares, sources hydrothermales et systèmes d'épuration des eaux d'égout. En 1969, *Thermus aquaticus* fut la première espèce thermophile prospérant à 75 °C, découverte par Brock et ses collaborateurs (Brock and Freeze, 1969). Le record de l'hyperthermophilie est jusqu'à présent détenu par *Pyrolobus fumarii* avec une croissance optimale à 113 °C (Blöchl et al., 1997). La température maximale au-delà de laquelle toute forme de vie serait inabordable reste inconnue et dépendante de la stabilité des biomolécules (Stetter, 1999), mais probablement pas trop distante des 113 °C. Les procaryotes, archées et bactéries, sont les représentants effectifs de la communauté hyperthermophile.

### 1.4.4. Adaptations moléculaires aux hautes températures

Non seulement ces organismes sont capables de survivre à des températures avoisinant la température de l'ébullition de l'eau, mais aussi possèdent une croissance optimale dans ces circonstances (Vieille and Zeikus, 2001). En conséquence, toutes les macromolécules biologiques de ces extrémophiles doivent être adaptées pour fonctionner dans de tels environnements. Ainsi, les enzymes des hyperthermophiles furent nommées thermozymes (Vieille et al., 1996). La thermostabilité et l'activité optimale sous des températures élevées sont des propriétés intrinsèques de la protéine. Il est communément admis que la thermostabilité des édifices macromoléculaires est associée à un cœur hydrophobe plus compact, à une réduction des boucles ainsi qu'à une optimisation du nombre des interactions ioniques, des ponts disulfures et des résidus hydrophobes (Cacciapuoti et al., 2012; Reed et al.,

## INTRODUCTION

---

2013). Des travaux qui comparent la structure et la dynamique moléculaire de diverses enzymes, montrent que l'adaptation thermique est acquise *via* des substitutions d'acides aminés dans des régions clés de la protéine (Lee et al., 2005), ce qui a pour effet de modifier les propriétés de dynamique moléculaire de l'édifice macromoléculaire. Ces modifications permettent, à température physiologique, d'empêcher le dépliement de la chaîne polypeptidique, tout en préservant une flexibilité nécessaire au fonctionnement biologique (Sinha and Smith-Gill, 2002; Tehei et al., 2004). Néanmoins, certains facteurs intracellulaires peuvent aussi contribuer à la thermostabilité comme les cofacteurs métalliques (Colombo et al., 2016), les substrats, les petits solutés (KCl, glycine, saccharose, glutamate), les chaperonnes, les coenzymes, les sels, ... (Vieille and Zeikus, 2001; Neelon et al., 2005; Trivedi et al., 2006; Liu et al., 2012).

### 1.4.5. Phylogénétique des archées

Actuellement, l'arbre phylogénétique des archées comporte presque six phyla : les Euryarchaeota, les Crenarchaeota, les Thaumarchaeota, les Aigarchaeota, les Korarchaeota et les Nanoarchaeota (**Figure 28**). La grande majorité des espèces décrites jusqu'à présent appartiennent aux Euryarchaeota et aux Crenarchaeota (Gribaldo and Brochier-Armanet, 2006; Brochier-Armanet et al., 2011). Le phylum des Euryarchaeota comporte des espèces hyperthermophiles, halophiles extrêmes et méthanogènes, et regroupe neuf ordres parmi lesquels les Methanococcales et les Thermococcales auxquels sont attribués les deux modèles de cette étude : *Methanocaldococcus jannaschii* et *Pyrococcus horikoshii*, respectivement. La plupart des espèces classifiées dans le phylum des Crenarchaeota sont des acidophiles terrestres et représentent les archées les plus abondantes dans les milieux marins. Les Thaumarchaeota comme les Aigarchaeota, les Korarchaeota et les Nanoarchaeota ont été identifiés plus récemment. Les premiers jouent un rôle important dans les cycles géochimiques de la Terre tels les cycles de l'azote et du carbone (Jarrell et al., 2011). Quant aux Aigarchaeota, certaines de leurs caractéristiques communes aux Thaumarchaeota rendent leur classification un sujet à débat, soit comme un nouveau phylum, soit comme un ordre faisant partie des Thaumarchaeota. Les Korarchaeota sont supposés s'être différenciés plus tôt par rapport aux Euryarchaeota et aux Crenarchaeota en référence aux études génétiques entreprises par Elkins et ses collègues (Elkins et al., 2008). Enfin, la classification des Nanoarchaeota, représentés à ce jour par une seule espèce connue pour avoir l'un des plus petits génomes (0,5 Mb) (Huber et al., 2002), est aussi un sujet à débat du fait que certaines



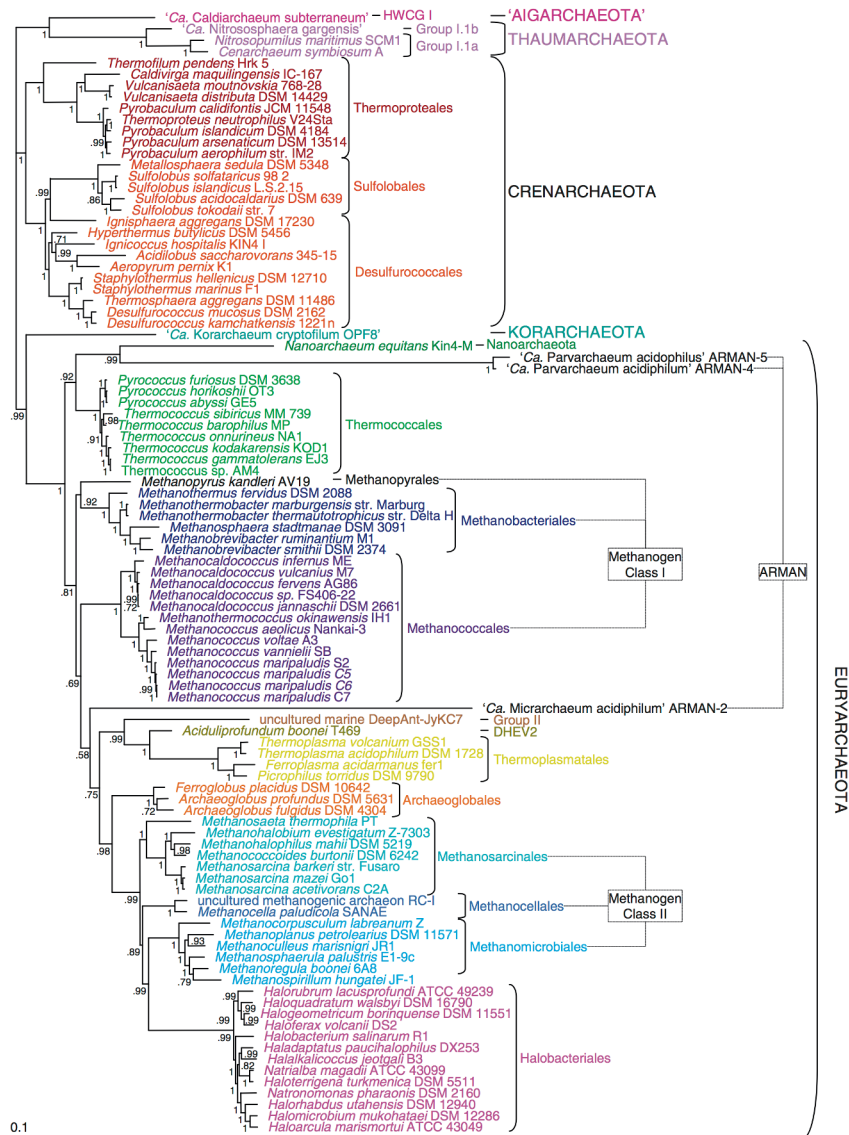
## INTRODUCTION

---

études les assignent au phylum des Euryarchaeota (Brochier et al., 2005). Un nouveau phylum provenant des résultats de métagénomique a été récemment proposé. C'est le cas des Lokiarchaeota révélant une étonnante proximité avec les eucaryotes, ce qui a conduit leurs auteurs à proposer qu'ils seraient les ancêtres directs des eucaryotes (Spang et al., 2015).

Les deux peptidases auxquelles je me suis intéressée durant ma thèse sont élaborées par deux espèces d'archées hyperthermophiles appartenant au phylum des Euryarchaeota : *Methanocaldococcus jannaschii* et *Pyrococcus horikoshii* qui feront l'objet d'une exposition dans la partie suivante.

# INTRODUCTION



**Figure 28 : Arbre phylogénétique des Archées comme proposé par Brochier-Armanet et ses collègues.** Le domaine des Archées comportent plus ou moins six phyla dont la nomenclature se termine par le terme « AEOTA ». Le phylum le plus développé est celui des Euryarchaeota comportant les deux ordres des Methanococcales et des Thermococcales auxquels appartiennent les deux espèces d'intérêt *M. jannaschii* et *P. horikoshii* (Brochier-Armanet et al., 2011).

## 1.4.5.1. *Methanocaldococcus jannaschii*

La découverte de *M. jannaschii* remonte à 1983 grâce aux travaux menés par Jones et ses collègues. Ceux-ci ont réussi à isoler à partir des échantillons collectés au niveau des sources hydrothermales (fumeur blanc) à 2600 m de profondeur au niveau de la dorsale Est Pacifique, une nouvelle espèce d'archée baptisée *Methanocaldococcus jannaschii* en l'honneur de Dr. Holger W. Jannasch (Jones et al., 1983).

*M. jannaschii* est une archée hyperthermophile anaérobie stricte, autotrophe utilisant le

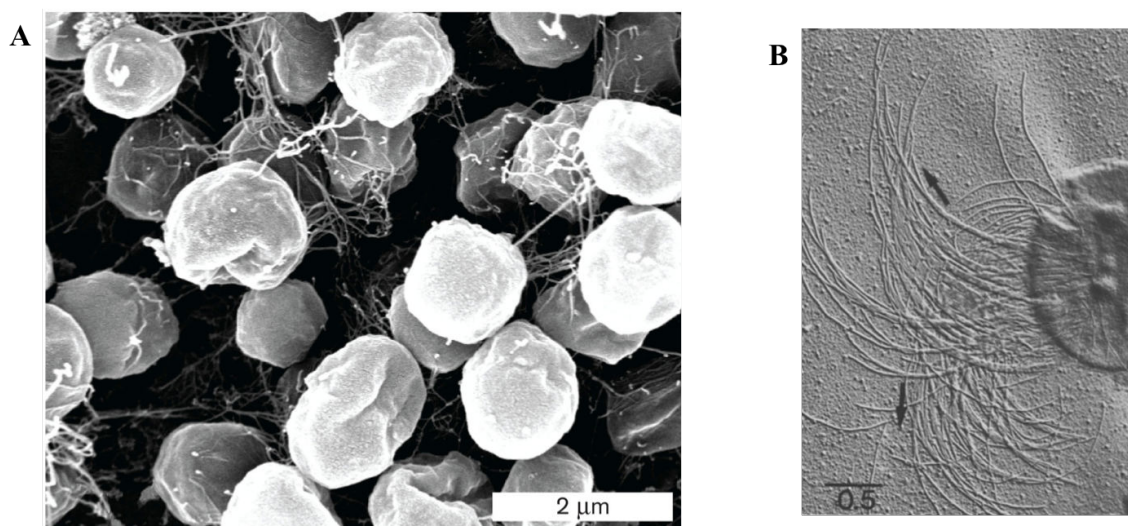
## INTRODUCTION

---

dioxyde de carbone (CO<sub>2</sub>) comme seule source de carbone pour la synthèse des composés organiques. C'est un méthanogène se servant du CO<sub>2</sub> et du dihydrogène (H<sub>2</sub>) pour la production du méthane (CH<sub>4</sub>) (Zhu et al., 2004). L'exploitation du dihydrogène comme source d'énergie pour la méthanogenèse permet de lui attribuer la propriété de chimioautotrophe.

*M. jannaschii* prospère à des températures comprises entre 50 et 86 °C et un pH entre 5,2 et 7 avec un optimum proche de 85 °C et pH 6. Une pression de 200 atmosphères et une salinité variant entre 1 et 5 % dominant l'habitat de cette archée. La cellule adopte la forme d'une coque irrégulière dont la paroi montre des plissures rappelant la structure d'un grain de raisin sec (**Figure 29A**). *M. jannaschii* est une archée flagellaire comportant deux faisceaux divergents d'archaellum insérés au niveau d'un même pôle cellulaire (**Figure 29B**). La membrane cytoplasmique est entourée par une enveloppe externe « *S-layer* » d'environ 20 nm d'épaisseur.

Elle fut la première archée dont le génome est complètement séquencé (Bult et al., 1996). Une prédiction des rôles attribués à certains gènes montre que 9 % sont impliqués dans le métabolisme énergétique et 80 % interviennent dans le métabolisme des acides gras et des phospholipides. D'autres dont les fonctions sont inconnues, sont supposés requis pour la croissance cellulaire (Zhu et al., 2004).



**Figure 29** : *Methanocaldococcus jannaschii* observé sous microscope électronique. **A**. Image de *M. jannaschii* obtenu par microscope électronique (Slonczewski and Foster, 2013). **B**. Observation de deux touffes flagellaires divergentes insérées au niveau du même pôle cellulaire. L'orientation de chacune est représentée par une flèche noire (Jones et al., 1983).

### 1.4.5.2. *Pyrococcus horikoshii*

Quatorze ans après la découverte de *M. jannaschii*, González et ses collègues parviennent à identifier une nouvelle espèce d'archée du genre *Thermococcus* : *Pyrococcus horikoshii* (nommée ainsi en l'honneur de Dr. K. Horikoshi) à 1395 m de profondeur dans la fosse d'Okinawa (González et al., 1998).

*P. horikoshii* dévoile une forme de coques irrégulières (**Figure 30**) entourées par une enveloppe « *S-layer* » et dont le diamètre varie entre 0,8 et 2  $\mu\text{m}$ . La cellule présente une touffe d'archaellum implantée au niveau d'un même pôle cellulaire. Il s'agit d'une espèce hyperthermophile vivant à une température comprise entre 80 et 102 °C avec un optimum de 98 °C. À ces températures, le dioxygène n'est plus soluble d'où le caractère anaérobique de cette souche. *P. horikoshii* se développe sur une large gamme de pH qui varie entre 5 et 8 et en présence de 1 à 5 % de NaCl en exhibant une croissance optimale à pH 7 et 2,4 % de concentration saline. *P. horikoshii* est une espèce hétérotrophe se développant sur des substrats protéiques comme la peptone, la caséine ainsi qu'un mélange d'acides aminés complexés à des vitamines (González et al., 1998). Ceux-ci sont utilisés comme source de carbone et d'énergie. Contrairement à ses voisins appartenant au même genre des *Pyrococcus* (tels que *P. abyssi* et *P. furiosus*), *P. horikoshii* est incapable de pousser dans un milieu non peptidique où par exemple le maltose ou l'amidon sont les sources de carbone.

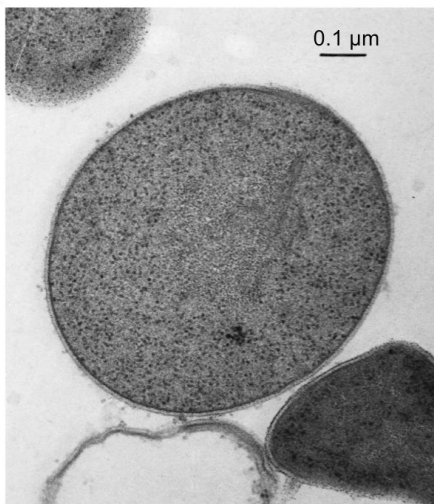


Figure 30 : Observation d'une fine section de *Pyrococcus horikoshii* par microscopie électronique (González et al., 1998).

### 1.4.6. Pourquoi des machineries archéennes sont-elles choisies comme modèles d'étude de la protéolyse ?

Le laboratoire ELMA « *Extremophiles and Large Molecular Assemblies* » à l'Institut de Biologie Structurale, utilise principalement les archées pour disséquer la fonction de protéolyse, rechercher de nouvelles peptidases à partir des génomes et étudier les mécanismes d'assemblage et le mode d'action des machineries cellulaires complexes comme le protéasome. Caractérisées par un petit génome et une position particulière dans l'évolution, les archées constituent de bons modèles pour étudier les processus vitaux qui se déroulaient dans les cellules primitives. Comme nous l'avons vu précédemment, les machineries archéennes peuvent être considérées comme des versions moins complexes de leurs équivalents eucaryotes. Produites par des espèces hyperthermophiles, ces macromolécules présentent un grand avantage pour les études biochimiques et structurales. En effet, leurs propriétés thermophiles autorisent l'ajout d'une étape supplémentaire de purification qui se présente par un choc thermique permettant de se débarrasser des protéines mésophiles d'*Escherichia coli*, généralement utilisée comme hôte de la surexpression protéique. De plus, elles favorisent la possibilité d'entreprendre la purification des protéines hyperthermophiles à température ambiante sans un risque d'agrégation. De même, la thermostabilité et la thermo-activation des telles protéines facilitent la préparation d'échantillons monodisperses et stables pour les études enzymatiques et structurales (cristallographie, diffusion aux petits angles, résonance magnétique nucléaire).

En ce qui concerne la microbiologie, on dispose actuellement d'une bonne base de génomes séquencés (Leigh et al., 2011). Des outils génétiques ont également été développés permettant de réaliser des knockouts pour l'étude fonctionnelle des protéines orphelines. Néanmoins, la manipulation des Methanococcales et Thermococcales reste difficile, du fait de leur caractère anaérobique et de leur capacité à dégager du sulfure d'hydrogène, H<sub>2</sub>S, connu pour sa toxicité. Toutefois, les importantes fluctuations et les conditions extrêmes que subissent ces microorganismes dans leurs environnements naturels, encouragent les chercheurs à explorer leurs capacités adaptatives et leur plasticité métabolique. Ainsi, les recherches menées au sein du laboratoire ELMA visent à répondre aux questions qui se posent sur les rôles des machineries protéolytiques archéennes dans l'adaptation extrêmophile. En effet, la protéolyse est une fonction clé pour l'adaptation environnementale des microorganismes. Elle contrôle l'élimination des protéines anormales, ou qui sont devenues

## INTRODUCTION

---

inutiles pour la cellule. Ou encore, elle participe au métabolisme énergétique *via* le catabolisme des protéines ou des peptides du « soi » (autophagie), ou de celles provenant du milieu environnant. De plus, un rôle régulateur de certaines fonctions physiologiques pourrait être attribué à la protéolyse et en particulier aux peptidases, notamment les aminopeptidases qui sont susceptibles de jouer un rôle clé dans l'homéostasie des acides aminés dont certains pourraient jouer un rôle dans des adaptation extrêmophiles comme la piézophilie (Jebbar et al., 2015).

## 2. PROBLÉMATIQUE ET OBJECTIFS DE LA THÈSE

Mon projet de thèse se situe dans la thématique de la protéolyse intracellulaire qui est une fonction cellulaire complexe. La destruction ou la modification des protéines et des peptides apparaît être essentielle pour la régulation de nombreuses fonctions cellulaires et pour lutter contre le stress. De nombreuses familles de protéases et de peptidases interviennent dans ce processus et beaucoup d'entre elles n'ont pas de rôle physiologique clairement défini. Dans ce contexte, mon travail de thèse s'intéresse aux peptidases ATP-indépendantes capables de former de grands assemblages et plus particulièrement aux métallo-aminopeptidases TETs de la famille M42. En effet, bien que de nombreuses études biochimiques et structurales aient été réalisées sur cette classe d'enzymes, notamment chez l'archée *P. horikoshii*, il reste de nombreuses questions concernant le(s) rôle(s) et le mode d'action de ces étonnantes machines cellulaires.

Le but de cette thèse était de révéler de nouveaux aspects de la biologie des TETs à travers l'étude de deux nouveaux systèmes que nous avons choisis sur la base des analyses phylogénétiques réalisées par S. Gribaldo à l'Institut Pasteur, basées sur les motifs structuraux identifiés par A. Appolaire (manuscrit en préparation). Ces analyses, dont les résultats principaux sont présentés de manière schématique sur la **Figure 31**, ont révélé que, chez les archées, les TETs possèdent deux origines différentes. Certaines d'entre-elles proviennent d'un transfert horizontal de gènes depuis les bactéries. C'est le cas des TETs exprimées par les Thermococcales, les Halobacteriales et les Korarchaeota. Ces TETs d'origines bactériennes sont présentes en plusieurs copies, comme l'illustrent les Thermococcales. D'autres ont conservé des TETs archéennes apparentées à celles de l'ancêtre commun des archées (*LACA, last archaeal common ancestor*). C'est le cas par exemple des TETs identifiées chez les Methanococcales et les Desulfurococcales. Celles-ci sont généralement présentes en une seule copie, hormis les Nanohaloarchaea qui en possèdent deux exemplaires (**Figure 31**). Pour mieux comprendre le système TET, nous avons choisi de caractériser une TET archéenne, présente en une seule version chez l'archée méthanogène hyperthermophile autotrophe *Methanocaldococcus jannaschii* et une quatrième TET « putative » d'origine bactérienne présente chez l'archée hyperthermophile hétérotrophe *P. horikoshii*.

L'étude comparative des propriétés enzymatiques et des structures d'une enzyme archéenne et d'une nouvelle enzyme d'origine bactérienne avait pour objectif de répondre aux

## PROBLÉMATIQUE ET OBJECTIFS

---

questions suivantes :

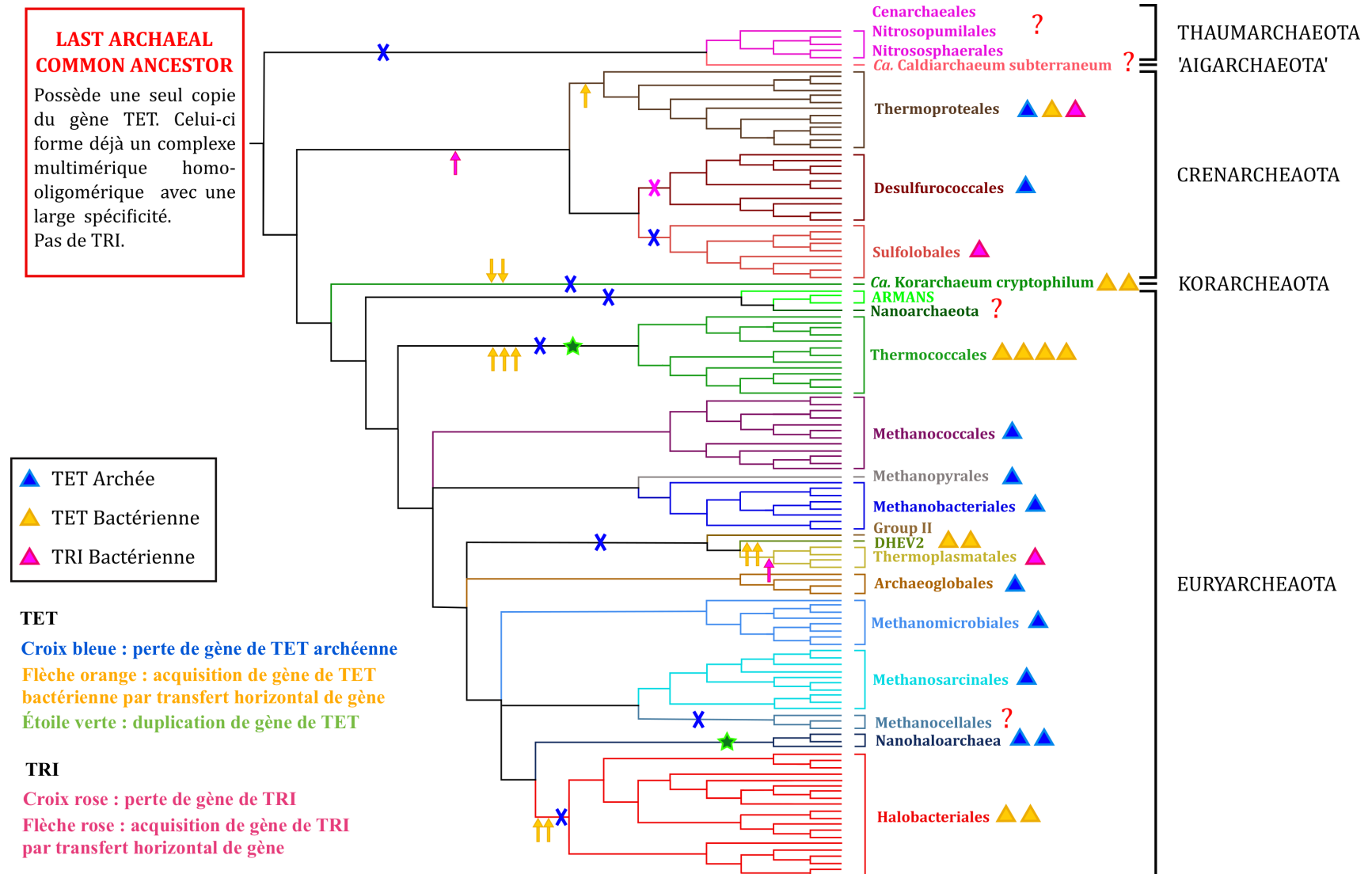
- Quels sont les points communs et les différences entre les systèmes TETs archéens et bactériens ?
- Quel(s) avantage(s) y a-t-il à acquérir plusieurs versions d'une machinerie cellulaire comme TET ?
- Pourquoi y a-t-il eu des transferts de gènes issus des bactéries et une perte du système archéen chez certaines archées ?
- Quel est le rôle physiologique du système TET « ancestral » ?
- Quels sont les déterminants structuraux de la spécificité de substrat chez les aminopeptidases TETs ?

Pour pouvoir répondre à ces questions, j'ai mis au point des protocoles de purification pour les deux enzymes, réalisé des tests enzymatiques utilisant des substrats commerciaux. J'ai également étudié par chromatographie en phase inverse, HPLC, le profil de dégradation de peptides de compositions diverses. Parallèlement, j'ai produit des cristaux de MjTET taguée du côté N-terminal d'une étiquette polyhistidine, qui m'ont permis de résoudre la structure de MjTET à 2,4 Å par remplacement moléculaire. De plus, en collaboration avec le groupe de G. Schoehn de l'IBS, nous avons élaboré une structure à basse résolution de MjTET à 4,1 Å. Grâce à la combinaison de la cristallographie et de la cryo-EM, nous avons pu construire un modèle atomique complet de MjTET et proposé un nouveau modèle de trafic de peptides à l'intérieur de la machine tétraédrique TET.

Les différentes techniques et méthodes employées durant cette thèse sont exposées dans la partie suivante.



## PROBLÉMATIQUE ET OBJECTIFS



**Figure 31 : Représentation schématique de la phylogénie des TETs et de TRI (Tricorne peptidase) chez les archées.** La phylogénie schématisée est basée sur une analyse bayésienne de protéines ribosomales concaténées (Appolaire, 2014).

### 3. MATÉRIELS ET MÉTHODES

#### 3.1. MATÉRIELS ET MÉTHODES BIOCHIMIQUES

Une protéine d'intérêt peut être isolée soit directement à partir de son organisme d'origine ou soit après une surproduction dans un organisme hôte ayant reçu le gène codant pour cette protéine. Compte tenu de la complexité de cultiver les espèces archéennes hyperthermophiles anaérobiques, telles que *Methanocaldococcus jannaschii* et *Pyrococcus horikoshii*, nous avons eu recours à la surproduction des deux protéines TETs dans des souches d'*Escherichia coli*. De plus cette méthode permet d'augmenter le rendement de protéines nécessaires pour les études envisagées.

##### 3.1.1. Production de protéines

###### 3.1.1.1. Synthèse de gènes et surproduction des protéines

La synthèse de gènes et la surproduction protéique avaient pour but de générer une quantité assez suffisante de MjTET et PhTET4 requise pour les études enzymatiques, cristallographiques et de microscopie électronique qui ont été entreprises durant ma thèse. Il est important de noter qu'une deuxième forme de MjTET taguée par une étiquette polyhistidine du côté N-terminal (MjTET-Nt) était aussi produite en vue de générer des cristaux indispensables pour la caractérisation structurale. En fait, les cristaux obtenus avec la forme non taguée n'étaient pas adéquats (forme irrégulière) pour la collecte des données cristallographiques. Ainsi au cours de ma thèse, les protocoles expérimentaux d'expression et de purification de trois protéines : MjTET, MjTET-Nt et PhTET4 ont été optimisés.

La synthèse des gènes codant pour les protéines MjTET, MjTET-Nt et PhTET4 a été entreprise par GeneCust Europe, Luxembourg. Les séquences correspondantes ont été optimisées pour les codons rares des souches d'*E. coli* utilisées pour l'expression protéique. MjTET et PhTET4 ont été clonées, séparément, dans pET-41c, alors que le gène codant pour MjTET-Nt a été cloné dans le vecteur pET28a+. L'expression protéique a été induite par de l'IPTG (0,1 mM pour les deux formes de MjTET et 0,05 mM pour PhTET4) une fois que la densité optique à 600 nm ayant atteint 0,6. Les deux formes de MjTET ont été exprimées dans *E. coli* ArcticExpress (DE3)-RIL (*Agilent Technologies*) pendant 7 heures à 16 °C dans un

## MATÉRIELS ET MÉTHODES

---

litre de milieu de culture (LB) auquel sont ajoutés de la kanamycine et la gentamycine à 50 µg/ml. Le choix de la souche ArcticExpress (DE3)-RIL était basé sur le fait que MjTET était produite sous forme insoluble dans *E. coli* BL21 (DE3)-RIL utilisée ordinairement pour la production des TETs dans notre laboratoire. Grâce à celle-ci, l'expression et le repliement protéiques peuvent être entrepris à basse température (ici 16 °C). PhTET4 a été surexprimée dans *E. coli* BL21 (DE3)-RIL pendant 4 heures à 37 °C dans un litre de LB contenant 50 µg/ml de la kanamycine.

Les cellules sont récupérées par centrifugation à 4000 rpm durant 20 min à 4 °C en utilisant un rotor JLA-8.1000 (Beckman) et sont ensuite conservées à -80 °C.

### 3.1.1.2. Purification

Les bactéries exprimant les deux protéines, MjTET et MjTET-Nt, sont solubilisées dans du tampon de lyse : 50 mM Tris-HCl, 200 mM NaCl, pH 8, ce mélange est supplémenté avec 50 mM d'imidazole dans le cas de MjTET-Nt. Quant à PhTET4, les bactéries sont resuspendues dans 25 ml Tris-HCl 50 mM pH 8, NaCl 90 mM, 0,1 % Triton X-100, 6,25 mg lysozyme (Euromedex), 1,25 mg DNase I grade II (Roche), 5 mg RNase (Roche), 25 mg Pefabloc SC (Roche) et 0,25 ml MgSO<sub>4</sub> à 2 M sont ajoutés à chacune de trois solutions de lyse. Les cellules sont lysées par 5 à 10 cycles de sonication (*Branson Sonifer*) (30 sec ON / 30 sec OFF) à 30 watts à 4 °C. Puisque les protéines étudiées sont isolées d'organismes hyperthermophiles, elles sont plus stables à haute température que les protéines mésophiles d'*E. coli*. Ainsi, une étape de purification supplémentaire, une clarification, est introduite. Elle consiste à chauffer 20 min à 85 °C et 75 °C les deux formes de MjTET et PhTET4, respectivement. Les lysats sont ensuite centrifugés à 12000 rpm pendant 1 heure à 4 °C, en utilisant un rotor JA.20 (Beckman) et les surnageants contenant les protéines sont récupérés.

L'étape de clarification permet ainsi de se débarrasser de la majorité des protéines mésophiles de l'hôte. Pour compléter cette première étape, une purification de MjTET, MjTET-Nt et PhTET4 est entreprise par deux types de chromatographies : une forte échangeuse d'anions (Resource Q) et une chromatographie d'exclusion de taille (Superose 6 ou Superdex 200) pour MjTET et PhTET4. MjTET-Nt est purifiée par chromatographie d'affinité (*HiTrap Chelating HP*), suivie d'une chromatographie d'exclusion de taille (Superose 6 ou Superdex 200). Les colonnes sont fournies par *GE Healthcare Life Sciences*.

## MATÉRIELS ET MÉTHODES

---

Les purifications de chacune de ces trois protéines sont détaillées dans la partie « *Experimental procedures* » des deux *ébauches* d'articles.

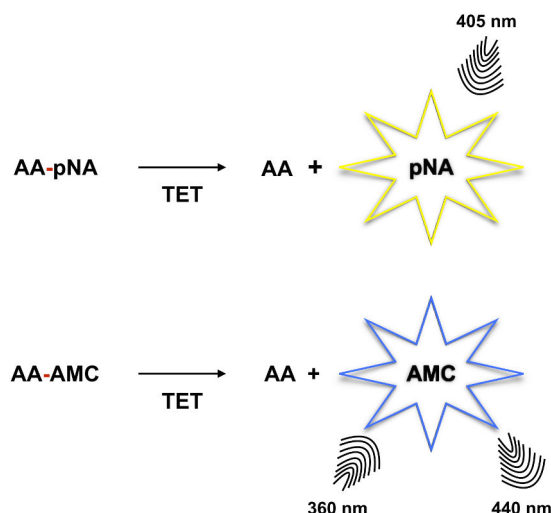
Le contrôle de la pureté des protéines produites est réalisé par électrophorèse sur gel de polyacrylamide en condition dénaturante (SDS-PAGE).

### 3.1.2. Activité enzymatique

#### 3.1.2.1. Substrats chromogènes et fluorogènes

Selon MEROPS, MjTET et PhTET4 sont classées parmi les métallopeptidases de la famille M42. C'est pourquoi leur activité enzymatique a été évaluée contre une diversité de substrats couplés par une liaison peptidique au pNA (*p-NitroAniline*), un chromophore, ou à l'AMC (*7-Amino-4-MethylCoumarin*), un fluorophore. Nous avons aussi utilisé des substrats peptidiques. Dans le cas des substrats couplés au pNA ou à l'AMC, nous avons suivi l'activité enzymatique en mesurant la quantité libérée de ces molécules en fonction du temps. La quantité de pNA produite est déterminée en mesurant son absorbance à 405 nm à l'aide d'un spectrophotomètre JASCO V-630. Le taux d'AMC libéré est déterminé en mesurant la fluorescence après une excitation et une émission à 360 et 440 nm, respectivement, à l'aide d'un spectrofluorimètre JASCO FP-8500 (**Figure 32**).

Le volume réactionnel de 390  $\mu$ l renferme 50 mM tampon (PIPES, CHES ou CAPS) pH x, 150 mM KCl, 0,1 mM sel métallique ( $XCl_2$ ) et 5 mM de substrat. Il est recouvert de 20  $\mu$ l d'huile minérale pour empêcher l'évaporation de l'eau et est chauffé à la température requise (entre 20 et 95 °C) pendant quelques secondes. 10  $\mu$ l d'enzyme y sont ajoutés, pour une concentration finale de 4  $\mu$ g/ml. La réaction enzymatique est suivie tout en chauffant (grâce à un bloc chauffant intégré à l'appareil de mesure) et en mesurant la quantité de pNA ou d'AMC produite durant 10 min. Il est important de noter que les concentrations du tampon, d'ions métalliques et du substrat sont les concentrations finales dans un volume de 400  $\mu$ l.



**Figure 32 : Représentation schématique de la réaction enzymatique exécutée par une TET (MjTET ou PhTET4) contre un substrat chromogénique (AA-pNA) et fluorogénique (AA-AMC).** La liaison en rouge représente une liaison peptidique et AA correspond à un acide aminé.

### 3.1.2.2. High performance liquid chromatography (HPLC)

Dans le but d'étudier la dégradation peptidique par MjTET et PhTET4, nous avons eu recours à la HPLC permettant de séparer des molécules de faible poids moléculaire, comme c'est le cas des produits générés par la digestion peptidique.

La HPLC en phase inverse est une forme hautement améliorée de la chromatographie permettant, sous une forte pression, de séparer les molécules en se basant sur leurs caractères hydrophobes (Aguilar, 2004). Elle est caractérisée par une phase stationnaire en silice modifiée sur laquelle sont greffées des chaînes d'hydrocarbures typiquement C2-C18 et d'une phase mobile polaire contenant un solvant organique, comme l'acétonitrile. Dans cette situation une forte interaction existe entre les molécules polaires et le solvant, contrairement aux molécules apolaires qui interagissent plus longtemps avec la phase stationnaire. Ainsi, les molécules polaires sont éluées plus rapidement que les solutés hydrophobes qui seront élués d'une façon isocratique ou par gradient continu au fur et à mesure de la diminution de la polarité de la phase mobile.

Un volume réactionnel de 98  $\mu$ l contenant 50 mM CHES, 150 mM KCl, 0,75 mM peptide et 0,05 mM  $\text{CoCl}_2$ , pH 8,5 (dans le cas de MjTET) ou 0,05 mM  $\text{NiCl}_2$ , pH 9,5 (dans le cas de PhTET4), est chauffé une minute à 85  $^{\circ}\text{C}$  après avoir couvert la totalité de ce volume par 20

## MATÉRIELS ET MÉTHODES

---

$\mu\text{l}$  d'huile minérale. Ensuite, 2  $\mu\text{l}$  d'enzyme y sont ajoutés pour une concentration finale de 2  $\mu\text{g}/\text{ml}$ . Après 6 min d'incubation à 85 °C, 80  $\mu\text{l}$  du volume réactionnel sont récupérés et la réaction est arrêtée en y ajoutant 220  $\mu\text{l}$  d'une solution d'acétonitrile (ACN) 2 % et d'acide trifluoroacétique (TFA) 0,1 %, à 4 °C. Le mélange est centrifugé par la suite 15 min à 13000 rpm à 4 °C afin de culoter les protéines ayant précipité lors de l'addition de l'ACN et du TFA. Finalement, 280  $\mu\text{l}$  sont récupérés et injectés sur une colonne C18 (60 Å, 4  $\mu\text{m}$ , 3,9 mm x 300 mm) (Waters) équilibrée par 100 % de phase A : 2 % ACN, 0,1 % TFA, et l'élution des peptides est entreprise avec un gradient croissant d'ACN variant de 2 à 33,2 %. Dans le cas des peptides incubés avec PhTET4, ceux-ci sont soumis à un séquençage Nt afin de vérifier la nature de l'acide aminé clivé.

### 3.2. MATÉRIELS ET MÉTHODES BIOPHYSIQUES

#### 3.2.1. Microscopie électronique après coloration négative

En vue d'évaluer l'état oligomérique des protéines purifiées MjTET, MjTET-Nt et PhTET4, 4  $\mu\text{l}$  de protéines à 0,05 mg/ml sont adsorbés sur la facette propre d'un film de carbone déposé sur une couche de mica. Les échantillons de MjTET et de PhTET4 sont colorés par 2 % de silicotungstate de sodium pH 7,5 et 2 % d'acétate d'uranyle, respectivement, puis transférés sur une grille en cuivre et séchés à l'air. Les images sont prises sous conditions à faible dose ( $<10 \text{ e}^-/\text{\AA}^2$ ), avec un microscope électronique *Tecnai 12 LaB6* à 120 kV, sous un grossissement de 23 et 49 kX en utilisant une caméra *CCD Gatan Orius 1000*.

