



**HAL**  
open science

# Approche biocatalytique pour la synthèse de lactames

Océane Touré

► **To cite this version:**

Océane Touré. Approche biocatalytique pour la synthèse de lactames. Chimie organique. Université Paris-Saclay, 2024. Français. NNT : 2024UPASF027 . tel-04701603

**HAL Id: tel-04701603**

**<https://theses.hal.science/tel-04701603v1>**

Submitted on 18 Sep 2024

**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.

# Approche biocatalytique pour la synthèse de lactames

*Biocatalytic approach for lactam synthesis*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n°571, Sciences Chimiques : Molécules, Matériaux, Instrumentation et  
Biosystèmes (2MIB)

Spécialité de doctorat : Chimie

Graduate School : Chimie. Référent : Université d'Évry Val d'Essonne

Thèse préparée dans l'unité de recherche **Génomique métabolique** (Université Paris-Saclay, Univ Évry, CNRS, CEA), sous la direction d'**Anne ZAPARUCHA**, Professeure des Universités, Université d'Évry-Val d'Essonne

Thèse soutenue à Paris-Saclay, le 1<sup>er</sup> juillet 2024, par

**Océane TOURÉ**

## Composition du Jury

Membres du jury avec voix délibérative

### **Delphine JOSEPH**

Professeure des Universités,  
Université Paris-Saclay

Présidente

### **Rénato FROIDEVAUX**

Professeur des Universités,  
Université de Lille

Rapporteur & Examineur

### **Claire MOULIS**

Maîtresse de conférences, HDR,  
INSA Toulouse

Rapporteuse & Examinatrice

### **Egon HEUSON**

Chargé de recherche, HDR, Centrale  
Lille

Examineur



## REMERCIEMENTS

Premièrement, je tiens à remercier les membres de mon jury d'avoir accepté d'évaluer mes travaux de thèse menés durant ces quatre années. Je remercie les rapporteurs Pr Renato FROIDEVAUX et Dr Claire MOULIS pour leurs examens de mon manuscrit de thèse dans de courts délais, ainsi que Pr Delphine JOSEPH (présidente du jury) et Dr Egon HEUSON pour leur qualité en tant d'examineurs et les discussions enrichissantes qui ont suivi. Je remercie également Dr Christine HÉLAINE d'avoir accepté d'être présente en tant qu'invitée dans ce jury, mais aussi pour son suivi et ses conseils tant que membre de mon comité de suivi de thèse aux côtés de Dr Luis BERMUDEZ-HUMARAN.

Ensuite, je veux remercier ma directrice de thèse Pr Anne ZAPARUCHA, qui est également la chef de l'équipe Biocatalyse ! Merci de m'avoir prise en thèse et de m'avoir guidé tout au long de cette aventure. Malgré les difficultés (techniques) qu'on a pu rencontrer, j'espère avoir pu amener ces travaux sur les CoA ligases encore plus loin ! Je garderai toujours en souvenir ton rire très communicatif que j'entendais depuis le bureau ☺. Je remercie aussi Dr. Carine VERGNE-VAXELAIRE et Aurélie FOSSEY-JOENNE pour leur aide et leurs conseils qui m'ont également aidé tout au long de ces années de thèse. J'ai été heureuse d'intégrer cette équipe 100% féminine et avec tellement de bienveillance envers ses membres doctorants et stagiaires. Par la même occasion, je veux remercier l'ensemble de l'équipe actuelle de Biocatalyse. Tout d'abord, les doctorants Josemarco MENDOZA-AVILA et Célestin GAMONET (avec qui je forme l'équipe des trois mousquetaires qui restent jusqu'à 20h, et parfois plus tard) et les stagiaires actuels Camille BALAND (le meilleur binôme de bureau), Alban ADRIEN et Lou AMARO. Je remercie également tous les anciens membres de l'équipe que j'ai côtoyés durant ces 4 années au Genoscope : Laurine (le petit rayon de soleil), Chloé L. dont j'ai repris ses travaux après son excellente thèse, Dipesh (I loved acting as your translator and making fun of France and French people), Alix (mon partenaire de karaoké), Jonathan, Alex et Luca. Longue vie à l'équipe de Biocatalyse du L2BMS !

Je remercie également Jean-Louis PETIT et Adrien DEBARD de la plateforme FunDivEx avec qui l'équipe travaille en étroite collaboration ! Merci Jean-Louis (alias Dieu pour les intimes) pour ton expertise en biologie moléculaire. Merci pour ton aide et celui d'Adrien sur ce projet des CoA ligases qui vous aura donné du fil à retordre jusqu'au bout. Je remercie à nouveau Egon HEUSON et sa jeune équipe de doctorants et post-doctorants de m'avoir accueillie pendant 3 semaines au sein de leurs locaux à Centrale Lille. J'ai une pensée particulière pour Lucas BISEL, qui m'a aidé pour la production de 20 CoA ligases en vain, je me souviendrai tout de même des bonnes journées bien longues que cela nous a pris ! Je remercie aussi Oumaima RABAOUI qui nous a guidée durant les premiers jours de production. Je me souviendrai aussi de la French Press qui nous aura pris beaucoup de temps à nous tous ! Je remercie tout le reste de l'équipe pour votre chaleureuse accueil et aide durant mon séjour : Diya, Richard, Abel, Wissal et Juan. Profitez tous de cette aventure qu'est la thèse pour ceux qui sont actuellement doctorants et j'espère que nos chemins se recroiseront ! Je remercie aussi tous les partenaires impliqués dans l'ANR SMALA avec qui de bonnes discussions ont pu éclore sur ce projet prometteur.

Je tiens aussi à remercier l'école doctorale 2MIB de l'Université Paris-Saclay et sa direction pour le financement de la thèse mais aussi leur soutien pour l'aide à la mobilité pour partir en congrès ou écoles d'été. J'ai apprécié travailler avec vous pour l'organisation de la journée de l'école doctorale en 2022. J'ai également aimé pouvoir faire partie du conseil de l'ED en tant que représentante des doctorants du site d'Évry. Cela fut intéressant de connaître les coulisses du travail de tous les professeurs pour aider les doctorants durant la thèse mais aussi de pouvoir échanger avec vous pour améliorer les conditions du doctorat pour les étudiants. Par la même occasion, je remercie le département de Chimie



de l'UEVE de m'avoir permis de prolonger ma thèse pour une 4<sup>ème</sup> année grâce au poste de demi-ATER, qui m'aura été bénéfique pour finir mes travaux de recherches dans les meilleures conditions. Ce fut également un plaisir de travailler avec l'équipe enseignante du département de chimie durant ces quatre années. Enfin, je remercie Patrick WINCKER et Véronique de BERARDINIS, directeur et directrice adjointe du Genoscope, de m'avoir accueillie dans leurs locaux ainsi qu'au sein de l'UMR Génomique Métabolique.

Maintenant je passe à des remerciements un peu plus personnels et ça risque d'être long, je préfère prévenir ! Tout d'abord, je veux remercier les doctorant(e)s du Genoscope, notamment ceux du 1<sup>er</sup> étage que j'ai pu côtoyé durant ces quatre années de thèse ! Je commence par toi, Josemarco, and I'm switching in English just for you. We've already talked about it a million times, so I'll try to keep it brief because if I had to express everything that I feel, it would take a lot of pages to write! I'll never be grateful enough to have you in my life. I can never thank life or fate enough for bringing us together at Genoscope, and for giving birth to this friendship. You were (and you are still) my greatest support until the end, even when we were in different countries, and for that, words can't describe how grateful I am right now. You're the best person I've met in the last four years, and I'm happy and proud to count you among my closest friends. We've lived a whole lifetime together, we've shared our happiness, a lot of laughs but also our doubts and our fears and we've always been there for each other, even when we were far apart. I love you so much and I can't wait to be there for you as you were for me in the last moments of your life as a PhD student. I look forward to making new memories with you and to seeing the great scientist you will become, my dearest friend. La deuxième personne que je tiens à remercier, c'est toi mon très cher Briec. C'est dur à admettre, mais tes moqueries envers ma personne vont me manquer, ainsi que nos chamailleries, à savoir qui a tort ou raison... Même si on sait très bien entre nous que j'avais raison la plupart du temps (pour ne pas dire dans la majorité des cas) ! Ça va me manquer de rire avec toi au quotidien, ou même se moquer d'une certaine personne qui nous aura tapé sur les nerfs pendant un petit moment ! Je suis contente d'être ton amie et j'espère que tu ne m'oublieras pas d'aussi vite et qu'on se reverra bien hors du Genoscope bien sûr ! Merci pour ton soutien et ta bonne humeur depuis ces 3 dernières années ! Et que tu me crois ou pas, j'ai bien réussi à trouver Guillaume Le Conquérant en 4 coups au Pedantix ! Et je n'oublie pas nos paris, hâte de recevoir mes petites boîtes de Pringles goût Spicy s'il-te-plaît ! Courage à toi pour cette fin de ta thèse, je te sais bien entouré avec Marie et Léa et j'ai hâte d'assister à ta soutenance. Je remercie aussi les doctorants actuels : Célestin (ou celui qui n'arrive pas au labo avant 9h30, mais c'est en voie d'amélioration haha), Marie P. (la spécialiste des cages supramoléculaires fluorescentes et la reine du crochet), Florent (le chauve beau gosse et surtout le meilleur organisateur des GenoPub), Tom, Yasmine, Antoine et Eva. J'ai une pensée aussi pour toutes les doctorantes déjà parties avant moi et dont je garde de très bons souvenirs : Laurine, Deborah, Humbeline, Aysis, Marion, Nina, Chloé B., Chloé L. et Marie B. Ensuite, je remercie tous les stagiaires anciens comme actuels, surtout ceux de cette année qui sont au top ! Je pense forcément à vous les filles, Léa et Camille avec qui nous formons avec Marie P. le fameux groupe des Spice Girls du Genoscope (et je suis fière de ce nom de groupe) ! J'ai adoré nos pauses du matin ou de fin de journée dans le bureau à discuter de nos vies ! Je suis contente d'avoir fait cette année supplémentaire et de vous avoir toutes rencontrées ! Je te remercie à nouveau Camille, mon binôme de bureau pour nos bavardages interminables mais aussi ton soutien surtout dans cette fin de thèse ! Je te souhaite une bonne continuation pour la suite et j'ai hâte aussi de voir la brillante chercheuse que tu deviendras dans les prochaines années. Et je n'oublie pas pour ta start-up, c'est quand tu veux que tu m'appelles pour monter ta boîte ! Je te remercie également Léa pour la joie et la bonne humeur que tu as apporté dans ces derniers de mois de thèse. Profites bien de ton aventure de future thésarde et je ne doute pas que ta thèse sera couronnée de succès ! Merci à Samuel, grand stagiaire bossueur de la team CLD et surtout binôme de Léa, pour ta gentillesse et tes bonnes pâtisseries ! Je te remercie aussi ma petite discrète

Samira, j'espère que tu as apprécié ton temps au Genoscope et je te souhaite pleins de bonheur pour la suite. Je remercie aussi tous les membres du labo du 1<sup>er</sup> étage avec qui j'ai passé de bons moments comme Ivan d'une gentillesse incommensurable, Jean-Louis, Peggy, Sébastien, Agnès, Célia, Alaksh, Mélodie, Eddy parmi tant d'autres. Je garderai toujours en souvenir les pique-niques et les repas de Noël en votre compagnie ! Enfin, pour finir avec cette partie labo, je remercie l'ancienne femme de ménage Nicole, qui nous aimait comme ses enfants et qui restait avec nous les doctorants jusqu'à tard dans notre open space, merci pour tes encouragements et j'espère que tu profites bien de ta retraite bien méritée.