#### 3.2.2. Cristallographie

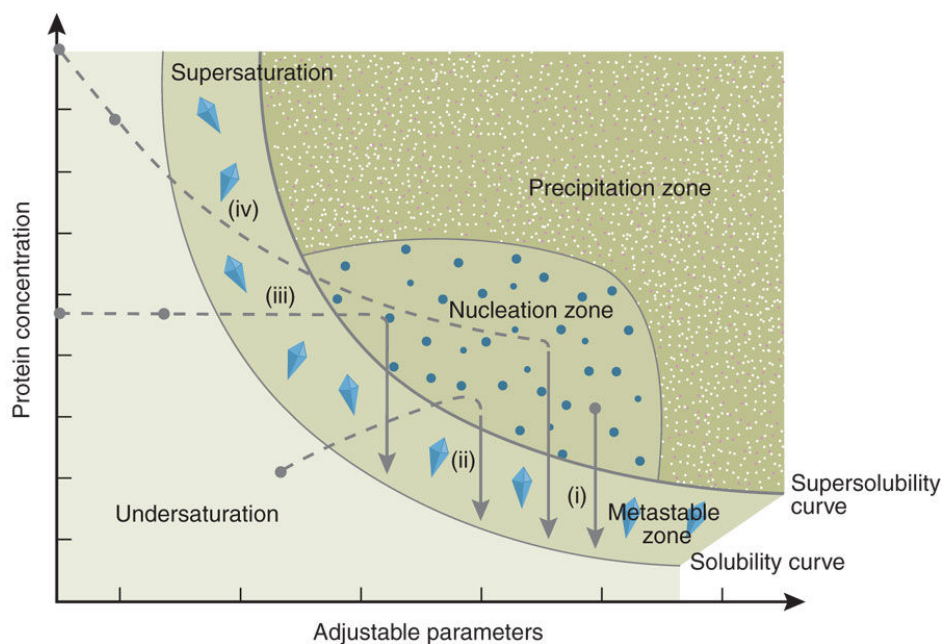
##### 3.2.2.1. Cristallisation

La détermination de la structure cristallographique d'une macromolécule est conditionnée par l'obtention des cristaux et par la qualité des données de diffraction obtenues après diffraction des rayons X par les électrons des macromolécules empilées dans le cristal.

Un cristal est l'arrangement périodique ordonné dans un espace tridimensionnel d'un motif se répétant identiquement à lui-même d'une maille à une autre. Cette dernière

correspond à l'unité de base permettant de reconstruire le cristal en y appliquant des translations.

La difficulté de la cristallogénèse est de trouver les conditions thermodynamiques et les facteurs cinétiques pouvant amener la macromolécule (ici une protéine) d'une phase liquide à un état soluble instable de sursaturation et induire la nucléation afin de passer à un état solide ordonné sans tomber dans la précipitation (**Figure 33**).



**Figure 33 : Représentation schématique du diagramme de phase de la cristallisation d'une protéine.** Les quatre méthodes de cristallisation sont représentées : (i) *microbatch*, (ii) diffusion de vapeur, (iii) dialyse et (iv) FID (*Free interface diffusion method*) (Chayen and Saridakis, 2008). Dans cette étude, la méthode de diffusion de vapeur, technique goutte suspendue est utilisée au cours de la cristallogénèse.

Dans le but de déterminer les conditions de cristallisation de MjTET, j'ai eu recours à la plateforme de criblage automatisé HTXLab de l'EMBL, Grenoble. Disposant d'un robot de cristallisation nanogoutte, cette plateforme permet de tester un grand nombre de conditions de cristallogénèse sans consommer énormément de protéines. Effectivement, 100  $\mu$ l de protéines à x mg/ml sont suffisants pour tester 576 conditions de cristallisation en gouttes assises (6 plaques à 96 puits chacune). Néanmoins, aucun cristal de forme régulière n'est obtenu pour MjTET. En conséquence, j'ai ajouté une étiquette polyhistidine du côté Nt de la protéine dans le but de tester son effet sur la cristallisation. Après reproduction manuelle en goutte suspendue des meilleures conditions de cristallisation, de beaux cristaux de  $\sim$ 80-110  $\mu$ m sont

obtenus en deux jours à 20 °C dans du citrate trisodique 0,1 M pH 5,6, 15 % tert-butanol, en utilisant 5,6 mg/ml de MjTET-Nt.

### 3.2.2.2. Enregistrement des données de diffraction

L'étape suivante consiste à collecter des données cristallographiques par diffraction des rayons X par les électrons des molécules présentes dans le cristal. Ainsi, pour atténuer l'endommagement de ce dernier par les radicaux libres dus aux irradiations engendrées par l'énergie intense du rayonnement synchrotron, les cristaux sont cryo-protégés à des températures inférieures à 100 K. À basse température, les cristaux sont protégés par un cryoprotectant comme le PEG400, l'éthylène glycol, le 2-méthylpentane-2,4-diol (MPD), le glycérol, ... Une fois trempés avec le cryoprotectant adéquat, les cristaux sont plongés dans de l'azote liquide pour éviter la formation de cristaux de glace pouvant altérer la qualité des données cristallographiques enregistrées.

Un cristal est formé d'une infinité de plans dans lesquels les atomes sont arrangés d'une manière périodique. Il s'agit des plans réticulaires caractérisés chacun par un indice appelé indice de Miller  $h$ ,  $k$  et  $l$ . À chaque plan du cristal (espace réel) correspond une tâche sur le cliché de diffraction (espace réciproque) appelée réflexion. D'après la loi de Bragg, la famille des plans réticulaires du cristal donne lieu à un faisceau diffracté si :

$$n \lambda = 2d_{hkl} \sin \theta$$

où  $d_{hkl}$  est la distance interréticulaire de la famille des plans  $[hkl]$ ,  $\theta$  est l'angle situé entre le faisceau incident de longueur d'onde  $\lambda$  et la famille des plans  $[hkl]$  (**Figure 34**).

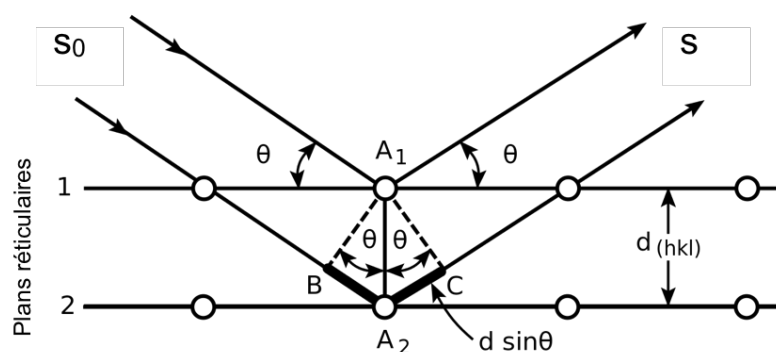


Figure 34 : Schéma décrivant la diffraction des rayons X selon la loi de Bragg.



À chaque réflexion dans le réseau réciproque est associé un facteur de structure  $F_{hkl}$ . Celui-ci est défini par une amplitude  $\| F_{hkl} \|$  et une phase  $\varphi_{hkl}$ . Ces deux composantes sont requises pour calculer la densité électronique en chaque point de l'espace. En effet, elle est calculée par transformée de Fourier inverse des facteurs de structure.

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} F_{hkl} e^{-2i\pi(hx+ky+lz)}$$

Le facteur de structure est relié à la nature des atomes par le facteur de diffusion atomique  $f_j$  et à la position  $(x,y,z)$  des atomes  $j$  dans la maille cristalline du réseau direct.

$$F_{hkl} = \| F_{hkl} \| e^{i\varphi_{hkl}} = \sum_{j=1}^N f_j e^{2i\pi(hx_j+ky_j+lz_j)}$$

Au cours de l'enregistrement des données cristallographiques, les ondes diffractées par le cristal ne peuvent pas être refocalisées. On ne peut donc pas remonter à l'information de phase : c'est le problème des phases. Ainsi, les données expérimentales permettent uniquement d'accéder aux amplitudes des facteurs de structure, qui seront calculées à partir des intensités ( $I_{hkl}$ ) des réflexions puisque :

$$I_{hkl} = \| F_{hkl} \|^2$$

Afin de résoudre le problème de phase, plusieurs méthodes dites de phasage ont été développées telles le remplacement moléculaire (MR), le remplacement isomorphe (SIR et MIR), la diffusion anormale (SAD et MAD) et les méthodes directes. La technique utilisée au cours de ce travail de thèse a été celle du remplacement moléculaire.

### 3.2.2.3. Traitement de données

Le traitement des images de diffraction comporte trois étapes essentielles : l'indexation, l'intégration et la mise à l'échelle.

Grâce à une estimation de l'orientation et de la symétrie du cristal, l'indexation consiste à attribuer les indices de Miller à chacune des réflexions (tâches de diffraction) présentes sur le cliché de diffraction. Cette étape est suivie par l'intégration qui permet de mesurer l'intensité de chacune des réflexions (en établissant un profil moyen des réflexions). À cause de la

## MATÉRIELS ET MÉTHODES

---

dégradation progressive du cristal dans le flux des rayons X, ou de la forme du cristal, les intensités des réflexions censées être identiques par symétrie (Loi de Friedel) et qui sont mesurées à différents intervalles de temps ou pour différentes orientations, ne le seront plus. Ainsi, la mise à l'échelle tient compte de tous ces paramètres et permet de corriger au mieux les intensités de ces tâches de diffraction.

Au cours de mon travail, le traitement des données de diffraction a été réalisé avec le logiciel XDS (Kabsch, 2010). Le groupe d'espace a été confirmé par XTRIAGE de la suite PHENIX (McCoy et al., 2007). Une mise à l'échelle complémentaire a été effectuée avec le programme SCALA de la suite CCP4 (Collaborative Computational Project, Number 4, 1994). Les valeurs moyennes du module du facteur de structure ont été obtenues avec le programme TRUNCATE (CCP4).

La méthode du remplacement moléculaire peut être choisie comme méthode de phasage si la macromolécule d'intérêt est apparentée structurellement à une autre dont la structure est déjà connue et quand les deux molécules, modèle et cible, partagent une similarité de séquence d'au moins 30 %. Cette méthode consiste à placer une ou plusieurs molécules modèles dans la maille cristalline en vue de déterminer les phases initiales des facteurs de structure.

Le modèle utilisé est un polyAla pour lequel les chaînes latérales sont éliminées afin de réduire le biais introduit sur la détermination des phases de la molécule étudiée. Afin de déterminer la structure de MjTET, j'ai généré un modèle polyAla de PhTET2 (PDB ID : 1XFO), une aminopeptidase exprimée par *P. horikoshii* et partageant une identité de séquence de ~45 % avec MjTET. L'étape suivante consiste à trouver l'orientation (fonction de rotation) puis la position de la molécule modèle (fonction de translation) dans l'unité asymétrique du cristal, grâce au logiciel PHASER de la suite CCP4. Pour juger de la qualité du phasage, PHASER produit des paramètres statistiques : les Z-scores des fonctions de rotation (RFZ) et de translation (TFZ), ainsi que le gain du logarithme de maximum de vraisemblance global (LLG). Une solution est considérée comme définitive lorsque son TFZ est supérieur à 8. L'amélioration des phases par aplatissage de solvant a été ensuite réalisée avec Parrot de CCP4. Le modèle produit par PHASER a été manuellement amélioré à l'aide du logiciel COOT (Emsley et al., 2010) en utilisant la carte de densité électronique issue de Parrot.

L'amélioration de la structure consistait en plusieurs cycles d'affinement en utilisant PHENIX (1.10-2155 version) et de reconstruction manuelle dans COOT.

### **3.2.2.4. Interfaces de dimérisation et d'oligomérisation**

Les analyses des interfaces de dimérisation et d'oligomérisation (ou de dodécamérisation) des aminopeptidases TETs sont effectuées grâce au serveur en ligne PDBePISA ([http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)) (Krissinel and Henrick, 2007). Cet outil permet de déterminer les propriétés biochimiques et structurales des interfaces entre les protéines d'intérêt. Les surfaces des interfaces d'interaction, les résidus impliqués dans la formation de ponts salins, de liaisons hydrogènes, de contacts hydrophobes, ... sont ainsi évalués.

### **3.2.2.5. Création des figures**

Les images de structures des protéines qui figurent dans ce manuscrit sont générées grâce à UCSF Chimera (version 1.10.1) (Pettersen et al., 2004) et PyMOL (version 1.8.2.1).

## **4. REVUE SUR LES TETs M42 PROCARYOTES**

Au cours de ma thèse, j'ai contribué à la revue ci-après décrivant les aminopeptidases tétraédriques TETs de la famille M42, chez les procaryotes.

**4.1. TET peptidases: A family of tetrahedral complexes conserved in prokaryotes**

Contents lists available at [ScienceDirect](#)

Biochimie

journal homepage: [www.elsevier.com/locate/biochi](http://www.elsevier.com/locate/biochi)

## Review

## TET peptidases: A family of tetrahedral complexes conserved in prokaryotes

Alexandre Appolaire <sup>a, b, c, 1</sup>, Matteo Colombo <sup>a, b, c, 1</sup>, Hind Basbous <sup>a, b, c</sup>, Frank Gabel <sup>a, b, c</sup>, E. Girard <sup>a, b, c</sup>, Bruno Franzetti <sup>a, b, c, \*</sup><sup>a</sup> CNRS, IBS, F-38027 Grenoble, France<sup>b</sup> CEA, DSV, IBS, F-38027 Grenoble, France<sup>c</sup> Univ. Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38027 Grenoble, France

## ARTICLE INFO

## Article history:

Received 1 July 2015

Accepted 2 November 2015

Available online xxx

## Keywords:

Aminopeptidase

M42

TET

Intracellular proteolysis

Large molecular assemblies

## ABSTRACT

The TET peptidases are large polypeptide destruction machines present among prokaryotes. They form 12-subunits hollow tetrahedral particles, and belong to the family of M42 metallo-peptidases. Structural characterization of various archaeal and bacterial complexes has revealed a unique mechanism of internal compartmentalization and peptide trafficking that distinguishes them from the other oligomeric peptidases. Different versions of the TET complex often co-exist in the cytosol of microorganisms. In depth enzymatic studies have revealed that they are non-processive cobalt-activated aminopeptidases and display contrasting substrate specificities based on the properties of the catalytic chambers. Recent studies have shed light on the assembly mechanism of homo and hetero-dodecameric TET complexes and shown that the activity of TET aminopeptidase towards polypeptides is coupled with its assembly process. These findings suggested a functional regulation based on oligomerization control *in vivo*. This review describes a current knowledge on M42 TET peptidases biochemistry and discuss their possible physiological roles. This article is a part of the Special Issue entitled: «*A potpourri of proteases and inhibitors: from molecular toolboxes to signalling scissors*»

© 2015 Published by Elsevier B.V.

## Contents

1. Introduction .....	00
2. Crystallographic structures of M42 peptidases forming TET dodecamers .....	00
3. The TET dodecamer: a unique type of quaternary structure .....	00
4. TET mode of action .....	00
5. Enzymatic properties .....	00
6. Functions and substrate specificities of the different TET complexes .....	00
7. Effect of metal ions on TET catalysis and oligomeric state .....	00
8. Assembling the TET system: a highly ordered process .....	00
9. Biological significance of oligomerization in TET peptidases .....	00
10. Physiological roles of the TET peptidases .....	00
11. Concluding remarks .....	00
Acknowledgements .....	00
References .....	00

\* Corresponding author. Univ. Grenoble Alpes, Institut de Biologie Structurale (IBS), F-38027 Grenoble, France.

E-mail address: [franzetti@ibs.fr](mailto:franzetti@ibs.fr) (B. Franzetti).

<sup>1</sup> These authors have equally contributed to this work.

## 1. Introduction

In prokaryotes, the ATP-independent peptidases represent a large family of cytosolic enzymes involved in a broad array of

cellular processes. The trimming of the N- and C-termini of polypeptidic chains is involved in the final steps of intracellular protein degradation, beyond the action of energy dependent proteases such as the 20S proteasome or the Clp system [1,2]. This action generates biologically active peptides such as antigenic peptides [3]. The combined action of the peptidases is also involved in the processing of exogenous peptides for energy purposes and for the metabolism of amino acids [4]. The proteins and associated endogenous peptides are extremely diversified through post-translational modifications that modify their biological activities and represent a regulatory mechanism for specific cellular functions and in microbial pathogenicity [5]. The importance of the peptidome composition in the regulation of specific cellular functions has recently been acknowledged [6]. The role of intracellular peptidases as regulators of the peptidome composition is not well understood. Therefore, the functional diversity of peptidases and the regulation of their activities represents an important field of research. Among the numerous intracellular energy-independent peptidases that co-exist in the cytosol, only a few have the capacity to self-assemble as large complexes. The peptidase activities in these systems are confined to inner cavities, accessible exclusively to unfolded polypeptides [7]. These complexes are potential regulators in peptide metabolism. Initially most of the complexes have been described in Archaea due to the high stability of thermophilic macromolecular edifices that facilitated structural studies. These include leucine aminopeptidases, bleomycin hydrolase, DppA, pab87, protease 1, Tricorn protease and the TET aminopeptidase [8–14]. The present review focuses on TET peptidases. They are bi-metallic metallopeptidases belonging to the M42 family according to the MEROPS peptidase classification [15]. The first TET complex was purified in its native state from *Haloarcula marismortui* cellular extracts [14]. The halophilic archaeal particle formed hollow quaternary structures with a peculiar tetrahedral shape [14]. The protein was named TET with reference to the tricorn protease called TRI. The latter, also identified in Archaea, possesses the shape of a tricorn hat [16]. Homologous protein from *Pyrococcus horikoshii*, a hyperthermophilic archaeon was also found to be able to fold as a 12-ssu tetrahedron. Three different TET complexes were found to co-exist in *Thermococcus* archaeal species and subsequently named as TET1, TET2 and TET3 [17–20]. Recent studies have described a mechanism of their controlled assembly and have shown that oligomerization represents a way to enhance the efficiency of polypeptide processing and also a way to assemble multicatalytic peptidosomes [21–23]. At the same time, several X-ray structures of M42 peptidases have revealed the existence of TET-like complexes in bacteria [24]. Comparison with their archaeal edifices has revealed common features that indicate a high degree of conservation between Archaeal and Bacterial TET systems. This suggests that TET systems play an important biological role. Functional studies have also been conducted mainly in the three TET systems from *P. horikoshii* to identify their substrate specificities and mode of action [14,17,19–29]. These works are reviewed in the chapters below.

## 2. Crystallographic structures of M42 peptidases forming TET dodecamers

The different TET monomers have a molecular weight of about 40 kDa. To date 5 crystallographic structures of M42 proteins assembled as dodecamers have been published: 3 from the hyperthermophilic archaeon *P. horikoshii* TET1, TET2 and TET3 (2WYR, 1YOR and 2WZN), 1 in *Desulfurococcus kamchatkensis* (4WVW) and 1 from the bacteria *Streptococcus pneumoniae* (3KL9) [17,19,24,30,31]. While these proteins display 30–40% similarity between each other, they share a common fold. Crystallographic

structures of the archaeal PhTET3 and Bacterial Pep A monomers are presented in Fig. 1. The proteins are arranged in two well-separated domains: a catalytic domain and a dimerization domain. In all M42 peptidases the catalytic domain displays a conserved backbone which consists of a three-layer  $\alpha\beta\alpha$  sandwich arranged around a central, mixed eight-stranded  $\beta$ -sheet. This structure is consistent with the canonical structure of M42 peptidases such as BsYsdC (PDB 1VHE). A mixed four-stranded  $\beta$ -sheet motif is present at the periphery of the catalytic domain and interacts with the dimerization domain of a second subunit to form the dimer.

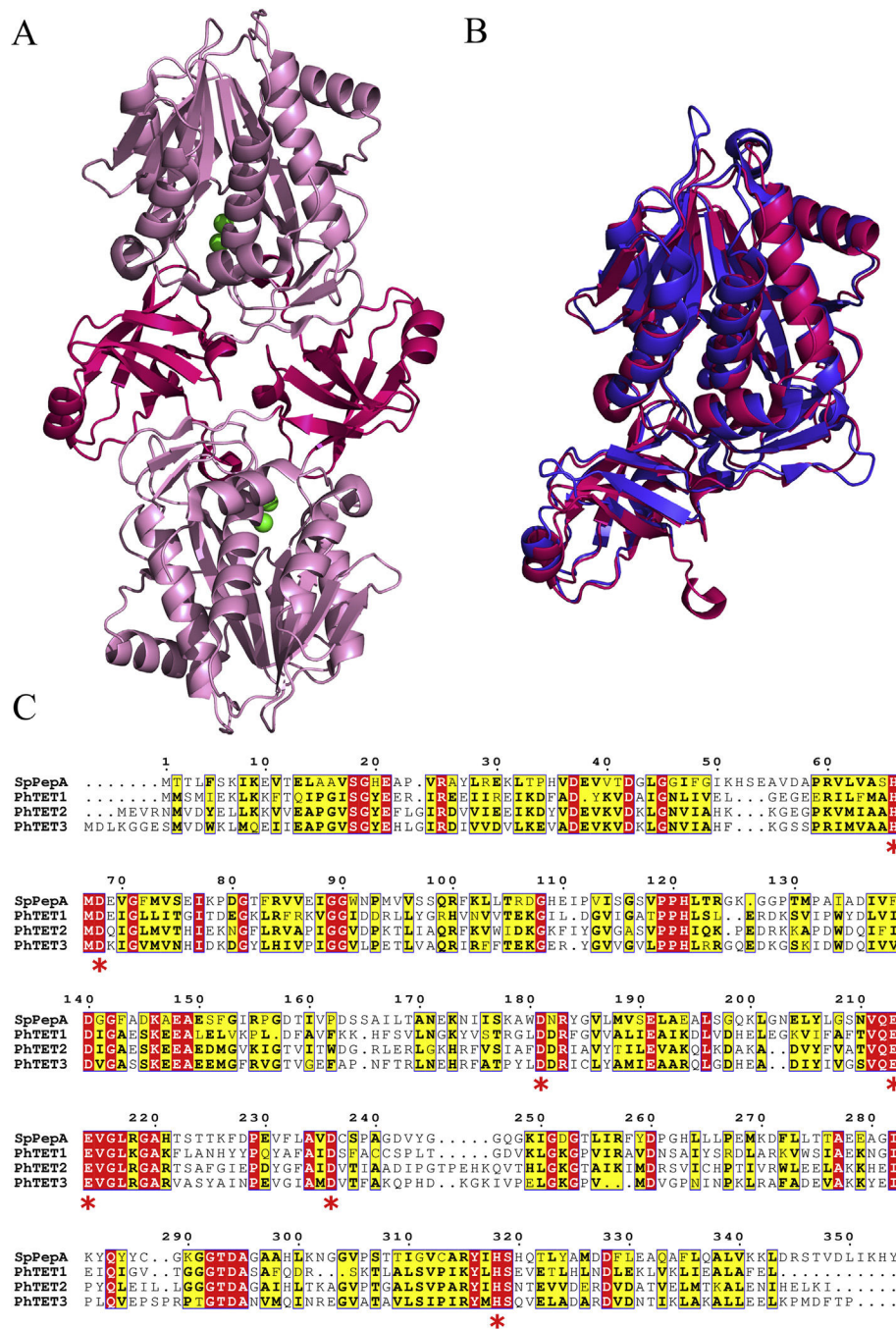
The catalytic sites of the M42 TET peptidases contain a dinuclear metal centre. One metal ion is located close to the entry of the active site, while the second metal ion is located in the vicinity of the specificity pocket that is responsible for substrate recognition. The metal ligand coordination of TET is well conserved amongst the M42 peptidases. In the case of PhTETs it is identical to the one of *Aeromonas proteolytica* aminopeptidase 1 (*Ap1*) indicating a similar catalytic mechanism of peptide bound hydrolysis [32] [33]. In TET peptidases crystal structures, the dinuclear metal site was modelled with two zinc ions or two cobalt ions based on the crystallization conditions employed. Although the  $\text{Co}^{2+}$  ions have been shown to activate TET aminopeptidases, the molecular basis of such behaviour has not been elucidated yet.

The fold of the catalytic domain of archaeal and bacterial TET is highly similar to the ones of AAP and *Streptomyces griseus* aminopeptidase, two peptidases from family M28 [34]. All the bi-metallic peptidases from this family are monomeric and are secreted by the cells. These peptidases, as well as some M42 peptidases such as the one of *Thermotoga maritima*, lack the dimerization domain that is present in all M42 peptidases forming dodecameric edifices. The dimerization interface is mediated by the two PDZ-like domains that form the dimeric building block as shown in Fig. 2. The dimerization domain is made up of a six-stranded antiparallel  $\beta$ -barrel, completed by two helical segments on the outside of the barrel. Due to a large interaction surface involving an extensive hydrogen bonds and putative salt bridges network, the dimers appear to be the building blocks in all TET peptidases.

## 3. The TET dodecamer: a unique type of quaternary structure

Most of the cytosolic compartmentalized peptidases adopt a barrel-shaped architecture. On the contrary, the TET aminopeptidases share a unique dodecameric edifice with a typical tetrahedral shape that is conserved between archaea and bacteria (Fig. 3) [17,19,24,30,31]. Moreover, while in other oligomeric peptidases, the active sites are usually lined up alongside a single central channel; the TET active sites are located in vast, funnel-shaped chambers enclosed in the apices of the particles [17]. These chambers represent catalytic compartments each containing three active sites arranged in a circular fashion.

In the quaternary arrangements of the TET complexes, structures that were determined, the dimeric building blocks form the edges of the tetrahedrons (Fig 4 A). On the tetrahedron faces, large openings (~20 Å) are formed by the junction of six subunits (Fig. 4B). The TET particles therefore possess 4 orifices that are considered to be the substrates' entry pores. On each tetrahedron vertex, three catalytic domains arising from 3 different subunits are arranged as trimers in a circular fashion. The junction found in each trimer delimits a small pore that is believed to represent the product exit way (Fig. 4 C). This small orifice is delimited by the hydroxyl groups of aromatic residues situated at the border of the hole of ~7 Å diameter. The pore has been described in most of the TET particle, except PhTET3 [19]. However, in this latter, there are two possible conformations for the Tyrosine groups, one of them



**Fig. 1.** Crystallographic structures of the archaeal PhTET3 (PDB code: 2WZN) and bacterial Pep A (PDB code: 3KL9). A) Crystallographic structure of the PhTET3 dimer. The catalytic domain of the aminopeptidase is represented in light pink, the dimerization domain is represented in hot pink, and the catalytic ions are represented as green spheres. B) Superposition of the crystallographic structures of the PhTET3 (hot pink) and PepA (deep blue) monomers. C) Sequence alignment of bacterial PepA and archaeal PhTET1, PhTET2 and PhTET3. The red asterisks represent the active site residues.

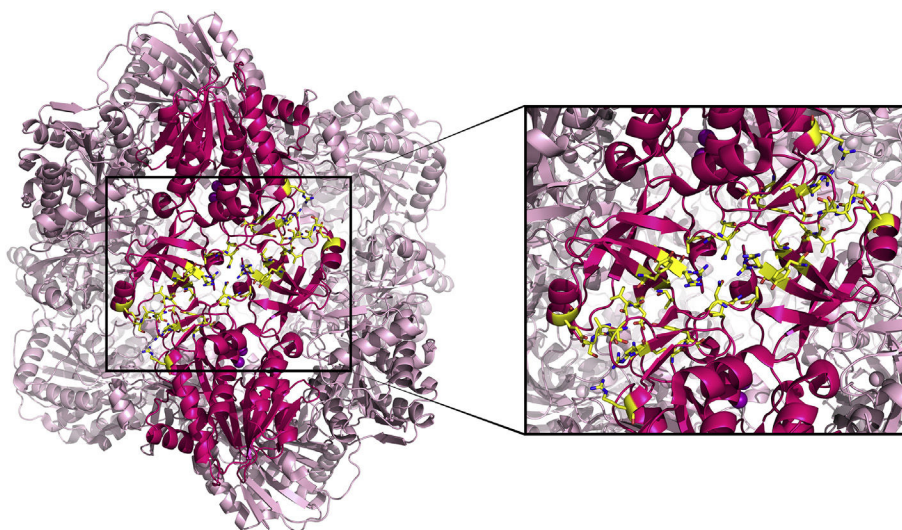
unmasking the hole.

#### 4. TET mode of action

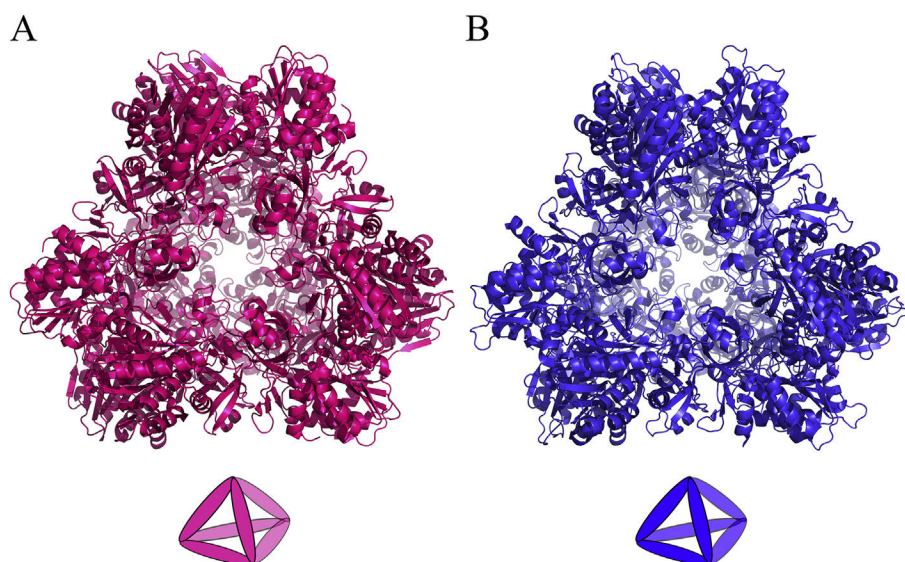
The mechanism of peptide trafficking in exopeptidases is poorly understood. In the case of TET, a model for peptide and product trafficking can be proposed, based on existing TET structures and on studies combining electron microscopy with X-ray crystallography (Fig. 5 A and B). In the case of PhTET1, the examination of the particle interior revealed a complex compartmentalization [17,19].

The size of openings situated on each face of the tetrahedron limits the access of the catalytic sites to unstructured polypeptides. The openings give access to four channels. In all TET, they exhibit a contrasted electrostatic charge distribution that is involved in the substrate peptide navigation and orientation toward the catalytic chambers [17,19,24,30]. A series of mobile loops would then allow addressing the peptides N-termini toward one of the four catalytic chambers that are mainly positively charged. Electrostatic repulsion would then orientate the positively charged N-termini toward one of the negatively charged active site pockets. The four funnel-





**Fig. 2.** The dimerization interface of PhTET3. A dimer of PhTET3 inside the dodecameric complex is represented in hot pink colour. The dimerization interface is highlighted in yellow, the residues involving the interactions between the monomers are represented as sticks. The interaction network involves 17 residues per monomer, there are 28 potential hydrogen bonds and 20 potential salt bridges.



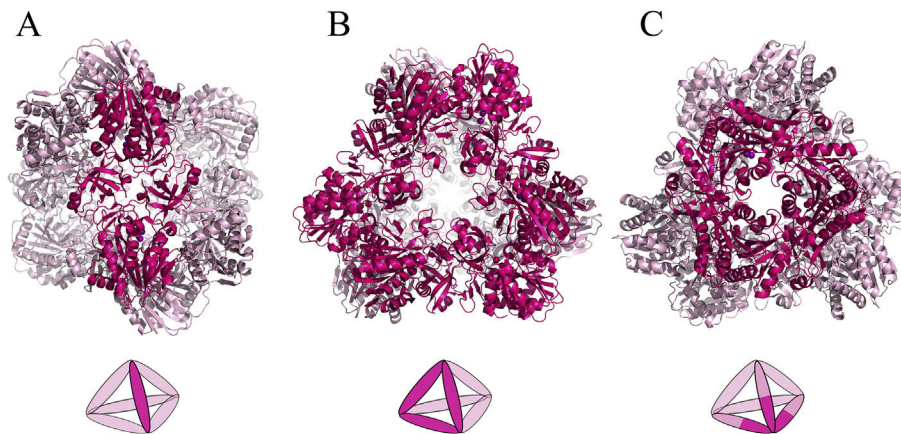
**Fig. 3.** Crystallographic structures of the PhTET3 (A) and PepA (B) aminopeptidases dodecameric edifices. Schematic representations of the position of the dimeric building blocks within the tetrahedral complexes are presented below each structure.

shaped catalytic chambers open to the exterior through narrow holes situated at each apex, close to the 3 bi-metallic active sites. These pores are not wide enough to allow the easy passage of an extended peptide. Accordingly, these orifices appear well adapted for the expulsion of the free amino acids generated by the aminopeptidase reaction [17]. Another mechanism has been proposed by Borissenko and Groll in which the small pores located around the PhTET2 facet's large opening, and that are also connected to the active site, would represent product exit systems [30]. However these orifices seem to be absent in the other TET structures. The non-processive nature of TET enzymes is consistent with a dynamical mechanism coupling the release of the cleaved substrates with the opening of the small pores located at the apices of the tetrahedrons to expel the release free aminoacids. This vectorial model is supported by the conservation of charge properties and specific residues involved in the proposed mechanisms in the

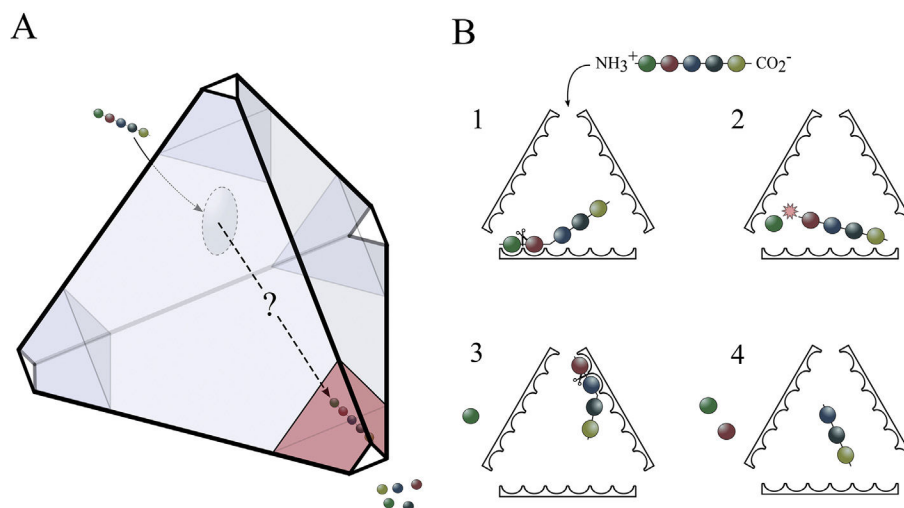
different available TET structures. There is however a lack of experimental data to fully validate this model. It is also not known whether or not the TET system can process several polypeptides at the same time.

### 5. Enzymatic properties

Exopeptidases activities depend on the length of the substrate and on its sequence composition. In some cases exopeptidases can combine aminopeptidases activities with endopeptidases or di/tripeptididyl peptidases activities [35–40]. The modes of action of the different TET particles were studied in depth in the case of the archaeal TET systems from *P. horikoshii* and *H. marismortui* [14,17,19–22,25,41]. A broad array of standard chromogenic and fluorogenic peptides corresponding to endopeptidases, aminopeptidases, carboxypeptidases and peptide deblocking activities



**Fig. 4.** Quaternary arrangement of the TET complexes. A schematic representation of each view of the crystallographic structure of the tetrahedral complex is depicted under each view of the tetrahedron. A) A dimeric building block, represented in hot pink, form each edge of the tetrahedron. B) View of the tetrahedron from one of the face of the complex, each face is formed by 6 subunits arranged around a large pore. C) View of the tetrahedron from one the vertex of the complex. Each vertex is formed by three subunits forming a catalytic pocket.



**Fig. 5.** Model depicting the mode of action and the non-processing nature of TET dodecameric complexes. Only one of 4 peptide-processing channels is represented. (A) Peptide trafficking model within the TET tetrahedron. Pores situated in each facets of the edifice capture the substrate peptide. The polypeptide is then guided across the edifice toward one of the 4 large catalytic chambers situated at each apex of the tetrahedron (represented in Red). Inside the particle, the peptide trafficking system involves contrasted electrostatic charge distribution and mobiles loops (not represented here). (B) Non-processive peptide hydrolysis process. 1) The N-terminal end of the substrate peptide is directed toward one of the 3 active sites present in the catalytic chamber and positioned in the active site pocket. 2) The N-ter amino acid is cleaved by a dinuclear metal centre. After each cleavage, the reaction product is expelled through a small pore situated at each apex of the edifice and the polypeptides N-termini are released from the active sites. 3) The new peptide substrate is then redirected toward one of the 3 neighbouring active sites. 4) The process continues until complete destruction of the polypeptide. In the case of hetero-oligomeric TET complex containing active sites with different substrate specificities, this allows to optimize the enzymatic efficiency toward complex and long peptides.

were tested. In addition, the kinetics of peptide degradations were studied by reverse phase HPLC using synthetic peptides of different sizes and compositions. It was found that TET could process polypeptides that are up to 40 aa in length, yielding free aminoacids. The *P. horikoshii* enzymes PhTET1 seem to be unable to cleave a peptidic bond when pNA or AMC is in the P1' position, making activity assays difficult [17]. In this case, the quantity of aminoacids generated can be measured according to the method of Roth (1971), modified to minimize sample dilution and increase sensitivity [42]. Alternatively, the release of free aminoacids can also be measured by the test developed by Frotin et al. in mesophilic conditions. Ando et al. identified a deblocking activity for the TET1 homologue of *Pyrococcus furiosus* (PfTET1) [43]. In the case of PhTET1, as well as the two other TET enzymes from *P. horikoshii*, no deblocking activity could be measured against the various N-terminal blocked

derivatives tested. Functional studies were performed with a few bacterial M42 peptidases forming TET edifices using reduced sets of mono-acyl chromogenic substrates confirming the aminopeptidase activity of these edifices in bacteria [24,27–29]. However it is important to underline that often these studies were performed with his-tagged proteins that were purified by nickel affinity, which, regarding the importance of metal ions for the activation and oligomeric state of the TET particle may affect results. The analysis of the different PhTETs actions on complex substrates showed that the levels of the released amino acids do not increase in parallel and that the ratio of their concentrations changes with time meaning that TET do not degrade peptides in a processive manner [17,19,20] Taken together, these studies indicated that the archaeal and bacterial TET peptidases degrade oligopeptides in a sequential manner and displayed strict aminopeptidase behaviour.

## 6. Functions and substrate specificities of the different TET complexes

Surprisingly, the functional characterizations of the different TET complexes revealed important differences from one cell type to the other and between the different versions of the enzymatic complexes that co-exist in a single cell [17,19,20,24,27–29,44]. However a hallmark for these systems is the weak cleavage activities or even the absence of activity toward aromatic and branched residues. In Bacteria, two M42 peptidases from *Clostridium thermocellum* and *T. maritima*, previously mis-annotated as cellulases, were assigned as M42 TET peptidases based on sequence homology and X-ray structure (3ISX) [28]. These enzymes are leucyl aminopeptidases. The PepA enzyme from *S. pneumoniae* displays Glutamyl-Aspartyl aminopeptidase activity [24]. In the pathogenic bacteria *Mycoplasma hyopneumoniae* a M42 protein, MHJ\_0125, was assigned as a TET complex by three-dimensional modelling using the structure of PhTET2 as a template [27]. In this case, the protein was found to be a glutamyl aminopeptidase, surprisingly devoid of activity against aspartic acid. Amongst the different TET peptidases characterized so far, the TET complex from the halophilic archaeon *H. marismortui* was found to have the broader specificities with a preference for neutral and basic residues [14]. Prokaryotic genomes often contain several M42 peptidases. In Archaea, the thermococcales contain 3 different versions of the TET peptidase. In *P. horikoshii*, PhTET1 is a Glutamyl-Aspartyl aminopeptidase apparently only raised against peculiar N-ter sequence composition [17,45]. The two other TET complex, PhTET2 and PhTET3, display broader specificities while still displaying clear specificities: PhTET2 is a leucyl aminopeptidase with a broad activity against neutral aminoacids [20] while PhTET3 is a Lysyl aminopeptidase with a clear preference for positively charged residues [19]. Because of their complementary specificities it was proposed that the different TET enzymes from *Pyrococcus* represent an integrated peptide degradation system in the cytosol. Recently it was shown that, after co-expression in *Escherichia coli*, the recombinant PhTET2 and PhTET3 proteins can form heterooligomeric assemblies [22]. The same particles were detected in *Pyrococcus* cells with a dominant iso-stoichiometric composition. Therefore, *in vivo*, PhTET2 and PhTET3 are in fact two subunits of a single enzymatic complex named PhTET2-3. The PhTET2-3 structure and its thermal stability were studied by small angle scattering and electron microscopy [22,23]. It showed that the heterooligomeric complex is a *bona fide* TET dodecamer. The biochemical activity of the PhTET2-3 complex was studied on chosen peptides containing neutral (PhTET2 substrates) and positively charged (PhTET3 substrates) aminoacids. These studies showed that the PhTET2-3 complex is more efficient than an equivalent mixture of PhTET2 and PhTET3 homomers. Moreover, the efficiency of the PhTET2-3 complex was assayed on complex peptides mixtures. It revealed that the hetero-oligomers not only process peptides more efficiently but can even break down peptides that remained intact in the presence of a mixture of PhTET2 and PhTET3 homooligomers [22]. Thus, the PhTET2-3 complex can be defined a multi-subunit peptidosome.

These studies indicated that TET peptidase are highly diverse in terms of substrate specificity and represent sophisticated machineries. When functional assignments exist together with atomic structures, the comparison of binding pocket properties and the modelling of substrates support a clear correlation between charges distribution and substrate specificities. In the case of enzymes for which the amidolytic activity is specialized in the elimination of negatively charged residues from peptides N-termini such as PhTET1, PepA or MHJ\_0125, a positive charge is conferred to the S1 pocket and the residues involved in the opening of the small

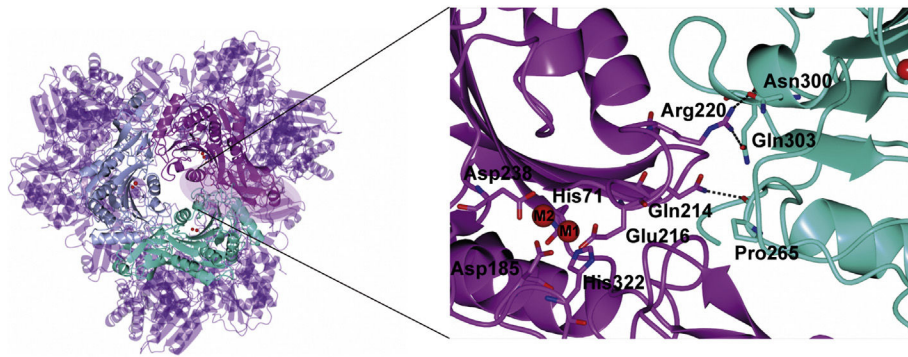
product exit pore are dominated by positively charged residues [17,24,27]. The opposite was found in the case of the lysyl-aminopeptidase PhTET3 with a negative electrostatic charge and a larger size of the pocket compatible with large and mainly positively charged substrate [19]. In leucyl-aminopeptidases with a broad specificity toward neutral or small hydrophobic residues, such as PhTET2 the specificity can be attributed to a smaller and hydrophobic specificity pocket [30]. Thus, despite highly conserved tertiary and quaternary structure, the TET complexes can host active sites possessing very different specificities.

## 7. Effect of metal ions on TET catalysis and oligomeric state

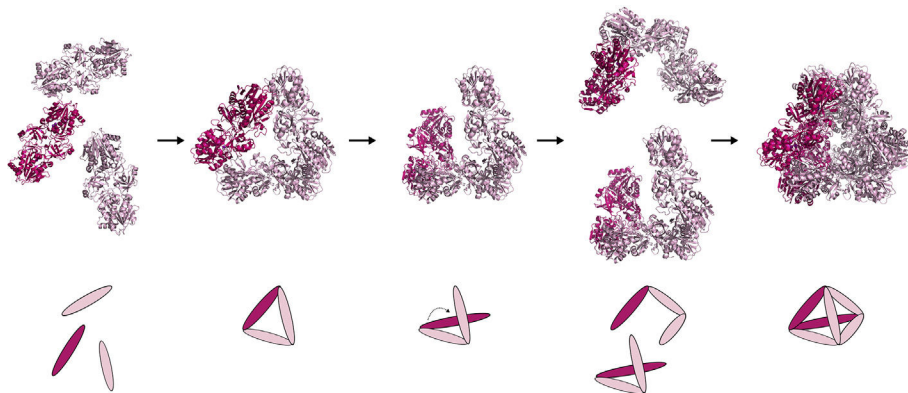
In each TET aminopeptidase there are two metal ions that are responsible for the enzymatic activity. The nature of the metal occupying the dinuclear metal site has been shown to vary the TET enzymatic rate [14,17,19–29].  $\text{Co}^{2+}$  appears the best catalyser for the archaeal PhTET1, PhTET2, PhTET3 and for the bacterial TET (SpPepA). In particular, the cobalt dependence for activity is more pronounced for PhTET1 compared to other TET-like enzymes that have been studied so far [45]. Based on nomenclature of Schechter and Berger site 1 metal ion (M1) is closest to the mouth of the active site while site 2 metal ion (M2) is relatively buried and next to the specificity binding pocket (P1) that is devoted to substrate recognition [46]. As mentioned above, the metals in the active site are close to the TET oligomeric interface and it has been shown that chelating agents induce the dissociation of the dodecameric particle towards lower oligomerization states [25]. In particular the M1 site is coordinated by a triade Glu/His/Asp, in which Glu lay on a loop hosting conserved residues (Arg220, Gln214 in PhTET3) among archaeal and bacterial TET aminopeptidases, that are implicated in a network of hydrogen bonds (Fig 6). Thus, the removal of the M1 site may destabilize the formation of the oligomeric interface leading to the TET dodecamer dissociation. An interesting difference among archaeal hyperthermophilic TETs and the mesophilic SpPepA is the response to  $\text{Zn}^{2+}$  addition. While the archaeal TET aggregate, SpPepA resulted slightly activated [24]. This is particularly relevant by remembering that archaeal and mesophilic dinuclear metal sites have the same protein ligands.

## 8. Assembling the TET system: a highly ordered process

Oligomeric enzymatic complexes involving at least 6 subunits represent more than 10% of the cellular proteomes [47]. Understanding the mechanisms underlying the self-assembling of subunits into highly organized structures as well as the evolutionary events responsible for the emergence of large molecular edifices represent an important challenge [48]. With their unique tetrahedral symmetry and sophisticated compartmentalization, the TET peptidases represent remarkable model edifices. The assembling of TET dimers into dodecamers involve principally subunit interactions located at the edge of the tetrahedrons, where the 3 proteolytic domains are held together forming the apexes of the edifice. An in depth analysis of the inter-subunits contacts present in different types of TET particles led to the design of mutant proteins in which the natural assembling sequence were slowed down. In the case of PhTET2, this strategy allowed to dissect the whole assembling process. The kinetics of intermediate oligomers species were determined and their structural properties were determined by combining various biophysical methods in solution with X-ray crystallography and electron microscopy [21]. It showed that the TET system does not result from the incremental addition of dimer but from a highly controlled process involving an obligatory



**Fig. 6.** Role of catalytic metals in TET assembling. On the left, the dodecameric assembly of PhTET3 (PDB: 2WZN). Three monomers are highlighted in ice blue, cyan and magenta, forming the proposed exit hole of the processed amino acids. Each highlighted monomer interacts through the oligomeric interface. On the right, a detail of the H-bond network occurring in the oligomeric interface involving Gln214 and Arg220 of one monomer and Asn300, Gln303 and Pro265 of the other subunit. Noteworthy, Gln214 and Arg220 lay on the same loop hosting Glu216 that is one of the ligand of the M1 site. The red spheres are the metal ions in the active site.



**Fig. 7.** Oligomerization pathway of the TET aminopeptidases. The building block of the assembly is a dimer of TET; 3 dimers assemble themselves to form a triangular hexamer in which the interactions at the vertices are not stable enough to remain close. Therefore, the triangular hexamer opens itself to form a so-called Z-form of the hexamer. Two Z-form hexamers interact to form a dodecameric TET complex. This assembling process is presented as Supplementary Data in Movie 1.

hexameric intermediate. The TET assembling process is presented in Fig. 7 and in Supplementary Movie 1. This model was experimentally validated by using small angle neutron scattering experiments that were performed on the hetero-oligomeric PhTET2-3 complex [22,23]. In this elegant approach, deuterated and hydrogenated dimers were mixed *in vitro* and the reconstituted TET edifice was purified. SANS data were collected in various buffers containing different H<sub>2</sub>O/D<sub>2</sub>O ratios. Contrast variation allowed to determine the internal topology of the PhTET2 and PhTET3 subunits within the PhTET2-3 dodecamers. The experiments revealed the existence of a limited number of conformers all involving the opening of the hexameric intermediates to adopt a Z-conformation in the final complex. Interestingly, this mechanism results in the formation of multicatalytic chambers at the apices of the tetrahedrons. The TET enzymes being non-processive, the polypeptides N-termini are released from the active sites after each cleavage. The co-localization of active sites possessing complementary specificities represents a great advantage and explains the observed efficiency of the PhTET2-3 complex compared to mixtures of homomeric complexes. Thus, in the case of the TET system, the subunits' interface seems to have evolved not only to control the TET assembly but also to favour the best possible internal topology of the respective subunits to promote for broad and efficient polypeptide destruction.

Supplementary data related to this article can be found online at <http://dx.doi.org/10.1016/j.biochi.2015.11.001>.

## 9. Biological significance of oligomerization in TET peptidases

The purified recombinant TET assemblies are stable molecular edifices. The different TET edifices from *P. horikoshii* are amongst the most robust natural nanoparticles *in vitro*. They show half-lives of about 8 h when incubated at 80 °C. The PhTET1, PhTET3 and PhTET2 edifices were studied by small angle scattering in extreme temperature, pHs and pressure conditions without showing significant signs of quaternary structures destabilization or interparticular subunits exchanges [17,19–21,23,26]. In a paradoxical way, it was noticed that, *in vivo*, a large part of the TET proteins are accumulated as stable dimers [21,22]. This was observed for the 3 TETs proteins of *P. horikoshii* as well as for the TET protein of the halophilic archaea *Haloarcula salinarum*. The catalytic property of the PhTET2 dimer was studied and compared to the one of the dodecamer. The dimer retains aminopeptidase activity on di or tripeptides but displayed a significant loss of activity as the size of the polypeptide increase [21]. Important differences in reaction speeds were observed and the  $K_m$  and  $k_{cat}$  measurements indicated that the TET reaction efficiency toward long peptides is compromised when the enzyme is not assembled as dodecamers. Oligomerization and active site compartmentalization serve different functions in cytosolic proteases. In the 20S proteasome, bleomycin hydrolase or tricorin protease, the sequestration of active sites into the particle prevents unwanted attacks on folded polypeptides in the cytosol [49–52]. In Protease 1 or Pab87, oligomerization is

requested to build up the active sites [11,12]. In bovine lens leucyl aminopeptidases it stabilizes the edifices while in tryptase or kallikrein-related peptidases it provides a way to inhibit the enzyme activity [8,38,53,54]. In the case of TET peptidase, the purpose of oligomerization is to trigger the activity of the system toward larger substrates. Since the dimeric precursors are stabilized *in vivo* it has been suggested that a functional regulation operates on the dimer–dodecamer equilibrium, *in vivo*, in response to varying physiological demands.

## 10. Physiological roles of the TET peptidases

The sequential elimination of the amino-terminal residues from polypeptidic chains by aminopeptidases fulfils a broad range of biological functions ranging from energy metabolism to the regulation of specific physiological processes [4,35,55]. Aminopeptidases substrates can be diverse. They are peptides obtained from the environment or produced by the breakdown of proteins by ATP dependent machines such as the proteasomes. In some cases, post-translational or co-translational N-terminal modifications of proteins or biogenic peptides trigger their activity or control their half-life in the cytosol. Aminopeptidases are therefore essential for housekeeping functions such as proteins turnover and amino acid homeostasis and to control specific cellular processes.

The distribution of TET peptidases in prokaryotic genomes is difficult to establish because members of the M42 aminopeptidases are often misannotated as endoglucanases or cellulases. This highlights the importance of experimental biochemical studies to refine the functional annotation of these enzymes. Genomic studies indicated however that M42 enzymes possessing a TET-like structure are widely present in prokaryotic genomes, highly diversified and that the existence of multiple copies of the TET system is not uncommon even in bacteria (in preparation). A close examination of the TET proteins in the archaeal kingdom also reveals that the number of TET could be particularly high in the genome of deep-sea organisms. This suggests a link between the TET physiological role and the lifestyle of marine extremophiles that remains to be clarified. It has also been reported that the TET and TRI systems seem to mutually exclude in most genomes, although there are some exceptions, thus suggesting that they represent two evolutionarily divergent degradation systems involved in the same function. However, given the conservation of the TET systems and their diversity in Bacteria and Archaea it is likely that the TET system is involved in important physiological functions. It has been suggested that the broad aminopeptidase activity of PhTET2 display housekeeping functions by trimming the peptides generated by the 20S proteasome endopeptidase activity [30]. However there are numerous other exopeptidases in the cytosol of *Pyrococcus* that could also fulfil this function. The broad activity of the hetero-oligomeric PhTET2-3 complex is also compatible with an unspecific degradation of exogenous or endogenous peptides for metabolic or energetic purpose. The production of amino acids can be utilized for processes such as protein synthesis, generation of metabolic energy, and recycling of reduced cofactors. On the other hand, the regulation of the dimer–dodecamer equilibrium within cells is consistent with a regulatory role of TET peptidases in the cells [21]. Moreover, the selective specificity of the PhTET1 archaeal system and some bacterial TET displaying Glutamyl or Leucine aminopeptidase activities may be indicative for a specific role of in the degradation or maturation of a special class of polypeptides or proteins that remains to be determined. In this respect it was reported that the PhTET1 complex is not assembled in the physiological conditions in which the PhTET2-3 complex is dodecameric and therefore activated. In bacteria, the aminopeptidases play important roles in pathogenesis not only by providing aminoacids

from exogeneous proteins but also by modifying host proteins and by the destruction of immunological antibacterial peptides. The Glutamyl aminopeptidases found in the pathogenic bacteria such as SpPepA may be involved in these processes [24]. Gene expression and proteomic studies as well as the analyses of gene knock-out mutant strains would be *required* to elucidate the physiological role(s) of the different types of TET peptidases.

## 11. Concluding remarks

Since the discovery of the TET complexes in Archaea, important progress has been made in determining their structures, activities, substrates' specificities and assembling process. However, we still do not fully understand the polypeptide trafficking processes as well as the functional dynamics within the TET edifices. Moreover, many different types of TET systems present in bacterial and archaeal genomes still await structural and mechanistic studies. Regarding the position of the TET systems in the proteolytic pathways and cell physiology it seems that we have just begun to scratch the surface in terms of an understanding of their regulation, natural substrate specificities and high genomic diversity.

## Acknowledgements

We acknowledge financial support from the Agence Nationale de la Recherche (Grant "Archelyse" ANR-12-BSV8-0019-01).

## References

- [1] T. Yao, R.E. Cohen, Giant proteases: beyond the proteasome, *Curr. Biol.* 9 (1999) R551–R553.
- [2] L. Saveanu, D. Fruci, P. van Endert, Beyond the proteasome: trimming, degradation and generation of MHC class I ligands by auxiliary proteases, *Mol. Immunol.* 39 (2002) 203–215.
- [3] M. Weimershaus, I. Evnouchidou, L. Saveanu, P. van Endert, Peptidases trimming MHC class I ligands, *Curr. Opin. Immunol.* 25 (2013) 90–96.
- [4] T. Gonzales, J. Robert-Baudouy, Bacterial aminopeptidases: properties and functions, *FEMS Microbiol. Rev.* 18 (1996) 319–344.
- [5] J.A. Cain, N. Solis, S.J. Cordwell, Beyond gene expression: the impact of protein post-translational modifications in bacteria, *J. Proteomics* 97 (2014) 265–286.
- [6] Z.W. Lai, A. Petrer, O. Schilling, The emerging role of the peptidome in biomarker discovery and degradome profiling, *Biol. Chem.* 396 (3) (2014).
- [7] A. Lupas, J.M. Flanagan, T. Tamura, W. Baumeister, Self-compartmentalizing proteases, *Trends Biochem. Sci.* 22 (1997) 399–404.
- [8] S.K. Burley, P.R. David, A. Taylor, W.N. Lipscomb, Molecular structure of leucine aminopeptidase at 2.7-Å resolution, *Proc. Natl. Acad. Sci. U S A* 87 (1990) 6878–6882.
- [9] L. Joshua-Tor, H.E. Xu, S.A. Johnston, D.C. Rees, Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6, *Science* 269 (1995) 945–950.
- [10] H. Remaut, C. Bompard-Gilles, C. Goffin, J.M. Frere, J. Van Beeumen, Structure of the *Bacillus subtilis* D-aminopeptidase DppA reveals a novel self-compartmentalizing protease, *Nat. Struct. Biol.* 8 (2001) 674–678.
- [11] V. Delfosse, E. Girard, C. Birck, M. Delmarcelle, M. Delarue, O. Poch, P. Schultz, C. Mayer, Structure of the archaeal pab87 peptidase reveals a novel self-compartmentalizing protease family, *PLoS One* 4 (2009) e4712.
- [12] X. Du, I.G. Choi, R. Kim, W. Wang, J. Jancarik, H. Yokota, S.H. Kim, Crystal structure of an intracellular protease from *Pyrococcus horikoshii* at 2-Å resolution, *Proc. Natl. Acad. Sci. U S A* 97 (2000) 14079–14084.
- [13] T. Tamura, N. Tamura, Z. Cejka, R. Hegerl, F. Lottspeich, W. Baumeister, Tricorn protease—the core of a modular proteolytic system, *Science* 274 (1996) 1385–1389.
- [14] B. Franzetti, G. Schoehn, J.F. Hernandez, M. Jaquinod, R.W. Ruigrok, G. Zaccari, Tetrahedral aminopeptidase: a novel large protease complex from archaea, *Embo J.* 21 (2002) 2132–2138.
- [15] N.D. Rawlings, A.J. Barrett, A. Bateman, MEROPS: the database of proteolytic enzymes, their substrates and inhibitors, *Nucleic Acids Res.* 40 (2012) D343–D350.
- [16] T. Tamura, N. Tamura, Z. Cejka, R. Hegerl, F. Lottspeich, W. Baumeister, Tricorn protease—the core of a modular proteolytic system, *Science* 274 (1996) 1385–1389.
- [17] G. Schoehn, F.M. Vellieux, M. Asuncion Dura, V. Receveur-Brechot, C.M. Fabry, R.W. Ruigrok, C. Ebel, A. Roussel, B. Franzetti, An archaeal peptidase assembles into two different quaternary structures: a tetrahedron and a giant octahedron, *J. Biol. Chem.* 281 (2006) 36327–36337.
- [18] S. Russo, U. Baumann, Crystal structure of a dodecameric tetrahedral-shaped

- aminopeptidase, *J. Biol. Chem.* 279 (2004) 51275–51281.
- [19] M.A. Dura, E. Rosenbaum, A. Larabi, F. Gabel, F.M. Vellieux, B. Franzetti, The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea, *Mol. Microbiol.* 72 (2009) 26–40.
- [20] M.A. Dura, V. Receveur-Brechot, J.P. Andrieu, C. Ebel, G. Schoehn, A. Roussel, B. Franzetti, Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus horikoshii*, *Biochemistry* 44 (2005) 3477–3486.
- [21] A. Appolaire, E. Rosenbaum, M.A. Dura, M. Colombo, V. Marty, M.N. Savoye, A. Godfroy, G. Schoehn, E. Girard, F. Gabel, B. Franzetti, *Pyrococcus horikoshii* TET2 peptidase assembling process and associated functional regulation, *J. Biol. Chem.* 288 (2013) 22542–22554.
- [22] A. Appolaire, M.A. Dura, M. Ferruit, J.P. Andrieu, A. Godfroy, S. Gribaldo, B. Franzetti, The TET2 and TET3 aminopeptidases from *Pyrococcus horikoshii* form a hetero-subunit peptidosome with enhanced peptide destruction properties, *Mol. Microbiol.* 94 (2014) 803–814.
- [23] A. Appolaire, E. Girard, M. Colombo, M.A. Dura, M. Moulin, M. Hartlein, B. Franzetti, F. Gabel, Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70 (2014) 2983–2993.
- [24] D. Kim, B.H. San, S.H. Moh, H. Park, D.Y. Kim, S. Lee, K.K. Kim, Structural basis for the substrate specificity of PepA from *Streptococcus pneumoniae*, a dodecameric tetrahedral protease, *Biochem. Biophys. Res. Commun.* 391 (2010) 431–436.
- [25] E. Rosenbaum, M. Ferruit, M.A. Dura, B. Franzetti, 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, *Biochim. Biophys. Acta* 1814 (2011) 1289–1294.
- [26] E. Rosenbaum, F. Gabel, M.A. Dura, S. Finet, C. Clery-Barraud, P. Masson, B. Franzetti, Effects of hydrostatic pressure on the quaternary structure and enzymatic activity of a large peptidase complex from *Pyrococcus horikoshii*, *Arch. Biochem. Biophys.* 517 (2012) 104–110.
- [27] M.W. Robinson, K.A. Buchtman, C. Jenkins, J.L. Tacchi, B.B. Raymond, J. To, P. Roy Chowdhury, L.K. Woolley, M. Labbate, L. Turnbull, C.B. Whitchurch, M.P. Padula, S.P. Djordjevic, MHJ\_0125 is an M42 glutamyl aminopeptidase that moonlights as a multifunctional adhesin on the surface of *Mycoplasma hyopneumoniae*, *Open Biol.* 3 (2013) 130017.
- [28] R. Dutoit, N. Brandt, C. Legrain, C. Bauvois, Functional characterization of two M42 aminopeptidases erroneously annotated as cellulases, *PLoS One* 7 (2012) e50639.
- [29] Y. Kumaki, M. Ogawa, T. Hirano, K. Yoshikawa, N. Iwasawa, T. Yagi, W. Hakamata, T. Oku, T. Nishio, Family M42 aminopeptidase from the syntrophic bacterium *Symbiobacterium thermophilum*: characterization using recombinant protein, *J. Biosci. Bioeng.* 111 (2011) 134–139.
- [30] L. Borissenko, M. Groll, Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes, *J. Mol. Biol.* 346 (2005) 1207–1219.
- [31] T.E. Petrova, E.S. Slutskaya, K.M. Boyko, O.S. Sokolova, T.V. Rikitina, D.A. Korzhenevskiy, M.A. Gorbacheva, E.Y. Bezsudnova, V.O. Popov, Structure of the dodecamer of the aminopeptidase APDKam598 from the archaeon *Desulfurococcus kamchatkensis*, *Acta Crystallogr. Sect. F. Struct. Biol. Commun.* 71 (2015) 277–285.
- [32] B. Chevrier, C. Schalk, H. D'Orchymont, J.M. Rondeau, D. Moras, C. Tarnus, Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family, *Structure* 2 (1994) 283–291.
- [33] C. Stamper, B. Bennett, T. Edwards, R.C. Holz, D. Ringe, G. Petsko, Inhibition of the aminopeptidase from *Aeromonas proteolytica* by L-leucinophosphonic acid. Spectroscopic and crystallographic characterization of the transition state of peptide hydrolysis, *Biochemistry* 40 (2001) 7035–7046.
- [34] H.M. Greenblatt, O. Almog, B. Maras, A. Spungin-Bialik, D. Barra, S. Blumberg, G. Shoham, *Streptomyces griseus* aminopeptidase: X-ray crystallographic structure at 1.75 Å resolution, *J. Mol. Biol.* 265 (1997) 620–636.
- [35] A. Taylor, Aminopeptidases: towards a mechanism of action, *Trends Biochem. Sci.* 18 (1993) 167–171.
- [36] H. Singh, G. Kalnitsky, alpha-N-benzoylarginine-beta-naphthylamide hydrolase, an aminoendopeptidase from rabbit lung, *J. Biol. Chem.* 255 (1980) 369–374.
- [37] R.P. Koldamova, I.M. Lefterov, M.T. DiSabella, J.S. Lazo, An evolutionarily conserved cysteine protease, human bleomycin hydrolase, binds to the human homologue of ubiquitin-conjugating enzyme 9, *Mol. Pharmacol.* 54 (1998) 954–961.
- [38] E. Geier, G. Pfeifer, M. Wilm, M. Lucchiari-Hartz, W. Baumeister, K. Eichmann, G. Niedermann, A giant protease with potential to substitute for some functions of the proteasome, *Science* 283 (1999) 978–981.
- [39] D. Turk, V. Janjic, I. Stern, M. Podobnik, D. Lamba, S.W. Dahl, C. Lauritzen, J. Pedersen, V. Turk, B. Turk, Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases, *EMBO J.* 20 (2001) 6570–6582.
- [40] D. Chandu, D. Nandi, PepN is the major aminopeptidase in *Escherichia coli*: insights on substrate specificity and role during sodium-salicylate-induced stress, *Microbiology* 149 (2003) 3437–3447.
- [41] S. Porciero, V. Receveur-Brechot, K. Mori, B. Franzetti, A. Roussel, Expression, purification, crystallization and preliminary crystallographic analysis of a deblocking amino peptidase from *Pyrococcus horikoshii*, *Acta Crystallogr.* 61 (2005) 239–242.
- [42] M. Roth, Fluorescence reaction for amino acids, *Anal. Chem.* 43 (1971) 880–882.
- [43] S. Ando, K. Ishikawa, H. Ishida, Y. Kawarabayasi, H. Kikuchi, Y. Kosugi, Thermostable aminopeptidase from *Pyrococcus horikoshii*, *FEBS Lett.* 447 (1999) 25–28.
- [44] E.S. Slutskaya, E.Y. Bezsudnova, A.V. Mardanov, V.M. Gumerov, T.V. Rikitina, V.O. Popov, V.M. Lipkin, Characterization of a novel M42 aminopeptidase from crenarchaeon *Desulfurococcus kamchatkensis*, *Doklady Biochem. Biophys.* 442 (2012) 30–32.
- [45] B. Franzetti, M.A. Durà, PhTET1 Aminopeptidase in, in: Neil D. Rawlings, Guy S. Salvesen (Eds.), *Handbook of Proteolytic Enzymes*, Academic Press, Oxford, 2013, pp. 1638–1645.
- [46] I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun.* 27 (1967) 157–162.
- [47] J.M. Matthews, M. Sunde, Dimers, oligomers, everywhere, *Adv. Exp. Med. Biol.* 747 (2012) 1–18.
- [48] N.J. Marianayagam, M. Sunde, J.M. Matthews, The power of two: protein dimerization in biology, *Trends Biochem. Sci.* 29 (2004) 618–625.
- [49] M.W. Thompson, J. Miller, M.R. Maurizi, E. Kempner, Importance of heptameric ring integrity for activity of *Escherichia coli* ClpP, *Eur. J. Biochem.* 258 (1998) 923–928.
- [50] M. Groll, L. Ditzel, J. Lowe, D. Stock, M. Bochtler, H.D. Bartunik, R. Huber, Structure of 20S proteasome from yeast at 2.4 Å resolution, *Nature* 386 (1997) 463–471.
- [51] H. Brandstetter, J.S. Kim, M. Groll, R. Huber, Crystal structure of the tricorn protease reveals a protein disassembly line, *Nature* 414 (2001) 466–470.
- [52] P.A. O'Farrell, F. Gonzalez, W. Zheng, S.A. Johnston, L. Joshua-Tor, Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease, *Structure* 7 (1999) 619–627.
- [53] U. Marquardt, F. Zettl, R. Huber, W. Bode, C. Sommerhoff, The crystal structure of human alpha1-tryptase reveals a blocked substrate-binding region, *J. Mol. Biol.* 321 (2002) 491–502.
- [54] M. Debela, V. Magdolen, V. Grimminger, C. Sommerhoff, A. Messerschmidt, R. Huber, R. Friedrich, W. Bode, P. Goettig, Crystal structures of human tissue kallikrein 4: activity modulation by a specific zinc binding site, *J. Mol. Biol.* 362 (2006) 1094–1107.
- [55] M. Matsui, J.H. Fowler, L.L. Walling, Leucine aminopeptidases: diversity in structure and function, *Biol. Chem.* 387 (2006) 1535–1544.

### 5. RÉSULTATS

Dans cette partie, je présente les résultats obtenus durant mes trois années de thèse, sur les aminopeptidases tétraédriques TETs isolées de deux archées hyperthermophiles abyssales : *Methanocaldococcus jannaschii* et *Pyrococcus horikoshii*. Une diversité de techniques ont été exploitées afin de caractériser les deux protéines MjTET et PhTET4, allant de l'optimisation de leur expression et purification, en passant par les études enzymatiques pour terminer par les investigations structurales entreprises par cristallographie. Les résultats obtenus sont exposés dans deux chapitres comprenant chacun une courte introduction, un manuscrit en préparation pour publication et une discussion incluant des résultats ou des hypothèses non présentées dans l'article. Le premier chapitre est consacré à la caractérisation enzymatique et structurale du complexe MjTET de *M. jannaschii* et le second aux études enzymatiques effectuées sur PhTET4 isolée à partir de *P. horikoshii*.

#### 5.1. CHAPITRE 1 : CARACTÉRISATIONS FONCTIONNELLES ET STRUCTURALES DE L'UNIQUE COMPLEXE TET ARCHÉEN DE *METHANOCALDOCOCCUS JANNASCHII*

##### 5.1.1. Présentation du travail

Les études phylogénétiques réalisées par S. Gribaldo à l'Institut Pasteur, ont révélé que les aminopeptidases TETs exprimées par les archées possèdent des origines différentes. Pour résumer, ce travail qui reste à publier montre que le système TET est une machinerie cellulaire très ancienne, déjà présente dans l'ancêtre des archées et des bactéries et peut-être même dans LUCA (*Last Universal Common Ancestor*). Il montre également que certaines archées comme celles appartenant au genre des *Thermococcus* dont fait partie *Pyrococcus horikoshii*, possèdent de multiples peptidases TETs qui sont issues de plusieurs transferts horizontaux de gènes provenant de différentes espèces de bactéries. Ces transferts sont associés à une perte du système TET archéen d'origine. Ce sont à titre d'exemple les *Methanococcales* qui possèdent de « vraies » TETs archéennes, généralement présentes en une seule copie dans la cellule, contrairement aux *Nanohaloarchaea* qui en possèdent deux exemplaires. Parmi ces « vraies » TETs archéennes, nous nous sommes intéressés à l'étude des TETs exprimées par les *Methanococcales*, en particulier l'unique TET de l'hyperthermophile *Methanocaldococcus jannaschii*, nommée ultérieurement MjTET.

## RÉSULTATS

---

L'étude entreprise sur MjTET répond à plusieurs motivations. Tout d'abord, il s'agissait de mieux comprendre le rôle physiologique fondamental du système TET à travers l'étude des propriétés enzymatiques et structurales d'une TET archéenne et unique. Il s'agissait également de savoir si le système archéen est plutôt destiné à la destruction des substrats spécifiques ou au contraire s'il représente un système de destruction à large spectre d'action, plus compatible avec une fonction de ménage. Ces propriétés ne sont pas aisément prédictibles à partir de la séquence peptidique. Il s'agissait corollairement de comprendre l'intérêt physiologique associé à l'acquisition de gènes issus de TET bactérienne par les archées et de déterminer ainsi le (ou les) avantage(s) acquis par certaines espèces d'archées après avoir substitué leur enzyme TET d'origine par plusieurs autres enzymes TETs d'origine bactérienne.

D'après le système de classification MEROPS, MjTET exprimée par *M. jannaschii* est une métallo-aminopeptidase de la famille M42 du clan MH. L'alignement de séquence de cette protéine avec les protéines décrites dans la littérature et appartenant à la même famille révèle une similarité de séquence de ~45 % avec PhTET2, la leucyl aminopeptidase TET de *P. horikoshii*. La protéine recombinante correspondante MjTET a pu être purifiée avec un bon degré de pureté en utilisant un protocole inspiré de ceux utilisés pour les TETs de *P. horikoshii* et optimisé pour MjTET. Celui-ci comprend une étape de choc thermique et deux chromatographies. La chromatographie d'exclusion de taille et une étude en microscopie électronique par coloration négative ont montré que MjTET présentait bien une organisation tétraédrique similaire à celle adoptée par les TETs. Afin de déterminer ses substrats, j'ai réalisé des études enzymatiques en utilisant une collection de substrats chromophores (acides aminés couplés au pNA) et fluorophores (couplés à AMC) qui ont indiqué que MjTET possède une activité aminopeptidase préférentielle des résidus neutres hydrophobes, notamment la leucine, mais que son action s'étend à beaucoup d'acides aminés, comme les résidus aromatiques tels que la phénylalanine et la tyrosine, une propriété absente chez les TETs de la famille M42 (MEROPS). Pour confirmer cette observation, j'ai réalisé des tests de dégradation sur des peptides de différentes compositions en analysant leur profil de dégradation par HPLC. Ces travaux ont bien confirmé que MjTET est une aminopeptidase très peu sélective comparativement aux autres TETs caractérisées à ce jour, avec une propriété améliorée vis-à-vis des résidus aromatiques. Pour le reste, la caractérisation des conditions optimales d'activité de MjTET révèle une capacité à fonctionner à haute température, sur une large gamme de pH, et avec le cobalt comme métal activateur ce qui est



## RÉSULTATS

---

assez proche des 3 TETs (PhTET1, TET2 et TET3) d'origine bactérienne identifiées chez *P. horikoshii*.

En se référant aux travaux publiés sur PhTET1, PhTET3 et à l'étude réalisée dans cette thèse sur PhTET4 de *Pyrococcus horikoshii*, les TETs des Thermococcales provenant des bactéries apparaissent très spécialisées dans l'hydrolyse de certains types d'acides aminés. Au contraire, les résultats obtenus pour la TET archéenne de *M. jannaschii* vont clairement montrer que nous avons à faire à une TET « non spécialisée ».

Plusieurs hypothèses ont été formulées par d'autres groupes et nous mêmes quant au mécanisme de capture et de navigation des peptides et des produits réactionnels dans la machinerie TET. Le système TET archéen de *M. jannaschii* étant capable de dégrader une large gamme de peptides, nous pouvons alors penser que son mode d'action fondamental sera plus facile à décrypter que dans des systèmes spécialisés. Ceci nous a conduit à en déterminer la structure cristallographique. Pour cela, j'ai dû ajouter une étiquette polyhistidine du côté N-terminal de la protéine et j'ai changé de protocole de purification. En effet, la forme non taguée n'a pas permis une genèse de cristaux ordonnés. Avec cette nouvelle forme, les cristaux ont diffracté à 2,4 Å permettant ainsi la résolution de la structure de MjTET. En raison du problème rencontré avec la cristallisation de la forme non taguée, nous avons eu recours à la cryo-microscopie électronique, en collaboration avec l'équipe de G. Schoehn à l'IBS, grâce à laquelle une structure à 4 Å a pu être obtenue. Bien qu'elle soit à basse résolution, cette nouvelle structure a apporté des informations précieuses et complémentaires aux données cristallographiques. En effet, ces dernières ne renfermaient pas la densité électronique correspondant à la boucle (115-124) portée par le domaine de dimérisation I, tout comme les autres TETs archéennes dont les structures cristallographiques ne montrent pas de densité électronique pour cette boucle. Conséquemment grâce à la carte de cryo-EM, j'ai pu construire les résidus de la boucle dans la structure cristallographique de MjTET. La flexibilité de cette boucle, son organisation à l'intérieur de la particule et son orientation vers le site actif du monomère adjacent dans un même dimère, nous ont permis de proposer son implication dans l'orientation du substrat vers l'un des sites actifs de l'un des trois apex situés aux alentours du pore d'entrée. Ces informations nous permettent de proposer un nouveau modèle pour expliquer le mode d'action des peptidases TETs.

### **5.1.2. Crystallographic and cryo-EM studies of an archaeal TET peptidase complex from *Methanocaldococcus jannaschii* reveal internal network of loops responsible for peptide processing**

### **Crystallographic and cryo-EM studies of an archaeal TET peptidase complex from *Methanocaldococcus jannaschii* reveal internal network of loops responsible for peptide processing**

Hind Basbous<sup>1</sup>, Gregory Effantin<sup>1</sup>, Guy Schoehn<sup>1</sup>, Eric Girard<sup>1</sup> \* and Bruno Franzetti<sup>1</sup> \*

<sup>1</sup> Université Grenoble Alpes, CNRS, Commissariat à l'Énergie Atomique, Institut de Biologie Structurale F-38044 Grenoble, France

\* Corresponding authors: [bruno.franzetti@ibs.fr](mailto:bruno.franzetti@ibs.fr) and [eric.girard@ibs.fr](mailto:eric.girard@ibs.fr)

#### **ABSTRACT**

Amino peptidases TET of M42 family are ~500 kDa hollow dodecameric particles present in the 3 kingdoms of life. The autotrophic archaeon *Methanocaldococcus jannaschii* contains single copy of TET metallopeptidase. The purified recombinant protein forms stable tetrahedral dodecamers and was therefore named MjTET. The substrate specificity of MjTET was studied using a range of amino-acyl and peptide substrates. MjTET is a thermophilic cobalt-activated enzyme and is the first reported leucine amino peptidase to demonstrate broad amino acid substrate specificity including aromatic residues. The crystal structure of MjTET was determined at 2.4 Å of resolution, revealing that MjTET monomers and dodecamers are built similarly to PhTET2 peptidase from *Pyrococcus horikoshii*. However, differences in the active site dimension and composition explain the extended MjTET amidolytic properties. Cryo-EM structure of MjTET dodecamer was obtained at resolution of 4.1 Å. It provided information on conserved mobile loops located inside TET cavities, allowing building the missing parts of the X-ray structure. The analysis of this complete atomic model indicated new routes for polypeptide trafficking within TET particles. In this new model, the four tetrahedron apexes could process simultaneously several substrates while avoiding peptide congestion inside the particle.

#### **INTRODUCTION**

Amino peptidases represent large class of enzymes that hydrolyze peptide bonds between amino acids in protein or peptide chain. They are involved in broad array of cellular functions, from protein destruction to post-translational modifications of proteins and peptides (1,2). During the last decade, it has been found that several enzymes from M42 and M18 metallopeptidase families self-assembled in half megadalton complexes made of twelve

## RÉSULTATS

---

subunits (3). They were named TET because of the tetrahedral shape adopted by these dodecameric edifices. TET complex was first purified from halophilic archaea during a search for protein quality control machines involved in low salt stress (4). This novel type of cage-forming protease attracted interest since it is present in many cell types and has homologues in all kingdoms of life. Previous structural studies revealed that all TET dodecamers share similar architectures: the twelve monomers are associated in a tetrahedral particle with four large substrate entry pores formed by the junction of six subunits situated on each facet and four catalytic chambers each comprising three active sites of three monomers which formed the apex of tetrahedron (3). This architecture is strikingly different from the barrel-shaped one adopted by other cytosolic compartmentalized peptidases. TET assembling mechanism from dimeric precursors was determined using a combination of biophysical techniques. They demonstrated that it is a highly ordered process involving predominantly hexameric precursors (5-7). Biochemical characterizations of TET complex *in vivo* and *in vitro* indicated that the purpose of self-compartmentalization in the M42 peptidases is to trigger their amidolytic activity toward long peptides (7). Different types of TET enzymes were found in archaea, bacteria and eukarya and different versions of enzymatic complexes often co-exist in the cytosol of the same organism (8,9) (10) (11) (12) (13,14) (15) (16). All M42 TET complexes characterized so far are cobalt-activated aminopeptidases able to degrade oligopeptides to single amino acids. However, enzymatic studies revealed that they displayed marked substrate preferences. This aspect of TET biochemistry was mostly studied in *Pyrococcus horikoshii*, a deep-sea hyperthermophilic Euryarchaeota that contains three different versions of TET complex: PhTET1, a glutamyl-aspartyl aminopeptidase, PhTET2, a leucyl aminopeptidase with broad activity against neutral amino acids, and PhTET3, a lysyl aminopeptidase (13,14) (15) (16). In these complexes, the comparison of the “S1” binding pocket properties supports a clear correlation between charge distribution and substrate specificities.