Maintenant je veux remercier tous mes amis qui m'ont soutenu durant ces quatre années de thèse ! Je vais commencer par vous remercier Abdoul, Oumou et Grégory ! Nous nous sommes connus dans le cadre de l'organisation de la JED 2022 et finalement une belle amitié est née, et j'en suis la plus heureuse ! Merci de m'avoir soutenue dans les derniers mois de rédaction et j'ai hâte de voir les futures carrières qui vous attendent ! Je remercie aussi mon Ordre du Phénix, Pauline et David, fidèles au poste depuis la licence ! Bien qu'on ne se voit plus aussi souvent que quand nous étions tous ensemble à SU, nous avons partagé de bons moments, notamment aller (enfin) aux studios Harry Potter à Londres avec vous ! C'est un moment qui restera gravé dans ma vie. Ensuite je remercie tous mes amis de master. Je commence par toi Edwyn, tu auras été d'un grand soutien depuis que nous nous connaissons, nous avons les mêmes objectifs et nous y voilà enfin, tous les deux docteurs ! Je suis contente de te compter parmi mes plus proches amis avec qui j'ai vécu mes plus grands délires (coucou la boîte de nuggets au goût de Fanta pour ne pas rater notre train) ! J'en profite pour remercier toute la bande des « idiots » : Irène et François, Laura et Gautier, Édouard et Thibaut (alias Titi pour les intimes) pour ces séjours à la ferme qui ne sont pas prêts de se finir ! Je suis contente de vous avoir dans ma vie et d'assister à chaque nouvelle étape de votre vie. Je remercie aussi Alicia la marraine et Noémie la globetrotteuse pour nos moments partagés depuis toutes ces années. Je te remercie aussi Christelle, tu es sûrement l'une de mes plus anciennes amies finalement, depuis le collège à Charles Péguy ! Je suis contente que nous ayons repris contact après tant d'années et de te compter parmi mes amis proches. Je remercie aussi mes anciennes colocataires, surtout toi Nourhène. Je n'aurais jamais cru que je ferais une aussi belle rencontre humaine en étant en colocation et je suis heureuse d'avoir pu partager le mois de Ramadan à tes côtés ainsi que tous les événements heureux qui se sont passés dans nos vies respectives. Je remercie aussi mes meilleures amies Aurélie, Elsa, Sonia et Solenn, de votre soutien et de tous les moments partagés durant ces dernières années. Notre amitié n'a fait que se renforcer bien que nous ayons toutes nos vies chacune de notre côté et je suis contente de vous avoir vu évoluer depuis toutes ces années.

Pour finir, je veux remercier ma famille. Je remercie infiniment mes parents, de m'avoir toujours soutenue moralement et financièrement et je me rends compte de la chance que j'ai d'avoir pu me consacrer pleinement à mes études supérieures durant ces dernières 9 années. Vous nous avez toujours poussés mes sœurs et moi à être ambitieuses, rêver en grand et surtout d'aller au plus loin dans nos études. Je me rends compte sans trop m'en rendre compte de ce que ça doit représenter pour vous, vous qui avez quitté la Côte d'Ivoire pour une vie meilleure en France. Ce diplôme de doctorat vous appartient également. Ce diplôme est aussi pour vous, mes sœurs Malika et Lola. Notre sororité n'a fait que se renforcer contre vents et marées et j'ai de la chance vous avoir comme sœurs. Je suis admirative des femmes que vous êtes devenues et que j'espère que nous resterons soudées pour toujours et que personne et rien n'arriveront à nous séparer. Je sais que je ne le dis pas souvent voire rarement et je m'en excuse si je n'ai pas su le montrer : sachez que je vous aime ma famille et je vous remercie pour votre soutien depuis toujours.



# SOMMAIRE

REMERCIEMENTS.....	1
SOMMAIRE .....	5
TABLE DES FIGURES .....	10
TABLE DES SCHEMAS.....	13
TABLE DES TABLEAUX.....	16
TABLE DES ANNEXES .....	18
LISTE DES ABREVIATIONS .....	20
INTRODUCTION GENERALE .....	27
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>31</b>
<b>CHAPITRE I : ETAT DE L'ART SUR LA SYNTHÈSE ENZYMATIQUE D'AMIDES ET DE LACTAMES</b> .....	<b>35</b>
<b>I.1. ETAT DE L'ART SUR LA SYNTHÈSE ENZYMATIQUE DE LIAISON AMIDE .....</b>	<b>36</b>
<b>I.1.1. FORMATION DE LIAISON AMIDE DANS LE VIVANT.....</b>	<b>36</b>
I.1.1.1. Synthèse peptidique ribosomale.....	36
I.1.1.2. Peptides synthétases non ribosomiques (NRPS).....	37
I.1.1.3. Autres enzymes formant des liaisons amides dans le vivant .....	39
I.1.1.3.1. Enzymes dépendantes de l'ATP .....	39
I.1.1.3.1.1. Enzymes capteuses d'ATP .....	40
I.1.1.3.1.2. Enzymes adénylantes .....	41
I.1.1.3.2. Transpeptidases.....	42
I.1.1.3.3. N-acyltransférases .....	43
I.1.1.3.4. Nitriles hydratases .....	44
<b>I.1.2. SYNTHÈSE BIOCATALYTIQUE DE LIAISON AMIDE .....</b>	<b>44</b>
I.1.2.1. Hydrolases .....	45
I.1.2.1.1. Lipases .....	46
I.1.2.1.1.1. Aminolyse d'esters .....	46
I.1.2.1.1.2. Amidation directe d'acides carboxyliques .....	47
I.1.2.1.2. Pénicilline acylases .....	48
I.1.2.1.3. Protéases à serine .....	49
I.1.2.2. Acyltransférases .....	50
I.1.2.3. Enzymes dépendantes de l'ATP .....	52
I.1.2.3.1. Formation de liaisons amides par catalyse de l'amidation .....	52
I.1.2.3.1.1. Enzymes capteuses d'ATP .....	52
I.1.2.3.1.2. Amide Bond Synthétases (ABS) .....	53
I.1.2.3.2. Détournement de l'activité de l'enzyme.....	54

I.1.2.3.2.1. Domaines d'adénylation autonomes des NRPS .....	55
I.1.2.3.2.2. Carboxylic acid reductases .....	55
I.1.2.3.2.3. CoA ligases .....	57
<b>I.1.3. CONCLUSION SUR L'ETAT DE L'ART POUR LA SYNTHÈSE ENZYMATIQUE DE LIAISON AMIDE .....</b>	<b>59</b>
<b>I.2. ETAT DE L'ART DE LA SYNTHÈSE ENZYMATIQUE DE LACTAMES .....</b>	<b>59</b>
<b>I.2.1. SYNTHÈSE DE LACTAMES DANS LE VIVANT .....</b>	<b>60</b>
<b>I.2.2. SYNTHÈSE BIOCATALYTIQUE DE LACTAMES .....</b>	<b>62</b>
I.2.2.1. Synthèse biocatalytique de lactames par fermeture de cycle par liaison C-N ou C-C .....	62
I.2.2.1.1. Galactose oxydases et alcool déshydrogénases .....	63
I.2.2.1.2. Laccases et peroxydases .....	64
I.2.2.1.3. Synthèse biocatalytique de lactames avec des réactions « new-to-nature » ...	65
I.2.2.1.3.1. Ène-réductases .....	65
I.2.2.1.3.2. Metalloenzymes naturelles et artificielles .....	67
I.2.2.2. Synthèse biocatalytique de lactames à partir d'acides $\omega$ -aminés .....	70
I.2.2.2.1. Cyclisation spontanée d'esters d'acides $\omega$ -aminés .....	70
I.2.2.2.1.1. Transaminases .....	70
I.2.2.2.1.2. Hydrolases .....	72
I.2.2.2.2. Les enzymes dépendantes de l'ATP .....	73
<b>1.3. CONCLUSION SUR L'ETAT DE L'ART POUR LA SYNTHÈSE ENZYMATIQUE DE LIAISON AMIDE ET DE LACTAMES .....</b>	<b>75</b>
<b>RÉFÉRENCES BIBLIOGRAPHIQUES .....</b>	<b>77</b>
<b>CONTEXTE DE LA THESE ET OBJECTIFS .....</b>	<b>87</b>
<b>RÉFÉRENCES BIBLIOGRAPHIQUES .....</b>	<b>89</b>
<b>CHAPITRE II : INGENIERIE PROTEIQUE D'UNE COA LIGASE POUR LA SYNTHÈSE DE LACTAMES : CAS DE L'ENZYME A4YDT1 .....</b>	<b>91</b>
<b>II.1. DESCRIPTION STRUCTURALE DE LA COA LIGASE DES COA LIGASES ET RESULTATS PRELIMINAIRES POUR LA FORMATION DE LACTAMES AVEC A4YDT1 .....</b>	<b>92</b>
<b>II.1.1. STRUCTURE GLOBALE DES COA LIGASES ET LEUR MECANISME .....</b>	<b>92</b>
<b>II.1.2. PRESENTATION DE L'ENZYME A4YDT1 ET RESULTATS PRELIMINAIRES POUR LA SYNTHÈSE DE LACTAMES .....</b>	<b>95</b>
<b>II.2. ETUDE STRUCTURALE DE LA COA LIGASE A4YDT1 .....</b>	<b>96</b>
<b>II.2.1. CRISTALLISATION ET DETERMINATION DE LA STRUCTURE D'A4YDT1 .....</b>	<b>96</b>
II.2.1.1. Cristallisation d'A4YDT1 .....	96
II.2.1.2. Structure globale d'A4YDT1 .....	97
II.2.1.2.1. Site de liaison de l'AMP du site actif de A4YDT1 .....	100
II.2.1.2.2. Site de liaison du substrat dans le site actif de A4YDT1 .....	100
<b>II.2.2. MODIFICATION DU SITE DE LIAISON DU SUBSTRAT DANS LE SITE ACTIF DE A4YDT1 .....</b>	<b>102</b>

II.2.2.1. Docking des intermédiaires acyl-AMP dans le site actif et choix des résidus à muter .....	102
II.2.2.2. Mutagenèse dirigée du site de liaison du substrat et surexpression des mutants .....	103
<b>II.3. SYNTHÈSE DE LACTAMES ET D'AMIDES AVEC LES MUTANTS DE A4YDT1 .....</b>	<b>104</b>
<b>II.3.1. ACTIVITÉ DES MUTANTS POUR LA SYNTHÈSE DE LACTAMES .....</b>	<b>104</b>
<b>II.3.2. SYNTHÈSE DE LACTAMES AVEC LA 2<sup>ÈME</sup> SÉRIE DE MUTANTS DE A4YDT1 .....</b>	<b>106</b>
<b>II.3.3. SPECTRE D'ACTIVITÉ DES MUTANTS D'A4YDT1 SUR DES SUBSTRATS ACIDES POUR LA SYNTHÈSE D'AMIDES .....</b>	<b>107</b>
<b>II.3.4. PARAMÈTRES CINÉTIQUES DES MEILLEURS MUTANTS .....</b>	<b>108</b>
II.3.4.1. Principe du suivi cinétique de l'adénylation par la méthode NADH-PPi.....	110
II.3.4.2. Paramètres cinétiques de A4YDT1 et des mutants W259G et W259G/V238T .....	111
<b>II.4. CONCLUSION ET DISCUSSION SUR L'INGÉNIERIE PROTÉIQUE DE L'ENZYME A4YDT1 .....</b>	<b>112</b>
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>114</b>
<b>CHAPITRE III : EXPLORATION DE LA BIODIVERSITÉ POUR LA RECHERCHE D'ENZYMES POUR LA SYNTHÈSE BIOCATALYTIQUE DE LACTAMES .....</b>	<b>117</b>
<b>III.1. CRIBLAGE DE LA 1<sup>ÈRE</sup> COLLECTION DE COA LIGASES (COA LIG1) .....</b>	<b>118</b>
<b>III.1.1. RECHERCHE DE NOUVELLES ENZYMES DANS LA BIODIVERSITÉ PAR APPROCHE GÉNOMIQUE. 118</b>	
III.1.1.1. Approche génomique pour la recherche de CoA ligases.....	119
III.1.1.2. Set de référence de CoA Lig1 .....	119
<b>III.1.2. CRIBLAGE DE COA LIG1 .....</b>	<b>120</b>
III.1.2.1. Choix des substrats pour le criblage.....	120
III.1.2.2. Méthodes de suivi de l'activité des CoA ligases en conditions de criblage.....	123
III.1.2.2.1. Méthode de suivi colorimétrique de l'adénylation : dosage au vert de malachite (MG assay).....	123
III.1.2.2.2. Cascade pyruvate kinase-lactate déshydrogénase (cascade PK-LDH).....	124
<b>III.1.3. RESULTATS DU CRIBLAGE DE COA LIG1 .....</b>	<b>125</b>
III.1.3.1. Identification des hits.....	125
III.1.3.2. Vérification par UHPLC-MS ou UHPLC-UV .....	131
<b>III.2. ANALYSE DES RESULTATS COA LIG1 ET NOUVEAU SET DE RÉFÉRENCE.....</b>	<b>135</b>
<b>III.2.1. ALIGNEMENT DE SÉQUENCES DES COA LIGASES SÉLECTIONNÉES.....</b>	<b>136</b>
<b>III.2.2. SUPERPOSITION DES SITES ACTIFS.....</b>	<b>137</b>
<b>III.3. CONSTRUCTION ET PRODUCTION DE COA LIG2.....</b>	<b>140</b>
<b>III.3.1. CONSTRUCTION DE LA DEUXIÈME COLLECTION DE COA LIGASES.....</b>	<b>140</b>
<b>III.3.2. PRODUCTION DE LA DEUXIÈME COLLECTION DE COA LIGASES .....</b>	<b>141</b>
<b>III.3.3. PRODUCTION DES 20 ENZYMES.....</b>	<b>145</b>
III.3.3.1. Production de la collection réduite de CoA ligase à l'UCCS (Centrale Lille) .....	146

III.3.3.2. Sous-clonage des 20 CoA ligases dans pET-22b(+)	146
<b>III.4. CRIBLAGE DE LA COLLECTION REDUITE DE 20 COA LIGASES</b>	<b>147</b>
<b>III.5. CONCLUSION ET PERSPECTIVES SUR L'EXPLORATION DE LA BIODIVERSITE POUR LA RECHERCHE DE NOUVELLES COA LIGASES POUR LA SYNTHESE DE LACTAMES</b>	<b>148</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b>	<b>149</b>
<b>CONCLUSION ET PERSPECTIVES</b>	<b>151</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b>	<b>153</b>
<b>EXPERIMENTAL SECTION</b>	<b>154</b>
<b>GENERALS</b>	<b>154</b>
<b>CHEMICALS</b>	<b>154</b>
<b>EQUIPMENT</b>	<b>154</b>
<b>ENZYME PRODUCTION</b>	<b>154</b>
<b>GENERAL PROCEDURE FOR ENZYME OVEREXPRESSION</b>	<b>154</b>
<b>GENERAL METHOD FOR SMALL-SCALE PURIFICATION OF ENZYMES</b>	<b>155</b>
<b>GENERAL METHOD FOR LARGE-SCALE PRODUCTION OF ENZYMES</b>	<b>155</b>
<b>GENERAL METHOD FOR 96-WELL PLATE PURIFICATION</b>	<b>155</b>
<b>GENERAL METHOD FOR SITE-DIRECTED MUTAGENESIS</b>	<b>155</b>
PCR primers	155
Site-directed mutagenesis protocol	156
Expression and purification of the mutants	156
<b>ANALYTICAL BIOCATALYTIC REACTIONS</b>	<b>156</b>
<b>SAMPLE PREPARATION AND REACTION</b>	<b>156</b>
Analytical biocatalytic reactions in Chapter II	156
Biocatalytic reactions for lactam synthesis with A4YDT1 and A4YDT1 mutants	156
Biocatalytic reactions for N-methylamide synthesis with A4YDT1 and A4YDT1 mutants	157
Analytical biocatalytic reactions for kinetic parameters and specific activity	157
Kinetic parameters of A4YDT1 and A4YDT1 variants W259G and W259G/V238T	157
Specific activities of A4YDT1 variants for N-methylamide synthesis	158
Analytical biocatalytic reactions in Chapter III	158
Screening of the collection CoA Lig1	158
Screening of the collection CoA Lig2	159
<b>SAMPLE TREATMENT</b>	<b>159</b>
For UHPLC-MS (ESI and APCI) analyses of lactams	159
For UHPLC-MS (ESI) analyses of N-methylamides	159
<b>UHPLC-UV/MS AND DERIVATIZATION PROCEDURE WITH BENZOYL CHLORIDE</b>	<b>159</b>
Benzoyl chloride derivatization	159

UPLC-UV/MS (ESI) analyses .....	160
<b>UHPLC-MS ANALYSES .....</b>	<b>160</b>
Non-functionalized lactams and functionalized lactams with non-polar groups (APCI) ..	160
Functionalized lactams with carboxylic group (ESI) .....	160
<b>SPECTROPHOTOMETRIC ASSAYS FOR ENZYMATIC ACTIVITIES .....</b>	<b>160</b>
<b>96-MICROWELL PLATE ENZYMATIC SCREENING ASSAYS .....</b>	<b>160</b>
Pyruvate Kinase – Lactate Dehydrogenase (PK-LDH) cascade: lactam formation monitoring .....	160
Malachite Green assay .....	161
<b>PYROPHOSPHATE REAGENT (SIGMA-ALDRICH®) FOR ENZYMATIC ACTIVITIES .....</b>	<b>161</b>
<b>STANDARD SYNTHESIS .....</b>	<b>162</b>
<b>SYNTHESIS OF N-METHYLHEXANAMIDE.....</b>	<b>162</b>
<b>SYNTHESIS OF N-METHYLHEPTANAMIDE.....</b>	<b>162</b>
<b>SYNTHESIS OF N-METHYLOCTANAMIDE.....</b>	<b>163</b>
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>170</b>
<b>ANNEXE.....</b>	<b>171</b>