Despite important progress in determining the structure-function relationship of TET machinery, many fundamental questions remain about the cellular function and the mode of action of TET system. In particular, we do not fully understand the detailed mechanism of polypeptide substrate filtering and guidance to the active site as well as the amino acid-product expulsion system remains elusive. A genomic survey on the occurrence of TET-like enzymes in prokaryotic genomes revealed that some archaea possess up to four versions of TET proteins while others only express one version. In order to specify the basic

physiological roles of TET system and the reasons underlying the multiplicity of TET cellular machines in some microorganisms and to bring further insights into their structure-function relationship, we studied the biochemical and structural properties of the unique TET complex from *Methanocaldococcus jannaschii*, a deep-sea hyperthermophilic archaea (17). The enzymatic characterization of MjTET revealed that this archaeal TET aminopeptidase displayed broad substrate specificity, thus indicating a housekeeping function devoted to the non-specific destruction of polypeptides. Structural analysis combining X-ray crystallography and cryo-electron microscopy allowed providing complete MjTET atomic model. In this model, internal network of mobile loops that were poorly defined in previous crystallographic studies could be observed. The topological analysis of these loops highlights an elegant peptide guiding system toward catalytic sites.

### EXPERIMENTAL PROCEDURES

#### *Protein expression and purification*

Genes encoding for MjTET and its tagged form MjTET-Nt (with N-terminal His<sub>6</sub> tag) were synthesized and cloned by GeneCust Europe in pET-41c and pET28a+ vectors, respectively. We used *Escherichia coli* ArcticExpress (DE3)-RIL strain (Agilent Technologies) for the over-expression of these two recombinant proteins. The cells were grown in 1 L of lysogeny broth (LB) medium containing gentamycin and kanamycine at 50 µg/ml. Then the cells were incubated with shaking at 37 °C until the A<sub>600</sub> reached 0.6-1.0, at this stage the induction was carried out by adding isopropyl β-D-thiogalactopyranoside at final concentration of 0.1 mM and proteins are overexpressed during 7 h at 16 °C. The induced cells were harvested by centrifugation, and the pellets were then conserved at -80 °C until use. These pellets were suspended in 50 ml of 50 mM Tris-HCl, 200 mM NaCl and 0,1 % Triton X-100, pH 7.5, supplemented with 12.5 mg of lysozyme (Euromedex), 2.5 mg of DNase I grade II, 10 mg of RNase, 50 mg of Pefabloc SC (Roche), 500 µl of 2 M MgSO<sub>4</sub> and 50 mM imidazole for cells expressing MjTET-Nt protein. The cell disruption was achieved by sonication in a Branson sonifier at 30 watts for 10 cycles on/off of 30 sec each, at 4 °C. The crude extracts were heated at 85 °C for 20 min and later clarified by centrifugation (JA20 rotor) at 12000 rpm for 1 h at 4 °C. The supernatant was diluted two times with 20 mM Tris-HCl pH 7.5, and loaded on anion exchange column (Resource Q) (GE Healthcare Life Sciences) for MjTET recombinant form. The protein was eluted with a linear NaCl gradient from 190-340 mM in 20 mM Tris-HCl pH 7.5. While MjTET-Nt purification was performed on HiTrap Chelating

## RÉSULTATS

---

HP column (GE Healthcare Life Sciences) with 200 and 300 mM imidazole in 20 mM Tris-HCl, 200 mM NaCl, pH 7.5. Exclusion chromatography for the two protein forms was done on Superose 6 column or Superdex 200 column (GE Healthcare Life Sciences) previously equilibrated with 20 mM Tris-HCl, 200 mM NaCl, pH 7.5. The fractions containing protein (38/40 kDa) according to SDS-PAGE were pooled and concentrated using Amicon cell (Millipore) with a molecular mass cutoff of 30 kDa and kept at 4 °C. 3 and 21.5 mg of MjTET and MjTET-Nt were produced from 1 L of culture, respectively.

### *Negative Electron Microscopy*

MjTET samples, at approximately 0.05 mg/mL in 20 mM Tris-HCl and 200 mM NaCl, pH 7.5, were adsorbed to the clean side of a carbon film on mica, stained with 2 % sodium silicotungstate pH 7.5 and transferred to a 400-mesh copper grid. The images were taken under low dose conditions ( $<10 \text{ e}^-/\text{\AA}^2$ ) at a magnification of 23 and 49 kX times with defocus values between 1.2 and 2.5  $\mu\text{m}$  on a Tecnai 12 LaB6 electron microscope at 120 kV accelerating voltage using CCD Camera Gatan Orius 1000.

### *Effect of pH on MjTET activity*

The effect of pH on MjTET activity was determined in pH range varying from 6 to 11. The buffers used were: PIPES, pH 6–7.5; CHES, pH 8.2–10 and CAPS, pH 10.5–11. All buffers were used at 50 mM concentration in presence of 150 mM KCl, 0.1 mM  $\text{CoCl}_2$ , 4  $\mu\text{g/mL}$  of MjTET and 5 mM Leu-pNA as substrate. The given pH values are at 75 °C. Incubation was done at 75 °C for 5 min and the quantity of produced pNA was evaluated by measuring the absorbance at 405 nm.

### *Effect of metal cations on MjTET activity*

The metal cations  $\text{XCl}_2$  were incubated at 0.1 mM concentration with 4  $\mu\text{g/mL}$  of MjTET and 5 mM of Leu-pNA in 50 mM CHES, 150 mM KCl, pH 8.5 for 5 min at 75 °C. To assess the effect of metal cations on MjTET activity, the enzymatic reaction was followed as described before.

### *Effect of temperature on MjTET activity and stability*

Different temperatures ranging from 20 to 90 °C were tested on MjTET activity. Incubation was done in 50 mM CHES, 150 mM KCl, 0.1 mM  $\text{CoCl}_2$ , pH 8.5 in presence of 4  $\mu\text{g/mL}$  of protein and 5 mM Leu-pNA for 5 min. Absorbance of released pNA was measured at 405 nm.

## RÉSULTATS

---

The effect of temperature on MjTET stability was assayed by incubating the enzyme (160 µg/ml) at 85 °C. Aliquot of 10 µl was taken at different time intervals, and the remaining aminopeptidase activity was assayed against 5 mM Leu-pNA in 390 µl of 50 mM CHES, 150 mM KCl, 0.1 mM CoCl<sub>2</sub>, pH 8.5 incubated at 85 °C for 5 min. Enzymatic stability of MjTET was evaluated by measuring the quantity of produced pNA at 405 nm.

### *Effect of inhibitors on MjTET activity*

The inhibitory effect of EDTA, Amastatine and Bestatine was tested on the MjTET enzymatic activity. 4 µg/mL of MjTET were incubated with each inhibitor separately in the presence of 5 mM Leu-pNA, 50 mM CHES, 150 mM KCl, 0.1 mM CoCl<sub>2</sub>, pH 8.5 at 85 °C for 5 min. Three replicates (with inhibitor) and two controls (without inhibitor) were assayed for each experimental point.

### *Determination of MjTET activity on both synthetic chromogenic, fluorogenic and peptide compounds*

The hydrolytic activity of MjTET on synthetic chromogenic and fluorogenic compounds was determined by using different aminoacyl-pNAs and aminoacyl-AMCs through the following procedure. Reactions were initiated by addition of 4 µg/mL of enzyme to a pre-warmed mixture containing 5 mM of the chromogenic or fluorogenic compound in 50 mM CHES, 150 mM KCl and 0.1 mM CoCl<sub>2</sub>, pH 8.2, in a total volume of 400 µl. Incubation was performed at 75 °C for 5 min. The quantity of liberated pNA was evaluated by measuring the absorbance at  $\lambda=405$  nm in a JASCO V-630 spectrophotometer, and for AMC by measuring the fluorescence using excitation and emission wavelengths of 360 and 440 nm, respectively in a JASCO FP-8500 spectrofluorometer. Three replicates and two enzyme blanks were assayed for each experimental point. Additional enzymatic studies were performed in order to investigate MjTET capacities to hydrolyze longer substrates (H-Trp-Phe-Tyr-Ser(PO<sub>3</sub>H<sub>2</sub>)-Pro-Arg-pNA, H-Met-Trp-Ala-Glu-Asn-Lys, H-Tyr-Asp-Thr-Ser-Ile-Val-Gln-Arg, H-Asp-Ala-Tyr-Pro-Ser and H-Glu-Gly-Ile). 2 µg/ml (final concentration) of MjTET were added to a pre-warmed mixture of 0.75 mM peptide, 50 mM CHES, 150 mM KCl, 0.1 mM CoCl<sub>2</sub>, pH 8.5. To avoid water evaporation, 20 µl of mineral oil were added on the top of the total volume. The reaction incubation was done at 85 °C during 6 min. Aliquot of 80 µl was then removed and added to 220 µl of 2 % acetonitrile (ACN), 0.1 % trifluoroacetic acid (TFA). Proteins were removed by centrifugation at 13,000 rpm during 15 min. 100 µl of the supernatant were retrieved and injected on Nova-Pak C18 column, (4 µm, 3.9 x 300 mm,

## RÉSULTATS

---

(Waters)) in a HPLC purifier system (PerkinElmer), equilibrated with 2 % ACN, 0.1% TFA. The elution of peptide products was achieved with a linear ACN gradient (2-33.2 %) and followed by measuring the absorbance at 214 nm. Chromatographic runs were carried out at room temperature.

### *MjTET-Nt crystallization, data collection, structure determination and refinement*

MjTET-Nt was purified as described above and concentrated up to 5.6 mg/ml using an Amicon Ultra 30 kDa cut/off filter. Crystals were grown at 20 °C in Trisodium citrate 0.1 M pH 5.6, Tertbutanol 15 % using the hanging drop vapor diffusion method. MjTET-Nt was mixed with an equal volume of the crystallization solution and crystals grew up to average 80  $\mu\text{m}$  size in two days. The crystals were cryo-protected with Trisodium citrate 0.1 M pH 5.6, Tertbutanol 15 % supplemented with 2-Methyl-2,4-pentanediol 25 %. The crystals were flash cooled in liquid nitrogen prior to data collection at 100 K. Data set was collected at ESRF beamline ID29. The diffraction extended up to 2.4 Å. Data were indexed and integrated by XDS (Kabsch, 2010). Integrated intensities were scaled and merged with SCALA and TRUNCATE from the CCP4 programs suite (19). MjTET-Nt structure was determined by molecular replacement using PHASER MR (20) and PhTET2 as searching model (PDB ID: 1XFO) depleted of side chains. A clear solution was determined with a Translation Function Z-score of 12.4. The asymmetric unit contains four monomers and the space group is cubic  $P2_13$ . The model was manually completed and improved in COOT prior to refinement with PHENIX (1.10-2155 version) software package (Adams et al., 2011). The model was then optimized through iterative rounds of refinement and model building resulting in final R-work= 0.1859 and R-free= 0.2291 (**Table 1**).

During the initial step of refinement, we realized that MjTET-Nt structure was depleted of  $\text{Zn}^{2+}$  in Site1. Consequently, in order to provide description of MjTET active site, MjTET-Nt crystals were soaked with  $\text{Zn}^{2+}$  excess solubilized in the cryo-protection solution used above. Data at 2.69 Å resolution were collected at ESRF beamline ID29. Data processing and refinement were conducted as described above, except that the 2.4 Å resolution MjTET-Nt(-\_Zn1) model was used as reference model to steer refinement of MjTET(Zn1\_Zn2) resulting in final R-work= 0.1755 and R-free= 0.2368 (**Table 1**).

### *Cryo-electron microscopy of MjTET dodecamers*



## RÉSULTATS

---

Four microliters of sample was applied to 2:1 Quantifoil holey carbon grid (Quantifoil Micro Tools GmbH, Germany) and the grid was frozen in liquid ethane with a Vitrobot Mark II (FEI, the Netherlands). The sample was observed with FEI Polara at 300 kV. Images were recorded on K2 summit direct detector (Gatan Inc., USA) in super resolution counting mode. Movies were recorded at a nominal magnification of 23,000 (0.82 Å/pixel at the camera level) for a total exposure of 4s and 100 ms per frame resulting in 40 frame movies with a total dose of  $\sim 40 \text{ e}^-/\text{Å}^2$ . 100 movies were manually recorded with Digital Micrograph (Gatan Inc., USA).

Movies were first motion corrected with unblur (21). Around 2000 particles were picked semi automatically with boxer from the EMAN suite (22) and 2D classified in 20 classes in RELION 1.4 (23). The best-defined 2D class averages were used to generate an initial model using the RICO server imposing tetrahedral symmetry. The resulting low-resolution reconstruction was used as a template for automatic particle picking using the Fast Projection Matching (FPM) method (24) as described in (25). To speed up calculation, the raw micrographs were binned by a factor of 8 (6.56 Å/pixel) and, during template matching, the resolution was limited to 3 nm. For each micrograph, only particles having cross correlation with the reference higher than the mean correlation of all the particles were kept. The output coordinate files from FPM were converted to boxer one (.box) and imported in RELION 1.4 where all subsequent steps were done with two-times binned images (final sampling of 1.64). CTF estimation, particle extraction and preprocessing were done in RELION. The data set was first cleaned by 2D classification. A first 3D auto-refine was then done using the *ab-initio* model obtained with the RICO server imposing tetrahedral symmetry. The obtained 3D model was used as input for 3D classification in five classes. The particles belonging to the classes showing the best resolved features were used to do a final 3D auto-refine. The final reconstruction includes 20000 out of 50000 and has a resolution of 4.1 Å at FSC = 0.143.

## RESULTS

### ***MJ0555 gene encoded for tetrahedral dodecameric protein***

To gain insight into the basic aspects of TET structure and function, we cloned and expressed in *E. coli* MJ0555 gene encoding for the single putative TET peptidase that is present in the genome of *M. jannaschii*. The corresponding recombinant protein was purified with protocols inspired from those developed for the 3 hyperthermophilic TET from *P. horikoshii*. When analyzed by gel filtration chromatography, MJ0555 protein displayed an apparent molecular mass similar to one determined for the homologous TET dodecameric proteins from *P.*

## RÉSULTATS

---

*horikoshii* and *P. furiosus* (13,15,16,26). Negative stain electron microscopy images of MJ0555 samples showed typical hollow-tetrahedral dodecameric particles with dimensions similar to those previously determined for the different M42 TET peptidases and MJ0555 was therefore named MjTET (**Figure 1**).

### ***Influence of pH, metal ions, and temperature on leucine aminopeptidase activity of MjTET***

The sequence alignment of MjTET with other TET peptidases from hyperthermophilic archaea revealed ~45% sequence identity with PhTET2 that was defined as leucine metallo-aminopeptidase (LAP). The optimal conditions for MjTET activity were therefore tested using Leu-pNA. To assess for the metallic dependent behavior of MjTET, we tested several inhibitors of M42 metallopeptidases. Under optimal conditions of MjTET activity (see bellow), adding 5 mM EDTA to the reaction volume completely inhibited the enzyme, thus confirming the metallic dependence of MjTET. Moreover the MjTET LAP activity was completely inhibited by 0.2 mM amastatine and 1 mM bestatine, respectively aminopeptidase and leucyl aminopeptidase inhibitors.

We investigated the effect of pH on MjTET LAP activity. **Figure 2A** shows that MjTET maintained high percentage of activity (between 88 to 100%) over a wide range of pH, from neutral (pH 7) to alkaline (pH 10). Moreover it maintains more than 50% of its activity in acidic and alkaline conditions such as pH 6 and 10.5. Thus as PhTET2, MjTET is an aminopeptidase that can work on broad range of pH conditions (between 6 and 9, for PhTET2) (15). As different divalent metal ions have been found to modulate the activity of aminopeptidases, we tested the effect of several metals on MjTET activity. The data presented on **Figure 2B** showed that the increase of LAP activity occurs only in the presence cobalt ions. By contrast, a significant inhibition was observed when 0.1 mM ZnCl<sub>2</sub> was added to the reaction mixture. Thus, MjTET can be defined as cobalt-activated enzyme.

The temperature dependence of MjTET activity was determined. As shown in **Figure 2C**, LAP activity was found to be maximal at 85 °C, corresponding to the optimal growth temperature of *M. jannaschii* species (17). When incubated at this temperature, MjTET showed a half-life of eight hours. No activity was detected in mesophilic conditions under 50 °C.

### ***MjTET is a leucine aminopeptidase with extended substrate specificity***

A common hallmark to M42 TET aminopeptidases studied so far is their marked preference

## RÉSULTATS

---

for a subset of amino acids. Hence, in *P. horikoshii*, PhTET1 exhibits narrow selectivity toward glutamyl and aspartyl residues while PhTET2 is a leucyl aminopeptidase with broad activity against neutral amino acids and PhTET3 is a lysyl aminopeptidase with clear preference for positively charged residues (14-16). In order to determine MjTET function, its amino acid selectivity was investigated using chromogenic and fluorogenic monoacyl substrates. As shown in **Figure 3**, the enzyme acted on large number of aminopeptidase substrates. The optimal amidolytic activity was observed against Leu-pNA, thus allowing classifying MjTET as leucine aminopeptidase. Interestingly, MjTET also displayed significant amidolytic activities, principally against hydrophobic residues such as phenylalanine, methionine and tyrosine. Reduced activities were also detected in the presence of polar and positively charged residues. Aspartate and glutamate, cysteine and proline were the only residues toward which no activity was detected. Finally, no activity was detected against blocked residues and residues in D-stereo configuration, as shown for Ac-Leu-pNA and D-Leu-pNA substrates, respectively. M42 aminopeptidases displayed poor amidolytic activity against aromatic substrates such as phenylalanine and tyrosine. PhTET2 was found to exhibit weak activity against aromatic amino acids, however, as shown in **Figure 4**, MjTET enzymatic activities appeared to be more efficient notably against phenylalanine. In addition to the broad activity spectrum, this represents another noticeable difference between MjTET and PhTET2.

The exopeptidase activities and specificities can be different in peptide context, compared to the one measured on monoacyl compounds. To test if MjTET maintains broad amino acid substrate specificity as well as the ability to cleave aromatic residues in peptide context, its action was studied on different substrates containing 6 to 8 residues. After incubation with MjTET enzyme, the peptide composition of the reaction mixture was investigated at different time points by reverse phase HPLC. **Figure 5A** shows the hydrolysis profile of the H-Trp-Phe-Tyr-Ser(PO<sub>3</sub>H<sub>2</sub>)-Pro-Arg-pNA peptide. After 5 minutes, the original peptide substrate was almost hydrolyzed and four peptide products were detected, thus confirming that the enzyme displays broad aminopeptidase activity, even on peptides containing aromatic residues. This was confirmed with other substrates: H-Asp-Ala-Tyr-Pro-Ser (**Figure 5B**), H-Tyr-Asp-Thr-Ser-Ile-Val-Gln-Arg, H-Met-Trp-Ala-Glu-Asn-Lys, and H-Glu-Gly-Ile (data not shown). Interestingly, these peptides also contained Asp and Glu residues. Since no activity was measured against Asp-pNA and Glu-pNA, it can be inferred that, these negatively charged amino acids can also be cleaved by MjTET in peptide context, indicating

## RÉSULTATS

---

that C-terminal end is required for better fitting of the P1 amino acid inside “S1” specificity pocket of the enzyme. Taken together these experiments showed that, compared to other M42 aminopeptidases, MjTET is characterized by unusual broad substrate specificity.

### *Topology of MjTET enzyme determined by X-ray crystallography*

Crystal structure of MjTET-Nt was resolved at 2.4 Å by molecular replacement using PhTET2 monomer structure (PDB ID: 1XFO) (27) as search probe. The asymmetric unit of the crystal encloses four monomers, each of ~40 kDa and made up of 366 amino acids. The monomer had typical structure of M42 aminopeptidases forming TET dodecamers. Comparison of the dimerization and oligomerization interfaces of MjTET and PhTET2, by using PISA server (<http://www.ebi.ac.uk/pdbe/pisa/>), revealed that they are very similar and involve the same monomer contacts with nearly the same number of hydrogen bonds and salt bridges.

MjTET monomer can be divided in two  $\alpha/\beta$  globular domains connected by two loops (residues 63-66 and 158-162) (**Figure 6A**). The bigger domain is the catalytic one formed by residues 1-66 and 158-349 and contains the N- and C-terminal ends. The smaller one is the dimerization domain (also named dimerization domain I) formed by residues 67-157. No electron density was observed for His-tag Nt neither for 115-124 residues. The catalytic domain contains central eight-stranded  $\beta$ -sheet surrounded by seven  $\alpha$ -helices and two  $3_{10}$  helices. The two  $\beta$ -strands located on each side of this  $\beta$ -sheet are antiparallel and shorter than the central four  $\beta$ -strands, which are much longer and parallel. Additionally, three-stranded antiparallel  $\beta$ -sheet (residues 162-171 and 325-328) is positioned on the surface of the catalytic domain and can be named dimerization domain II. The dimerization domain I slightly exhibits cylindrical  $\beta$ -barrel shape made up of six antiparallel  $\beta$ -strands and two  $\alpha$ -helices and one  $3_{10}$  helix.

### *MjTET binuclear active sites*

The cleft located between the catalytic and dimerization (I) domains harbors the binuclear active site of the metallo-aminopeptidase. In untreated MjTET-Nt crystals, only one metal ion per monomer of protein was detectable (MjTET-Nt(-\_Zn2)). The second exchangeable metal (M1) was detected after soaking MjTET-Nt crystals with excess of zinc ions during cryoprotection step. In this condition, two  $Zn^{2+}$  ions are modeled in the active site of the enzyme (MjTET-Nt(Zn1\_Zn2)).  $Zn_1^{+2}$  has a trigonal bipyramidal geometry and is coordinated

## RÉSULTATS

---

by Asp175O<sup>δ2</sup>, Glu208O<sup>ε1,ε2</sup>, His321N<sup>ε2</sup> and Citrate O<sup>β2</sup>. However, Zn<sub>2</sub><sup>+2</sup> is tetrahedrally coordinated by His62 N<sup>ε2</sup>, Asp175O<sup>δ1</sup> and Asp230O<sup>δ1,δ2</sup> (**Figure 6B**). The distance between the metal ions is ~3.87 Å comparable to that one of all other M42 TET peptidases and the leucyl aminopeptidase AAP of *Vibrio proteolyticus* (formerly known as *Aeromonas proteolytica*) used as prototype to describe the enzymatic mechanism of peptide bond hydrolysis by binuclear aminopeptidase (28).

### ***Specificity pocket properties explain MjTET broad substrate specificities***

Behind the active site is located the specificity pocket “S1” that was suggested to receive the side chain of the residue that will be cleaved (29). A mainly large and hydrophobic specificity pocket (residues 231-233, 254, 256, 259-262, 264, 289, 291, 293-296, 320) was found in MjTET monomer, which is in agreement with the preference of this aminopeptidase to hydrolyze hydrophobic residues such as leucine and methionine as described above. Interestingly, this pocket is characterized by the presence of glycine residues (Gly256, 262, 264, 291, 293 and 294) that may provide important flexibility to this region, thus allowing the reception of large side chains such as those of aromatic residues especially phenylalanine and tyrosine. Furthermore, negatively charged amino acids, such as Asp259 and Glu289, also present in the “S1” pocket could interact with the positive side chains of histidine, lysine and arginine for potential cleavage by MjTET (**Figure 7**).

### ***Combining crystallographic and cryo-EM data allowed the determination of MjTET complete atomic model***

Previous structural studies showed that, despite their similar architectures, the different types of TET complexes showed differences in their secondary and tertiary structures as well as pores and channels dimensions (3). Noticeable changes were also detected in the surface charge properties of the different TET edifices, in particular with respect to the electrostatic charge distribution in the specificity pocket. These differences were proposed to be responsible for the contrasted substrate specificities between the different TET enzymes and allowed to propose models for peptide trafficking inside the particle (13,14,16,30). However, there is a lack of structural information on the TET particle interior, apparently due to flexible regions that could not be completely resolved by X-ray crystallography, which preclude validating the proposed models. In this work, we combined X-ray crystallography and Cryo-EM experiments to provide better insights into the general internal organization of TET particles.

### *Internal structural organization of MjTET revealed by cryo-EM*

Intriguingly, electron density was usually absent for the loops harbored by the dimerization domain (I) in crystal structures of TET complexes, especially those belonging to M42 family. This observation indicated the mobile behavior of these loops and may suggest an important role in substrate trafficking. In MjTET cryo-EM structure, electron densities of the twelve loops (residues 114-124) are clearly visible allowing us to build the X-ray structure missing parts (**Figure 8A**). In fact, to get the atomic structure of X-ray missed loops, we fitted the X-ray structure of MjTET dimer (building block) in cryo-EM map resolved at 4.1 Å, wherein we built up the internal loop residues (115-124) of the dimers, by using Coot software (31). Then, we generated automatically the complete tetrahedral particle for which we checked the position of all constructed internal loops. The analysis of complete particle structure achieved by combining crystallography and cryo-EM, revealed new features regarding internal organization of tetrahedron. Indeed, we showed here that MjTET entire loops established internal bulk inside the dodecameric macromolecule (**Figure 8B**). MjTET electrostatic potential showed strong positive charge engendered by these internal loops. They appeared to disrupt peptide trafficking inside the particle. Interestingly, these positive patches are located behind negatively charged one created by acidic residues of three building blocks forming the opening pore of the particle (**Figure 9**). These negative charges were described previously to captivate positively charged N-termini of peptide substrate. The location of internal positively charged loops behind the negative patches might avoid peptide navigation from entry pore to the exit hole located on opposite side of the tetrahedron. Furthermore, the organization of the dynamic loops in the same building block, showed that each loop extended near the active site of adjacent monomer, with a peculiar orientation of His116 toward this neighboring site (**Figure 10**). Together, the newly defined positive network that obstructed the internal cavity of MjTET particle and the loop orientation toward active site provide a significantly shorter route from the outside of TET to one of the three catalytic chambers present near the entry hole compared to the alternative route that crosses the central cavity to catalytic chamber situated at the opposite side of the tetrahedron (**Figure 11**).

## DISCUSSION

### *Single TET peptidase from Methanocaldococcus jannaschii is a non-specialized peptidasome*

Genomic survey and functional characterization indicate that in prokaryotic cells the number

## RÉSULTATS

---

of TET proteins can vary from one to up to four different versions. In the deep-sea hyperthermophilic archaeon *P. horikoshii*, three different versions of TET complexes having contrasted substrate preferences have been described (13,15,16,30). According to MEROPS classification, *M. jannaschii*, another hyperthermophilic archaeon isolated from a natural environment similar to the one of *P. horikoshii*, contains only one gene encoding for a putative M42 family TET enzyme of MH clan. In this work, we showed that the corresponding recombinant protein, called here MjTET, is indeed a TET aminopeptidase. The search for its physiological function using monoacyl and peptide substrates revealed the enzyme is a leucine aminopeptidase able to cleave hydrophobic, neutral as well as aromatic and positively charged residues such as phenylalanine, tyrosine, lysine and histidine. MjTET can therefore be qualified as non-specialized peptidase. Unlike *P. horikoshii*, *M. jannaschii* is an autotrophic organism that does not rely on the hydrolysis of external peptide to support its metabolism. Thus, we can speculate that the presence of single TET system, “a priori” less efficient than the multiple specialized TET enzymes found in Thermococcales, could be related to autotrophy.

PhTET2 was also described as leucine aminopeptidase (15). It shares ~45% of sequence identity with MjTET. However, its specificity range is not as wide as the one of MjTET and it is devoid of hydrolytic activity toward N-terminal phenylalanine residues. Therefore, the comparison of the two enzyme 3D structures can help to determine the structural determinants responsible for the remarkable MjTET polyvalence. This analysis showed that the “S1” specificity pocket of MjTET enzyme is slightly larger than PhTET2 one, and is enriched by glycine residues that are not present in other specialized TET enzymes. Compared to PhTET2, MjTET “S1” pocket contains three additional residues and three glycine residues that could confer more flexibility thus allowing accommodating phenylalanine and tyrosine large side chains. Compared to MjTET, the reduce activity of PhTET2 against positively charged substrates could be explained by the presence of Lys261 (in reference to PhTET2 sequence) that can disrupt positive side chain fitting in “S1” pocket. In peptide context, MjTET also displayed significant cleavage capacity of acidic residues. Colombo *et al.* revealed the presence of third metal ion in the specificity pocket of PfTET3, a lysyl aminopeptidase of *P. furiosus* (26). This cation is coordinated by Thr232 and Glu281 residues and facilitate the accommodation of negative charged side chain of substrates in “S1” pocket and subsequently their degradation (26). In similar fashion, the “S1” pocket of MjTET encloses Thr232 and Glu289, which are localized in the same position as the former ones, allowing the possibility

to receive a third metal, thus extending MjTET cleavage capacity to acidic residues.

### *TET system consists in 4 independent peptide-processing modules*

The cryo-EM experiments allowed the edification of the twelve internal dynamic loops carried by the dimerization domains inside the M42 MjTET particle cavity. The loop positions with respect to the active sites, entry holes and their topology within the central cavity suggest that they are of pivotal importance for the functioning of the TET dodecamers. In particular, the cryo-EM map revealed that some internal loops residues, particularly His116, are orientated toward the binuclear active site. Sequence alignments reveal a strict conservation of His116 amongst M18 and M42 aminopeptidases exhibiting tetrahedral architecture (**Figure 12**). In eukaryotic M18, His170 of hDNPEP (human aspartyl aminopeptidase) and His166 of bDNPEP (bovine aspartyl aminopeptidase) are localized near the active site and were proposed to be involved in substrate binding and in the stabilisation of tetrahedral intermediate of enzymatique reaction (8,9). Moreover, the His170Phe mutation completely abolished the enzymatic activity of hDNPEP (32). Thus in M42 TET, histidine residue located at the internal loop extremity is suggested to play important role in peptides hydrolysis. Interestingly in cryo-EM structure, the distance that separates the His116 from the active site is much longer than the one measured in the crystal structures of M18 family TET (characterized by longer amino acid sequence than MjTET one). This may reflect a dynamic movement of MjTET internal loops to be closed to the active site. Thus, this motion allows the orientation of N-ter polypeptide end toward the binuclear catalytic site where hydrolysis will occur.

The position of the internal loops as determined the MjTET particle leads to drastically change our vision regarding the peptide navigation pathway within M42 family TET particles. The twelve internal loops of MjTET built in the electron density of cryo-EM map obstructed the internal cavity. They define a strong positively charged cluster generated by Lys113, His116, Arg117, Lys119, Lys123, Lys125 and Lys128 residues. These features would prevent polypeptide substrates to cross the particle from one entry hole to the opposite apex where three active sites are present, as we proposed before for the three TET from *P. horikoshii* (13,16). Thus another route for peptide navigation must be considered. The analysis of MjTET surface electrostatic potential showed three well-defined negative patches located at the entry hole of the tetrahedron. The comparison of the electrostatic potential of the different thermophilic M42 family TET complexes revealed that this feature is well conserved. In



## RÉSULTATS

---

MjTET, each acidic patch is made up by six acidic amino acids, Asp74-75, Glu243, Asp244 and Asp248 of each dimer loop encompassing the large opening of the particle. We noticed that these loops situated at the entrance hole and which harbored precedent negatively charged residues are located in front the internal positive loops at a correct distance and position allowing their interaction. Therefore, the function of the entry loops may be polypeptide capture in a such manner that the substrate N-terminal will be pre-orientated toward the internal cavity and one of the twelve internal positively charged loop devoted to substrate guidance and positioning in the catalytic chamber.

In conclusion, our results indicated that TET combines four peptide-processing modules each consisting in one entry hole in connection with three adjacent catalytic chambers. Three structural loops form a chain to guide the substrate N-termini toward one of three active sites of three adjacent catalytic chambers. The building of these modules also involves three other loops that participate to another functional module, which underline the functional relevance of TET dodecamerization. In this new model, and unlike to the previously proposed one, the four tetrahedron apexes could process simultaneously, and in an independent manner, at least four peptides while avoiding peptide congestion inside the particle.

## ACKNOWLEDGMENTS

We acknowledge financial support from the Agence Nationale de la Recherche (grant “Archelyse” ANR-12-BSV8-0019-0). Hind Basbous was supported by the CEA PhD program IRTELIS. This work used the platforms of the Grenoble Instruct Center (ISBG: UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The electron microscope facility is supported by the Rhône-Alpes Region, the Fondation Recherche Médicale (FRM), the Fonds FEDER and the GIS-Infrastructures en Biologie Santé et Agronomie (IBISA).

FIGURES

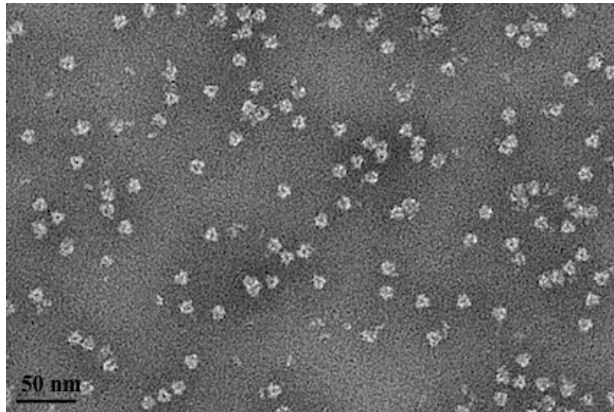


Figure 1: MjTET displays hollow tetrahedral-shaped particles observed by negative stain electron microscopy at 49000X magnification.

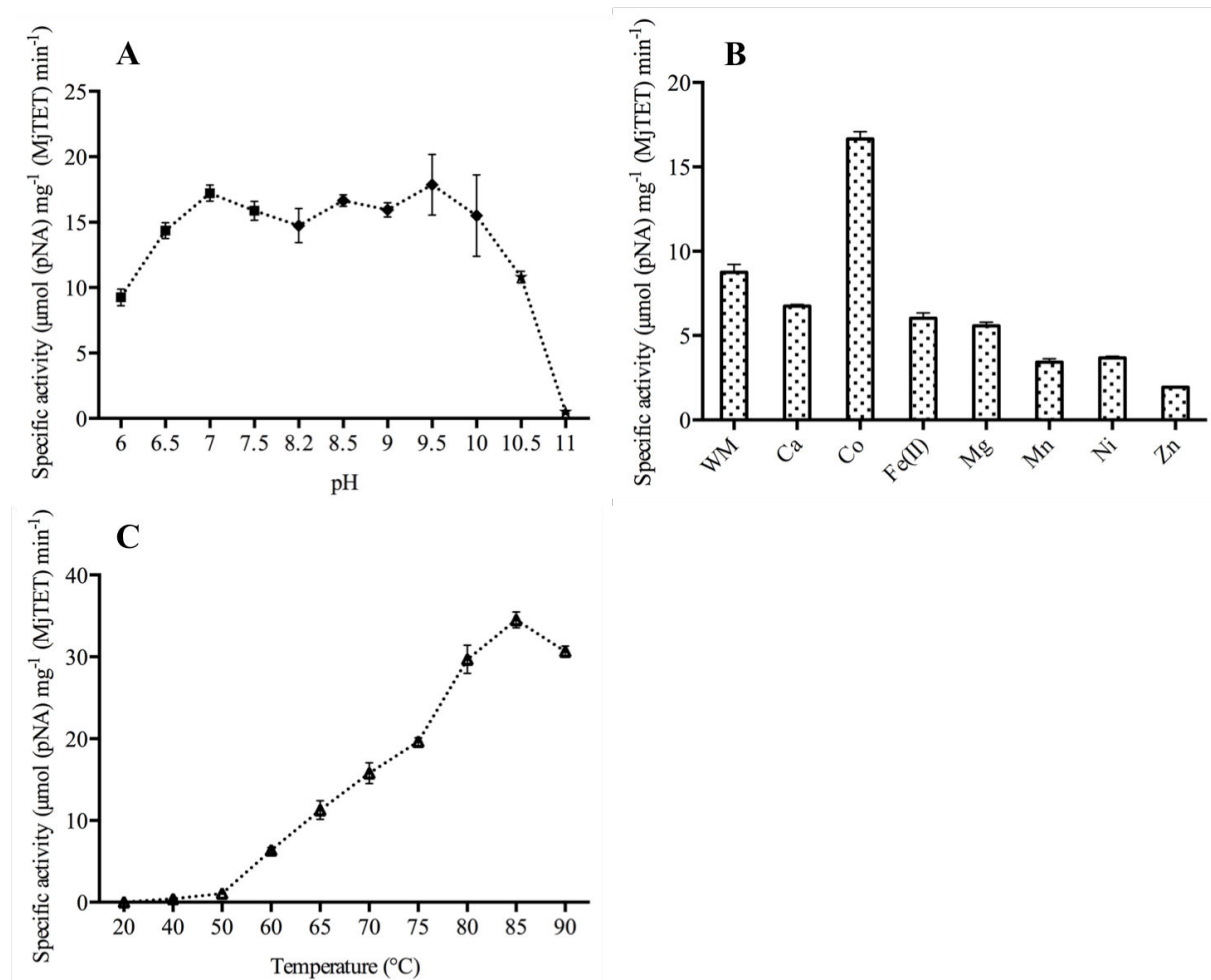
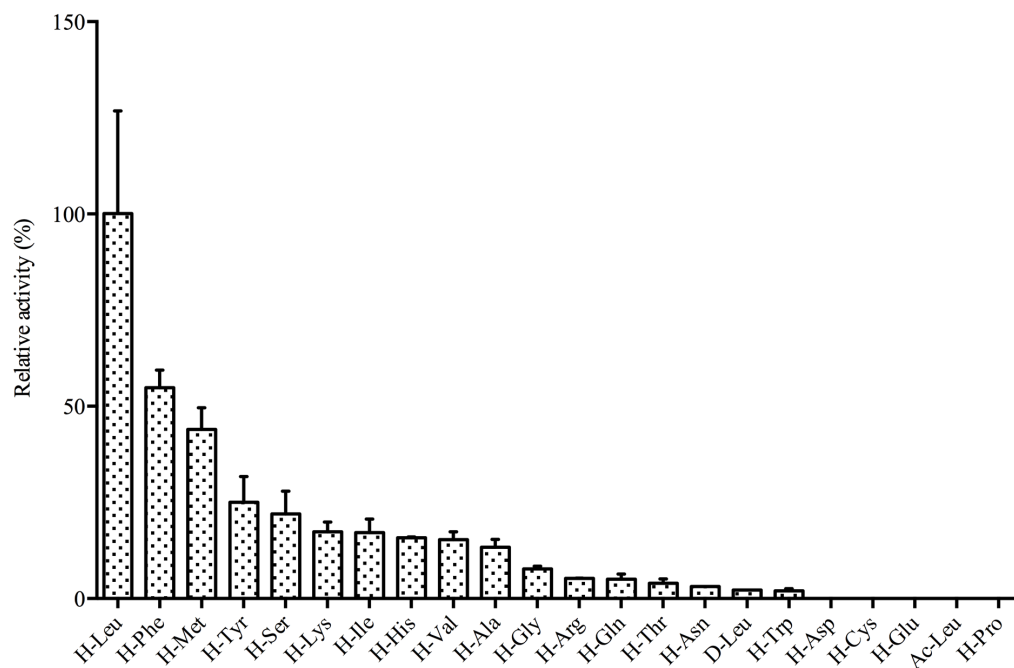
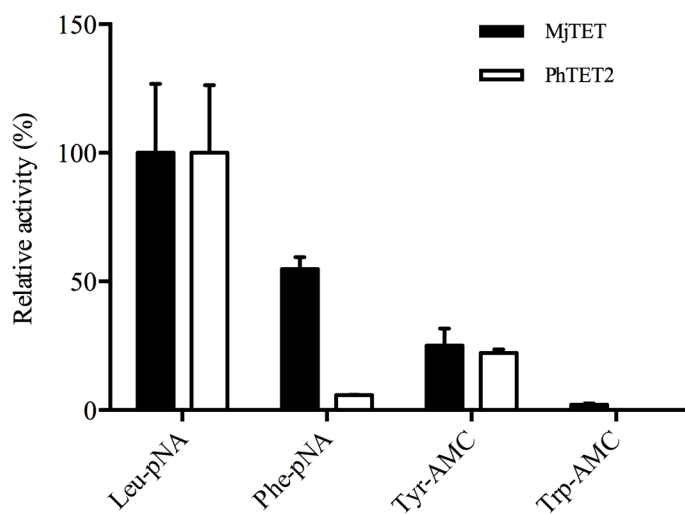


Figure 2: Influence of pH, metal ion and temperature variations on MjTET specific activities. **A.** Evolution of MjTET specific activity as function of pH in PIPES (■), CHES (◆) and CAPS (★) buffers. **B.** Specific activity modifications of MjTET in absence (WM: without metal) or in presence of different metal ions. **C.** Effect of temperature increasing on MjTET leucyl aminopeptidase activity.

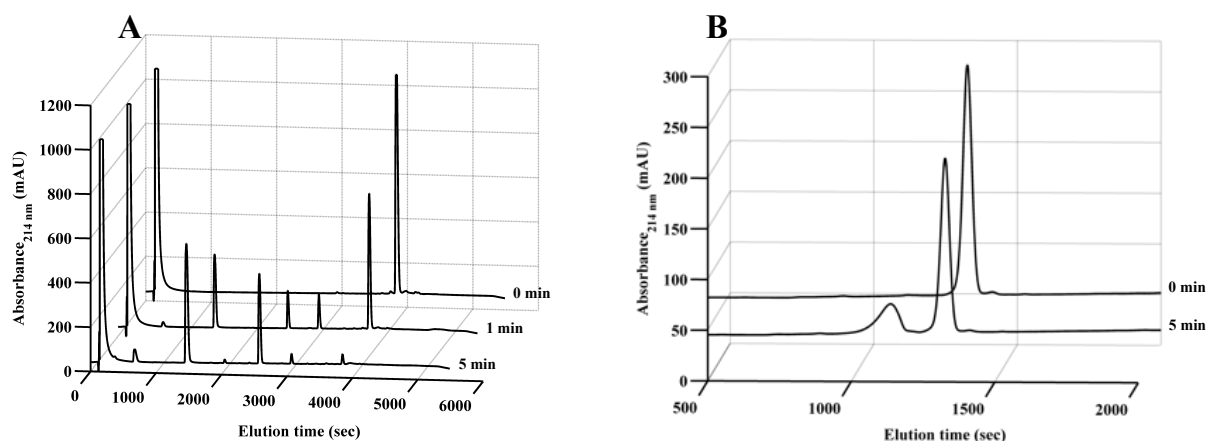
## RÉSULTATS



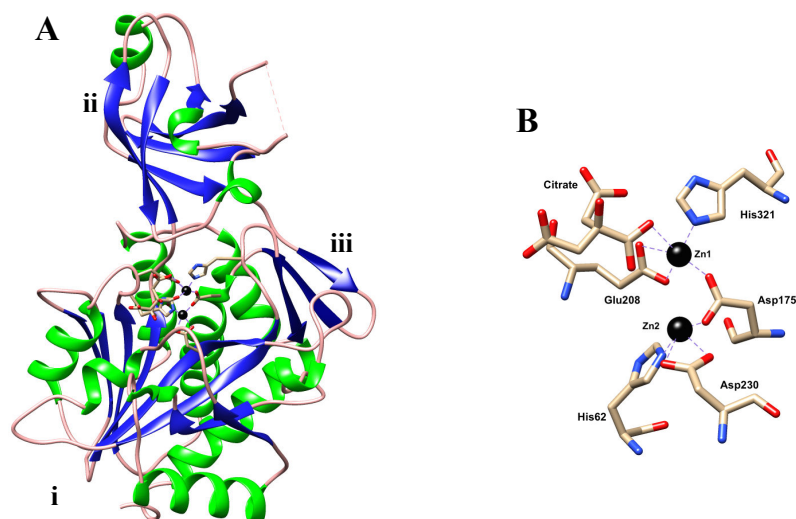
**Figure 3: MjTET substrate specificities.** MjTET displays broad substrate preference and higher activity was measured against hydrophobic residues: leucine, phenylalanine and methionine.



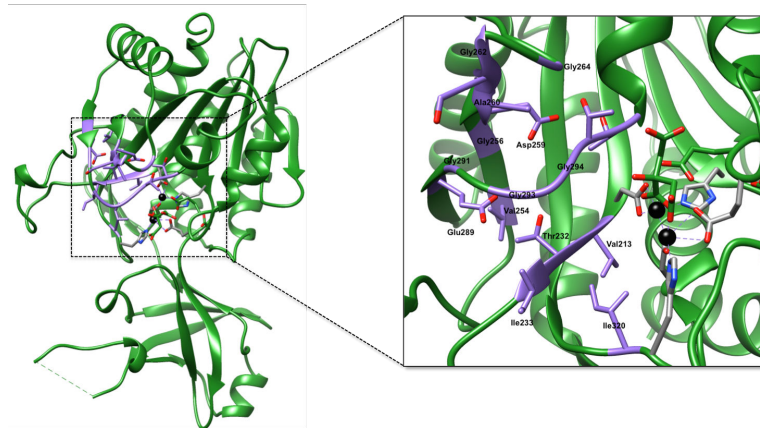
**Figure 4: Comparison of MjTET and PhTET2 amidolytic activities against aromatic residues.** MjTET shows higher capacity to hydrolyze phenylalanine and tyrosine in comparison to PhTET2. However these enzymes are unable to cleave tryptophan from substrate Nt end.



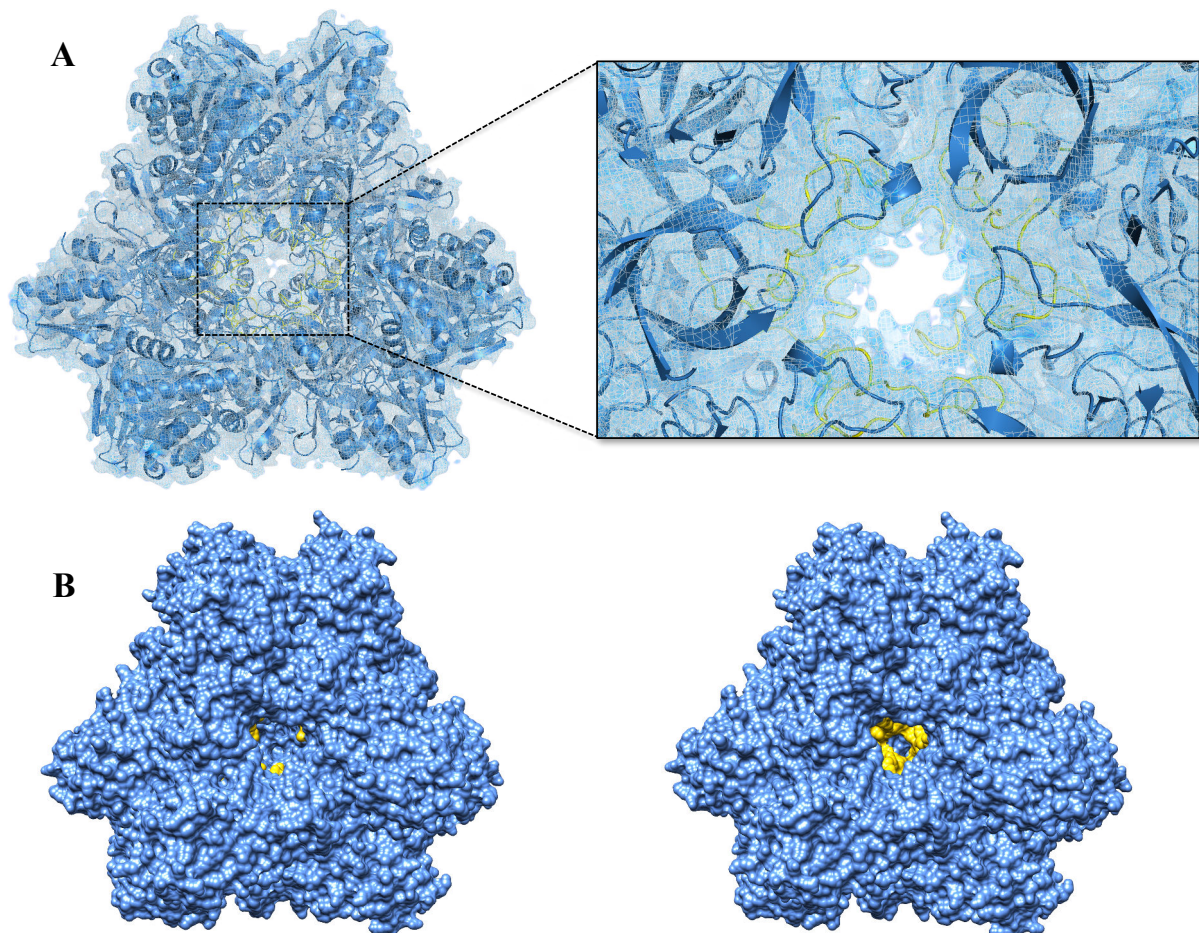
**Figure 5: Chromatographic profiles of peptide degradations by MjTET, obtained from reverse phase HPLC. A.** Chromatographic representation of H-Trp-Phe-Tyr-Ser(PO<sub>3</sub>H<sub>2</sub>)-Pro-Arg-pNA hydrolysis by MjTET after 1 min and 5 min of incubation. **B.** Chromatographic profile of H-Asp-Ala-Tyr-Pro-Ser degradation by MjTET after 5 min of incubation.



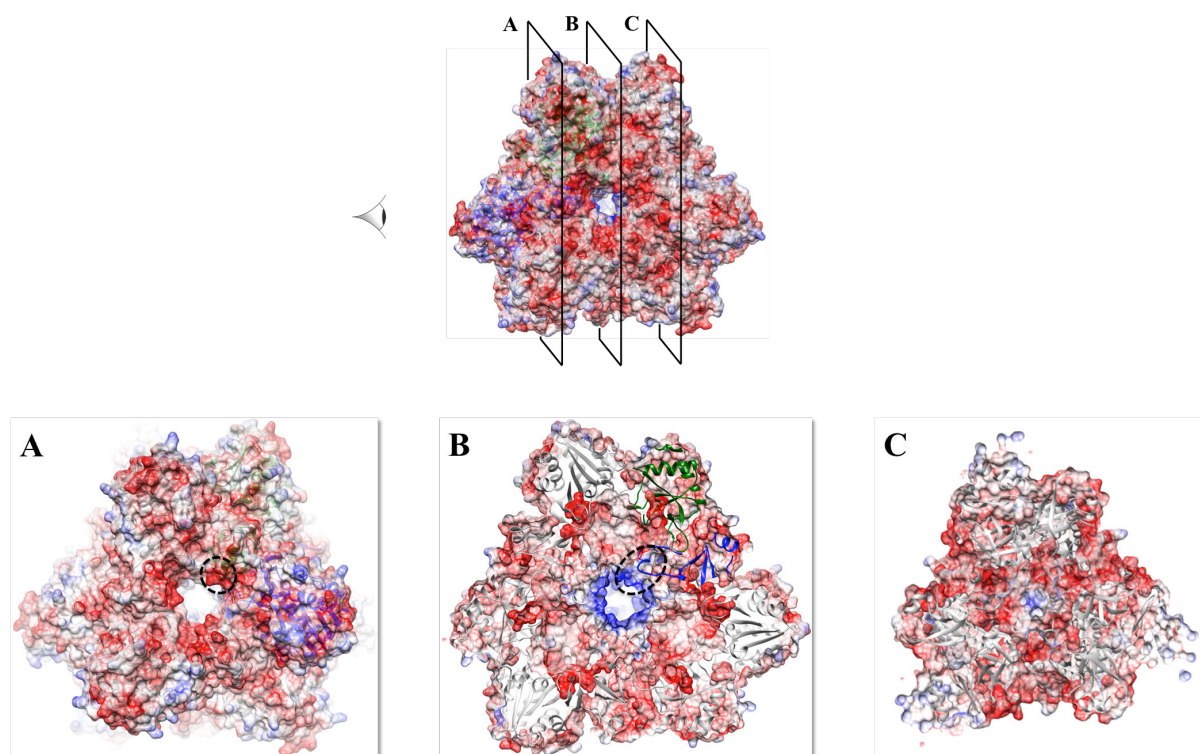
**Figure 6: X-ray structure of MjTET monomer. A.** Tertiary structure of MjTET monomer. i. Catalytic domain formed by the bigger  $\alpha/\beta$  globular domain. ii. Smaller  $\alpha/\beta$  globular that formed dimerization domain (I). Residues 115-124 carried by the dimerization domain I, are absent. iii. Three-stranded antiparallel  $\beta$ -sheet that formed the dimerization domain II. **B.** Close-up view of MjTET bimetallic active site. Two zinc ions (black spheres) are modeled in the active. Zn<sub>1</sub> is coordinated by Asp175, Glu208, His321 and citrate molecule, while His62, Asp175 and Asp230 coordinate Zn<sub>2</sub>.



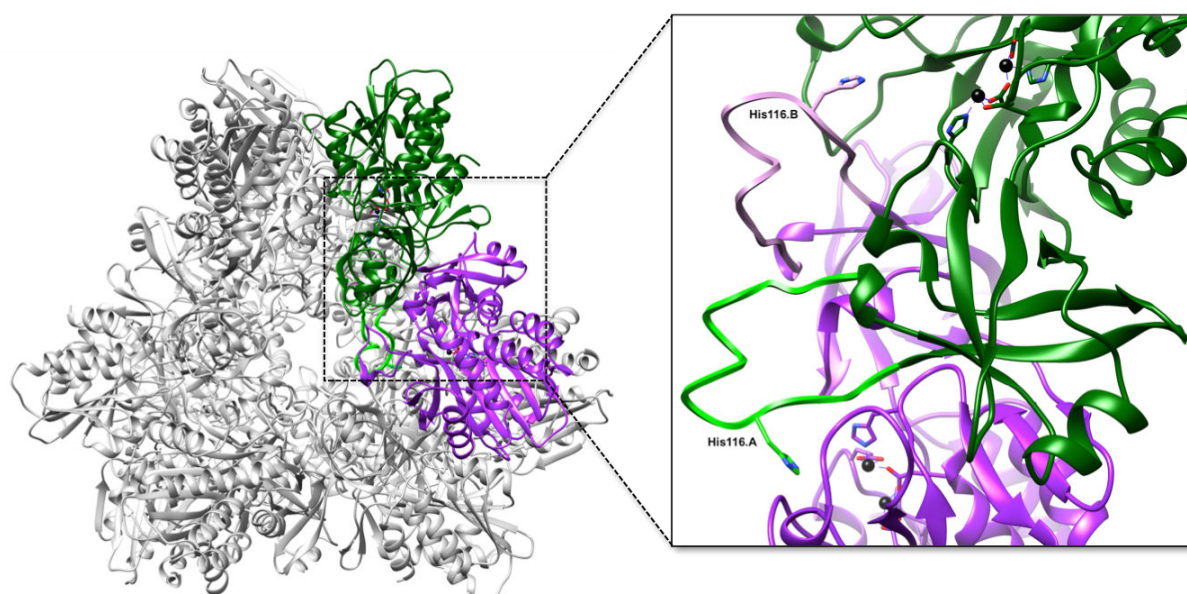
**Figure 7: View of MjTET specificity pocket near the active site.** Hydrophobic and neutral residues (represented by purple sticks) define MjTET specificity pocket allowing the reception of side chains of hydrophobic and neutral residues. Furthermore, two negative amino acids (Asp259, Glu289) are also present in this pocket, thus explaining the capacity of MjTET to hydrolyze positive substrates.



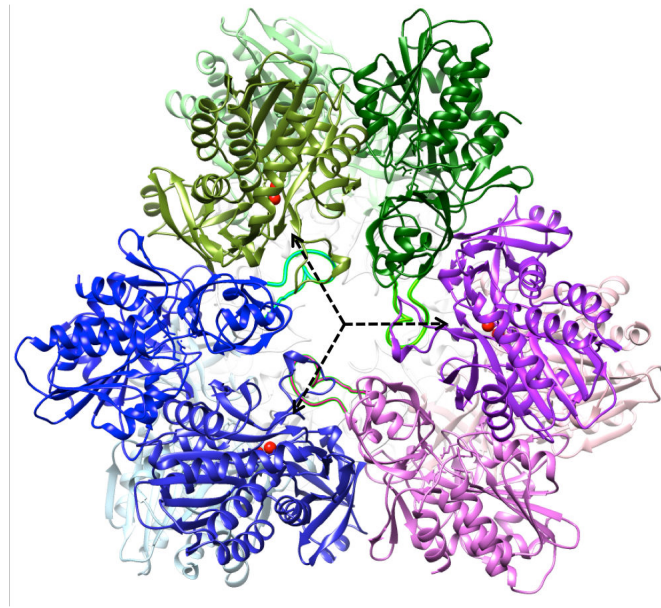
**Figure 8: Representation of the X-ray missing loops in the cryo-EM map and inside the surface of MjTET particle.** **A.** View of the internal electron density obtained by cryo-EM, wherein the X-ray atomic model of MjTET (blue) is fitted inside and the X-ray missing loops (yellow) were built. **B.** Surface representation of X-ray structure of MjTET without (left) and with (right) internal loops (yellow) built from cryo-EM map. Clear obstruction of the internal cavity is observed for MjTET model enclosing these mobile loops.



**Figure 9: MjTET electrostatic potential observed at three different levels of slicing from the facet harboring the opening hole toward the exit hole localized at the opposite side. A.** MjTET entry hole is lined by three negative patches. One of these patches made up by dimer (colored in green and blue), is highlighted by dashed circle. **B.** Deeply slice, inside MjTET particle, shows positive patches formed by internal loops. At this entrance, one loop, more precisely that one of blue monomer, is engaged in the positive observed charge and is highlighted by dashed circle. While the second loop of the green monomer is orientated toward the adjacent opening hole (not shown here). **C.** Observation of the opposite exit pore where three small channels made up by side chain of Arg215 and Phe219 of three monomers, can be observed.



**Figure 10: Topological location of two loops carried by two monomers forming the dimeric building block of the tetrahedral particle.** Each loop of each monomer (A or B, colored in green and purple, respectively) is inserted near the active site of the adjacent monomer in the same building block. Histidine residue (His116) is orientated in the direction of the active site, where black spheres represent two  $Zn^{2+}$  ions.



**Figure 11: Topological localization of three internal loops behind the opening hole.** Each opening hole is built by three dimers, and each monomer of dimer participates with two other monomers (of different dimers) in the genesis of the four particle apices. Here, three apices are represented with different nuance colors of green, purple and blue. Three internal loops (light green, magenta and cyan) are localized behind the entrance pore; each one is carried by one monomer of a building block. These loops guide peptide toward one of the three active sites, where zinc ions are represented as red spheres, as shown by dashed arrow.

	90	100	110	120
MjTET	.KIGG...IYDPTI	.FNQKVVHGSKCDL	...IGVIGSKPPH	RMK.EEEK
PhTET1	.KVG...IDDRLL	.YGRHVNVT.EKCL	...DGVIGATPPH	LSL..ERD
PhTET2	.PIGG...VDPKTL	.IAQRFKWWI.DKCFI	...YGVGASVPPH	IQK.PEDR
PhTET3	.PIGG...VLPETL	.VAQRIRFFT.EKER	...YGVVGLLPPH	LRGQEDK
PfTET3	.PIGG...VLPETL	.IAQKIRFFT.EKER	...YGVVGLLPPH	LRREAKDQ
DkamTET	.PIGG...VLPETL	.LYQRVVRT.RDCL	...RGVIGLKPPH	VIK.PEAA
SpPepA	.EIGG...WNPMVV	.SSQRFKLLT.RDCEI	...PVISGSVPPH	LRG.KGG
PaAP	EVYGCALFAPWFRDLS	LAGRVVTFR..ANG	KLESRLVDFRKAIA	VIPNLALINRAANE
hDNPEP	ETYGGGIWSTWFRDRL	LAGRVIVKCPST	GRLEQQLVHVERPIL	RIPHLALILQRNINEN
bDNPEP	ETYGGGIWSTWFRDRL	LAGRVIVKCPST	GRLEQQLVHVRPIL	RIPHLALILQRNVNEN

	130	140	150
MjTET	TKIIKYEDMFDIGAESREPAI	.....	EMGVNIGTWVSE
PhTET1	KSVIPWYDLVIDIGAESKEPAL	.....	ELVK.PLDFAVF
PhTET2	KKAPDWQIFIDIGAESKEPAE	.....	DMGVKIGTVITW
PhTET3	GSKIDWDQIVVDVGASSKEPAE	.....	EMGFRVGTVGEF
PfTET3	GKIDWDSILVDVGASSKEPAE	.....	EMGFRIGTIGEF
DkamTET	QKVPRELRELFIDVGASSKEPAE	.....	KMGTRVGDIAVF
SpPepA	PTMPATADIVFDGGFADKAPAE	.....	SFGTRPGDTIVP
PaAP	WFINAQNELPPIIAQLAPGPAE	.....	DFRLLLDEQLLREHG
hDNPEP	FGPNTEMHLVPIIATATIQEELKGTPEPGLNAVDERHHSVLMSSLCAH	LGLSPKDIVEM	
bDNPEP	FGPNMEMHLVPIIATATIQEELKGTPEPGLNAVDERHHSVLTSLCAH	LGLSPKDIVEM	

**Figure 12: Sequence alignment of M18 and M42 dodecameric aminopeptidases.** Internal loop residues are marked by blue triangles according to those of MjTET. This alignment shows conservation of histidine residues (red triangle) amongst M18 and M42 aminopeptidases. PhTET1, PhTET2, PhTET3, PfTET3 and DkamTET are archaeal aminopeptidases expressed by *Pyrococcus horikoshii*, *P. furiosus* and *Desulfurococcus kamchatkensis*, respectively. SpPepA and PaAP are bacterial aminopeptidases expressed by *Streptococcus pneumoniae* and *Pseudomonas aeruginosa*, respectively. hDNPEP and bDNPEP are human and bovine aminopeptidases, respectively. MjTET, PhTET1, 2 and 3, PfTET3, DkamTET and SpPepA belong to M42 family, while PaAP, hDNPEP and bDNPEP belong to M18 family according to MEROPS classification. Sequence alignment was done by using ESPrpt 3.0 server (Robert and Gouet, 2014).

## RÉSULTATS

Structure	MjTET-Nt(-_Zn2)	MjTET-Nt(Zn1_Zn2)
Beam line	ID29 (ESRF)	ID29 (ESRF)
Space group	P2 <sub>1</sub> 3	P2 <sub>1</sub> 3
Unit cell constants (Å)	a= 166.54 b= 166.54 c= 166.54	a= 166.15 b= 166.15 c= 166.15
Resolution range (Å)	48.08 – 2.40 (2.53 – 2.40)	47.96 – 2.69 (2.84 – 2.69)
Wavelength (Å)	1.2398	1.0332
R-merge (%)	10.1 (80.9)	9.2 (72.4)
R-pim	6.2 (49.4)	4.9 (39.1)
CC1/2	97.2 (55.5)	99.7 (72.5)
Mean I/σ(I)	6.0 (1.5)	9.1 (1.6)
Completeness (%)	99.3 (99.1)	99.2 (96.0)
Redundancy	4.3 (4.2)	4.9 (5.0)
Unique reflections	59713 (8617)	42082 (5899)
<b>Refinement</b>		
R-work (%)	18.59	17.55
R-free (%)	22.91	23.68
RMSD (angles)	1.18	1.56
RMSD (bonds)	0.011	0.016
<b>Ramachandran plot</b>		
Most favored region (%)	95.52	94.95
Allowed region (%)	3.73	4.53
Outliers (%)	0.75	0.52
<b>Number of atoms</b>		
Protein	10898	10554
Water	430	0
Heteroatoms:	4	60
Zinc	4	8
Citrate	0	52
<b>Mean isotropic B-factors</b>		
Protein	42.80	72.45
Chain A	37.92	63.62
Chain B	41.30	71.81
Chain C	46.28	77.02
Chain D	45.57	77.29
Water	43.45	0
Zinc	59.43	67.56
Citrate	0	77.37

**Table 1: Data collection and refinement statistics of MjTET-Nt(-\_Zn<sub>2</sub>) and MjTET(Zn<sub>1</sub>\_Zn<sub>2</sub>).** Values in parenthesis correspond for the highest resolution shell.

R- merge =  $\frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$  where  $I_{hkl,j}$  is the  $j^{\text{th}}$  intensity measurement of reflection  $hkl$  and  $\langle I \rangle$  is the average intensity from multiple observations.

R- pim =  $\frac{\sum_{hkl} \sqrt{\frac{1}{n-1} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl} \rangle|}}{\sum_{hkl} \sum_j I_{hkl,j}}$  with  $I_{hkl,j}$  is the  $j^{\text{th}}$  intensity measurements of reflection  $hkl$  and  $\langle I \rangle$  is the average intensity from multiple observations.  $n$  represents the multiplicity of the measurements.

CC<sub>1/2</sub>= correlation coefficient between random half datasets.

R<sub>work</sub> =  $\frac{\sum_{hkl} ||F_o| - |F_c||}{\sum_{hkl} |F_o|}$  for all data except 4% which were used for R<sub>free</sub> calculation.



### REFERENCES

1. Lowther, W. T., and Matthews, B. W. (2002) Metalloaminopeptidases: common functional themes in disparate structural surroundings. *Chemical reviews* **102**, 4581-4608
2. Gonzales, T., and Robert-Baudouy, J. (1996) Bacterial aminopeptidases: properties and functions. *FEMS Microbiol Rev* **18**, 319-344
3. Appolaire, A., Colombo, M., Basbous, H., Gabel, F., Girard, E., and Franzetti, B. (2015) TET peptidases: A family of tetrahedral complexes conserved in prokaryotes. *Biochimie*
4. Franzetti, B., Schoehn, G., Hernandez, J. F., Jaquinod, M., Ruigrok, R. W., and Zaccai, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *Embo J* **21**, 2132-2138.
5. Appolaire, A., Dura, M. A., Ferruit, M., Andrieu, J. P., Godfroy, A., Gribaldo, S., and Franzetti, B. (2014) The TET2 and TET3 aminopeptidases from *Pyrococcus horikoshii* form a hetero-subunit peptidasome with enhanced peptide destruction properties. *Mol Microbiol* **94**, 803-814
6. Marty, V., Jasnin, M., Fabiani, E., Vauclare, P., Gabel, F., Trapp, M., Peters, J., Zaccai, G., and Franzetti, B. (2013) Neutron scattering: a tool to detect in vivo thermal stress effects at the molecular dynamics level in micro-organisms. *Journal of the Royal Society, Interface / the Royal Society* **10**, 20130003
7. Appolaire, A., Rosenbaum, E., Dura, 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. *J Biol Chem* **288**, 22542-22554
8. Chen, Y., Farquhar, E. R., Chance, M. R., Palczewski, K., and Kiser, P. D. (2012) Insights into substrate specificity and metal activation of Mammalian tetrahedral aspartyl aminopeptidase. *J Biol Chem* **287**, 13356-13370
9. Chaikuad, A., Pilka, E. S., Riso, A. D., Delft, F. V., Kavanagh, K. L., Venien-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 Struct Biol* **12**, 14
10. Dutoit, R., Brandt, N., Legrain, C., and Bauvois, C. (2012) Functional characterization of two M42 aminopeptidases erroneously annotated as cellulases. *PloS one* **7**, e50639
11. 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. *Biochem Biophys Res Commun* **391**, 431-436
12. Robinson, M. W., Buchtman, K. A., Jenkins, C., Tacchi, J. L., Raymond, B. B., To, J., Roy Chowdhury, P., Woolley, L. K., Labbate, M., Turnbull, L., Whitchurch, C. B., Padula, M. P., and Djordjevic, S. P. (2013) MHJ\_0125 is an M42 glutamyl aminopeptidase that moonlights as a multifunctional adhesin on the surface of *Mycoplasma hyopneumoniae*. *Open biology* **3**, 130017
13. Schoehn, G., Vellieux, F. M., Asuncion Dura, M., Receveur-Brechot, V., Fabry, C. M., Ruigrok, R. W., Ebel, C., Roussel, A., and Franzetti, B. (2006) An archaeal peptidase assembles into two different quaternary structures: A tetrahedron and a giant octahedron. *J Biol Chem* **281**, 36327-36337
14. Franzetti, M. A. D. a. B. (2013) PhTET1 Aminopeptidase. in *In Neil D. Rawlings and Guy S. Salvesen, editors: Handbook of Proteolytic Enzymes, Oxford: Academic Press.* pp pp. 1638 - 1645

15. Dura, 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
16. Dura, M. A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F. M., 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. *Mol Microbiol* **72**, 26-40
17. Tsao, J. H., Kaneshiro, S. M., Yu, S. S., and Clark, D. S. (1994) Continuous culture of *Methanococcus jannaschii*, an extremely thermophilic methanogen. *Biotechnol Bioeng* **43**, 258-261
18. Kabsch, W. (2010) Xds. *Acta crystallographica. Section D, Biological crystallography* **66**, 125-132
19. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760-763
20. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674
21. Grant, T., and Grigorieff, N. (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife* **4**, e06980
22. Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**, 82-97
23. Scheres, S. H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530
24. Estrozi, L. F., and Navaza, J. (2008) Fast projection matching for cryo-electron microscopy image reconstruction. *J Struct Biol* **162**, 324-334
25. Effantin, G., Estrozi, L. F., Aschman, N., Renesto, P., Stanke, N., Lindemann, D., Schoehn, G., and Weissenhorn, W. (2016) Cryo-electron Microscopy Structure of the Native Prototype Foamy Virus Glycoprotein and Virus Architecture. *PLoS pathogens* **12**, e1005721
26. 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
27. Russo, S., and Baumann, U. (2004) Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J Biol Chem* **279**, 51275-51281
28. 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
29. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157-162
30. Borissenko, L., and Groll, M. (2005) Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J Mol Biol* **346**, 1207-1219
31. 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
32. Wilk, S., Wilk, E., and Magnusson, R. P. (2002) Identification of histidine residues important in the catalysis and structure of aspartyl aminopeptidase. *Arch Biochem Biophys* **407**, 176-183

### 5.1.3. Discussion et Perspectives

Le travail décrivant la caractérisation enzymatique et structurale de MjTET, montre que cette unique TET M42 de *M. jannaschii* présente une activité enzymatique à large spectre avec une préférence prononcée contre la leucine et une capacité à cliver les résidus aromatiques. Ceci a permis de lui attribuer la propriété de leucine aminopeptidase, tout comme PhTET2, une autre leucine aminopeptidase M42 isolée de *P. horikoshii*. Ces deux aminopeptidases sont des enzymes hyperthermophiles dont les optimums d'activités sont de 85 °C et 100 °C, pour MjTET et PhTET2, respectivement. Les analyses et les comparaisons des interfaces de dimérisation (entre deux monomères d'un même *building block*) et d'oligomérisation (entre trois monomères localisés au sommet du tétraèdre), montrent une grande conservation de ces régions maintenues par un nombre similaire de liaisons hydrogènes et de ponts salins entre les deux TETs. De surcroît, les résidus Val209, Lys212, Arg215, Phe219 (en référence à MjTET) sont conservés avec ceux de PhTET2 décrits comme impliqués dans les interactions hydrophobes fondamentales pour maintenir l'état dodécamérique de la particule (Appolaire et al., 2014a). Ceci pourrait expliquer alors le fait que ces deux TETs possèdent des demi-vies voisines, de 8 et 10 heures en chauffant MjTET et PhTET2 à 85 et 80 °C, respectivement. Cette grande similarité entre les déterminants clés responsables de l'assemblage du complexe TET entre système bactérien et archéen, indique que le tétraèdre TET était probablement présent dans l'ancêtre commun aux archées et aux bactéries.

Les deux aminopeptidases MjTET et PhTET2 sont caractérisées par leur large spectre d'activité. Pourtant, la différence majeure existante réside dans la capacité de MjTET à dégrader certains résidus non ou faiblement hydrolysables par PhTET2. Ce trait distinctif est illustré par l'hydrolyse des résidus aromatiques tels que la phénylalanine et la tyrosine par MjTET qui clive aussi plus efficacement les acides aminés positifs, lysine et histidine, que PhTET2. Ces comportements différents sont attribués aux différences des poches de spécificité « S1 » de chacune. Effectivement, la poche « S1 » de MjTET renferme trois résidus supplémentaires par rapport à celle de PhTET2 et est caractérisée par la présence de six résidus glycines lui conférant probablement une grande flexibilité pour la réception de larges chaînes latérales comme celles des résidus aromatiques. De plus, la plupart des acides aminés bordant la poche « S1 » de MjTET sont de courtes chaînes latérales comme la thréonine, ce qui crée alors un bon environnement pour le positionnement de grosses chaînes

## RÉSULTATS

---

latérales. Quant à PhTET2, un résidu lysine (absent chez MjTET) pointe au sein de la poche réduisant probablement son volume ainsi que sa capacité à recevoir les résidus basiques. Des mutations des glycines de la poche « S1 » de MjTET pourraient être proposées dans le but d'explorer leur implication dans la capacité de MjTET à hydrolyser les résidus aromatiques.

Cette large spécificité de MjTET pourrait être expliquée par le fait qu'elle soit la seule TET exprimée par *M. jannaschii*, qui ainsi prend en charge la dégradation d'une grande diversité de peptides. De ce fait, le qualificatif de « peptidasome » non spécifique peut lui être attribué. Celui-ci serait probablement impliqué dans des fonctions de « ménage » cellulaire. De plus, *M. jannaschii* est une espèce d'archée autotrophe qui n'utilise pas les composés organiques extracellulaires tels que les peptides et les protéines comme source de carbone et d'énergie. Ceci suggère que MjTET ne jouerait pas un rôle métabolique aussi prépondérant que les 4 TETs de *Pyrococcus horikoshii* qui est un organisme hétérotrophe. Nous pouvons ainsi penser que la spécialisation des TET des Thermococcales, acquises *via* le transfert de gènes permet de remplacer un système archéen moins efficace. Il serait intéressant de pouvoir comparer l'efficacité du système bactérien vis-à-vis du système archéen et de déterminer le rôle précis de l'unique aminopeptidase des souches *M. jannaschii* par une approche génétique. Des souches déficientes en MjTET pourraient être obtenues pour évaluer (si elles sont viables) l'impact de l'absence de MjTET sur la survie, le métabolisme et le peptidome des souches concernées. D'ailleurs dans un même contexte, il serait aussi important d'étudier à quel(s) stade(s) de croissance MjTET est exprimée et sous quel(s) état(s) oligomérique(s) se trouve-t-elle.

Dans cette étude, la combinaison de la cristallographie et de la cryo-EM nous a permis de construire les boucles localisées à l'intérieur de l'édifice dodécamérique. Contrairement aux aminopeptidases tétraédriques M18, les densités électroniques qui correspondent à ces boucles sont déjà absentes des données structurales de toutes les TETs M42. Dans MjTET, l'organisation de cette boucle montre qu'elle est orientée vers le site actif d'un monomère adjacent dans un même dimère. De plus, celle-ci est caractérisée par une forte charge positive. Sa position est précisément située derrière le pore d'entrée du tétraèdre bordé par des charges négatives décrites pour leur capacité à attirer les extrémités Nt positives des substrats peptidiques (Schoehn et al., 2006). Cette organisation nous a permis de supposer que cette boucle positive contrôle le trafic du substrat depuis son entrée dans la particule vers l'un des trois sites actifs situés au niveau des sommets présents de part et d'autre du pore d'entrée, et

non plus vers l'apex situé en face de ce dernier. Cette hypothèse est consolidée par le fort potentiel positif créé par cette boucle et qui empêche la circulation du peptide du pore d'entrée vers le trou de sortie en face. Cette capacité à diriger le peptide vers le site actif a été évoquée précédemment par Chaikuad et ses collègues pour hDNPEP, une aspartyl aminopeptidase dodécamérique humaine (Chaikuad et al., 2012). Pour comprendre plus précisément le rôle des boucles internes dans la particule, il serait intéressant de co-cristalliser MjTET avec un substrat non hydrolysable, afin d'identifier si celui-ci serait maintenu par les boucles internes au niveau du site actif. En outre, des mutations des résidus positifs de la boucle interne par des résidus neutres à chaîne latérale courte, tout en conservant l'oligomérisation de MjTET, pourraient être envisageables pour vérifier la veracité du modèle proposé concernant l'implication des boucles internes dans le trafic des peptides dans la particule.

## 5.2. CHAPITRE 2 : CARACTÉRISATIONS FONCTIONNELLES DE LA QUATRIÈME TET, PhTET4 DE *PYROCOCCLUS HORIKOSHII*

### 5.2.1. Présentation du travail

La présence d'une quatrième TET a été mise en évidence, grâce aux études phylogénétiques, chez l'archée hyperthermophile abyssale *Pyrococcus horikoshii*. Cette nouvelle enzyme est classée dans la banque MEROPS comme étant une peptidase putative appartenant à la famille M42 des métallopeptidases du clan MH. Tout comme ses trois homologues, PhTET1, TET2 et TET3, PhTET4 est une aminopeptidase d'origine bactérienne dont le gène codant *PH0737* a été acquis par transfert horizontal depuis une espèce bactérienne. Les alignements de séquences réalisés entre PhTET4 et ses homologues PhTETs de *P. horikoshii* montrent qu'elle présente la structure primaire la plus éloignée. En effet, elle partage une identité de séquence de ~21, 23 et 22% avec PhTET1, PhTET2 et PhTET3, respectivement.

En se référant aux résultats préliminaires suscités, plusieurs questions se sont posées :

- Est-ce que la protéine codée par *PH0737*, nommée ensuite PhTET4, adopte-t-elle la même structure quaternaire que ses homologues PhTET1, 2 et 3 ?

## RÉSULTATS

---

- PhTET4 serait-elle une métallo-aminopeptidase ? Si oui, alors que serait la signification biologique de sa séquence peptidique distincte de celles des autres PhTETs ?
- Que serait l'intérêt biologique de l'expression de quatre différentes TETs par *P. horikoshii* et les Thermococcus, le genre auquel appartient cette espèce ?