## TABLE DES FIGURES

Figure 1. Exemples de composés contenant une liaison amide. ....	27
Figure 2. Exemples de composés naturels et synthétiques avec le motif lactame. ....	28
Figure 3. <b>(A)</b> Schéma de la structure d'un ribosome avec les domaines A, P et E. La chaîne violette correspond à la séquence peptidique en cours de synthèse par le ribosome. Créée avec BioRender.com. <b>(B)</b> Réaction d'activation de l'acide aminé par les aa-ARNt synthétases. <b>(C)</b> Formation de la liaison peptidique au sein du ribosome. ....	37
Figure 4. Exemples de peptides non ribosomiques biosynthétisés par les NRPSs. ....	38
Figure 5. <b>(A)</b> Représentation schématique des modules constituant les NRPSs. Inspiré de l'article de Süssmuth et Mainz (2017). A correspond au domaine d'adénylation, PCP au peptidyl carrier protein, C au domaine de condensation et TE à la thioestérase. <b>(B)</b> Mécanisme de la formation de la liaison peptidique par les NRPSs. ....	39
Figure 6. Stratégies employées pour la formation de liaisons amides par les enzymes dépendantes de l'ATP. ....	52
Figure 7. <b>(A)</b> Représentation schématique d'une CAR avec ses trois domaines : d'adénylation (A), de thiolation (T) et de réduction (R). <b>(B)</b> Mécanisme de la réaction catalysée par les CARs. ....	56
Figure 8. <b>(A)</b> Représentation des modules du gène TAS1 pour la biosynthèse de l'acide tenuazonique. <b>(B)</b> Mécanisme de la biosynthèse de l'acide tenuazonique par le système enzymatique codé par le gène TAS1 avec le domaine de condensation (C), le domaine peptidyl carrier protein (PCP) du NRPS, et le domaine cétyosynthase (KS) du PKS. ....	62
Figure 9. Objectifs de thèse pour la recherche CoA ligases pour la synthèse de lactames. ....	88
Figure 10. Structure tridimensionnelle de l'acétyl-CoA ligase Acs de <i>Salmonella enterica</i> en conformation de thioestérification (PDB : 1PG4) (Gulick <i>et al.</i> , 2003). Le domaine N-terminal est représenté en bleu-violet et le domaine C-terminal en rose clair. Structure co-cristallisée avec l'intermédiaire adénylé propyl-AMP représenté en vert, le HSCoA en jaune-orange et le cation Mg <sup>2+</sup> représenté par une sphère cyan. ....	92
Figure 11. Structure tridimensionnelle de la CoA ligase Acs cristallisée en conformation de thioestérification (PDB : 1PG4). Les motifs A3 et A10 impliqués dans l'adénylation sont indiqués en rouge et bleu marine respectivement. Le motif A8 contenant la région charnière à l'interface des deux domaines est représenté en jaune. L'intermédiaire propyl-AMP est représenté en vert tandis que le HSCoA est représenté en jaune-orange. Le cation Mg <sup>2+</sup> est représenté par une sphère cyan. ....	93
Figure 12. Représentation tridimensionnelle des deux conformations d'une 4-chlorobenzoate CoA ligase (4-CBL) de <i>Alcaligenes sp. AL3007</i> (Wu <i>et al.</i> , 2008 ; Reger <i>et al.</i> , 2008). Le domaine N-ter est représenté en cyan et le domaine C-ter en rose. <b>(A)</b> Conformation d'adénylation avec l'intermédiaire adénylé du substrat, 4-chlorobenzoate-AMP, représenté en vert (PDB : 3CW8). <b>(B)</b> Conformation de thioestérification avec le produit thioester 4-chlorobenzoate-CoA représenté en vert, l'AMP représenté en violet et Mg <sup>2+</sup> représenté par une boule jaune (PDB : 3CW9). ....	94
Figure 13. Rendement en $\gamma$ -butyrolactame, $\delta$ -valérolactame et $\epsilon$ -caprolactame avec la CoA ligase A4YDT1. Conditions réactionnelles : 5 mM substrat $\omega$ -acide aminé, 5 mM ATP, 5 mM MnCl <sub>2</sub> , 50 mM tampon MOPS pH 8.5, 2 mg/mL A4YDT1 purifié. 24h, 60°C. Analyse par UHPLC-MS (APCI). ....	96
Figure 14. Structures cristallisées de la CoA ligase A4YDT1 en conformation de thioestérification, en présence d'AMP représenté en vert (PDB : 8BIQ) et en présence d'acétyl-AMP représenté en vert et du HSCoA représenté en rose clair (PDB : 8BIT). Les domaines N-ter et C-ter sont représentés en cyan et beige respectivement et la région charnière en rose. ....	98
Figure 15. Alignement des séquences protéiques ClustalOmega d'A4YDT1 et des CoA ligases servant pour la reconstruction structurale de A4YDT1. Les motifs structuraux sont encadrés et indiqués sur la figure. ....	99

Figure 16. Structure tridimensionnelle de A4YDT1 en conformation de thioestérification avec les motifs A3 et A10 impliqués dans l'adénylation, représentés en rouge et bleu marine respectivement. Le motif A8 correspondant à la région charnière est coloré en rose foncé. L'acétyl-AMP et le HSCoA sont colorés en vert et en rose pâle respectivement. .... 99

Figure 17. **(A)** Diagramme schématique montrant les interactions entre le HSCoA, l'AMP et le substrat avec les résidus présents dans le site actif de A4YDT1, basé sur l'analyse de sa structure. Les résidus appartenant au domaine N-ter sont indiqués en bleu et vert et les résidus du domaine C-ter sont indiqués en orange. **(B)** Zoom sur les résidus en interaction avec l'AMP (en jaune) dans la structure 3D de A4YDT1. Figures faites par Dr Nikolas Capra. .... 100

Figure 18. Zoom sur les résidus constituant le site de liaison du substrat dans la partie N-ter de A4YDT1 (PDB : 8BIT), colorés en jaune. L'intermédiaire acétyl-AMP est coloré en blanc. .... 101

Figure 19. **(A)** Docking de l'intermédiaire 4-aminobutyryl-AMP (en gris) dans le site de liaison du substrat de A4YDT1. **(B)** Docking des intermédiaires 4-aminobutyryl-AMP (en rose), 5-aminopentanoyl-AMP (vert pomme) et 6-aminohexanoyl-AMP (cyan) dans le site de liaison de substrat de A4YDT1. Les résidus constituant le site de liaison du substrat sont colorés en bleu. .... 102

Figure 20. Docking des intermédiaires 5-aminopentanoyl-AMP en orange **(A)** et 6-aminohexanoyl-AMP **(B)** en jaune en présence de la mutation W259G représentée par une sphère bleue. .... 103

Figure 21. Rendements en  $\gamma$ -butyrolactame,  $\delta$ -valérolactame et  $\epsilon$ -caprolactame. Conditions réactionnelles : 5 mM substrat acide  $\omega$ -aminé, 5 mM ATP, 5 mM  $MnCl_2$ , 50 mM tampon MOPS pH 8.5 ; 2 mg/ml d'enzyme purifiée. 24h, 60°C, 400 rpm. Analyse par UHPLC-MS, mode APCI. .... 105

Figure 22. Rendements de la 2<sup>ème</sup> série de mutants en relatif par rapport à A4YDT1 WT pour la synthèse de  $\gamma$ -butyrolactame (62% de rendement), et W259G pour la synthèse de  $\delta$ -valérolactame (13% de rendement) et de  $\epsilon$ -caprolactame (30% de rendement). Conditions réactionnelles : 5 mM substrat acide  $\omega$ -aminé, 5 mM ATP, 5 mM  $MnCl_2$ , 50 mM tampon MOPS pH 8.5, 2 mg/ml d'enzyme purifiée. 60°C, 24h, 400 rpm. Analyse par UHPLC-MS (APCI). .... 106

Figure 23. Comparaison du spectre d'activité de A4YDT1 et des mutants A4YDT1 avec divers substrats acides par détection du produit N-méthylamide correspondant. Conditions réactionnelles : 5 mM substrat acide, 5 mM ATP, 5 mM  $MgCl_2$ , 50 mM méthylamine, 50 mM tampon phosphate pH 8, 0.2 mg/ml d'enzyme purifiée. 60°C, 24h, 400 rpm. .... 108

Figure 24. Récapitulatif des différentes stratégies possibles pour l'approche génomique pour la recherche de nouvelles enzymes. .... 118

Figure 25. Approche génomique mis en place au laboratoire pour la recherche de nouvelles enzymes au sein de la biodiversité. .... 119

Figure 26. Liste des 17 acides  $\omega$ -aminés répartis par pool pour le criblage de la collection de CoA ligases CoA Lig1. .... 122

Figure 27. Heat map des activités des CoA ligases sélectionnées pour le pool 1 suite à la détection des lactames attendus (**1b-3b**) par UHPLC-MS (APCI). Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM  $MnCl_2$ , 1,5 mM  $MgCl_2$ , 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Analyse par UHPLC-MS (APCI). Rendements déduits d'une gamme de calibration avec les standards commerciaux. Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. .... 132

Figure 28. Heat map des activités des CoA ligases sélectionnées pour le pool 2 suite à la détection des lactames attendus (**4b-6b**) par UHPLC-UV. Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 2 mM ATP, 2 mM  $MnCl_2$ , 2 mM  $MgCl_2$ , 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. Légende : n.d. pour non déterminé. .... 133

Figure 29. Heat map des activités des CoA ligases sélectionnées pour le pool 3 suite à la détection des lactames attendus ( <b>8b-10b</b> ) par UHPLC-UV. Conditions réactionnelles : 0,5 mM de chaque acide $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl <sub>2</sub> , 1,5 mM MgCl <sub>2</sub> , 50 mM tampon MOPS pH 8,5, 10 $\mu$ L d'enzyme purifiée. 60°C sans agitation à l'étuve, 5h. Analyse par UHPLC-MS (ESI). Aires détectées à partir du m/z attendus pour les lactames correspondants. Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. Légende : n.d. pour non déterminé. ....	134
Figure 30. Heat map des activités des CoA ligases sélectionnées pour le pool 6 suite à la détection des lactames attendus ( <b>1b-3b</b> ) par UHPLC-MS (APCI). Conditions réactionnelles : 0,5 mM de chaque acide $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl <sub>2</sub> , 1,5 mM MgCl <sub>2</sub> , 50 mM tampon MOPS pH 8,5, 10 $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Détection des aires correspondants aux lactames attendus par UHPLC-MS (APCI). Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. Légende : ++ pour aire détectée forte. ....	135
Figure 31. Matrice d'identité avec les CoA ligases sélectionnées pour former le nouveau set de référence. ....	137
Figure 32. Alignement des séquences peptidiques des CoA ligases constituant le nouveau set de référence par ClustalOmega. Les motifs A3 et A10 essentiels pour l'adénylation sont indiqués en vert et la région charnière A8 est encadrée en vert. ....	137
Figure 33. Comparaison des résidus constituant le site de liaison du substrat et de l'AMP pour les CoA ligases ayant les plus fortes homologies de séquence : ( <b>A</b> ) D1A9C5 et K6P2L2 avec 62% d'identité et ( <b>B</b> ) Q97UW7 et O28246 avec 50% d'identité. Identification des résidus effectuée par superposition des structures avec A4YDT1 comme CoA ligase de référence sur le logiciel PyMOL. ....	139
Figure 34. Résidus formant le site de liaison du substrat et de l'AMP dans le site actif des CoA ligases Q3ABN2 et K6Q029. Identification des résidus effectuée par superposition des structures avec A4YDT1 comme CoA ligase de référence sur le logiciel PyMOL. ....	139
Figure 35. Résidus formant le site de liaison du substrat et de l'AMP dans le site actif de la CoA ligase F2NPU2. Identification des résidus effectuée par superposition des structures avec A4YDT1 comme CoA ligase de référence sur le logiciel PyMOL. ....	140
Figure 36. Identification des résidus constituant les sites actifs suite à l'analyse ASMC en utilisant la structure de la CoA ligase A4YDT1 comme enzyme référente. En rouge est indiqué les résidus situés en fond de poche pour chaque CoA ligase. ....	142
Figure 37. Matrice d'identité des 20 CoA ligases sélectionnées pour former la collection réduite de CoA Lig2 et les 7 CoA ligases de référence. ....	144
Figure 38. Proportions des résidus présents chez les 20 CoA ligases sélectionnées en position équivalente aux résidus présents dans le site actif d'A4YDT1 suite à l'alignement de séquences avec le logiciel ASMC. Figure faite sur Web Logo 3 par Dr Eddy Elisée (LABGeM). ....	145
Figure 39. Heat map des activités des 8 CoA ligases de la collection réduite de CoA Lig2 suite à la cascade PK-LDH. Les enzymes du set de référence sont indiquées en bleu. Expérience réalisée en duplicat. Figure réalisée par Jean-Louis Petit de la plateforme FunDivEx. ....	147