Pour pouvoir répondre à ces questions, plusieurs investigations biochimiques, enzymatiques et biophysiques ont été entreprises. Tout d'abord, la chromatographie d'exclusion de taille et l'observation des échantillons purifiés de PhTET4 en microscopie électronique par coloration négative démontrent que celle-ci, identiquement aux PhTETs, adopte aussi une organisation en tétraèdre creux. Ensuite, les caractérisations enzymatiques effectuées contre les vingt acides aminés, en utilisant une diversité des monoacyls couplés au pNA et à l'AMC, dévoilent une activité aminopeptidase à spectre très restreint avec une capacité d'hydrolyser uniquement les résidus glycines. Contrairement aux homologues PhTETs, PhTET4 est efficacement activée en présence du nickel et non pas du cobalt. Néanmoins, le cobalt, comme le manganèse, active faiblement PhTET4. En outre l'étude des conditions optimales d'activation de cette glycine aminopeptidase, démontre une activité amidolytique maximale à pH basique, pH 9,5 et à une température de 95 °C, étant donné que *P. horikoshii* est une espèce hyperthermophile prospérant à 98 °C (González et al., 1998). La spécificité stricte de PhTET4 a été validée en utilisant divers peptides dont le profil de digestion est évalué par HPLC. En effet, aucune hydrolyse n'est observée que lorsqu'un résidu glycine est situé en position P1 du substrat peptidique.

Avant mon arrivée au laboratoire, des données cristallographiques avaient été enregistrées pour PhTET4 à 3 Å de résolution. J'ai effectué une analyse préliminaire qui conduit à un groupe d'espace orthorhombique P22<sub>1</sub>2<sub>1</sub> et à la présence d'un dodécamère entier dans l'unité asymétrique. Ces données ont été collectées au seuil du zinc ce qui permettra d'envisager plusieurs approches pour déterminer la structure de PhTET4 :

- Phasage *de novo* par la méthode SAD,
- Remplacement moléculaire, malgré la faible similarité de séquence de TET4 avec les autres PhTETs,

## RÉSULTATS

---

- Approche combinée remplacement moléculaire-SAD (MR-SAD).

Compte tenu de la présence d'un dodécamère dans l'unité asymétrique et la résolution moyenne des données, l'affinement pourra ne pas être simple à conduire. En conséquence, au cours de ma thèse et en parallèle de la détermination de la spécificité de PhTET4, j'ai entrepris d'améliorer la diffraction des cristaux, afin d'obtenir des données à plus haute résolution, et j'ai cherché à obtenir d'autres formes cristallines qui conduiraient à un nombre plus faible de molécules dans l'unité asymétrique. Toutes les tentatives n'ont, pour l'instant, pas été couronnées de succès.

**5.2.2. The hyperthermophilic archaeon *Pyrococcus horikoshii* possesses a fourth version of TET aminopeptidase complex with strict glycyI-specificity and nickel-activation properties**



### **The hyperthermophilic archaeon *Pyrococcus horikoshii* possesses a fourth version of TET aminopeptidase complex with strict glycyI-specificity and nickel-activation properties**

Hind Basbous<sup>1</sup>; Alexandre Appolaire<sup>1</sup>; Eric Girard<sup>1</sup> and Bruno Franzetti<sup>1\*</sup>

<sup>1</sup> Université Grenoble Alpes, CNRS, Commissariat à l'Énergie Atomique, Institut de Biologie Structurale F-38044 Grenoble, France

\* Corresponding authors: [bruno.franzetti@ibs.fr](mailto:bruno.franzetti@ibs.fr)

#### **ABSTRACT**

The TET peptidases are large self-compartmentalized complexes that form dodecameric particles. These members of the M42 family metallopeptidases are widely distributed in prokaryotes. Three different versions of TET complexes with different substrate specificities were found to co-exist in the cytosol of the hyperthermophilic archaeon *Pyrococcus horikoshii*. In the present work we have identified of a novel type of TET complex that we named PhTET4. The recombinant PhTET4 enzyme was found to self-assemble as a tetrahedral edifice similar to other TET complexes. We determined PhTET4 substrate specificity using a broad range of mono-acyl chromogenic and fluorogenic compounds. In addition, HPLC peptide degradation assays were also performed. These experiments demonstrated that PhTET4 is a strict glycyI aminopeptidase (GAP), devoid of amidolytic activity toward other type of amino acids. PhTET4 catalytic efficiency was studied under various conditions. The protein was found to be a thermophilic, neutral aminopeptidase activated by nickel ions.

#### **INTRODUCTION**

In all cell types, metallo-aminopeptidases play crucial roles in energy metabolism, protein maturation and degradation and in the regulation of biologically active peptides by removing the N-terminal amino acid from proteins and oligopeptides (1,2). In most of case, they operate beyond the action of an endoprotease and their activity is limited to small peptides. Several aminopeptidases were found to self assemble as large dodecameric particles (for a review see (3)). These half-megadalton molecular machines were first discovered in archaea and were named TET due to their peculiar tetrahedral shapes (4). The 13 nm hollow dodecahedrons enclose 12 active sites distributed in 4 funnel-shaped chambers located in the apices of the

## RÉSULTATS

---

particles and four large access holes, formed by the junction of six subunits, situated in the facets (3). This organization strongly distinguishes TET from the other cytosolic compartmentalized peptidases that mostly adopt a barrel-shaped architecture. Biochemical and structural studies of TET peptidases have accumulated over the past 10 years (3). They revealed that TET dodecamers represent a common scaffold for an efficient polypeptide capture and processing system (5). In all TET, the aminopeptidase activity is based on co-catalytic dinuclear metal active site belonging to M18 or M42 peptidase families according to MEROPS classification system (6). Bounding peptides are cleaved following common mechanism involving water molecule and glutamate residue (7). The nature of the metal occupying the bimetallic active site has been shown to modulate TET enzymatic activity and  $\text{Co}^{2+}$  ions appear to be the best activators for almost all archaeal and bacterial M42 TET (8-13).

TET machines are widespread and were found in the three life domains. Interestingly, in prokaryotes, 1 to 4 different types of TET complexes can co-exist in the cytosol depending upon the cell type. These enzymes can be categorized according to their preferences for the chemical structure of the N-terminal amino acid residues present in the polypeptide chains. Three main categories have been identified so far: glutamyl/aspartyl aminopeptidases, lysine aminopeptidases and leucine aminopeptidases, these later exhibiting the broader specificities. In eukarya, M18 TET complexes displayed aspartyl aminopeptidase activity (14,15). In bacteria, M42 TET peptidases, from *Clostridium thermocellum* and *Thermotoga maritima*, were assigned as leucyl aminopeptidases (12). Two M42 enzymes from the pathogens *Streptococcus pneumoniae* and *Mycoplasma hyopneumoniae* displayed glutamyl-aminopeptidase activities (11),(16). In archaea, the common TET structural scaffold can harbor disparate functions. So far, the unique TET complex described in the halophilic archaeon *Haloarcula marismortui* displayed the broader specificities with a preference for neutral and basic residues (4). *P. horikoshii*, and other related hyperthermophilic Thermococcales are distinguished by the fact that they possess three different TET complexes: PhTET1, a glutamyl/aspartyl aminopeptidase (9,17), PhTET2, a leucyl aminopeptidase with a broad activity against neutral amino-acids (8) and PhTET3, a lysyl aminopeptidase with a clear preference for positively charged residues (10). The analysis of their activities on synthetic peptides of different sizes and compositions using reverse phase HPLC indicated that the TET peptidases degrade oligopeptides in a sequential manner and displayed strict aminopeptidase behavior.

During the search for novel peptidase genes from the genome sequence of *P. horikoshii*, we found *PH0737* gene that encodes for new protein displaying about ~20% sequence identity with the three M42 self-compartmentalized peptidases PhTET1, PhTET2 and PhTET3 that co-exist in the cytosol of Thermococcales. The biophysical and biochemical characterizations presented in this work revealed that the *PH0737* gene product self-assembled as dodecameric complex with a macromolecular shape similar than the previously described TET complexes. Functional assays with monoacyl and peptide substrates revealed that it displayed aminopeptidase activity. The protein was therefore named PhTET4. To understand the biological function of this novel TET complex, we conducted biochemical characterizations of its optimal enzymatic conditions, preferred substrate and metal cofactor specificities. It revealed that, unlike other M42 aminopeptidases, PhTET4 is activated by nickel ions and is a strict glycyl aminopeptidase.

### EXPERIMENTAL PROCEDURES

#### *PhTET4 expression and purification*

The gene encoding for *PH0737* was cloned in pET-41c vector GeneCust Europe, Luxembourg. Recombinant protein was overexpressed in *Escherichia coli* BL21 (DE3)-RIL strain during 4 hours at 37 °C by induction with 0.1 mM IPTG in 1 L of lysogeny broth medium. The cell pellet was conserved at -80 °C until using. The cells were re-suspended in 25 ml Tris-HCl 50 mM pH 8, NaCl 90 mM and Triton X-100 0.1 %, supplemented with 6.25 mg lysozyme (Euromedex), 1.25 mg DNase I grade II (Roche), 5 mg RNase (Roche), 25 mg Pefabloc SC (Roche) and 0.25 ml MgSO<sub>4</sub> at 2 M. Cells were disrupted by sonication at 30 watts with 5 cycles on/off of 30 sec each at 4 °C, and then heated for 20 min at 75 °C to eliminate most mesophilic proteins of host strain. The lysate was clarified by centrifugation at 12,000 rpm, 4 °C for 1 hour with JA20 rotor (Beckman), and the supernatant was loaded on ResourceQ column (GE healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 100 mM NaCl. The 30 ml of the flow-through were retained and diluted with 20 ml Tris-HCl pH 7.5 to a final concentration of 105 mM NaCl. The new supernatant volume was loaded a second time on ResourceQ column. After washing the column with 20 mM Tris-HCl pH 7.5, 100 mM NaCl, bound proteins were eluted with linear salt gradient (154-290 mM NaCl). The protein-contained fractions were pooled together and loaded on Superdex200 10/300 GL size exclusion column (GE healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl. PhTET4 protein was eluted at 11 ml of column exclusion volume. The peak fractions were

## RÉSULTATS

---

combined and concentrated with Amicon Ultra 30 kDa cut-off and stored at 4 °C. The purity of prepared protein was checked by SDS-PAGE. 3 mg of PhTET4 were produced from 1 L of culture.

### *PhTET4 negative stain electron microscopy*

After the size exclusion chromatography step, 4 µl of PhTET4 at 0.05 mg/ml were deposited onto carbon-coated 400-mesh copper grids. The samples were stained using uranyl acetate 2 % and air-dried. Images were taken under low-dose conditions in a T12FEI electron microscope working at 120 kV and with a nominal magnification of 40,000 using an Orius SC1000 CCD camera.

### *PhTET4 substrate specificity*

The hydrolytic activity of PhTET4 was determined by using chromogenic and fluorogenic compounds: aminoacyl-pNAs (p-Nitroaniline) and aminoacyl-AMCs (7-Amino-4-methylcoumarin), respectively. Reactions were initiated by addition of 4 µg/ml (final concentration) of PhTET4 to 400 µl of pre-warmed mixture containing 5 mM chromogenic or fluorogenic substrate in 50 mM PIPES, 150 mM KCl, pH 7.5 at 80 °C. Since all tetrahedral aminopeptidases of *P. horikoshii* are activated by cobalt, we started by using 0.1 mM CoCl<sub>2</sub> as metal activator for PhTET4 enzymatic activity. To avoid water evaporation, the total volume was covered by a layer of mineral oil. Catalytic activity was followed during 10 min by measuring the absorbance of released pNA at 405 nm or the AMC fluorescence using excitation and emission wavelengths of 355 and 460 nm, respectively.

### *PhTET4 metallic cofactor*

In presence of cobalt low activity was identified against Gly-pNA. So, in order to enhance the enzymatic activity of PhTET4, several divalent metals were tested at 0.1 mM final concentration in a reaction volume containing 4 µg/ml PhTET4, 50 mM CHES, 150 mM KCl and 5 mM Gly-pNA, pH 9.5. The reaction was followed at 80 °C by measuring the absorbance of released pNA at 405 nm during 10 min.

### *PhTET4 optimal pH*

The effect of pH on PhTET4 enzymatic activity was studied by using different buffers: PIPES, pH 6-7.5; CHES, pH 8.2-10 and CAPS, pH 10.5-11. All buffers were used at 50 mM final concentration in presence of 4 µg/ml PhTET4, 150 mM KCl, 0.1 mM NiCl<sub>2</sub> and 5 mM Gly-

## RÉSULTATS

---

pNA at 80 °C. The incubation was done for 10 min by measuring the absorbance of released pNA at 405 nm.

### *PhTET4 optimal temperature*

The temperature impact on the enzymatic activity of PhTET4 was measured in a range from 20 to 95 °C. In all cases, 4 µg/ml of PhTET4 were incubated with 50 mM CHES, 150 mM KCl, 0.1 mM NiCl<sub>2</sub> and 5 mM Gly-pNA, pH 9.5 during 10 min. The reaction was assessed as described previously.

### *Peptide substrates*

More enzymatic studies were performed to decipher PhTET4 cleavage activity by using peptide substrates (GI; GL; GLM; GMDSLAFSGGL; LGG). 6 µg/ml (final concentration) of PhTET4 were added to a pre-warmed mixture of 3 mM peptide, 50 mM CHES, 150 mM KCl, 0.1 mM NiCl<sub>2</sub>, pH 9.5. To avoid water evaporation, 20 µl of mineral oil were added on the top of the total volume. The reaction incubation was done at 85 °C during 6 min. Aliquot of 80 µl was then removed and added to 220 µl of 2 % acetonitrile (ACN), 0.1 % trifluoroacetic acid (TFA). Proteins were removed by centrifugation at 13,000 rpm during 15 min. 100 µl of the supernatant were retrieved and injected on Nova-Pak C18 column, (4 µm, 3.9 x 300 mm, (Waters)) in a HPLC purifier system (PerkinElmer), equilibrated with 2 % ACN, 0.1% TFA. The elution of peptide products was achieved with a linear ACN gradient (2-33.2 %) and followed by measuring the absorbance at 214 nm. Chromatographic runs were carried out at room temperature. The separated fragments were collected and submitted to N-terminal sequence analysis.

### *PhTET4 inhibition*

The EDTA, Bestatin and Amastatin capacities to inhibit PhTET4 activity were studied. Respectively 5 mM, 1 mM or 0.2 mM of these inhibitors were added in optimal conditions (50 mM CHES, 150 mM KCl, 0.1 mM NiCl<sub>2</sub> and 5 mM Gly-pNA, pH 9.5, 85 °C) to PhTET4 reaction volume. PhTET4 activity was measured as described above. In this case, the heating temperature used was 85 °C instead of 95 °C for technical reasons.

Later PhTET4 was incubated in presence of 5 mM EDTA, 50mM CHES, 150 mM KCl, pH 9.5 at 85 °C during 15 min. Then, the oligomeric state of inhibited protein was evaluated by

using exclusion chromatography column (Superdex200) equilibrated with the same incubation buffer: 50mM CHES, 150 mM KCl, pH 9.5.

## RESULTS

### ***PH0737 gene product displays homology with the large dodecameric peptidases TET from *P. horikoshii****

An analysis of the genomes of Thermococcales revealed the existence of a conserved gene encoding for an unassigned peptidase of MH clan in MEROPS database (<http://merops.sanger.ac.uk/>). The protein encoded by *PH0737* shares 20.6 %, 22.5 % and 22 % of sequence identity with the three characterized aminopeptidases PhTET1, PhTET2 and PhTET3 respectively (**Figure 1**). The residues involved in the coordination of metal ions in M42 peptidase family are well conserved between PhTET1, 2, 3 and the new protein encoded by *PH0737* gene (6). Interestingly, two regions (the catalytic domain and the little  $\beta$  sheet domain localized on the top of the former) conferring the ability of M42 peptidases to form large multimers, are also present in the sequence of *PH0737* (9). These observations suggested that the product of *PH0737* gene could represent a paralogue to the three other TET peptidase complexes.

### ***PH0737 gene product forms tetrahedral dodecameric edifice***

Recombinant *PH0737* protein was produced in *E. coli*. The cellular extract was clarified by heat shock precipitation and the recombinant protein was purified by ion-exchange and size-exclusion chromatographies. At the final step, *PH0737* was eluted at the same exclusion volume (11 ml) than PhTETs homologues, suggesting that it forms large molecular weight assembly similar to the ~500 kDa TET dodecameric complexes (**Figure 2A**). Negative stain electron microscopy micrographs realized on *PH0737* peak fractions, showed that the protein self-assembled in a hollow tetrahedral-shaped complex of homogenous size (**Figure 2B**). The dimensions and the shape of *PH0737* are analogous to the three *P. horikoshii* tetrahedral TET edifices: PhTET1, PhTET2 and PhTET3 (9,10,18). The *PH0737* protein was therefore named PhTET4.

### ***PhTET4 is a nickel-activated peptidase***

To determine PhTET4 functional identity we tested its amidolytic activity toward the 20 amino acids by using a broad array of chromogenic p-nitroaniline (pNA) or fluorogenic 7-

## RÉSULTATS

---

amino-4-methylcoumarin (AMC) conjugated aminoacyl compounds. Since the 3 TET peptidases from *P. horikoshii* were found to be hyperthermophilic cobalt-activated aminopeptidases, we first assayed PhTET4 activity in the presence of 0.1 mM CoCl<sub>2</sub> as enzyme cofactor, at 80 °C and in identical buffer conditions (5 mM substrate, 50 mM PIPES, 150 mM KCl, pH 7.5). Surprisingly, no hydrolysis was observed against all tested substrates with the notable exception of Gly-pNA toward which PhTET4 exhibited a weak catalytic activity. Since metal cofactors have been shown to be essential to control the activity and the oligomeric state of the various TET edifices characterized so far (19), we tested the influence of various metal ions on PhTET4 glycylic aminopeptidase activity (**Figure 3A**). We revealed that Ni<sup>2+</sup> has an important stimulating effect on PhTET4 cleavage activity, with 12 times greater activity compared with control assay where no metallic ion (WM) was added to the reaction volume. Co<sup>2+</sup> and Mn<sup>2+</sup> also stimulated PhTET4 activity but less efficiently than Ni<sup>2+</sup> (3 and 1.4 fold activation, respectively). Interestingly, Zn<sup>2+</sup>, Ca<sup>2+</sup>, Fe<sup>2+</sup> and Mg<sup>2+</sup> were found to inhibit PhTET4 hydrolytic activity, and total inhibition was observed in presence of Ca<sup>2+</sup> ions. This is the first time that Ni<sup>2+</sup> ions have been described as an essential activating cofactor of an aminopeptidase from the M42 family.

### *PhTET4 is an alkaline hyperthermophilic peptidase*

In order to determine the influence of pH on PhTET4 enzymatic behavior, the amidolytic activity was measured between pH 6 and pH 11, against Gly-pNA at 80 °C (**Figure 3B**). The optimal activity was found at pH 9.5 and a significant percentage of activity was maintained at elevated pH. These experiments revealed that, compared to the other 3 PhTETs enzymes, PhTET4 can be defined as an alkaline peptidase.

To assess the hyperthermophilic properties of the enzyme, the temperature dependence of PhTET4 activity was studied at different temperatures varying from 20 to 95 °C. PhTET4 enzymatic activity increases in parallel with the augmentation of heating temperature, with a maximal activity observed at 95 °C in the measurable temperature range (**Figure 3C**). Thus PhTET4 displays a high hyperthermophilic behavior comparable to the ones reported for the 3 other PhTETs aminopeptidases present in *Pyrococcus horikoshii* cells (8-10).

### *PhTET4 is a strict glycylic aminopeptidase*

The initial characterization of PhTET4 aminopeptidase activity indicated that the enzyme displayed narrow substrate specificity with a strong preference for glycine residues. In order

## RÉSULTATS

---

to consolidate this finding, the experiments were repeated in the presence of nickel and in the optimal temperature and pH conditions defined above (0.1 mM NiCl<sub>2</sub>, pH 9.5 and at 85 °C). The result showed unambiguously that the enzyme acted only on Gly-pNA (**Figure 4**). No hydrolytic activity could be detected toward all other amino acids, even with long incubation times. We also tried to investigate if PhTET4 exhibits high D-stereospecificity against D-Alanine as shown for *Aspergillus oryzae* glycine aminopeptidase (20). For this, D-Ala-pNA was used as chromogenic substrate in optimal activity conditions (0.1 mM NiCl<sub>2</sub>, pH 9.5 and 85 °C). The experiment showed that PhTET4 is unable to cleave residue alanine in D-conformation, thus demonstrating that PhTET4 is devoid from D-stereospecificity.

In aminopeptidases, the catalytic activities and specificities can be affected by the length and N-terminal amino acid composition of the peptide substrates (21-25). Consequently, we tested if PhTET4 maintains its narrow specificity toward glycine residues in a peptide context. For this, we measured PhTET4 capacity to cleave N-terminal residue of the following peptides: GI; GL; GLM; GMDSLAFSGGL and LGG. After incubation of PhTET4 with the peptide substrates in optimal activity conditions, the reaction products were separated by reverse phase HPLC and identified by N-terminal sequencing. These experiments showed that no enzymatic activity was detected against peptide that do not start by glycine residue even if a glycine residue is present at P1' position as demonstrated with the LGG tripeptide. The HPLC chromatographic profiles of the degradation of GLM and GMDSLAFSGGL peptides are shown on **Figure 5**. The sequences of the detected accumulating peptides were determined and the results clearly demonstrated that PhTET4 does not exert amidolytic activity beyond the N-ter glycine in a peptide context. To assess if PhTET4 can process consequently several glycine residues in a peptide sequence, we tested the enzymatic activity against the chromogenic peptide Gly-Gly-pNA in the same conditions as described previously. PhTET4 displayed significant activity against this substrate, corresponding to 10 % of the total activity exhibited in presence of monoacyl compound Gly-pNA. Taken together, these experiments clearly mark PhTET4 as an aminopeptidase strictly specialized in the hydrolysis of N-terminal glycine residues.

### ***PhTET4 is a metallo-aminopeptidase***

In case of PhTET1, 2 and 3 aminopeptidases, it has been shown that the two metal ions present in the catalytic site are essential both for catalysis and for assembling the dodecameric edifice (19,26). For this reason treatment with the divalent ions chelating agent EDTA leads



## RÉSULTATS

---

to the dissociation of the TET quaternary structures and to enzymes inactivation. Indeed, in the optimal conditions for PhTET4 activity, the addition of 5 mM EDTA completely inhibited the glycine aminopeptidase activity. However, analysis of the oligomeric state of EDTA-treated PhTET4 samples by size-exclusion chromatography revealed that the PhTET4 dodecameric edifice remained unaffected by the EDTA treatment unlike what was reported for all the other TET enzymes (**Figure 6**). This suggests that the contribution of the metals ions situated in position M1 and/or M2 for PhTET4 oligomerization is not as important as for the other TET peptidases.

### *Effects of Amastatin and Bestatin on PhTET4 activity*

Amastatin and Bestatin represent typical aminopeptidase and leucyl aminopeptidase inhibitors, respectively. The addition of 0.2 mM Amastatin and 1 mM Bestatin reduced the amidolytic activity of PhTET4 against Gly-pNA to 78 and 73 % of the activity detected in the control experiment. These experiments showed that these inhibitors have a moderate effect on PhTET4 compared to what was reported for the other TET peptidases, especially PhTET3 which is more sensible to these inhibitors, but like PhTET4 is more inhibited by Amastatin rather than Bestatin.

## DISCUSSION

*Pyrococcus horikoshii* is a deep-sea hyperthermophilic archaeon that was isolated at 1395 m depth from a hydrothermal vent (27). As other hyperthermophilic organisms, its genome encodes for numerous peptidases (28). In particular, three different versions of the ~500 kDa dodecameric M42 TET peptidase were found to co-exist in its cytosol (8-10). Here, we characterized a fourth member of the TET family in the same organism that we named PhTET4. Despite a weak sequence similarity with other PhTETs enzymes, the protein complex was found to share the same typical tetrahedral architecture with the other TET dodecamers previously characterized in archaea, bacteria and eukaryotes. However, its biochemical characterization revealed important, unsuspected properties that clearly distinguished PhTET4 from all known TET complexes.

The dinuclear metal site is a key factor for the regulation of TET function. Previous studies revealed that it is involved both in the stability of the dodecameric edifice and in the fine-tuning of TET catalytic activity (19). The M42 canonical residues involved in metal ions binding are well conserved in PhTET4 sequence. This suggests that PhTET4 has a similar

## RÉSULTATS

---

catalytic mechanism to other members of the M42 family. X-ray crystallography and absorption (EXAFS) experiments have shown that the dinuclear metal center of TET can be occupied by  $Zn^{2+}$ ,  $Co^{2+}$  and  $Mn^{2+}$  ions (10,15,18). In previously characterized archaeal, bacterial and eukaryotic TET aminopeptidases,  $Co^{2+}$  and  $Mn^{2+}$  ions strongly activate the amidolytic activities when they are added to the reaction mixture (8-13). In archaea,  $Co^{2+}$  is the main activating metal of the 3 TET enzymes from *P. horikoshii*, while  $Mn^{2+}$  only activated the DkamTET enzyme from *Desulfurococcus kamchatkensis* (29). Here we report for the first time an aminopeptidase from the M42 family to be strongly activated by nickel. This is an interesting finding since, in metallo-aminopeptidases, nickel is rarely the main activation activator. To our knowledge this was only reported in methionine aminopeptidases such as the MtbMAP A from *Mycobacterium tuberculosis* (30).

We previously elucidated the molecular basis of the metal activation process in the TET3 enzyme from *P. furiosus* (19). The molecular basis of  $Co^{2+}$  activation was depicted by combining X-ray crystallography, UV-spectroscopy combined and enzymatic assays in the presence and absence of  $Zn^{2+}$ ,  $Co^{2+}$  and  $Mn^{2+}$ . In TET monomers, the active site hosts two metal ions (M1 and M2) forming dinuclear metal center. Our studies revealed that  $Co^{2+}$  activates PfTET3 by specifically replacing the M1 catalytic site, and not the M2 site, thus increasing the activity of PfTET3 through the stiffening of the active site. In PhTET4 it can be inferred that nickel activation is due to a similar mechanism. The dinuclear metal center was also found to play a structural role in TET (26). In the presence of high concentration of metal chelating agents such as EDTA and in basic pH conditions, TET dodecamers dissociate into monomers (26,31). The destabilization was attributed to the M1 site, as the residues implicated in the coordination of M1 metal ion, lay on loops that contribute in hydrogen bonds between adjacent subunits and mediate dimeric and oligomeric interfaces. Surprisingly, while the enzymatic activity of PhTET4 is inhibited upon EDTA treatments, its quaternary structure remained unaffected. This behavior suggests that the PhTET4 interfaces must be different than the ones of the other TET complexes to compensate a less important contribution of the metal center in the TET assembling process. The determination of PhTET4 crystal structure will be necessary to demonstrate this and to identify the structural determinants underlying the M1 site nickel specificity.

The most striking feature of PhTET4 revealed in this work is its exclusive amidolytic activity against glycine residues that justified categorizing this novel M42 peptidase as a glycyI aminopeptidase (GAP). This unexpected result underscore the fact that peptidase

## RÉSULTATS

---

specificity are hardly predicable from sequence analysis and that experimental work is important to unravel novel enzymatic properties arising from microbial biodiversity. Examples of GAP enzymes are rare. While aminopeptidases from different families are capable of hydrolyzing glycine residues, their activities remain weak by comparison with their preferred substrates. To our knowledge, only three aminopeptidases were found to exhibit clear preference for glycine residues. The first one is a Zn-dependent metallopeptidase from M61 family secreted by the gram-negative bacteria *Sphingomonas capsulata* (32). The second one is a eukaryotic S12 family serine peptidase found in the cytosol of *Actinomucor oryzae* (20). The third one is the cytosolic glycyl aminopeptidase of *Actinomucor elegans*, for which the residues implicated in enzymatic mechanism still ambiguous. PhTET4 is therefore the first GAP enzyme identified in archaea. Moreover, we showed that it displayed interesting thermophilic and alkaliphilic properties. It is also noticeable that, unlike PhTET4, the bacterial and eukaryotic TET enzymes did not display strict specificity for glycine residues. PhTET4 is consequently the first known aminopeptidase displaying an apparent exclusive activity toward N-terminal glycine residues. There are no available structures of GAP enzyme to explain the structural bases of glycine specificity. In TET complexes, the substrate specificity is mostly due to differences in charge surface properties in the “S1” specificity pocket. In PhTET1, TET2 and TET3 the crystal structures revealed that the electrostatic potential is adapted to welcome negatively charged, neutral or positively charged side chain residues, respectively (9,10,18). Glycine is a peculiar amino acid due to its small size and the absence of lateral chains. Therefore, one can hypothesize that the PhTET4 specificity pocket would display appropriate dimensions to exclude any type of amino acid except glycine.

The biological significance of specialized TET aminopeptidase is still not well understood. In the case of Thermococcales, the presence of different TET complexes displaying contrasted substrate preferences can be interpreted in diverse ways. As already suggested, the TET may act cooperatively to hydrolyze efficiently intracellular polypeptides arising from the ATP-dependent protein degradation machines such as the 20S proteasome or Lon (10,18). In the case of heterotrophic archaea, such as *P. horikoshii*, a combination of streamlined aminopeptidase activities could also provide an advantage to hydrolyze peptides imported from the extracellular environment as a carbon source (10). In this context, GAP activity may represent an advantage, as glycine is an arduous residue for hydrolysis due to the flexibility that it introduced in the polypeptide chains and to the absence of lateral chains (33). This could be particularly important in the physiological extreme temperature conditions that may

## RÉSULTATS

---

provide additional flexibility. However, the presence of GAP enzymes cannot be a crucial condition for living at high temperature as other hyperthermophilic strains only contain one type of TET. It is interesting to remark that hyperthermophilic organisms containing a unique TET protein, such as *Methanocaldococcus jannaschii*, are autotrophic thus suggesting a link between heterotrophy and TET diversity.

However, there are also arguments supporting a specific physiological role for PhTET4 in *P. horikoshii*. This was already proposed for PhTET1, as this enzymatic complex specialized in the hydrolysis of aspartate and glutamate, is not accumulated in cells at the same time as PhTET2 and PhTET3 dodecamers (34). Moreover the hetero-oligomeric complex PhTET2-3 is capable of hydrolyzing a wide variety of substrate, including glycine, without the help of PhTET4 (34). This raises the prospect that the PhTET4 GAP activity could be devoted to peculiar cellular functions. Indeed, glycine represents a thermoprotectant for enzymes and is also an important precursor amino acid residue for the synthesis of osmoprotectant such as trimethylglycine (35). Beside this, glycine is a precursor for glutathion, an important antioxidant anti metal detoxification precursor in eukaryotic cells (36), and for glyoxylate that, when metabolized, generates oxalate and formate, two antioxidants molecules (37). *P. horikoshii* lives in a strictly anaerobic condition but the high temperature combined with the presence of heavy metals are sources of free radicals that are responsible for intense DNA damages. It is therefore tempting to speculate that PhTET4 could participate in stress response under certain conditions encountered by the archaea.

## ACKNOWLEDGMENTS

We acknowledge financial support from the Agence Nationale de la Recherche (grant “Archelyse” ANR-12-BSV8-0019-0). Hind Basbous was supported by the CEA PhD program IRTELIS. This work used the platforms of the Grenoble Instruct Center (ISBG: UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

FIGURES

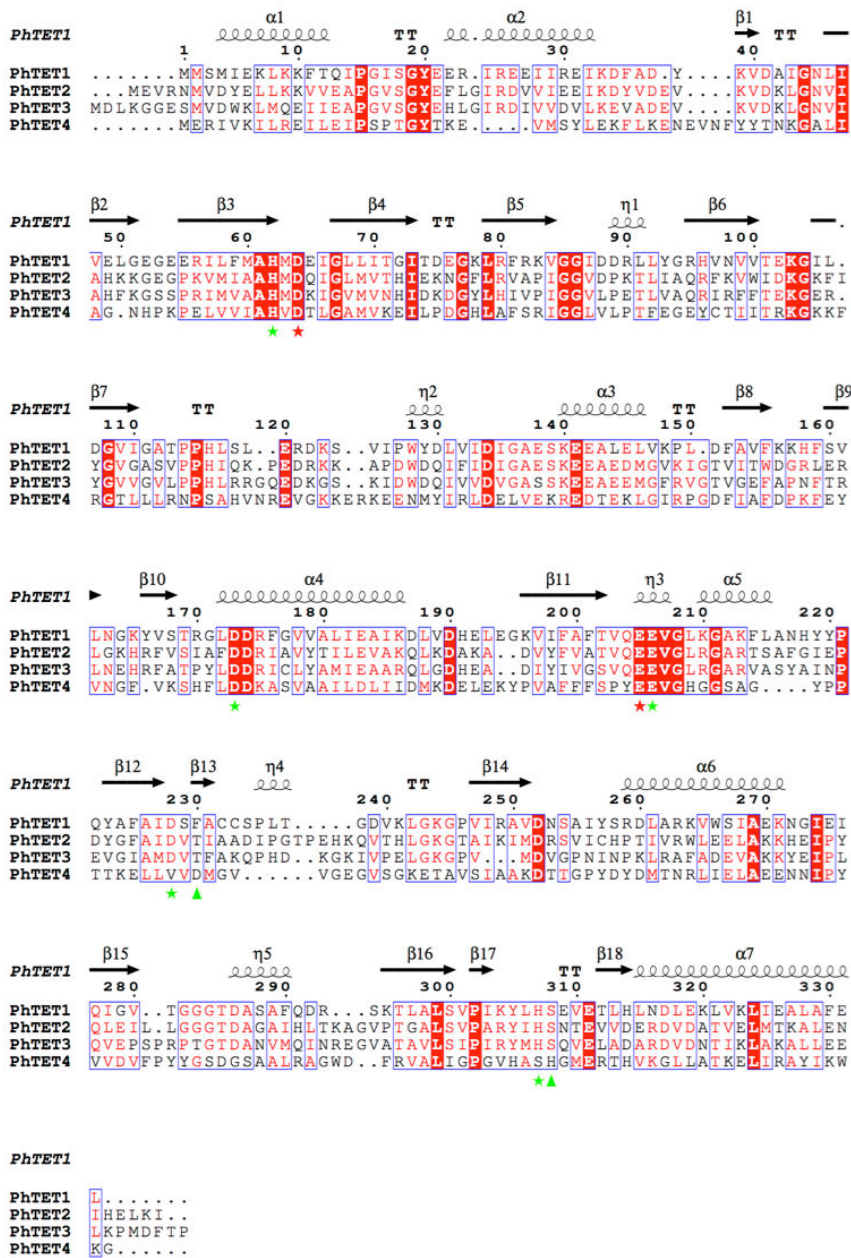
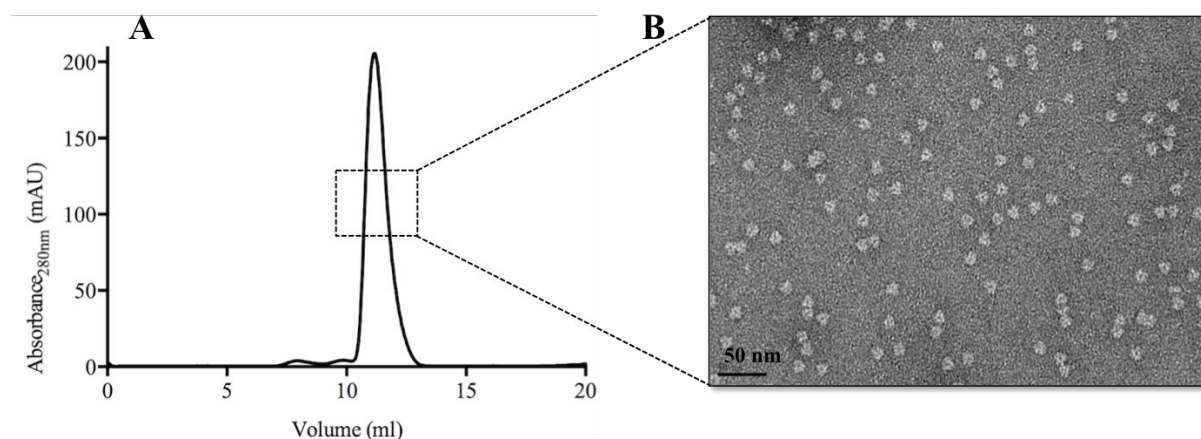
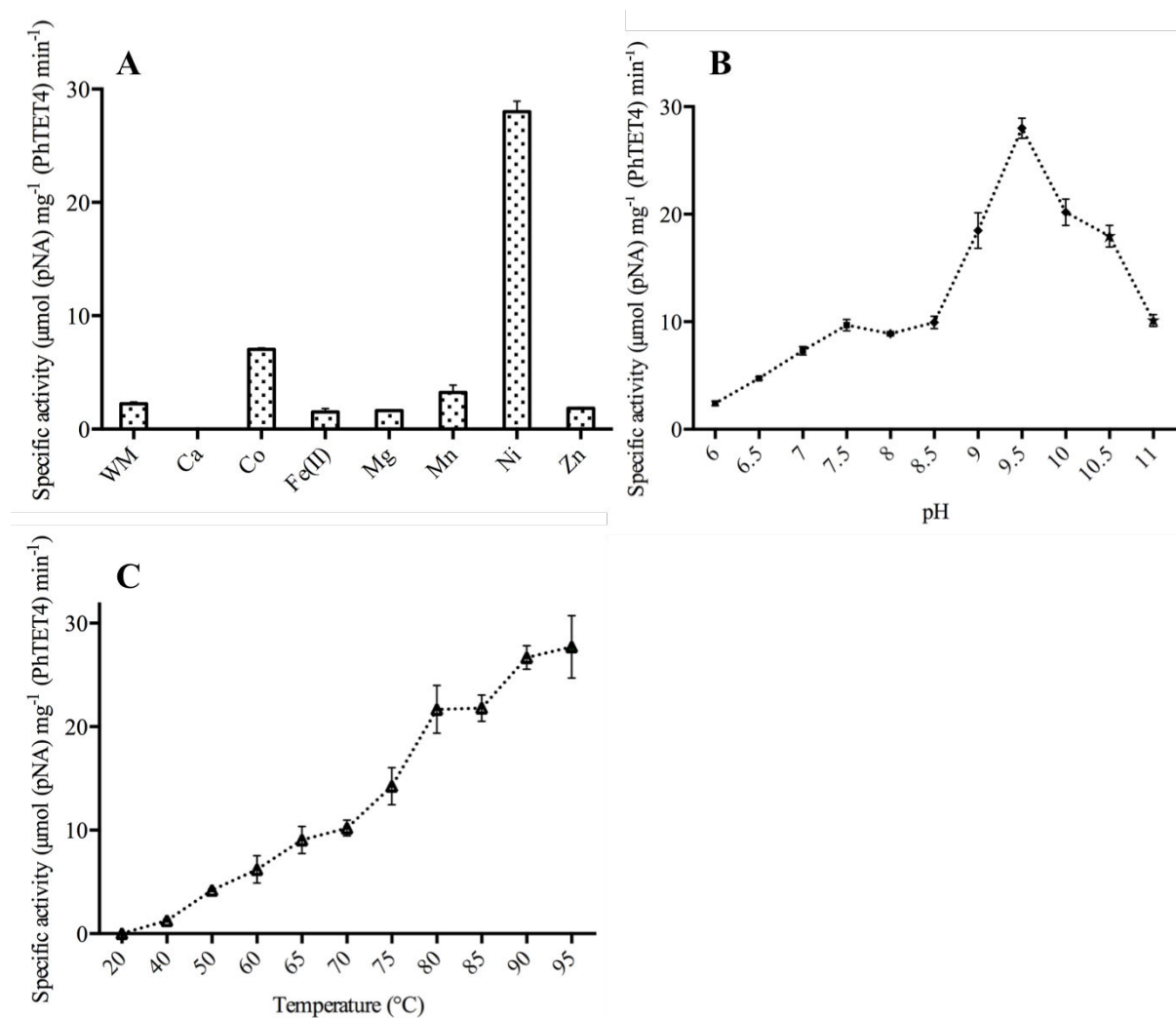


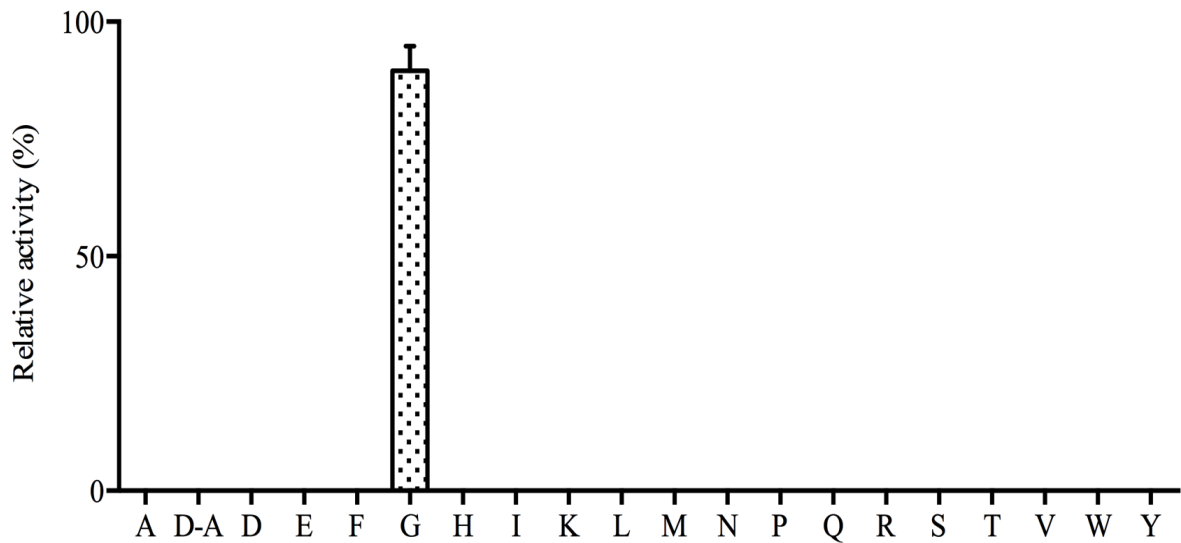
Figure 1: Multiple sequence alignment of the four *P. horikoshii* PhTETs: PhTET1, PhTET2, PhTET3 and PhTET4 as determined by ESPrpt 3.0 (38). The numbering and the secondary structure elements are those of PhTET1 (PDB ID: 2WYR) (9). Green and red stars highlight the metal-binding and active residues, respectively. Most likely, a shift was occurred for the two latest putative ligands of PhTET4, Asp231 and His311, which are highlighted by green triangle in comparison with the conserved position of PhTET1, 2 and 3 ligands.