## TABLE DES SCHEMAS

Schéma 1. Schéma de la délocalisation électronique de la liaison amide (phénomène de résonance).....	27
Schéma 2. (A) Schéma des deux stratégies d'amidation stœchiométrique et (B) son mécanisme réactionnel avec l'attaque nucléophile de l'amine.....	35
Schéma 3. Exemple d'amidation stœchiométrique à partir de chlorure de thionyle.....	35
Schéma 4. Exemples d'amidation catalytique par un organocatalyseur à base de bore (A) et avec un métal de transition, le zirconium (B).....	36
Schéma 5. Mécanisme d'activation de la fonction acide carboxylique par les enzymes ATP-dépendantes : les enzymes capteurs de l'ATP (A) et les enzymes adénylantes (B). Schéma inspiré de l'article Goswami <i>et al.</i> (2015). Légende : GP pour groupe partant et Ad pour adénosine.....	40
Schéma 6. Exemple de formation de liaison amide catalysées par des enzymes capteuses d'ATP.....	40
Schéma 7. Exemples de réactions métaboliques catalysées par des Amide Bond Synthétases (ABSs). (A) CouL, (B) NovL et CloL, (C) SimL et (D) McbA.....	42
Schéma 8. Mécanisme réactionnel de la formation de la liaison peptidique par (A) des transpeptidases à cystéine ou (B) à sérine.....	43
Schéma 9. Mécanisme réactionnel de la formation de liaison amide par les NATs.....	43
Schéma 10. (A) Mécanisme réactionnel de la formation de la liaison amide avec la NAT CapW. (B) Structure du cofacteur SAM.....	44
Schéma 11. Mécanisme réactionnel de formation de l'amide primaire par les NHases.....	44
Schéma 12. Mécanisme général pour la formation d'amides avec les hydrolases (A) et le mécanisme réactionnel détaillé avec la triade catalytique His-Asp-Ser (B). Reproduction du schéma de l'article de Lima <i>et al.</i> (Lima, dos Anjos et al. 2019).....	45
Schéma 13. Formation de liaison amide avec la CCL dans un solvant organique.....	46
Schéma 14. Formation d'un amide chirale en cascade avec une transaminase mutante (ATA) et CAL B immobilisés.....	46
Schéma 15. Exemples d'amidation directe en partant d'acide carboxylique avec CAL B. (A) Avec utilisation d'éthanol comme solvant et réactif. (B) Dans un solvant ionique ou sans solvant avec retrait de l'eau formée sous atmosphère réduite.....	47
Schéma 16. Synthèse de N-benzyl-3-phénylpropanamide par la lipase SpL en présence d'eau dans le n-hexane.....	47
Schéma 17. Synthèse de la N-(4-pyrdinylméthyl)acétamide (dérivé de N-picolineamide) par CAL B.....	48
Schéma 18. Exemple de synthèse de N-acylglycine avec la lipase mutée de <i>Rhizomucor miehei</i> (RML) dans un milieu réactionnel eau-glycérol.....	48
Schéma 19. Exemple de synthèse biocatalytique de pénicillines semi-synthétiques par des pénicillines acylases (PAs) en partant de l'hydrolyse de l'ampicilline. Légende : <b>6-APA</b> : acide 6-aminopénicillanique ; <b>D-PGA</b> : D-phénylglycine amide ; <b>D-HPGA</b> : D- <i>p</i> -hydroxyphénylglycine amide ; <b>7-ADCA</b> : acide 7-aminodésacétocéphalosporanique.....	49
Schéma 20. (A) Synthèse d'un amide intermédiaire dans la synthèse du Losmapimod par 66CaaT. (B) Structure du Losmapimod.....	50
Schéma 21. Synthèse d'amides avec MsAcT. (A) Synthèse de l'acide suberalinique (SA), un intermédiaire du SAHA et (B) synthèse d'amides en cascade avec une transaminase en partant de benzaldéhyde. Légende : <b>CHES</b> : acide N-cyclohexyl-2-aminoéthanesulfonique, <b>Sp-ATA</b> : transaminase de <i>Silicibacter pomeroyi</i> .....	51

Schéma 22. (A) Réaction naturellement catalysée par la transglutaminase et (B) celle utilisée pour des applications biocatalytiques.....	51
Schéma 23. Exemples de formation de liaison peptidiques par les L-amino acid ligases ywfE (A) et TabS (B).....	53
Schéma 24. Synthèse du moclobémide par l'ABS McbA, couplé à un système enzymatique de régénération de l'ATP avec les polyphosphates kinases de <i>Sinorhizobium meliloti</i> (SmPPK <sub>2</sub> -I) et de <i>Acinetobacter johnsonii</i> (A <sub>J</sub> PPK <sub>2</sub> -II). .....	54
Schéma 25. Synthèse du cinépezide (vasodilatateur) catalysée par ShABS. Légende : IPase pour phosphatase inorganique.....	54
Schéma 26. (A) Formation de liaison amide avec DhE entre de l'acide 2,3-dihydroxybenzoïque (DHB) et de la cystéine. (B) Formation de liaison amide avec SfaB entre l'acide heptanoïque et l'arginine comme donneur d'amine.....	55
Schéma 27. Exemples de formation de liaison amide par le domaine d'adénylation de CAR <sub>mm</sub> (CAR <sub>mm</sub> -A) couplé à un système de régénération de l'ATP. (A) Synthèse de cinépezide (vasodilatateur) et (B) de tensioactifs MEGAs. ....	57
Schéma 28. Réaction naturelle catalysée par les CoA ligases.....	57
Schéma 29. Exemples de synthèse biocatalytique de liaison amide avec les CoA ligases AcsA (A) et At4CL2 (B).....	58
Schéma 30. Formation de liaison amide avec les CoA ligases thermophiles A4YDT1 ou A4IBN3, couplée à un système de régénération enzymatique de l'ATP avec la PPK2-III de <i>Deinococcus geothermalis</i> (DgPPK2-III) et la PPase de <i>Pyrococcus horikoshii</i> (PhPPase). .....	59
Schéma 31. β-lactamisation par β-lactame synthétase (A) et la cabanepam synthase (B).....	60
Schéma 32. Synthèse du noyau lactame via une oxydation par l'isopénicilline synthase (IPNS) et (B) son mécanisme réactionnel. Le schéma du mécanisme réactionnel est une image issue de l'article de Chapman (Chapman and Rutledge 2021). .....	61
Schéma 33. Synthèse biocatalytique de lactames par une GOase en cascade avec une xanthine déshydrogénase (XDH) ou une aldéhyde oxydase (PaoABC). .....	63
Schéma 34. Synthèse biocatalytique de lactames par une ADH de foie de cheval (HLADH pour <i>Horse Liver ADH</i> en anglais). .....	63
Schéma 35. Mécanisme général de formation de l'oxazolidinone par la peroxydase de racine de Raifort (HRP) couplée à un système de régénération enzymatique du peroxyde d'hydrogène par une glucose oxydase (GOx). .....	64
Schéma 36. Synthèse d'oxazolidinone par la peroxydase de racine de Raifort (HRP) couplée à un système de régénération du H <sub>2</sub> O <sub>2</sub> par une glucose oxydase (GOx), ou par la laccase CotA. ....	64
Schéma 37. Cascade enzymatique pour la synthèse de 1-hydroxypyrrolidine par la lipase CAL B et la peroxydase HRP.....	65
Schéma 38. Mécanisme réactionnel de réduction des alcènes par des EREDs.....	66
Schéma 39. Réaction "new-to-nature" pour la synthèse de lactames à partir de de α-chloroamides par les EREDs. ....	66
Schéma 40. Synthèse de γ-butyrolactames et 2-azocanones avec les mutants de ERED de <i>Gluconobacter oxydans</i> (GluER). .....	67
Schéma 41. Synthèse de dérivé de γ-butyrolactame par un CYP450 avec une activité « new-to-nature » de cyclisation radicalaire par transfert d'atomes (P450 <sub>ATRC<sub>ase3</sub></sub> ). .....	68
Schéma 42. Synthèse de dérivés de γ-butyrolactames avec une CYP450 par une activité "new-to-nature" de cyclase de radicaux aromatiques (P450 <sub>arc</sub> ). .....	68
Schéma 43. Synthèse d'un dérivé de δ-valérolactame par une ArM à base complexe de rhodium dans la structure biotine-streptavidine. ....	69

Schéma 44. Synthèse d'un dérivé de $\gamma$ -butyrolactame à partir de dioxazolone par une ArM à base de complexe d'iridium dans la structure biotine-steptavidine.....	69
Schéma 45. Synthèse d'un spirolactame (A) et d'un dérivé de $\delta$ -valérolactame (B) avec une myoglobine mutée pour catalyser l'amidation.....	70
Schéma 46. Réaction naturellement catalysée par les $\omega$ -transaminases (TAs).....	71
Schéma 47. Synthèse du ( <i>R</i> )-4-phénylpyrrolidin-2-one par formation de l'acide $\omega$ -aminé par la $\omega$ -TA de <i>Vibrio fluvialis</i> .....	71
Schéma 48. Synthèse du ( <i>R</i> )-4-(4-fluoro)phénylpyrrolidin-2-one par la transaminase ATA-115... ..	71
Schéma 49. Synthèse du 3-( <i>S</i> )-éthylpipéridin-2-one par $\omega$ -TABm mutante.....	72
Schéma 50. Synthèse <i>in vivo</i> de lactames en partant de cycloalkylamines par une cascade enzymatique avec la transaminase de <i>Phaebacter porticola</i> (SPTA) et la cyclase mutée de <i>Enterobacter kobei</i> (Cy-EK <sub>V152A</sub> ). Légende : n.d. pour non déterminé.....	72
Schéma 51. Synthèses de lactames par la PPL.....	72
Schéma 52. Synthèse de lauro lactame par la $\omega$ -lauro lactame hydrolase.....	73
Schéma 53. Synthèse de lactames et d'un dérivé de fluorolactame avec CAL B.....	73
Schéma 54. Synthèse biocatalytique de lactames avec les enzymes dépendantes de l'ATP.....	73
Schéma 55. Synthèse <i>in vivo</i> de $\gamma$ -butyrolactame et de $\delta$ -valérolactame avec la CoA ligase ORF27 en partant de l'acide glutamique ( <b>A</b> ) ou de la lysine ( <b>B</b> ) respectivement.....	74
Schéma 56. Synthèse <i>in vivo</i> pour la synthèse de $\gamma$ -butyrolactame en employant la CoA ligase msed ou ORF27 pour l'activation de l'acide $\omega$ -aminé.....	75
Schéma 57. Synthèse de lactames chiraux par dédoublement cinétique avec des CARs.....	75
Schéma 58. Synthèse de N-méthylbutyramide avec les CoA ligases A4YDT1 ou A4IBN3 couplée à un système de régénération enzymatique de l'ATP par une PPK-III de <i>Deinococcus geothermalis</i> (DgPPK2-III) et la PPase de <i>Pyrococcus horikoshii</i> (PhPPase).....	87
Schéma 59. Mécanisme enzymatique simplifié des deux réactions catalysées par les CoA ligases. Figure inspirée de l'article de Wu <i>et al</i> (Wu <i>et al</i> , 2008).....	94
Schéma 60. Réaction naturelle catalysée par la CoA ligase A4YDT1 de <i>Metallosphaera sedulla</i> .....	95
Schéma 61. Protocole de la méthode « QuikChange® II site-directed mutagenesis » (Agilent) appliquée à la CoA ligase A4YDT1. Schéma issu du manuel d'instruction du kit.....	104
Schéma 62. Synthèse de N-méthylamides avec A4YDT1 WT ou les mutants A4YDT1.....	107
Schéma 63. Méthodes de suivi cinétique de l'étape d'adénylation des CoA ligases et des enzymes adénylantes : par marquage de l'ATP ( <b>A</b> ), détection de l'intermédiaire aminoacyl-AMP ( <b>B</b> ) et par détection du pyrophosphate libéré ( <b>C</b> ). Figure inspirée de l'article de Kittila (Kittila <i>et al</i> , 2016).....	109
Schéma 64. Cascade enzymatique du test NADH-PPi pour le suivi cinétique de l'étape d'adénylation des CoA ligases.....	110
Schéma 65. Stratégie pour la recherche de nouvelles CoA ligases actives sur des divers acides $\omega$ -aminés fonctionnalisés. Schéma inspiré du résumé de stage de Camille Baland (stagiaire M2 dans l'équipe). Fait avec BioRender.com.....	117
Schéma 66. Schéma simplifié de la formation de lactames avec des CoA ligases et dosage de la libération de pyrophosphate.....	123
Schéma 67. ( <b>A</b> ) Dosage du PPi par le vert de malachite. ( <b>B</b> ) Structure du vert de malachite.....	124
Schéma 68. Dosage de l'AMP libéré lors de la cyclisation spontanée par la cascade PK-LDH.....	125

## TABLE DES TABLEAUX

Tableau 1. CoA ligases homologues à A4YDT1 dont la structure a été cristallisée en conformation de thioestérification.....	98
Tableau 2. Comparaison des résidus constituant le site de liaison de substrat de A4YDT1 avec les autres CoA ligases homologues, identifiés par superposition des structures dans le logiciel PyMOL. .	101
Tableau 3. Paramètres cinétiques de l'enzyme native A4YDT1 avec l'acide 4-aminobutyrique (GABA) et des mutants W259G et W259/V328T avec l'acide hexanoïque (6-AHA). Détermination des données cinétiques en utilisant le kit Pyrophosphate Reagent® en méthode de point final. Expériences faites en duplicat.....	111
Tableau 4. Données cinétiques de A4YDT1 d'après l'article de Hawkins <i>et al.</i> (Hawkins, Han et al. 2013) et de quelques CoA ligases décrites dans la littérature (Wu, Cao et al. 2008, Zhang, Barajas et al. 2017, Sofo, Hart et al. 2019).....	111
Tableau 5. Activité spécifique des mutants de A4YDT1, le simple mutant W259G et les doubles mutants W259G/V238T et W259G/F350Y. Expériences faites en duplicat avec le kit Pyrophosphate Reagent®.....	112
Tableau 6. Set de référence composé de CoA ligases issues de la littérature pour construire la collection CoA Lig1. ....	120
Tableau 7. Liste des CoA ligases retenues pour une vérification de l'activité par UHPLC-MS pour le pool 1 suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les enzymes du set de référence de CoA Lig1 sont indiquées en gras, les enzymes déjà identifiées lors du 1 <sup>er</sup> criblage de CoA Lig1 sont indiquées en bleu.....	125
Tableau 8. Liste des CoA ligases retenues pour une vérification de l'activité par UHPLC-UV pour le pool 2 suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les enzymes du set de référence de CoA Lig1 sont indiquées en gras. Les CoA ligases retenues pour d'autres pools de substrats sont indiqués en vert. ....	126
Tableau 9. Liste des CoA ligases retenues pour le pool 3 pour une vérification en UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert. ....	127
Tableau 10. Liste des CoA ligases retenues pour le pool 4 pour une vérification en UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert. ....	128
Tableau 11. Liste des CoA ligases retenues pour le pool 5 pour une vérification par UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert. ....	129
Tableau 12. Liste des CoA ligases retenues pour le pool 6 pour une confirmation en UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert. ....	130
Tableau 13. Liste des CoA ligases constituant le nouveau set de référence pour la 2 <sup>e</sup> collection de CoA ligases et le pool de substrats acides ω-aminés sur lequel elles sont actives. ....	136
Tableau 14. Identification des résidus importants pour l'étape d'adénylation dans les sites actifs de chaque CoA ligase du nouveau set de référence. Les résidus ont été identifiés par superposition leur structure à celle de A4YDT1 comme enzyme référente (PDB : 8BIQ et 8BIT) dans le logiciel PyMOL. Les résidus sont colorés selon l'annotation ClustalOmega : en rouge pour les résidus petits et hydrophobes ; en bleu pour les résidus acides ; en magenta pour les résidus basiques ; en vert pour les résidus portant un groupement hydroxyl, thiol, amine et le résidu glycine. Les résidus conservés sont indiqués en gras. ....	138
Tableau 15. Liste des 20 CoA ligases sélectionnées de la 2 <sup>ème</sup> collection CoA Lig2. ....	143

Tableau 16. Liste des 8 CoA ligases obtenues sous forme purifiée.....147



## TABLE DES ANNEXES

Annexe 1. Gel SDS-PAGE de la CoA ligase A4YDT1 WT et des mutants. M : Marqueur de poids moléculaire SeeBlue® Plus2 pre-strained standard (6 µl déposé), I : Induction (15 µg déposé), L : lysat (surpernageant, 15 µg déposé), FT : Flow Through NiNTA (15 µg déposé) and P : enzyme purifié (5 µg déposé). A. Purification à petite échelle. B. Purification à grande échelle. ....	173
Annexe 2. Chromatogrammes UHPLC-MS (APCI) des standards lactames et leur spectre de masse associé. ....	174
Annexe 3. Chromatogrammes UHPLC-MS (ESI) des N-méthylamides détectés de milieux réactionnels avec les enzymes A4YDT1 WT ou les mutants. ....	176
Annexe 4. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques de A4YDT1 WT. Les points correspondent à la moyenne des duplicats. ....	179
Annexe 5. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques du mutant A4YDT1 W259G. Les points correspondent à la moyenne des duplicats. ....	179
Annexe 6. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques du mutant A4YDT1 W259G/V238T. Les points correspondent à la moyenne des duplicats. ....	179
Annexe 7. Tableaux des paramètres cinétiques d'A4YDT1 et des mutants. ....	180
Annexe 8. Rendements analytiques de la formation de lactames à partir des acides ω-aminés avec A4YDT1 et ses mutants. ....	183
Annexe 9. Comparaison des aires relatives pour divers substrats par rapport au meilleur mutant pour chaque substrat, par détection du N-méthylamide correspondant par UHPLC-MS (ESI). ....	183
Annexe 10. Droite de calibration des standards lactames commerciaux utilisée pour calculer les rendements analytiques avec A4YDT1 et les mutants. ....	184
Annexe 11. Droites de calibration des standards amides synthétisés au laboratoire. ....	185
Annexe 12. Schéma du clonage par ligation indépendante (LIC). ....	186
Annexe 13. Gels SDS-PAGE de la production de la première collection de CoA ligase CoA Lig1. ....	187
Annexe 14. Concentration (en mg/ml) des CoA ligases purifiées de la collection CoA Lig1. Concentrations déterminées selon la méthode Bradford. ....	188
Annexe 15. Résultats du criblage de la collection CoA Lig1 suite à l'analyse avec la cascade PK-LDH. Légende : vert si la CoA ligase est active sur ce pool et rouge si la CoA ligase est inactive ce pool. Une CoA ligase est considérée comme active si son activité est supérieure ou égale à 2% d'activité par rapport au seuil de détection. ....	191
Annexe 16. Résultats de la vérification des CoA ligases sélectionnées par analyse UHPLC. Légende : vert si l'activité a été vérifiée et rouge si l'activité n'a pas été vérifiée. ....	192
Annexe 17. Chromatogrammes des lactames attendus par cyclisation des acides ω-aminés et leur spectre de masse associé. ....	193
Annexe 18. Rendements analytiques de la synthèse de lactames du pool 1 avec les CoA ligases sélectionnées suite à l'analyse avec la cascade PK-LDH. Conditions réactionnelles : 0,5 mM de chaque acide ω-aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl <sub>2</sub> , 1,5 mM MgCl <sub>2</sub> , 50 mM tampon MOPS pH 8,5, 10 µL d'enzyme purifiée. 60°C, sans agitation, 5h. Analyse par UHPLC-MS (APCI). Rendements déduits d'une gamme de calibration avec les standards commerciaux. ....	195
Annexe 19. Détection des aires correspondant aux lactames <b>4b</b> et <b>5b</b> attendus pour les substrats du pool 2 avec les CoA ligases sélectionnées suite à la cascade PK-LDH. Conditions réactionnelles : 0,5 mM de chaque acide ω-aminé dans le pool, 2 mM ATP, 2 mM MnCl <sub>2</sub> , 2 mM MgCl <sub>2</sub> , 50 mM tampon MOPS pH 8,5, 10 µL d'enzyme purifiée. 60°C, sans agitation, 5h. ....	195