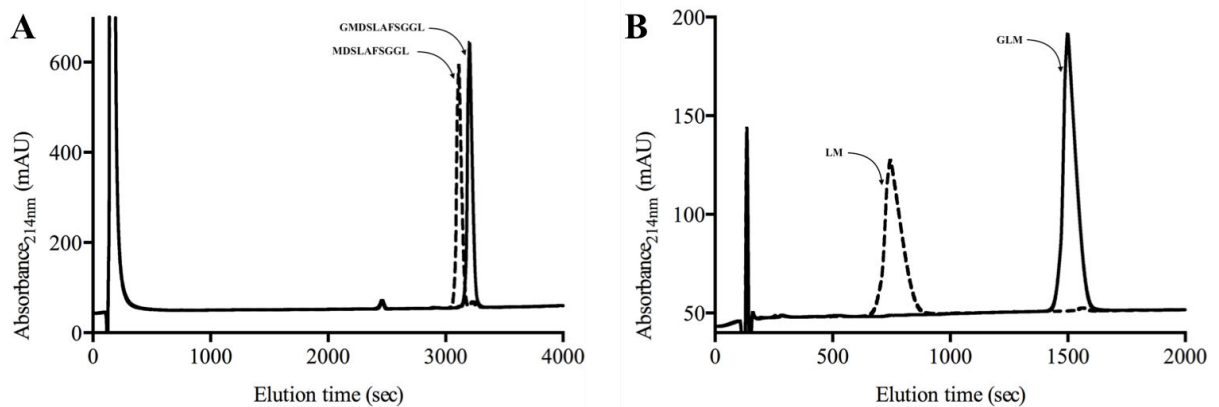
**Figure 2: PhTET4 exhibits dodecameric state in solution.** **A.** Elution profile of PhTET4 on size-exclusion chromatography column (Superdex200 10/300 GL). The recombinant protein was eluted at 11 ml of column exclusion volume. **B.** PhTET4 micrographs observed by negative stain electron microscopy. Homogenous populations of hollow tetrahedral-shaped particles are observed at 49000X.



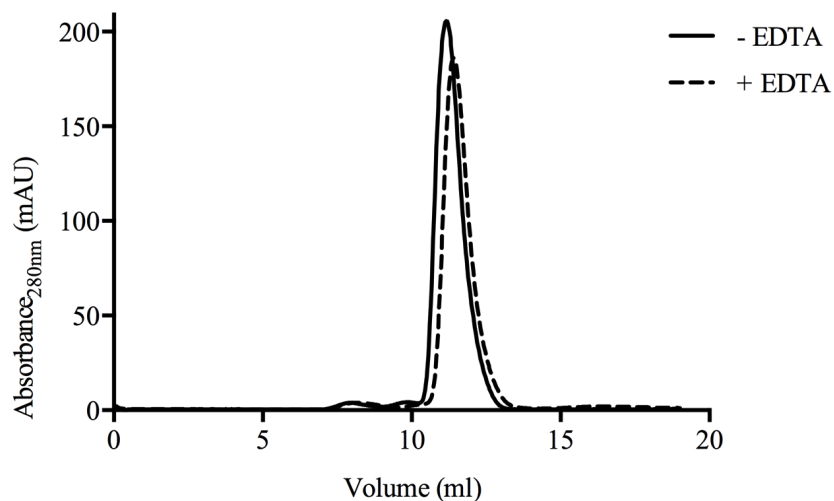
**Figure 3: Effect of metals, pH and temperatures on PhTET4 activity against Gly-pNA used as substrate.** **A.** Variation of PhTET4 specific activity in absence (WM: without metal) or presence of several metal ions. **B.** Evolution of PhTET4 specific activity as function of pH ( $\blacksquare$ ) PIPES, ( $\blacklozenge$ ) CHES, ( $\star$ ) CAPS). **C.** Evolution of PhTET4 specific activity in response to temperature variation.



**Figure 4: PhTET4 is a strict glycyl aminopeptidase.** Only glycyl aminopeptidase activity was found. In these conditions, PhTET4 activity against (H-Cys-pNA)<sub>2</sub> can't be measured because it precipitated in our enzymatic conditions.



**Figure 5: Chromatographic profiles of peptide degradation by PhTET4.** The sequences of peptides are indicated. In presence of PhTET4 (dashed line) only glycine residue was removed from the original peptide (full line). **A.** Degradation profile of GMDSLAFSGGL. **B.** Degradation profile of GLM.



**Figure 6: Dodecameric state of PhTET4 remains unchanged even in presence of EDTA.** Size exclusion chromatography of PhTET4 incubated at optimal conditions of enzymatic activity (0.1 mM Ni<sup>2+</sup>, pH 9.5, 85 °C) without EDTA (full line) in comparison with PhTET4 supplemented with 5 mM EDTA (dashed line). In both conditions, PhTET4 is present as dodecameric assemblies.



### REFERENCES

1. Lowther, W. T., and Matthews, B. W. (2002) Metalloaminopeptidases: common functional themes in disparate structural surroundings. *Chemical reviews* **102**, 4581-4608
2. Gonzales, T., and Robert-Baudouy, J. (1996) Bacterial aminopeptidases: properties and functions. *FEMS Microbiol Rev* **18**, 319-344
3. 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
4. Franzetti, B., Schoehn, G., Hernandez, J. F., Jaquinod, M., Ruigrok, R. W., and Zaccai, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *Embo J* **21**, 2132-2138.
5. Appolaire, A., Rosenbaum, E., Dura, 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. *J Biol Chem* **288**, 22542-22554
6. Rawlings, N. D., Barrett, A. J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **40**, D343-350
7. Gomis-Ruth, F. X. (2008) Structure and mechanism of metallo-carboxypeptidases. *Crit Rev Biochem Mol Biol* **43**, 319-345
8. Dura, 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
9. Schoehn, G., Vellieux, F. M., Asuncion Dura, M., Receveur-Brechot, V., Fabry, C. M., Ruigrok, R. W., Ebel, C., Roussel, A., and Franzetti, B. (2006) An archaeal peptidase assembles into two different quaternary structures: A tetrahedron and a giant octahedron. *J Biol Chem* **281**, 36327-36337
10. Dura, M. A., Rosenbaum, E., Larabi, A., Gabel, F., Vellieux, F. M., 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. *Mol Microbiol* **72**, 26-40
11. 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. *Biochem Biophys Res Commun* **391**, 431-436
12. Dutoit, R., Brandt, N., Legrain, C., and Bauvois, C. (2012) Functional characterization of two M42 aminopeptidases erroneously annotated as cellulases. *PloS one* **7**, e50639
13. 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. *Biochem Biophys Res Commun* **447**, 101-107
14. Chen, Y., Farquhar, E. R., Chance, M. R., Palczewski, K., and Kiser, P. D. (2012) Insights into substrate specificity and metal activation of Mammalian tetrahedral aspartyl aminopeptidase. *J Biol Chem* **287**, 13356-13370
15. Chaikuad, A., Pilka, E. S., Riso, A. D., Delft, F. V., Kavanagh, K. L., Venien-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 Struct Biol* **12**, 14
16. Robinson, M. W., Buchtmann, K. A., Jenkins, C., Tacchi, J. L., Raymond, B. B., To, J., Roy Chowdhury, P., Woolley, L. K., Labbate, M., Turnbull, L., Whitchurch, C. B.,

- Padula, M. P., and Djordjevic, S. P. (2013) MHJ\_0125 is an M42 glutamyl aminopeptidase that moonlights as a multifunctional adhesin on the surface of *Mycoplasma hyopneumoniae*. *Open biology* **3**, 130017
17. Franzetti, M. A. D. a. B. (2013) PhTET1 Aminopeptidase. in *In Neil D. Rawlings and Guy S. Salvesen, editors: Handbook of Proteolytic Enzymes, Oxford: Academic Press.* pp pp. 1638 - 1645
  18. Borissenko, L., and Groll, M. (2005) Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J Mol Biol* **346**, 1207-1219
  19. 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
  20. Marui, J., Matsushita-Morita, M., Tada, S., Hattori, R., Suzuki, S., Amano, H., Ishida, H., Yamagata, Y., Takeuchi, M., and Kusumoto, K. (2012) Enzymatic properties of the glycine D-alanine [corrected] aminopeptidase of *Aspergillus oryzae* and its activity profiles in liquid-cultured mycelia and solid-state rice culture (rice koji). *Applied microbiology and biotechnology* **93**, 655-669
  21. Singh, H., and Kalnitsky, G. (1980) alpha-N-benzoylarginine-beta-naphthylamide hydrolase, an aminoendopeptidase from rabbit lung. *J Biol Chem* **255**, 369-374
  22. Koldamova, R. P., Lefterov, I. M., Gadjeva, V. G., and Lazo, J. S. (1998) Essential binding and functional domains of human bleomycin hydrolase. *Biochemistry* **37**, 2282-2290
  23. Geier, E., Pfeifer, G., Wilm, M., Lucchiari-Hartz, M., Baumeister, W., Eichmann, K., and Niedermann, G. (1999) A giant protease with potential to substitute for some functions of the proteasome. *Science* **283**, 978-981.
  24. Turk, D., Janjic, V., Stern, I., Podobnik, M., Lamba, D., Dahl, S. W., Lauritzen, C., Pedersen, J., Turk, V., and Turk, B. (2001) Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J* **20**, 6570-6582
  25. Chandu, D., and Nandi, D. (2003) PepN is the major aminopeptidase in *Escherichia coli*: insights on substrate specificity and role during sodium-salicylate-induced stress. *Microbiology* **149**, 3437-3447
  26. Rosenbaum, E., Ferruit, M., Dura, 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. *Biochim Biophys Acta* **1814**, 1289-1294
  27. Gonzalez, J. M., Masuchi, Y., Robb, F. T., Ammerman, J. W., Maeder, D. L., Yanagibayashi, M., Tamaoka, J., and Kato, C. (1998) *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles* **2**, 123-130
  28. Ward, D. E., Shockley, K. R., Chang, L. S., Levy, R. D., Michel, J. K., Connors, S. B., and Kelly, R. M. (2002) Proteolysis in hyperthermophilic microorganisms. *Archaea* **1**, 63-74
  29. Petrova, T. E., Slutskaia, 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
  30. Narayanan, S. S., and Nampoothiri, K. M. (2012) Biochemical characterization of recombinant methionine aminopeptidases (MAPs) from *Mycobacterium tuberculosis* H37Rv. *Molecular and cellular biochemistry* **365**, 191-202

31. Appolaire, A., Girard, E., Colombo, M., Dura, M. A., Moulin, M., Hartlein, 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
32. Jamdar, S. N. (2009) A novel aminopeptidase from Burkholderia cepacia specific for acidic amino acids. *FEMS Microbiol Lett* **295**, 230-237
33. Byun, T., Tang, M., Sloma, A., Brown, K. M., Marumoto, C., Fujii, M., and Blinkovsky, A. M. (2001) Aminopeptidase from Sphingomonas capsulata. *J Biol Chem* **276**, 17902-17907
34. Appolaire, A., Dura, M. A., Ferruit, M., Andrieu, J. P., Godfroy, A., Gribaldo, S., and Franzetti, B. (2014) The TET2 and TET3 aminopeptidases from Pyrococcus horikoshii form a hetero-subunit peptidasome with enhanced peptide destruction properties. *Mol Microbiol* **94**, 803-814
35. Canovas, D., Vargas, C., Csonka, L. N., Ventosa, A., and Nieto, J. J. (1996) Osmoprotectants in Halomonas elongata: high-affinity betaine transport system and choline-betaine pathway. *J Bacteriol* **178**, 7221-7226
36. Briat, J. F., and Lebrun, M. (1999) Plant responses to metal toxicity. *C R Acad Sci III* **322**, 43-54
37. Alhasawi, A., Castonguay, Z., Appanna, N. D., Auger, C., and Appanna, V. D. (2015) Glycine metabolism and anti-oxidative defence mechanisms in Pseudomonas fluorescens. *Microbiol Res* **171**, 26-31
38. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324

### 5.2.3. État des connaissances sur les glycyL aminopeptidases

La peptidase PhTET4 caractérisée dans ce travail de thèse s'est révélée être une Glycyl aminopeptidase stricte (GAP). Il m'a paru important de consacrer un chapitre à l'état des connaissances sur ces enzymes encore très mal connues.

Beaucoup d'aminopeptidases sont capables d'hydrolyser de façon non spécifique les résidus glycines. Néanmoins, chez toutes ces aminopeptidases, l'efficacité d'hydrolyse de la glycine reste très faible (aux alentours de 18 %) en comparaison avec leurs substrats préférentiels. Citons par exemple : la leucine aminopeptidase MHJ\_0461 de *Mycoplasma hyopneumoniae* (Jarocki et al., 2015), l'aminopeptidase de *Bacillus subtilis* (BSAP) (Gao et al., 2013), la méthionine aminopeptidase de *Mycobacterium tuberculosis* (Narayanan and Nampoothiri, 2012), l'aminopeptidase (Pf-M1AAP pour alanine aminopeptidase) de *Plasmodium falciparum* (Poreba et al., 2012), la phénylalanine aminopeptidase (APF1) de *Schizophyllum commune* (Bilbrey et al., 1996) ainsi que les deux TETs de *Pyrococcus horikoshii* (PhTET2 et PhTET3) (Durá et al., 2005, 2009)... Il existe cependant 3 aminopeptidases qui ont été décrites comme des glycyL aminopeptidases (GAP), c'est à dire des peptidases possédant une forte activité ou une préférence pour les résidus glycines : une sécrétée dans le milieu externe par la bactérie à Gram négatif, *Sphingomonas capsulata* (Byun, 2001) et deux autres localisées dans le cytosol de deux espèces de champignons : *Actinomucor elegans* et *Aspergillus oryzae* (Ito et al., 2003; Ma et al., 2004; Marui et al., 2012).

C'est en criblant une diversité de microorganismes pour rechercher des activités aminopeptidases d'intérêt que, Ito et ses collègues ont réussi à détecter une activité glycyL aminopeptidase chez *A. elegans* (Ito et al., 2003), une espèce de champignon utilisée pour la production du tofu. La protéine (MEROPS ID : MER0401795, à vérifier) a été produite et purifiée à partir des extraits totaux de *A. elegans* IFO 6408 (Ma et al., 2004). La séquence peptidique de cette protéine reste cependant inconnue. Ceci n'est pas le cas pour l'autre type de glycine aminopeptidase eucaryote de *A. oryzae* (GdaA). Elle a ainsi pu être produite sous forme recombinante (MER032669) après une surexpression dans un mutant de *A. oryzae* (Marui et al., 2012). La glycyL aminopeptidase de la bactérie *S. capsulata* a été produite aussi bien sous forme native que sous forme recombinante (MER014496) après une surexpression dans *Bacillus subtilis* (Byun, 2001).

### 5.2.3.1. États oligomériques

Les états oligomériques des GAP eucaryotes de *A. elegans* et *A. oryzae* ont été évalués par chromatographie d'exclusion de taille (Ito et al., 2003; Marui et al., 2012). La première semble être organisée sous la forme d'un hexamère alors que la seconde est un dimère. Quant à GAP bactérienne de *S. capsulata*, son état oligomérique n'est pas décrit.

### 5.2.3.2. Spécificités enzymatiques

Les trois aminopeptidases ScGAP, AeGAP et GdaA de *S. capsulata*, *A. elegans* et *A. oryzae*, respectivement, présentent une spécificité catalytique accrue contre la glycine (activité relative de 100 %). Chez GAP de *A. elegans*, l'activité glycyll aminopeptidase est associée à une activité beaucoup plus réduite contre l'alanine (L-Ala) (18 %), la méthionine et l'arginine (6 %), la serine (3,5 %) et leucine (2,6 %) (Ito et al., 2003). Quant à GAP de *A. oryzae* (ou GdaA), cette aminopeptidase exprime une stéréospécificité élevée pour la D-Ala (93 %), contre une spécificité plus réduite pour la L-Ala (12 %) (Marui et al., 2012). La glycyll aminopeptidase de *S. capsulata* possède une spécificité catalytique particulière. Sur des substrats monoacyls couplés au 4-Nitroaniline (pNA), elle possède une préférence pour l'alanine (100 %) et ne présente que 14 % d'activité contre la glycine. En revanche, cette préférence est complètement renversée dans un contexte peptidique où la glycine est efficacement clivée par rapport à l'alanine (Byun, 2001). De plus, AeGAP et GdaA sont aussi capables de libérer les résidus glycines à partir des extrémités Nt des substrats peptidiques (di et tripeptides) tels que Gly-Gly où la liaison peptidique est connue pour être résistante à la protéolyse. L'efficacité de ce clivage peut être affectée par la nature de l'acide aminé situé en position P1' (Marui et al., 2012). ScGAP présente un optimum d'activité à pH neutre (7,4), alors que AeGAP et GdaA expriment une activité maximale à pH alcalin (8 et 8,5 respectivement). La température optimale pour ces trois glycyll aminopeptidases est de l'ordre de 40 °C.

### 5.2.3.3. Propriétés catalytiques

La caractérisation de la GAP bactérienne, ScGAP a permis l'identification d'un atome de zinc par molécule de protéine associé à un motif HEXXH, caractéristique des métallo-enzymes à zinc. Du plus, l'ajout de la 1,10-phénantroline, un chélateur d'ions métalliques, au milieu réactionnel induit une inhibition de cette aminopeptidase. Sur la base de ces résultats

## RÉSULTATS

---

Byun et ses collègues ont proposé que ScGAP est une métallo-enzyme à zinc (Byun, 2001), ce qui est en accord avec la classification MEROPS qui la considère d'après les homologies de séquences et la conservation du site actif comme étant une enzyme appartenant à la famille M61 du clan MA des métallopeptidases.

Les études enzymatiques entreprises sur GdaA de *A. oryzae* ont montré qu'il s'agit d'une aminopeptidase à sérine. En effet, elle possède un motif Ser-Xaa-Xaa-Lys au niveau de sa séquence protéique et son activité est abolie par les inhibiteurs des protéases à sérine, notamment le AEBSF (*4-(2-AminoEthyl) BenzeneSulfonyl Fluoride hydrochloride*) (Marui et al., 2012). Ceci est en accord avec sa classification dans MEROPS comme aminopeptidase appartenant à la famille S12 du clan SE des peptidases à sérine (<http://merops.sanger.ac.uk>).

Le mécanisme catalytique adopté par l'autre GAP eucaryote, AeGAP, reste quant à lui encore ambigu du fait de l'absence de séquence protéique. Les tests d'activité réalisés par Ito, Ma et leurs collègues montrent une inhibition de l'enzyme par des inhibiteurs des thiols tels que l'acide iodoacétique et le PCMB (*p-ChloroMercuriBenzoic acid*), suggérant la présence de groupement thiol dans le site actif. Cela a permis de supposer que la glycine aminopeptidase de *A. elegans* est une aminopeptidase à cystéine (Ito et al., 2003; Ma et al., 2004). Dans MEROPS, cette espèce ne présente que deux peptidases, une de 106 kDa et l'autre de 49 kDa. Pourtant, le poids moléculaire de la peptidase comme décrit par Ito et ses collègues, est de 56,5 kDa, légèrement supérieur à celui de la protéine présente dans MEROPS et qui est une peptidase de la famille S8 du clan SB des peptidases à sérine.

### 5.2.3.4. Rôles supposés des glycyll aminopeptidases

L'existence d'aminopeptidases dirigées plus spécifiquement contre les résidus glycines peut avoir différentes raisons. Tout d'abord, la glycine, avec la proline, constitue un acide aminé difficilement hydrolysable (Byun, 2001). Ceci est dû à une absence de chaîne latérale et à une flexibilité accrue. Comme c'est le cas pour la proline, on peut penser que des peptidases qui reconnaissent la glycine ont été développées par certains organismes dans le but de faciliter une dégradation totale des peptides riches en glycine en acides aminés libres (Marui et al., 2012). Par exemple, l'hydrolyse de la gélatine, dont la glycine représente l'acide aminé dominant (~21 %), par la protéinase alcaline de *Bacillus coagulans* est améliorée lorsque la glycyll aminopeptidase de *A. elegans* est ajoutée au milieu réactionnel (Ma et al., 2004). Marui et ses collègues ont montré que l'activité cytosolique de GdaA de *A. oryzae*

## RÉSULTATS

---

dépend de la phase de croissance cellulaire et de la disponibilité des acides aminés. En effet, le dosage de l'activité GAP indique que la quantité de Gda augmente dans des conditions de carence d'azote et de carbone, notamment au cours de la phase stationnaire, fournissant ainsi à cet organisme les acides aminés requis pour la croissance (Marui et al., 2012).

Il est possible que les GAPs puissent également jouer un ou plusieurs rôles physiologiques plus spécifiques. En effet, la glycine est un acide aminé précurseur important dans de nombreuses voies métaboliques. Par exemple, la triméthylglycine, qui est connue comme un osmoprotectant chez certaines bactéries, dérive de la méthylation de la glycine (Csonka, 1989). La glycine joue encore un rôle important dans la thermo-protection de l'activité des enzymes de *Methanocaldococcus jannaschii* (Neelon et al., 2005). Elle intervient aussi dans la cascade de la réponse au stress oxydatif. En particulier, elle est utilisée comme précurseur pour la synthèse du glyoxylate, un antioxydant dont le métabolisme génère du NADPH, de l'ATP, de l'oxalate (piégeur des ROS (Reactive Oxygen Species)) et du formate, un agent réducteur chez *Pseudomonas fluorescens* (Alhasawi et al., 2015). De même, la glycine est requise pour la synthèse du glutathion, l'antioxydant majeur des cellules eucaryotes (Hall, 1998), et d'autres protéines comme la créatine, l'hème,... (Wang et al., 2013). Ainsi, il est possible que les GAPs soient des enzymes régulatrices importantes dans tous ces processus.

Pour conclure on retiendra que, en absence d'études génétiques et structurales, les connaissances sur les GAPs restent limitées, tant sur le plan de leurs rôles physiologiques que sur les mécanismes moléculaires responsables de leurs spécificités. Dans le cadre de cette thèse nous avons entrepris l'étude structurale de la GAP de *Pyrococcus horikoshii* (PhTET4) qui est la première GAP hyperthermophile connue et la première GAP caractérisée dans la famille des aminopeptidases M42.

### 5.2.4. Discussion et Perspectives

Les études enzymatiques précédemment réalisées dans notre laboratoire sur les PhTETs montrent que PhTET1 est spécifique des résidus acides (Schoehn et al., 2006; Durá and Franzetti, 2013), PhTET2 est une leucine aminopeptidase à spectre large (Durá et al., 2005) et PhTET3 agit en préférence sur les résidus positifs (Durá et al., 2009). Ces métallo-aminopeptidases activées par le cobalt sont cependant aptes à hydrolyser moins efficacement une petite variété d'acides aminés considérés comme non préférables par l'enzyme. En revanche, la caractérisation enzymatique de la quatrième TET de *P. horikoshii*, PhTET4, montre une activité spécialisée et unique contre la glycine, en utilisant soit des monoacyls, soit des peptides comme substrats. Ce caractère distinctif de PhTET4 est d'autant plus prononcé qu'elle s'est révélée être plus efficacement activée par le nickel que le cobalt. Ces deux traits particuliers permettent de proposer PhTET4 comme la première métallo-aminopeptidase de la famille M42 disposant d'une activité amidolytique contre la glycine et possédant le nickel comme cofacteur activateur.

Une troisième caractéristique se révèle aussi singulière pour PhTET4. Celle-ci se traduit par son état dodécamérique qui est insensible à l'addition de l'EDTA. En effet, l'alignement de séquence de PhTET4 avec ses homologues PhTET1, 2 et 3, montre que l'acide aspartique Asp231 et l'histidine His311 ne seraient pas alignés exactement avec les acides aspartiques et les histidines des autres TETs. Effectivement, ce sont une valine et une sérine qui prennent la place respective de l'acide aspartique et de l'histidine. Théoriquement et en se basant sur la structure du site bimétallique des TETs homologues, Asp231 et His311 devraient coordonner M2 et M1, respectivement. Cette observation nous amène alors à proposer que l'organisation du site actif de PhTET4 serait différente de ses homologues en raison d'une implication probable de Val229 et Ser310. D'autre part, il se peut que les interfaces de dimérisation (entre deux dimères) et d'oligomérisation (entre trois monomères) ne soient pas proches du site métallique qui pourrait être encore plus enfoncé dans la structure tertiaire du monomère. Ainsi, l'enfouissement du site actif pourrait protéger les métaux contre la chélation et expliquer la capacité unique de PhTET4 à hydrolyser la glycine, étant donné que cette dernière n'a pas besoin d'un grand volume pour se positionner au niveau du site actif. Pour avoir plus de précision sur cette insensibilité, il est nécessaire de résoudre la structure de PhTET4 et d'étudier les particularités portées par son site actif.



## RÉSULTATS

---

*P. horikoshii* est une espèce hétérotrophe utilisant alors les composés organiques comme les peptides comme source de carbone. Contrairement à *Methanocaldococcus jannaschii* qui est une espèce autotrophe exprimant une seule TET, *P. horikoshii* en possède quatre versions, chacune présentant une activité enzymatique différente de l'autre. La coexistence de quatre copies TETs peut être reliée au caractère hétérotrophe de la souche qui les exprime, lui permettant ainsi d'exploiter efficacement les substrats peptidiques au profit de la croissance cellulaire. Ainsi pour évaluer cette hypothèse, il serait nécessaire d'étudier l'évolution du taux d'expression des PhTETs (1, 2, 3 et 4) dans différents milieux de culture, en présence ou en absence de peptides comme source de carbone. Ou encore des souches déficientes de *P. horikoshii* en TETs pourraient être élaborées en vue d'explorer l'effet de leur absence sur la croissance et le métabolisme cellulaire.

La spécialisation des TETs pourrait être accompagnée d'une augmentation de leur efficacité catalytique en comparaison avec les TETs à très large spectre mais ceci reste à démontrer. On peut ainsi penser que, au cours de l'évolution, une sélection des TETs bactériennes spécialisées s'est produite chez les espèces d'archées hétérotrophes conduisant à l'abandon de la TET archéenne.

À la fin de ma thèse, j'ai tenté d'étudier l'état oligomérique de PhTET4 *in vivo*. En effet, une régulation de l'état d'assemblage de TET, en fonction de l'état physiologique des cellules, a été montrée au laboratoire (Marty, 2006) (résultats non publiés). Pour cela, j'ai réalisé des tests d'immunodétection sur des extraits totaux de *P. horikoshii* cultivés à l'IFREMER de Brest (Laboratoire LM2E). Les études préliminaires montrent que, à la fin de la phase exponentielle, PhTET1 est dimérique alors que PhTET2, PhTET3 et PhTET4 sont présentent sous forme dimérique et dodécamérique (active). Ces résultats sont à confirmer mais nous pouvons supposer que ces trois TETs peuvent intervenir lors d'une carence nutritive en hydrolysant les peptides intracellulaires en acides aminés libres exploitables par la cellule. En effet, la glycine est connue comme un acide aminé difficilement clivable du fait de l'absence de chaîne latérale. Ainsi, la présence d'une aminopeptidase spécialisée contre tel résidu est requise pour promouvoir la dégradation peptidique dans la cellule.

### 6. CONCLUSION GÉNÉRALE

L'objectif principal de mon projet de thèse était de révéler de nouveaux aspects biologiques concernant les métallo-aminopeptidases tétraédriques TETs. Durant cette thèse, des caractérisations enzymatiques et structurales ont été entreprises sur deux aminopeptidases TETs, MjTET de *M. jannaschii* et PhTET4 de *P. horikoshii*.

Nous avons montré que MjTET et PhTET4 possèdent des activités amidolytiques très différentes. L'enzyme de *M. jannaschii* possède une activité catalytique à large spectre tandis que la quatrième enzyme TET de *P. horikoshii* présente une activité hydrolytique uniquement contre la glycine. Ces deux enzymes représentent donc des modèles intéressants pour comprendre les bases structurales qui conditionnent la spécificité de substrat chez les aminopeptidases M42. Notre travail sur MjTET indique que, en plus des propriétés de charge, la dimension et la dynamique moléculaire de la poche catalytique jouent un rôle important. Pour ce qui est de PhTET4, nous espérons que la détermination de la structure du site actif apportera un éclairage sur les causes de l'hyper-spécificité de cette enzyme vis-à-vis des résidus glycines. Des travaux de mutagenèse seront probablement nécessaires pour tester ces hypothèses. Il semble néanmoins très difficile de parvenir un jour à déduire de manière certaine la spécificité d'une aminopeptidase à partir de sa séquence peptidique.

Les travaux de S. Gribaldo indiquent que les 4 enzymes de *P. horikoshii* sont toutes issues de transfert de gènes provenant de bactéries, peut-être pour permettre à l'ancêtre des Thermococcales d'améliorer son métabolisme hétérotrophe. Ces quatre enzymes présentent toutes des préférences de substrat plus marquées que l'unique enzyme TET archéenne caractérisée chez *M. jannaschii*. Bien qu'on ne connaisse à ce jour aucune aminopeptidase M42 bactérienne présentant une spécificité pour la glycine aussi exclusive que celle de PhTET4, la question se pose de savoir si la spécialisation des aminopeptidases de la famille M42 est une caractéristique bactérienne ancestrale ou si elle a été développée par les archées.

Parmi les métallo-aminopeptidases M42 bactériennes dont les activités enzymatiques et l'organisation dodécamérique ont été caractérisées, deux sont des peptidases spécifiques des résidus acides et une autre est une leucyl aminopeptidase capable d'hydrolyser les résidus hydrophobes (Kim et al., 2010; Dutoit et al., 2012; Robinson et al., 2013). D'autres, appartenant aussi à la famille M42 mais pour lesquelles il n'existe pas de structure publiée,

## CONCLUSION GÉNÉRALE

---

sont décrites comme étant des aminopeptidases spécifiques des résidus basiques (Kumaki et al., 2011) ou encore des résidus neutres (Dutoit et al., 2012). Ces activités enzymatiques agencées par les métallo-aminopeptidases bactériennes rappellent celles exécutées par PhTET1, 2 et 3 de *P. horikoshii*. Ceci indique que chez les bactéries contemporaines les TETs sont des enzymes spécialisées et que l'ancêtre des Pyrococcus aurait acquis des enzymes déjà spécialisées. Des études supplémentaires sur des nouvelles TETs bactériennes et archéennes contemporaines ainsi que sur des protéines issues de gènes "ancestraux" proposés par des travaux de paléo-génomique seraient néanmoins requises pour consolider cette hypothèse. Il serait également intéressant de comparer les profils de spécificité de TETs issues d'organismes hétérotrophes et autotrophes afin de consolider notre hypothèse concernant le lien entre spécialisation des TET et l'hétérotrophie.

Les travaux réalisés sur MjTET ont aussi permis de proposer un nouveau mécanisme de navigation des substrats peptidiques depuis le pore d'entrée de la machine tétraédrique TET vers l'un des sites actifs des chambres catalytiques situées aux alentours du pore d'accès. En effet, MjTET est la première aminopeptidase de la famille M42 pour laquelle on a pu obtenir des informations structurales sur l'organisation interne de la particule et notamment sur les boucles mobiles. Ceci est dû à l'apport de la cryo-microscopie électronique. On peut se demander si ce mode d'action est universel chez toutes les TETs. Des boucles internes complètes de complexes TETs ont été observées chez deux aminopeptidases dodécamériques hDNPEP et bDNPEP appartenant à la famille M18 (Chaikuad et al., 2012; Chen et al., 2012). Toutefois, ces dernières présentent des structures primaires beaucoup plus longues que celles des M42, notamment des aminopeptidases archéennes. De plus, leurs structures quaternaires, toujours dodécamériques, présentaient des différences notables surtout en ce qui concerne la structure tertiaire de leurs domaines de dimérisation (I) qui portent les boucles. Effectivement, le domaine de dimérisation (I) des M42 adopte une structure similaire à un tonneau  $\beta$ , cependant celui des M18 est plus tordu et la structure du tonneau est brisée par des hélices et montre une conformation en *butterfly-shaped domain*. Pour cela, il est difficile de généraliser aux M18 le modèle proposé pour le mode d'action des M42 archéennes. Par contre, il serait intéressant de savoir si ce mécanisme est commun à toutes les TETs M42. Pour cela, il faudra comparer l'organisation interne de MjTET avec celles d'autres TETs M42 bactériennes ou d'origine bactérienne comme l'enzyme PhTET2. En effet, la structure de PhTET2 est très similaire à celle de MjTET et PhTET2 possède une activité enzymatique à large spectre qui, bien que plus restreinte, se rapproche de celle de MjTET. Dans ce contexte, il serait

## CONCLUSION GÉNÉRALE

---

intéressant d'utiliser la cryo-microscopie électronique pour étudier la topologie des boucles internes dans la structure de PhTET2 comme nous l'avons fait pour MjTET.

### 7. RÉFÉRENCES BIBLIOGRAPHIQUES

Adams, M.W. (1998). The biochemical diversity of life near and above 100°C in marine environments. *J. Appl. Microbiol.* *85 Suppl 1*, 108S – 117S.

Aguilar, M.-I. (2004). *HPLC of Peptides and Proteins: Methods and Protocols* (Totowa, New Jersey: Methods in Molecular Biology).

Alhasawi, A., Castonguay, Z., Appanna, N.D., Auger, C., and Appanna, V.D. (2015). Glycine metabolism and anti-oxidative defence mechanisms in *Pseudomonas fluorescens*. *Microbiol. Res.* *171*, 26–31.

Allen, M.P., Yamada, A.H., and Carpenter, F.H. (1983). Kinetic parameters of metal-substituted leucine aminopeptidase from bovine lens. *Biochemistry (Mosc.)* *22*, 3778–3783.

Ando, S., Ishikawa, K., Ishida, H., Kawarabayasi, Y., Kikuchi, H., and Kosugi, Y. (1999). Thermostable aminopeptidase from *Pyrococcus horikoshii*. *FEBS Lett.* *447*, 25–28.

Appolaire, A. (2014). Étude des grands assemblages protéolytiques de la famille TET : processus d'oligomérisation et régulation fonctionnelle associée.

Appolaire, A., Rosenbaum, E., Dura, M.A., Colombo, M., Marty, V., Savoye, M.N., Godfroy, A., Schoehn, G., Girard, E., Gabel, F., et al. (2013). *Pyrococcus horikoshii* TET2 Peptidase Assembling Process and Associated Functional Regulation. *J. Biol. Chem.* *288*, 22542–22554.

Appolaire, A., Durá, M.A., Ferruit, M., Andrieu, J.-P., Godfroy, A., Gribaldo, S., and Franzetti, B. (2014a). The TET2 and TET3 aminopeptidases from *P. yrococcus horikoshii* form a hetero-subunit peptidasome with enhanced peptide destruction properties: TET aminopeptidase multi-subunit complex. *Mol. Microbiol.* *94*, 803–814.

Appolaire, A., Girard, E., Colombo, M., Durá, M.A., Moulin, M., Härtle, M., Franzetti, B., and Gabel, F. (2014b). Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase. *Acta Crystallogr. D Biol. Crystallogr.* *70*, 2983–2993.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science* *234*, 179–186.

Banegas, I., Prieto, I., Vives, F., Alba, F., Gasparo, M. de, Segarra, A.B., Hermoso, F., Durán, R., and Ramírez, M. (2006). Review: Brain Aminopeptidases and Hypertension. *J. Renin Angiotensin Aldosterone Syst.* *7*, 129–134.

Barrett, A.J., and Rawlings, N.D. (2007). “Species” of peptidases. *Biol. Chem.* *388*, 1151–1157.

Barrett, A.J., Rawlings, N.D., and Woessner, J.F. (2004). *Handbook of proteolytic enzymes* (Elsevier Academic press).

Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell* *92*, 367–380.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Beninga, J., Rock, K.L., and Goldberg, A.L. (1998). Interferon-gamma can stimulate post-proteasomal trimming of the N terminus of an antigenic peptide by inducing leucine aminopeptidase. *J. Biol. Chem.* *273*, 18734–18742.
- Berlett, B.S., and Stadtman, E.R. (1997). Protein oxidation in aging, disease, and oxidative stress. *J. Biol. Chem.* *272*, 20313–20316.
- Bilbrey, R.E., Penheiter, A.R., Gathman, A.C., and Lilly, W.W. (1996). Characterization of a novel phenylalanine-specific aminopeptidase from *Schizophyllum commune*. *Mycol. Res.* *100*, 462–466.
- Blöchl, E., Rachel, R., Burggraf, S., Hafenbradl, D., Jannasch, H.W., and Stetter, K.O. (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C. *Extrem. Life Extreme Cond.* *1*, 14–21.
- Borissenko, L., and Groll, M. (2005). Crystal Structure of TET Protease Reveals Complementary Protein Degradation Pathways in Prokaryotes. *J. Mol. Biol.* *346*, 1207–1219.
- Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., and Forterre, P. (2005). Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.* *6*, R42.
- Brochier-Armanet, C., Forterre, P., and Gribaldo, S. (2011). Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* *14*, 274–281.
- Brock, T.D., and Freeze, H. (1969). *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *J. Bacteriol.* *98*, 289–297.
- Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., et al. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* *273*, 1058–1073.
- Buttner, K., Wenig, K., and Hopfner, K.-P. (2006). The exosome: a macromolecular cage for controlled RNA degradation. *Mol. Microbiol.* *61*, 1372–1379.
- Byun, T. (2001). Aminopeptidase from *Sphingomonas capsulata*. *J. Biol. Chem.* *276*, 17902–17907.
- Cacciapuoti, G., Fuccio, F., Petraccone, L., Del Vecchio, P., and Porcelli, M. (2012). Role of disulfide bonds in conformational stability and folding of 5'-deoxy-5'-methylthioadenosine phosphorylase II from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochim. Biophys. Acta* *1824*, 1136–1143.
- Cameron, V., House, C.H., and Brantley, S.L. (2012). A First Analysis of Metallome Biosignatures of Hyperthermophilic Archaea. *Archaea* *2012*.
- Carpenter, F.H., and Vahl, J.M. (1973). Leucine aminopeptidase (Bovine lens). Mechanism of activation by Mg<sup>2+</sup> and Mn<sup>2+</sup> of the zinc metalloenzyme, amino acid composition, and sulfhydryl content. *J. Biol. Chem.* *248*, 294–304.
- Carroll, R.K., Robison, T.M., Rivera, F.E., Davenport, J.E., Jonsson, I.-M., Florczyk, D., Tarkowski, A., Potempa, J., Koziel, J., and Shaw, L.N. (2012). Identification of an

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

intracellular M17 family leucine aminopeptidase that is required for virulence in *Staphylococcus aureus*. *Microbes Infect.* *14*, 989–999.

Carroll, R.K., Veillard, F., Gagne, D.T., Lindenmuth, J.M., Poreba, M., Drag, M., Potempa, J., and Shaw, L.N. (2013). The *Staphylococcus aureus* leucine aminopeptidase is localized to the bacterial cytosol and demonstrates a broad substrate range that extends beyond leucine. *Biol. Chem.* *394*.

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 Struct. Biol.* *12*, 14.

Chang, Y.H., and Smith, J.A. (1989). Molecular cloning and sequencing of genomic DNA encoding aminopeptidase I from *Saccharomyces cerevisiae*. *J. Biol. Chem.* *264*, 6979–6983.

Chayen, N.E., and Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nat. Methods* *5*, 147–153.

Chen, S.-L., Marino, T., Fang, W.-H., Russo, N., and Himo, F. (2008). Peptide Hydrolysis by the Binuclear Zinc Enzyme Aminopeptidase from *Aeromonas proteolytica*: A Density Functional Theory Study. *J. Phys. Chem. B* *112*, 2494–2500.

Chen, Y., Farquhar, E.R., Chance, M.R., Palczewski, K., and Kiser, P.D. (2012). Insights into Substrate Specificity and Metal Activation of Mammalian Tetrahedral Aspartyl Aminopeptidase. *J. Biol. Chem.* *287*, 13356–13370.

Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.M., Moras, D., and Tarnus, C. (1994a). Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Struct. Lond. Engl.* *1993* *2*, 283–291.

Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.M., Moras, D., and Tarnus, C. (1994b). Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Struct. Lond. Engl.* *1993* *2*, 283–291.

Chevrier, B., D'orchymont, H., Schalk, C., Tarnus, C., and Moras, D. (1996). The Structure of the *Aeromonas proteolytica* Aminopeptidase Complexed with a Hydroxamate Inhibitor. *Eur. J. Biochem.* *237*, 393–398.

Chu, L., Lai, Y., Xu, X., Eddy, S., Yang, S., Song, L., and Kolodrubetz, D. (2008). A 52-kDa Leucyl Aminopeptidase from *Treponema denticola* Is a Cysteinyglycinase That Mediates the Second Step of Glutathione Metabolism. *J. Biol. Chem.* *283*, 19351–19358.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* *50*, 760–763.

Colombo, M., Girard, E., and Franzetti, B. (2016). Tuned by metals: the TET peptidase activity is controlled by 3 metal binding sites. *Sci. Rep.* *6*, 20876.

Csonka, L.N. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiol. Rev.* *53*, 121–147.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Cubonova, L., Sandman, K., Hallam, S.J., DeLong, E.F., and Reeve, J.N. (2005). Histones in Crenarchaea. *J. Bacteriol.* *187*, 5482–5485.
- Dalal, S., and Klemba, M. (2007). Roles for Two Aminopeptidases in Vacuolar Hemoglobin Catabolism in *Plasmodium falciparum*. *J. Biol. Chem.* *282*, 35978–35987.
- DeLong, E.F. (1998). Everything in moderation: archaea as “non-extremophiles.” *Curr. Opin. Genet. Dev.* *8*, 649–654.
- DeLong, E.F., and Pace, N.R. (2001). Environmental diversity of bacteria and archaea. *Syst. Biol.* *50*, 470–478.
- Del-Val, M., and López, D. (2002). Multiple proteases process viral antigens for presentation by MHC class I molecules to CD8(+) T lymphocytes. *Mol. Immunol.* *39*, 235–247.
- Desmarais, W., Bienvenue, D.L., Bzymek, K.P., Petsko, G.A., Ringe, D., and Holz, R.C. (2006). The high-resolution structures of the neutral and the low pH crystals of aminopeptidase from *Aeromonas proteolytica*. *JBIC J. Biol. Inorg. Chem.* *11*, 398–408.
- Durá, M.A., and Franzetti, B. (2013). PhTET1 Aminopeptidase. In *Handbook of Proteolytic Enzymes*, (Elsevier), pp. 1638–1645.
- 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*<sup>†</sup>. *Biochemistry (Mosc.)* *44*, 3477–3486.
- 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. *Mol. Microbiol.* *72*, 26–40.
- Dutoit, R., Brandt, N., Legrain, C., and Bauvois, C. (2012). Functional Characterization of Two M42 Aminopeptidases Erroneously Annotated as Cellulases. *PLoS ONE* *7*, e50639.
- Elkins, J.G., Podar, M., Graham, D.E., Makarova, K.S., Wolf, Y., Randau, L., Hedlund, B.P., Brochier-Armanet, C., Kunin, V., Anderson, I., et al. (2008). A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 8102–8107.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 486–501.
- Ercan, A., Tay, W.M., Grossman, S.H., and Ming, L.-J. (2010). Mechanistic role of each metal ion in *Streptomyces* dinuclear aminopeptidase: PEPTIDE hydrolysis and 7x10(10)-fold rate enhancement of phosphodiester hydrolysis. *J. Inorg. Biochem.* *104*, 19–29.
- Forterre, P. (2015). The universal tree of life: an update. *Microb. Physiol. Metab.* *717*.
- Franzetti, B., Schoehn, G., Hernandez, J.-F., Jaquinod, M., Ruigrok, R.W.H., and Zaccai, G. (2002). Tetrahedral aminopeptidase: a novel large protease complex from archaea. *EMBO J.* *21*, 2132–2138.



## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Gao, X., Cui, W., Ding, N., Liu, Z., Tian, Y., and Zhou, Z. (2013). Structure-based approach to alter the substrate specificity of *Bacillus subtilis* aminopeptidase. *Prion* 7, 328–334.
- Gavioli, R., Frisan, T., Vertuani, S., Bornkamm, G.W., and Masucci, M.G. (2001). c-myc overexpression activates alternative pathways for intracellular proteolysis in lymphoma cells. *Nat. Cell Biol.* 3, 283–288.
- Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S., and Shoham, G. (2001). Interactions of *Streptomyces griseus* aminopeptidase with amino acid reaction products and their implications toward a catalytic mechanism. *Proteins* 44, 490–504.
- Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. *Nature* 426, 895–899.
- Gomis-Rüth, F.X., Grams, F., Yiallourous, I., Nar, H., Küsthardt, U., Zwilling, R., Bode, W., and Stöcker, W. (1994). Crystal structures, spectroscopic features, and catalytic properties of cobalt(II), copper(II), nickel(II), and mercury(II) derivatives of the zinc endopeptidase astacin. A correlation of structure and proteolytic activity. *J. Biol. Chem.* 269, 17111–17117.
- Gonzales, T., and Robert-Baudouy, J. (1996). Bacterial aminopeptidases: properties and functions. *FEMS Microbiol. Rev.* 18, 319–344.
- González, J.M., Masuchi, Y., Robb, F.T., Ammerman, J.W., Maeder, D.L., Yanagibayashi, M., Tamaoka, J., and Kato, C. (1998). *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extrem. Life Extreme Cond.* 2, 123–130.
- Gribaldo, S., and Brochier-Armanet, C. (2006). The origin and evolution of Archaea: a state of the art. *Philos. Trans. R. Soc. B Biol. Sci.* 361, 1007–1022.
- Gribaldo, S., Forterre, P., and Brochier-Armanet, C. (2011). Editorial: Archaea and the Tree of Life. *Res. Microbiol.* 162, 1–4.
- Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H.D., and Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* 386, 463–471.
- Hall, J.C. (1998). Glycine. *JPEN J. Parenter. Enteral Nutr.* 22, 393–398.
- Hartung, S., and Hopfner, K.-P. (2009). Lessons from structural and biochemical studies on the archaeal exosome. *Biochem. Soc. Trans.* 37, 83–87.
- Hickey, D.A., and Singer, G.A.C. (2004). Genomic and proteomic adaptations to growth at high temperature. *Genome Biol.* 5, 117.
- Howarth, J., and Lloyd, D.G. (2000). Simple 1,2-aminoalcohols as strain-specific antimalarial agents. *J. Antimicrob. Chemother.* 46, 625–628.
- Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C., and Stetter, K.O. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 417, 63–67.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Ito, K., Ma, X., Azmi, N., Huang, H.-S., Fujii, M., and Yoshimoto, T. (2003). Novel aminopeptidase specific for glycine from *Actinomucor elegans*. *Biosci. Biotechnol. Biochem.* *67*, 83–88.
- Jabs, A., Weiss, M.S., and Hilgenfeld, R. (1999). Non-proline cis peptide bonds in proteins. *J. Mol. Biol.* *286*, 291–304.
- Jaenicke, R. (1981). Enzymes under extremes of physical conditions. *Annu. Rev. Biophys. Bioeng.* *10*, 1–67.
- Jannasch, H.W. (1984). Deep-sea microbiology. *Annu. Rev. Microbiol.* *38*, 487–514.
- Jarocki, V.M., Santos, J., Tacchi, J.L., Raymond, B.B.A., Deutscher, A.T., Jenkins, C., Padula, M.P., and Djordjevic, S.P. (2015). MHJ\_0461 is a multifunctional leucine aminopeptidase on the surface of *Mycoplasma hyopneumoniae*. *Open Biol.* *5*, 140175–140175.
- Jarrell, K.F., Walters, A.D., Bochiwal, C., Borgia, J.M., Dickinson, T., and Chong, J.P.J. (2011). Major players on the microbial stage: why archaea are important. *Microbiol. Read. Engl.* *157*, 919–936.
- Jebbar, M., Franzetti, B., Girard, E., and Oger, P. (2015). Microbial diversity and adaptation to high hydrostatic pressure in deep-sea hydrothermal vents prokaryotes. *Extrem. Life Extreme Cond.* *19*, 721–740.
- Jia, B., Lee, S., Pham, B.P., Kwack, J.M., Jin, H., Li, J., Wang, Y., and Cheong, G.-W. (2011). Biochemical Characterization of Deblocking Aminopeptidases from the Hyperthermophilic Archaeon *Thermococcus kodakarensis* KOD1. *Biosci. Biotechnol. Biochem.* *75*, 1160–1166.
- Jones, W.J., Leigh, J.A., Mayer, F., Woese, C.R., and Wolfe, R.S. (1983). *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Arch. Microbiol.* *136*, 254–261.
- Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 125–132.
- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* *22*, 2577–2637.
- Karl, D.M. (2002). Nutrient dynamics in the deep blue sea. *Trends Microbiol.* *10*, 410–418.
- Kessler, B.M., Glas, R., and Ploegh, H.L. (2002). MHC class I antigen processing regulated by cytosolic proteolysis-short cuts that alter peptide generation. *Mol. Immunol.* *39*, 171–179.
- Kim, H., and Lipscomb, W.N. (1993). Differentiation and identification of the two catalytic metal binding sites in bovine lens leucine aminopeptidase by x-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.* *90*, 5006–5010.
- 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. *Biochem. Biophys. Res. Commun.* *391*, 431–436.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Kim, J.S., Groll, M., Musiol, H.-J., Behrendt, R., Kaiser, M., Moroder, L., Huber, R., and Brandstetter, H. (2002). Navigation inside a protease: Substrate selection and product exit in the tricorn protease from *Thermoplasma acidophilum*. *J. Mol. Biol.* *324*, 1041–1050.
- Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y., and Shimizu, N. (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* *392*, 605–608.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* *372*, 774–797.
- Krugliak, M., Zhang, J., and Ginsburg, H. (2002). Intraerythrocytic *Plasmodium falciparum* utilizes only a fraction of the amino acids derived from the digestion of host cell cytosol for the biosynthesis of its proteins. *Mol. Biochem. Parasitol.* *119*, 249–256.
- Kumaki, Y., Ogawa, M., Hirano, T., Yoshikawa, K., Iwasawa, N., Yagi, T., Hakamata, W., Oku, T., and Nishio, T. (2011). Family M42 aminopeptidase from the syntrophic bacterium *Symbiobacterium thermophilum*: Characterization using recombinant protein. *J. Biosci. Bioeng.* *111*, 134–139.
- Lauterbach, S.B., and Coetzer, T.L. (2008). The M18 aspartyl aminopeptidase of *Plasmodium falciparum* binds to human erythrocyte spectrin in vitro. *Malar. J.* *7*, 161.
- Lazdunski, A.M. (1989). Peptidases and proteases of *Escherichia coli* and *Salmonella typhimurium*. *FEMS Microbiol. Rev.* *5*, 265–276.
- Le Bris, N. (2013). Géochimie des sources hydrothermales.
- Lee, C.-F., Makhatadze, G.I., and Wong, K.-B. (2005). Effects of charge-to-alanine substitutions on the stability of ribosomal protein L30e from *Thermococcus celer*. *Biochemistry (Mosc.)* *44*, 16817–16825.
- 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. *J. Biosci. Bioeng.* *104*, 188–194.
- Leigh, J.A., Albers, S.-V., Atomi, H., and Allers, T. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol. Rev.* *35*, 577–608.
- Lepage, P., Leclerc, M.C., Joossens, M., Mondot, S., Blottière, H.M., Raes, J., Ehrlich, D., and Doré, J. (2013). A metagenomic insight into our gut's microbiome. *Gut* *62*, 146–158.
- Liu, Y.-F., Zhang, N., Liu, X., Wang, X., Wang, Z.-X., Chen, Y., Yao, H.-W., Ge, M., and Pan, X.-M. (2012). Molecular mechanism underlying the interaction of typical Sac10b family proteins with DNA. *PLoS One* *7*, e34986.
- Lowther, W.T., and Matthews, B.W. (2002). Metalloaminopeptidases: Common Functional Themes in Disparate Structural Surroundings. *Chem. Rev.* *102*, 4581–4608.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Ma, X., Zhou, X., and Yoshimoto, T. (2004). Purification and properties of a novel glycine amino peptidase from *Actinomucor elegans* and its potential application. *J. Appl. Microbiol.* *97*, 985–991.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* *39*, D225–D229.
- Maric, S., Donnelly, S.M., Robinson, M.W., Skinner-Adams, T., Trenholme, K.R., Gardiner, D.L., Dalton, J.P., Stack, C.M., and Lowther, J. (2009). The M17 Leucine Aminopeptidase of the Malaria Parasite *Plasmodium falciparum* : Importance of Active Site Metal Ions in the Binding of Substrates and Inhibitors. *Biochemistry (Mosc.)* *48*, 5435–5439.
- Martin, R.E., and Kirk, K. (2007). Transport of the essential nutrient isoleucine in human erythrocytes infected with the malaria parasite *Plasmodium falciparum*. *Blood* *109*, 2217–2224.
- Marty, V. (2006). Adaptation de l'archaea halophile *Halobacterium salinarum* aux stress environnementaux : mécanismes de survie et rôle de la protéolyse intracellulaire.
- Marui, J., Matsushita-Morita, M., Tada, S., Hattori, R., Suzuki, S., Amano, H., Ishida, H., Yamagata, Y., Takeuchi, M., and Kusumoto, K.-I. (2012). Enzymatic properties of the glycine d-alanine aminopeptidase of *Aspergillus oryzae* and its activity profiles in liquid-cultured mycelia and solid-state rice culture (rice koji). *Appl. Microbiol. Biotechnol.* *93*, 655–669.
- Matsui, M., Fowler, J.H., and Walling, L.L. (2006). Leucine aminopeptidases: diversity in structure and function. *Biol. Chem.* *387*.
- Maupin-Furlow, J. (2011). Proteasomes and protein conjugation across domains of life. *Nat. Rev. Microbiol.*
- Maurizi, M.R., Singh, S.K., Thompson, M.W., Kessel, M., and Ginsburg, A. (1998). Molecular properties of ClpAP protease of *Escherichia coli*: ATP-dependent association of ClpA and clpP. *Biochemistry (Mosc.)* *37*, 7778–7786.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.
- McGowan, S. (2013). Working in concert: the metalloaminopeptidases from *Plasmodium falciparum*. *Curr. Opin. Struct. Biol.* *23*, 828–835.
- Min, L., Sim, M.K., and Xu, X.G. (2000). Effects of des-aspartate-angiotensin I on angiotensin II-induced incorporation of phenylalanine and thymidine in cultured rat cardiomyocytes and aortic smooth muscle cells. *Regul. Pept.* *95*, 93–97.
- Mitić, N., Miraula, M., Selleck, C., Hadler, K.S., Uribe, E., Pedroso, M.M., and Schenk, G. (2014). Chapter Three - Catalytic Mechanisms of Metallohydrolases Containing Two Metal Ions. In *Advances in Protein Chemistry and Structural Biology*, C.Z. Christov, ed. (Academic Press), pp. 49–81.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Mori, K., and Ishikawa, K. (2005). New deblocking aminopeptidases from *Pyrococcus horikoshii*. *Biosci. Biotechnol. Biochem.* *69*, 1854–1860.
- Narayanan, S.S., and Nampoothiri, K.M. (2012). Biochemical characterization of recombinant methionine aminopeptidases (MAPs) from *Mycobacterium tuberculosis* H37Rv. *Mol. Cell. Biochem.* *365*, 191–202.
- Neelon, K., Schreier, H.J., Meekins, H., Robinson, P.M., and Roberts, M.F. (2005). Compatible solute effects on thermostability of glutamine synthetase and aspartate transcarbamoylase from *Methanococcus jannaschii*. *Biochim. Biophys. Acta BBA - Proteins Proteomics* *1753*, 164–173.
- Nelson, D.L., and Cox, M.M. (2000). *Lehninger PRINCIPLES OF BIOCHEMISTRY* (New York Worth Publishers).
- Ng, W.V., Kennedy, S.P., Mahairas, G.G., Berquist, B., Pan, M., Shukla, H.D., Lasky, S.R., Baliga, N.S., Thorsson, V., Sbrogna, J., et al. (2000). Genome sequence of *Halobacterium* species NRC-1. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 12176–12181.
- 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*. *Biochem. Biophys. Res. Commun.* *447*, 101–107.
- Niedermann, G., Geier, E., Lucchiari-Hartz, M., Hitziger, N., Ramsperger, A., and Eichmann, K. (1999). The specificity of proteasomes: impact on MHC class I processing and presentation of antigens. *Immunol. Rev.* *172*, 29–48.
- Peretó, J., López-García, P., and Moreira, D. (2004). Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem. Sci.* *29*, 469–477.
- Petrova, T.E., Slutskaia, 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 Crystallogr. Sect. F Struct. Biol. Commun.* *71*, 277–285.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* *25*, 1605–1612.
- Porciero, S., Receveur-Bréchet, V., Mori, K., Franzetti, B., and Roussel, A. (2005). Expression, purification, crystallization and preliminary crystallographic analysis of a deblocking aminopeptidase from *Pyrococcus horikoshii*. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* *61*, 239–242.
- Poreba, M., McGowan, S., Skinner-Adams, T.S., Trenholme, K.R., Gardiner, D.L., Whisstock, J.C., To, J., Salvesen, G.S., Dalton, J.P., and Drag, M. (2012). Fingerprinting the Substrate Specificity of M1 and M17 Aminopeptidases of Human Malaria, *Plasmodium falciparum*. *PLoS ONE* *7*.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Preston, C.M., Wu, K.Y., Molinski, T.F., and DeLong, E.F. (1996). A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc. Natl. Acad. Sci. U. S. A.* *93*, 6241–6246.
- Pritchard, G.G., and Coolbear, T. (1993). The physiology and biochemistry of the proteolytic system in lactic acid bacteria. *FEMS Microbiol. Rev.* *12*, 179–206.
- Pulido-Cejudo, G., Conway, B., Proulx, P., Brown, R., and Izaguirre, C.A. (1997). Bestatin-mediated inhibition of leucine aminopeptidase may hinder HIV infection. *Antiviral Res.* *36*, 167–177.
- Querellou, J., Alain, K., and Cambon-Bonavita, M.A. (2001). Thermophiles from deep-sea hydrothermal vents. *Vie Milieu* *51*, 161–172.
- Radzicka, A., and Wolfenden, R. (1996). Rates of Uncatalyzed Peptide Bond Hydrolysis in Neutral Solution and the Transition State Affinities of Proteases. *J. Am. Chem. Soc.* *118*, 6105–6109.
- Rawlings, N.D., and Barrett, A.J. (1993). Evolutionary families of peptidases. *Biochem. J.* *290 (Pt 1)*, 205–218.
- Rawlings, N.D., Tolle, D.P., and Barrett, A.J. (2004). MEROPS: the peptidase database. *Nucleic Acids Res.* *32*, D160–D164.
- Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* *40*, D343–D350.
- Reaux, A., Fournie-Zaluski, M.C., David, C., Zini, S., Roques, B.P., Corvol, P., and Llorens-Cortes, C. (1999). Aminopeptidase A inhibitors as potential central antihypertensive agents. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 13415–13420.
- Reed, C.J., Lewis, H., Trejo, E., Winston, V., and Evilia, C. (2013). Protein Adaptations in Archaeal Extremophiles. *Archaea* *2013*.
- Reeve, J.N., and Schleper, C. (2011). Archaea: very diverse, often different but never bad? *Curr. Opin. Microbiol.* *14*, 271–273.
- Rizzo, A.A. (1967). The possible role of hydrogen sulfide in human periodontal disease. I. Hydrogen sulfide production in periodontal pockets. *Periodontics* *5*, 233–236.
- Robinson, M.W., Buchtman, K.A., Jenkins, C., Tacchi, J.L., Raymond, B.B.A., To, J., Roy Chowdhury, P., Woolley, L.K., Labbate, M., Turnbull, L., et al. (2013). MHJ\_0125 is an M42 glutamyl aminopeptidase that moonlights as a multifunctional adhesin on the surface of *Mycoplasma hyopneumoniae*. *Open Biol.* *3*, 130017–130017.
- Rock, K.L., and Goldberg, A.L. (1999). Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* *17*, 739–779.
- Rock, K.L., York, I.A., Saric, T., and Goldberg, A.L. (2002). Protein degradation and the generation of MHC class I-presented peptides. *Adv. Immunol.* *80*, 1–70.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Rohrwild, M., Coux, O., Huang, H.C., Moerschell, R.P., Yoo, S.J., Seol, J.H., Chung, C.H., and Goldberg, A.L. (1996). HslV-HslU: A novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome. *Proc. Natl. Acad. Sci. U. S. A.* *93*, 5808–5813.
- Roncari, and Zuber (1970). Thermophilic aminopeptidase: API from *Bacillus stearothermophilus*. *Methods Enzymol.* *544–552*.
- 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. *Biochim. Biophys. Acta BBA - Proteins Proteomics* *1814*, 1289–1294.
- Rowell, S., Pauptit, R.A., Tucker, A.D., Melton, R.G., Blow, D.M., and Brick, P. (1997). Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure* *5*, 337–347.
- Russo, S., and Baumann, U. (2004). Crystal Structure of a Dodecameric Tetrahedral-shaped Aminopeptidase. *J. Biol. Chem.* *279*, 51275–51281.
- Samanta, U., Pal, D., and Chakrabarti, P. (2000). Environment of tryptophan side chains in proteins. *Proteins* *38*, 288–300.
- Saric, T., Beninga, J., Graef, C.I., Akopian, T.N., Rock, K.L., and Goldberg, A.L. (2001). Major histocompatibility complex class I-presented antigenic peptides are degraded in cytosolic extracts primarily by thimet oligopeptidase. *J. Biol. Chem.* *276*, 36474–36481.
- Schechter, I., and Berger, A. (1967). On the Size of the Active Site in Proteases. I. Papain. *Biochem. Biophys. Res. Commun.* *27*, 157–162.
- Schoehn, G., Vellieux, F.M.D., Asunción Durá, M., Receveur-Bréchet, 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. *J. Biol. Chem.* *281*, 36327–36337.
- Schut, G.J., Brehm, S.D., Datta, S., and Adams, M.W.W. (2003). Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on carbohydrates or peptides. *J. Bacteriol.* *185*, 3935–3947.
- Sherman, I.W. (1977). Amino acid metabolism and protein synthesis in malarial parasites. *Bull. World Health Organ.* *55*, 265–276.
- Sherman, M.Y., and Goldberg, A.L. (2001). Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases. *Neuron* *29*, 15–32.
- Sinha, N., and Smith-Gill, S.J. (2002). Protein structure to function via dynamics. *Protein Pept. Lett.* *9*, 367–377.
- 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., et al. (2012). X-ray Crystal Structure and Specificity of the *Plasmodium falciparum* Malaria Aminopeptidase PfM18AAP. *J. Mol. Biol.* *422*, 495–507.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Skinner-Adams, T.S., Lowther, J., Teuscher, F., Stack, C.M., Grembecka, J., Mucha, A., Kafarski, P., Trenholme, K.R., Dalton, J.P., and Gardiner, D.L. (2007). Identification of phosphinate dipeptide analog inhibitors directed against the Plasmodium falciparum M17 leucine aminopeptidase as lead antimalarial compounds. *J. Med. Chem.* *50*, 6024–6031.
- Sleytr, U.B., Egelseer, E.M., Ilk, N., Pum, D., and Schuster, B. (2007). S-Layers as a basic building block in a molecular construction kit. *FEBS J.* *274*, 323–334.
- Slonczewski, J.L., and Foster, J.W. (2013). *Microbiology: An Evolving Science: Third International Student Edition* (W. W. Norton & Company).
- Slutskaya, E.S., Bezsudnova, E.Y., Mardanov, A.V., Gumerov, V.M., Rakitina, T.V., Popov, V.O., and Lipkin, V.M. (2012). Characterization of a novel M42 aminopeptidase from crenarchaeon *Desulfurococcus kamchatkensis*. *Dokl. Biochem. Biophys.* *442*, 30–32.
- Soffer, R. (1981). *Biochemical regulation of blood pressure* (Wiley interscience).
- Spang, A., Saw, J.H., Jørgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., and Ettema, T.J.G. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* *521*, 173–179.
- Stack, C.M., Lowther, J., Cunningham, E., Donnelly, S., Gardiner, D.L., Trenholme, K.R., Skinner-Adams, T.S., Teuscher, F., Grembecka, J., Mucha, A., et al. (2007). Characterization of the Plasmodium falciparum M17 Leucyl Aminopeptidase: A PROTEASE INVOLVED IN AMINO ACID REGULATION WITH POTENTIAL FOR ANTIMALARIAL DRUG DEVELOPMENT. *J. Biol. Chem.* *282*, 2069–2080.
- Stahlberg, H., Kutejová, E., Suda, K., Wolpensinger, B., Lustig, A., Schatz, G., Engel, A., and Suzuki, C.K. (1999). Mitochondrial Lon of *Saccharomyces cerevisiae* is a ring-shaped protease with seven flexible subunits. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 6787–6790.
- Stetter, K.O. (1982). Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105 °C. *Nature* *300*, 258–260.
- Stetter, K.O. (1996). Hyperthermophilic prokaryotes. *FEMS Microbiol. Rev.* *18*.
- Stetter, K.O. (1999). Extremophiles and their adaptation to hot environments. *FEBS Lett.* *452*, 22–25.
- Stoll, E., Hermodson, M.A., Ericsson, L.H., and Zuber, H. (1972). Subunit structure of the thermophilic aminopeptidase I from *Bacillus stearothermophilus*. *Biochemistry (Mosc.)* *11*, 4731–4735.
- Stoll, E., Ericsson, L.H., and Zuber, H. (1973a). The Function of the Two Subunits of Thermophilic Aminopeptidase I. *Proc. Natl. Acad. Sci.* *70*, 3781–3784.
- Stoll, E., Ericsson, L.H., and Zuber, H. (1973b). The function of the two subunits of thermophilic aminopeptidase I. *Proc. Natl. Acad. Sci. U. S. A.* *70*, 3781–3784.
- Stoltze, L., Schirle, M., Schwarz, G., Schröter, C., Thompson, M.W., Hersh, L.B., Kalbacher, H., Stevanovic, S., Rammensee, H.G., and Schild, H. (2000). Two new proteases in the MHC class I processing pathway. *Nat. Immunol.* *1*, 413–418.



## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Story, S.V., Shah, C., Jenney, F.E., and Adams, M.W.W. (2005). Characterization of a Novel Zinc-Containing, Lysine-Specific Aminopeptidase from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *J. Bacteriol.* *187*, 2077–2083.
- Talon, R., Kahn, R., Durá, M.A., Maury, O., Vellieux, F.M.D., Franzetti, B., and Girard, E. (2011). Using lanthanoid complexes to phase large macromolecular assemblies. *J. Synchrotron Radiat.* *18*, 74–78.
- Tamura, N., Lottspeich, F., Baumeister, W., and Tamura, T. (1998). The role of tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell* *95*, 637–648.
- Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996). Tricorn protease--the core of a modular proteolytic system. *Science* *274*, 1385–1389.
- Taylor, A. (1993). Aminopeptidases: structure and function. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *7*, 290–298.
- Taylor, A., Daims, M., Lee, J., and Surgenor, T. (1982). Identification and quantification of leucine aminopeptidase in aged normal and cataractous human lenses and ability of bovine lens LAP to cleave bovine crystallins. *Curr. Eye Res.* *2*, 47–56.
- Tehei, M., Franzetti, B., Madern, D., Ginzburg, M., Ginzburg, B.Z., Giudici-Ortoni, M.-T., Bruschi, M., and Zaccai, G. (2004). Adaptation to extreme environments: macromolecular dynamics in bacteria compared in vivo by neutron scattering. *EMBO Rep.* *5*, 66–70.
- Teuscher, F., Lowther, J., Skinner-Adams, T.S., Spielmann, T., Dixon, M.W.A., Stack, C.M., Donnelly, S., Mucha, A., Kafarski, P., Vassiliou, S., et al. (2007). The M18 Aspartyl Aminopeptidase of the Human Malaria Parasite *Plasmodium falciparum*. *J. Biol. Chem.* *282*, 30817–30826.
- Thompson, G.A., and Carpenter, F.H. (1976). Leucine aminopeptidase (bovine lens). The relative binding of cobalt and zinc to leucine aminopeptidase and the effect of cobalt substitution on specific activity. *J. Biol. Chem.* *251*, 1618–1624.
- Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. *Science* *254*, 1374–1377.
- Trivedi, S., Gehlot, H.S., and Rao, S.R. (2006). Protein thermostability in Archaea and Eubacteria. *Genet. Mol. Res. GMR* *5*, 816–827.
- Turner, A.J., Hyde, R.J., Lim, J., and Hooper, N.M. (1997). Structural Studies of Aminopeptidase P. In *Cellular Peptidases in Immune Functions and Diseases*, S. Ansorge, and J. Langner, eds. (Springer US), pp. 7–16.
- Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N., and Tsukihara, T. (2002). The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Struct. Lond. Engl.* *10*, 609–618.
- Vieille, C., and Zeikus, G.J. (2001). Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol. Mol. Biol. Rev.* *65*, 1–43.

## RÉFÉRENCES BIBLIOGRAPHIQUES

---

- Vieille, C., Burdette, D.S., and Zeikus, J.G. (1996). Thermozyms. *Biotechnol. Annu. Rev.* *2*, 1–83.
- Wang, W., Wu, Z., Dai, Z., Yang, Y., Wang, J., and Wu, G. (2013). Glycine metabolism in animals and humans: implications for nutrition and health. *Amino Acids* *45*, 463–477.
- Ward, D.E., Shockley, K.R., Chang, L.S., Levy, R.D., Michel, J.K., Conners, S.B., and Kelly, R.M. (2002). Proteolysis in hyperthermophilic microorganisms. *Archaea Vanc. BC* *1*, 63–74.
- Werner, F. (2008). Structural evolution of multisubunit RNA polymerases. *Trends Microbiol.* *16*, 247–250.
- Wilk, S., Wilk, E., and Magnusson, R.P. (1998). Purification, characterization, and cloning of a cytosolic aspartyl aminopeptidase. *J. Biol. Chem.* *273*, 15961–15970.
- Wilk, S., Wilk, E., and Magnusson, R.P. (2002). Identification of histidine residues important in the catalysis and structure of aspartyl aminopeptidase. *Arch. Biochem. Biophys.* *407*, 176–183.
- Woese, C.R., and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* *74*, 5088–5090.
- Wright, J.W., and Harding, J.W. (1997). Important role for angiotensin III and IV in the brain renin-angiotensin system. *Brain Res. Brain Res. Rev.* *25*, 96–124.
- Yang, Y., Fang, S., Jensen, J.P., Weissman, A.M., and Ashwell, J.D. (2000). Ubiquitin protein ligase activity of IAPs and their degradation in proteasomes in response to apoptotic stimuli. *Science* *288*, 874–877.
- Yao, T., and Cohen, R.E. (1999). Giant proteases: beyond the proteasome. *Curr. Biol. CB* *9*, R551–R553.
- Yokoyama, R., Kawasaki, H., and Hirano, H. (2006). Identification of yeast aspartyl aminopeptidase gene by purifying and characterizing its product from yeast cells. *FEBS J.* *273*, 192–198.
- York, I.A., Goldberg, A.L., Mo, X.Y., and Rock, K.L. (1999). Proteolysis and class I major histocompatibility complex antigen presentation. *Immunol. Rev.* *172*, 49–66.
- Zhang, T., Ozbil, M., Barman, A., Paul, T.J., Bora, R.P., and Prabhakar, R. (2015). Theoretical Insights into the Functioning of Metallopeptidases and Their Synthetic Analogues. *Acc. Chem. Res.* *48*, 192–200.
- Zhu, W., Reich, C.I., Olsen, G.J., Giometti, C.S., and Yates, J.R. (2004). Shotgun proteomics of *Methanococcus jannaschii* and insights into methanogenesis. *J. Proteome Res.* *3*, 538–548.



## ABSTRACT

---

Aminopeptidases represent a group of enzymes displaying key cellular function in physiological and pathological mechanisms. They are involved in the enzymatic cascade beyond the action of endoproteases, in homeostasis through the renewal of the amino acid pool, in the energy metabolism, in the regulation of bioactive peptide activities, in the antigen presentation and in a diversity of pathological mechanisms such as neurological diseases as well as viral and parasitic infections. Aminopeptidases TET are able of forming tetrahedral macro-assemblies built by twelve subunits. In order to better understand their biological function and their mode of action, we studied the functional and structural properties of two novel TET complexes derived from hyperthermophilic archaea.

The hyperthermophilic archaeon *Methanocaldococcus jannaschii* has only one version of TET (MjTET) that was produced in *Escherichia coli* and purified as dodecameric macromolecule. The search for its enzymatic activity and peptide substrates by using chromogenic/fluorogenic assays and reverse phase HPLC studies, demonstrated that this enzyme is a cobalt-activated leucine aminopeptidase, discriminated from other M42 aminopeptidases by its very broad activity spectrum, that extends to aromatic residues. Complete structure of this aminopeptidase was determined by combining X-ray crystallography (2.4 Å) and cryo-electron microscopy (4.1 Å). Analysis of MjTET specificity pocket indicated possible molecular bases for substrate discrimination in TET peptidases. In depth investigation of the particle internal structure allowed to propose a novel peptide trafficking mechanism for the TET family tetrahedral particles.

Three types of TET complexes are present in the hyperthermophilic archaea, *Pyrococcus horikoshii*. The study of an unassigned protein displaying ~20% identity with the PhTETs systems allowed us to identify a fourth version of TET complex in this organism: PhTET4. The recombinant protein was purified. It formed tetrahedral dodecameric complex. Biochemical studies indicated that the enzyme has a very narrow hydrolytic specificity directed exclusively toward the peptide N-terminal glycine residues. In addition, this enzyme is activated by nickel ions. These features allowed proposing that, in archaea, the multiplicity of specialized TET systems could be associated with heterotrophy while unique TET system displaying “housekeeping” function is present in autotrophic organisms.

## RÉSUMÉ

---

Les aminopeptidases représentent un groupe d'enzymes qui possèdent une fonction cellulaire clef dans les mécanismes physiologiques et pathologiques. Elles interviennent dans la cascade enzymatique après l'action des endoprotéases, dans l'homéostasie au travers le renouvellement du *pool* d'acides aminés, dans le métabolisme énergétique, la régulation de l'activité des peptides bioactifs, la présentation antigénique ainsi dans une diversité de mécanismes pathologiques tels que les maladies neurologiques et les infections virales et parasitaires. Les aminopeptidases TETs sont capables de former des macro-assemblages tétraédriques comprenant douze sous-unités. En vue de mieux comprendre leur fonction biologique et leur mode d'action, nous avons étudié les propriétés fonctionnelles et structurales de deux nouveaux complexes TETs issus d'archées hyperthermophiles.

L'archée hyperthermophile *Methanocaldococcus jannaschii* ne possède qu'une version de TET (MjTET) qui a été produite dans *Escherichia coli* et purifiée sous forme de dodécamère. La recherche de son activité enzymatique et de ses substrats peptidiques par des tests chromogéniques et fluorogéniques, ainsi que des études par HPLC en phase inverse, montre que cette enzyme est une leucine aminopeptidase activée par le cobalt se distinguant des autres aminopeptidases M42 par son très large spectre d'action qui s'étend aux résidus aromatiques. Une structure complète de cette aminopeptidase a été résolue en combinant la cristallographie (2,4 Å) et la cryo-EM (4,1 Å). L'analyse de la poche de spécificité de MjTET permet de mieux comprendre les bases structurales de la discrimination de substrat chez les TETs. De plus, l'analyse de la structure interne de la particule permet de proposer un nouveau mécanisme de navigation des peptides à l'intérieur des particules tétraédriques de la famille TET.

L'archée hyperthermophile *Pyrococcus horikoshii* comporte trois types de complexes TETs. L'étude d'une protéine présentant ~20 % d'identité avec ces systèmes, nous a permis d'identifier une quatrième version du système TET dans cet organisme : PhTET4. La protéine recombinante a été purifiée. Elle forme un complexe dodécamérique tétraédrique. Les études biochimiques révèlent que l'enzyme possède une spécificité très étroite dirigée exclusivement vers l'hydrolyse des résidus glycines de l'extrémité N-terminale des peptides. De plus, elle est activée par le nickel. Ces caractéristiques permettent de proposer que, chez les archées, la multiplication et la spécialisation des enzymes TETs seraient associées au caractère hétérotrophes alors que le système des archées autotrophes se réduirait à une TET unique apte à assurer une fonction de « ménage ».