Annexe 20. Rendements analytiques pour la synthèse du 2-aminocaprolactame ( <b>6b</b> ) avec les CoA ligases sélectionnées pour le pool 2. Rendements déterminés à partir d'une gamme de calibration avec le standard commercial après dérivatisation au BzCl.....	196
Annexe 21. Détection des aires correspondant aux lactames <b>9b</b> et <b>10b</b> .....	196
Annexe 22. Rendements analytiques pour la synthèse de l'acide pyroglutamique ( <b>8b</b> ) avec les CoA ligases sélectionnées pour le pool 3. Rendements déterminés à partir d'une gamme de calibration avec le standard commercial.....	197
Annexe 23. ( <b>A</b> ) Détection des aires correspondant à la synthèse de N-méthylbutyrolactame ( <b>16b</b> ) et de N-éthylbutyrolactame ( <b>17b</b> ) avec les CoA ligases sélectionnées pour le pool 6. ( <b>B</b> ) Même histogramme en enlevant la forte réponse de la CoA ligase D1A9C5.....	198
Annexe 24. Droites de calibration des standards lactames pour le criblage de la collection CoA Lig1.....	199
Annexe 25. Liste des CoA ligases de la collection CoA Lig2 en fonction de la CoA ligase de référence pour laquelle elles ont été rapatriées.....	201
Annexe 26. Comparaison des résidus constituant le site actif des 20 CoA ligases sélectionnées pour former la collection réduite de CoA Lig2. Légende : résidus à chaîne latérale courte et hydrophobe en rouge (sauf Y), résidus acides en bleu, résidus basique en magenta et résidus à chaîne latérale polaire et résidu glycine en vert.....	205
Annexe 27. Plan de plaque des gènes des 20 CoA ligases de la collection réduite.....	206
Annexe 28. Gel d'agarose de la vérification de l'amplification PCR des gènes des 20 CoA ligases sélectionnées. Légende : de gauche à droite, les puits correspondent à A1 à H1 puis de A2 à H2 et A3 à D3.....	207
Annexe 29. Gels SDS-PAGE et concentration des CoA ligases de la production de la collection réduite de la deuxième collection CoA Lig2.....	208

## LISTE DES ABREVIATIONS

<b>β-LS</b>	β-lactame synthétase
<b>3-HP</b>	Acide 3-hydroxypropionate
<b>4-HB</b>	Acide 4-hydroxybutanoïque
<b>5-APA ou AVA</b>	Acide 5-aminopentanoïque ou acide aminovalérique
<b>66CaaT</b>	Sérotonine hydroxycinnamoyltransférase de <i>Capsicum annum</i>
<b>6-AHA ou ACA</b>	Acide 6-aminohexanoïque ou acide aminocaproïque
<b>6-APA</b>	Acide 6-aminopénicillanique
<b>7-ADCA</b>	Acide 7-aminodésacétocéphalosporanique
<b>aa-ARNt</b>	aminoacyl-ARNt
<b>ABS</b>	Amide Bond Synthetase
<b>ACV</b>	δ-(L-α-amino adipoyl)-L-cysteinyl-D-valine
<b>AcsA</b>	Acyl-CoA ligase de <i>Pseudomonas Chlororaphis</i>
<b>ADH</b>	Alcool déshydrogénase
<b>ADN</b>	Acide désoxyribonucléique
<b>AHR</b>	Aldéhyde réductase
<b>ALA</b>	Acide δ-aminolévulinique
<b>AMP</b>	Adénosine monophosphate
<b>AMPPNP</b>	Adénylyl-imidodiphosphate, analogue inhibiteur de l'ATP
<b>API ou IPA</b>	<i>Active Pharmaceutical Ingredient</i> ou Ingrédient pharmaceutique actif en français
<b>APCI</b>	<i>Atmospheric pressure chemical ionization</i>
<b>Arc</b>	Cyclase de radicaux aromatiques
<b>ArM</b>	Métalloenzyme artificielle
<b>ARN</b>	Acide ribonucléique
<b>ARNm</b>	Acide ribonucléique messager
<b>ARNt</b>	Acide ribonucléique de transfert
<b>ASMC</b>	<i>Active Site Modeling and Clustering</i> (logiciel)
<b>At4CL2</b>	4-coumarate CoA ligase de <i>Arabidopsis thaliana</i>
<b>ATase</b>	Acyltransférase

<b>ATP</b>	Adénosine triphosphate
<b>ATRC</b>	Cyclisation radicalaire par transfert d'atome, de l'anglais <i>atom transfer radical cyclisation</i>
<b>BLASTP</b>	<b>B</b> asic <b>L</b> ocal <b>A</b> lignment <b>S</b> earch <b>T</b> ool <b>P</b> rogram
<b>BRENDA</b>	BRAunschweig ENzyme DAtabase (base de données)
<b>BSA</b>	Sérum d'albumine bovin ou <i>Bovine Serum albumin</i> en anglais
<b>BzCl</b>	Chlorure de benzoyle
<b>C-ter</b>	Partie C-terminale
<b>CAL B</b>	Lipase de <i>Candida antarctica B</i>
<b>CAR</b>	Carboxylic acid reductase
<b>CAR<sub>mm</sub> ou MmCAR</b>	Carboxylic Acid Reductase de <i>Mycobacterium marinum</i>
<b>CAR<sub>mm</sub>-A</b>	Apoenzyme de la CAR <sub>mm</sub>
<b>CAR<sub>sr</sub>-A</b>	Domaine d'adénylation de CAR de <i>Segnilliparus rugosus</i>
<b>CBL</b>	4-chlorobenzoate CoA ligase
<b>CCL</b>	Lipase de <i>Candida cylindracea</i>
<b>CEA</b>	N <sup>2</sup> -(carboxylethyl)-L-arginine
<b>CHES</b>	Acide N-cyclohexyl-2-aminoéthanesulfonique
<b>CHMO</b>	Cyclohexanone monooxygénase
<b>CHU</b>	Polyphosphate kinase (PolyP) de classe III de <i>Cytophaga hutchinsonii</i>
<b>CPT</b>	Centre peptidyltransférase
<b>CoA ou CoASH</b>	Coenzyme A
<b>CoA Lig1</b>	Première collection de CoA ligases
<b>CoA Lig2</b>	Deuxième collection de CoA ligases
<b>DES</b>	Solvants eutectiques profonds, de l'anglais <i>Deep Eutectic Solvant</i>
<b>DGPC</b>	Acide deoxyguanidinoproclavaminique
<b>DHB</b>	Acide 2,3-dihydroxybenzoïque
<b>DIPE</b>	Ether diisopropylique
<b>ee</b>	Excès énantiomérique
<b>EHS (score)</b>	Environnement, Santé et Sécurité ou <i>Environment, Health and Safety</i> en anglais
<b>epPCR</b>	Error-prone PCR

<b>eq</b>	Equivalent molaire
<b>er</b>	Ratio énantiomérique
<b>ERED</b>	Ène-réductase
<b>ESI</b>	Electrospray ionization
<b>FACS</b>	<i>Fluorescent-activated cell sorting</i> ou triage de cellules activées par fluorescence en français
<b>FAIR (données)</b>	<b>F</b> acile à (re)trouver, <b>A</b> ccessible, <b>I</b> nteropérable et <b>R</b> éutilisable, de l'acronyme anglais <i>F</i> indable, <i>A</i> ccessible, <i>I</i> nteroperable, <i>R</i> eusable data
<b>FAS</b>	Acide gras synthases, de l'anglais <i>Fatty Acid synthases</i>
<b>FDA</b>	Food and Drug Administration
<b>FMN</b>	Flavine mononucléotide
<b>FunDivEx</b>	<b>F</b> unctional <b>D</b> iversity <b>E</b> xploration
<b>GABA</b>	Acide $\gamma$ -aminobutyrique ou acide 4-aminobutanoïque
<b>GDPH</b>	Glycérophosphate déshydrogénase
<b>GOase</b>	Galactose oxydase
<b>GOx</b>	Glucose oxydase
<b>GSK</b>	GlaxoSmithKline
<b>H<sub>2</sub>O</b>	Eau
<b>HCl</b>	Acide chlorhydrique
<b>HPGA</b>	<i>p</i> -hydroxyphénylglycine amide
<b>HRP</b>	Peroxydase de racine de Raifort de l'anglais <i>Horseradish peroxidase</i>
<b>IA</b>	Intelligence artificielle
<b>IMoPA (UMR)</b>	Ingénierie moléculaire, cellulaire & physiopathologie (UMR)
<b>IPN</b>	Isopénicilline N
<b>IPNS</b>	Isopénicilline N synthase
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>K<sub>M</sub></b>	Constante de Michaelis
<b>k<sub>cat</sub></b>	Constante catalytique
<b>KS</b>	Domaine cétoynthase des PKS
<b>L2BMS</b>	Laboratoire de Biocatalyse, Biodégradation et Métabolisme Synthétique
<b>LA</b>	Lactamase

<b>LABGeM</b>	Laboratoire d'Analyses Bioinformatiques pour la Génomique et le Métabolisme
<b>LCOB</b>	Laboratoire de Chimie Organique et Biocatalyse
<b>LGBM</b>	Laboratoire de Génomique et Biologie Métabolique
<b>LIC</b>	Clonage par ligation indépendante
<b>Mb</b>	Myoglobine
<b>MEGA</b>	N-alkanoyl-N-méthylglucamide
<b>MESG</b>	2-amino-6-mercapto-7-méthylpurine riboside
<b>ML</b>	<i>Machine learning</i>
<b>MOPS</b>	Acide 3-morpholino-1-propanesulfonique
<b>MsAcT</b>	N-acyltransférase de <i>Mycobacterium smegmatis</i>
<b>mSav</b>	Monomère de streptavidine
<b>MT-ACS</b>	Acétyl-CoA ligase de <i>Methanothermobacter thermautrophicus</i> Z245
<b>NaDES</b>	Solvants eutectiques profonds naturels, de l'anglais <i>Natural Deep Eutectic Solvent</i>
<b>NAD</b>	Nicotinamide adénine dinucléotide
<b>NADH</b>	Nicotinamide adénine dinucléotide réduit
<b>NADP</b>	Nicotinamide adénine dinucléotide phosphate
<b>NADPH</b>	Nicotinamide adénine dinucléotide phosphate réduit
<b>NaOH</b>	Hydroxyde de sodium ou soude
<b>NAT</b>	N-acyltransférase
<b>NCBI</b>	National Center for Biotechnology Information
<b>NHase</b>	Nitrile hydratase
<b>NIS</b>	<i>NRPS-independent siderophores</i>
<b>NMP</b>	N-méthyl-2-pyrrolidone
<b>NRP</b>	Peptides non ribosomiques, de l'anglais <i>Non-Ribosomal Peptides</i>
<b>NRPS</b>	Peptides synthétases non ribosomiques, de l'anglais <i>Non-Ribosomal Peptide Synthetases</i>
<b>N-ter</b>	Partie N-terminale
<b>OYE</b>	Old Yellow Enzyme
<b>PA</b>	Pénicilline acylase

<b>PaoABC</b>	Aldéhyde oxydase périplasmique de <i>E. coli</i>
<b>PCR</b>	Réaction en chaîne par polymérase, de l'anglais <i>Polymerase Chain Reaction</i>
<b>PCP</b>	Protéine porteuse de peptide, de l'anglais <i>peptidyl carrier protein</i>
<b>PDB</b>	Protein Data Bank
<b>PEP</b>	Phosphoénol pyruvate
<b>PEG</b>	Polyéthylène glycol
<b>PGA (composé chimique)</b>	Phénylglycine amide
<b>PGA (enzyme)</b>	Pénicilline G acylase
<b>Pi</b>	Phosphate inorganique
<b>PK</b>	Pyruvate Kinase
<b>PKS</b>	Polycétide synthase, de l'anglais <i>Polyketide synthase</i>
<b>PLP</b>	Pyridoxal 5'-phosphate
<b>PPan</b>	4-phosphopantéthéine
<b>PPi</b>	Pyrophosphate inorganique
<b>PPi-PFK</b>	Fructose-6-phosphate kinase, dépendante du pyrophosphate
<b>PolyP</b>	Polyphosphate
<b>PPK-II</b>	Polyphosphate kinase 2
<b>PPL</b>	Lipase pancréatique porcine
<b>RiPPs</b>	<i>Ribosomally synthesized and post-translationally modified peptides</i>
<b>SA</b>	Acide suberanilique
<b>SAAD</b>	Domaine autonome d'adénylation, de l'anglais <i>Stand Alone Adenylation Domain</i>
<b>SAH</b>	S-adénosylhomocystéine
<b>SAHA</b>	Acide suberoylanique hydroxamique
<b>SAM</b>	S-adenosylmethionine (cofacteur)
<b>Sav</b>	Streptavidine
<b>ShABS</b>	ABS provenant de <i>Streptoalloteichus hindustanus</i>
<b>Sp-ATA</b>	Transaminase de <i>Silicibacter pomeroyi</i>
<b>SpL</b>	Lipase de <i>Sphingomonas sp. HXN-200</i>
<b>T4 Pol</b>	T4 Polymérase

<b>TA ou ATA</b>	Transaminase ou aminotransférase
<b>TBI</b>	Toulouse Biotechnology Institute
<b>TE</b>	Thioestérase
<b>TGM</b>	Transglutaminase
<b>TPI</b>	Triose phosphate isomérase
<b>UHPLC</b>	Chromatographie liquide à ultra haute performance de l'anglais <i>Ultra High Performance Liquid Chromatography</i>
<b>UHPLC-MS</b>	Chromatographie liquide à ultra haute performance couplée un spectromètre de masse
<b>UHPLC-UV</b>	Chromatographie liquide à ultra haute performance couplée à un spectrophotomètre ultraviolet
<b>UMR</b>	Unité Mixte de Recherche
<b>URD (AgroParisTech)</b>	<b>ABI</b> Unité de Recherche et de Développement Agro-Biotechnologies Industrielles
<b>RML</b>	Lipase de <i>Rhizomucor miehei</i> (RML)
<b>RT</b>	Température ambiante, de l'anglais <i>room temperature</i>
<b>WT</b>	Wild-type (enzyme native)
<b>XDH</b>	Xanthine déshydrogénase

Les acides aminés sont nommés selon leur code à 1 lettre ou à 3 lettres et à une seule lettre lorsque cela réfère à la position d'un résidu dans la structure d'une enzyme. Les mutants sont notés sous AXB avec A correspondant au résidu acide aminé natif, X la position de ce résidu et B l'acide aminé par lequel le résidu natif a été muté.





# INTRODUCTION GENERALE

La liaison amide est une liaison chimique formée par condensation d'une fonction acide carboxylique et d'une fonction amine. Cette liaison est plane avec une délocalisation du doublet non liant de l'azote vers le carbonyle, ce qui empêche toute rotation autour de la liaison C-N et lui confère une forte stabilité (Schéma 1) (Clayden, Greeves et al. 2012).

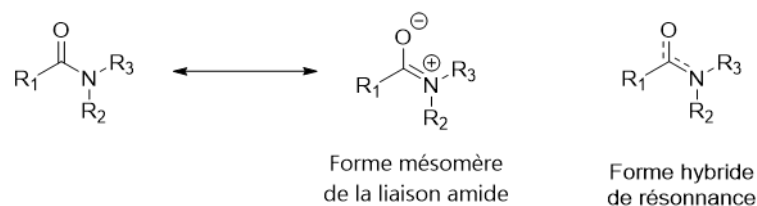


Schéma 1. Schéma de la délocalisation électronique de la liaison amide (phénomène de résonance).

La liaison amide est une fonction très présente aussi bien dans les produits naturels que dans les composés synthétiques. Dans la nature, cette liaison amide se retrouve notamment dans les peptides et protéines, on l'appelle alors liaison peptidique. Nous retrouvons des molécules possédant une liaison amide dans des composés synthétisés en chimie pharmaceutique, mais aussi en chimie des polymères, dans les pesticides et en agro-alimentaire (Figure 1). En chimie médicinale, la synthèse de liaison amide est l'une des réactions les plus courantes, représentant 25% des réactions en 2016 et étant mentionnée dans environ 50% des articles publiés en 2014 (Brown and Boström 2016, Boström, Brown et al. 2018).

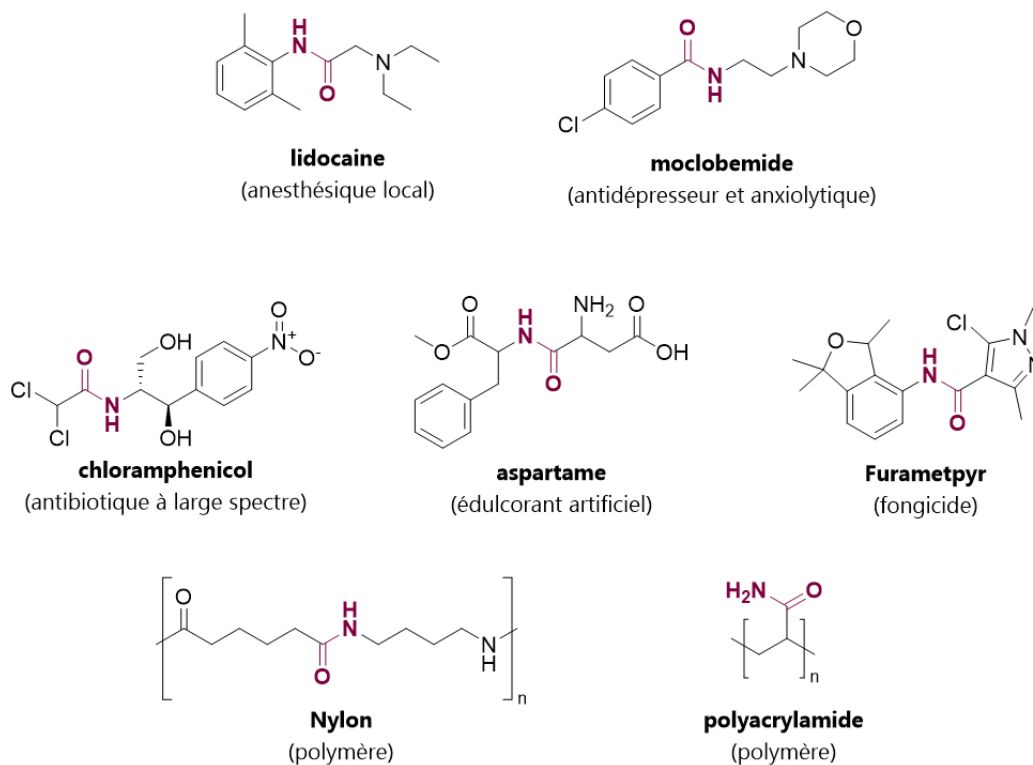


Figure 1. Exemples de composés contenant une liaison amide.

Lorsque cette fonction amide est comprise dans un cycle, nous parlons alors de lactames. Ils sont omniprésents dans le vivant (champignons, bactéries, plantes et animaux) et ont diverses propriétés biologiques. Ils sont principalement connus pour leur activité antibiotique avec la classe des  $\beta$ -lactames découverts par Fleming avec la pénicilline en 1928, mais d'autres lactames présentent un intérêt grandissant face à l'émergence de la résistance aux antibiotiques, notamment les  $\gamma$ -lactames (Caruano, Muccioli et al. 2016). Ce sont également des ingrédients actifs pharmaceutiques (APIs pour *Active Pharmaceutical Ingredients* en anglais) pour leur propriétés antitumorales, anti-inflammatoires ou antidépresseurs par exemple (Saldívar-González, Lenci et al. 2019, Kumari, Carmona et al. 2020). Par ailleurs, les lactames servent aussi, par ouverture de cycle, de précurseurs de polymères comme le nylon (Figure 2).

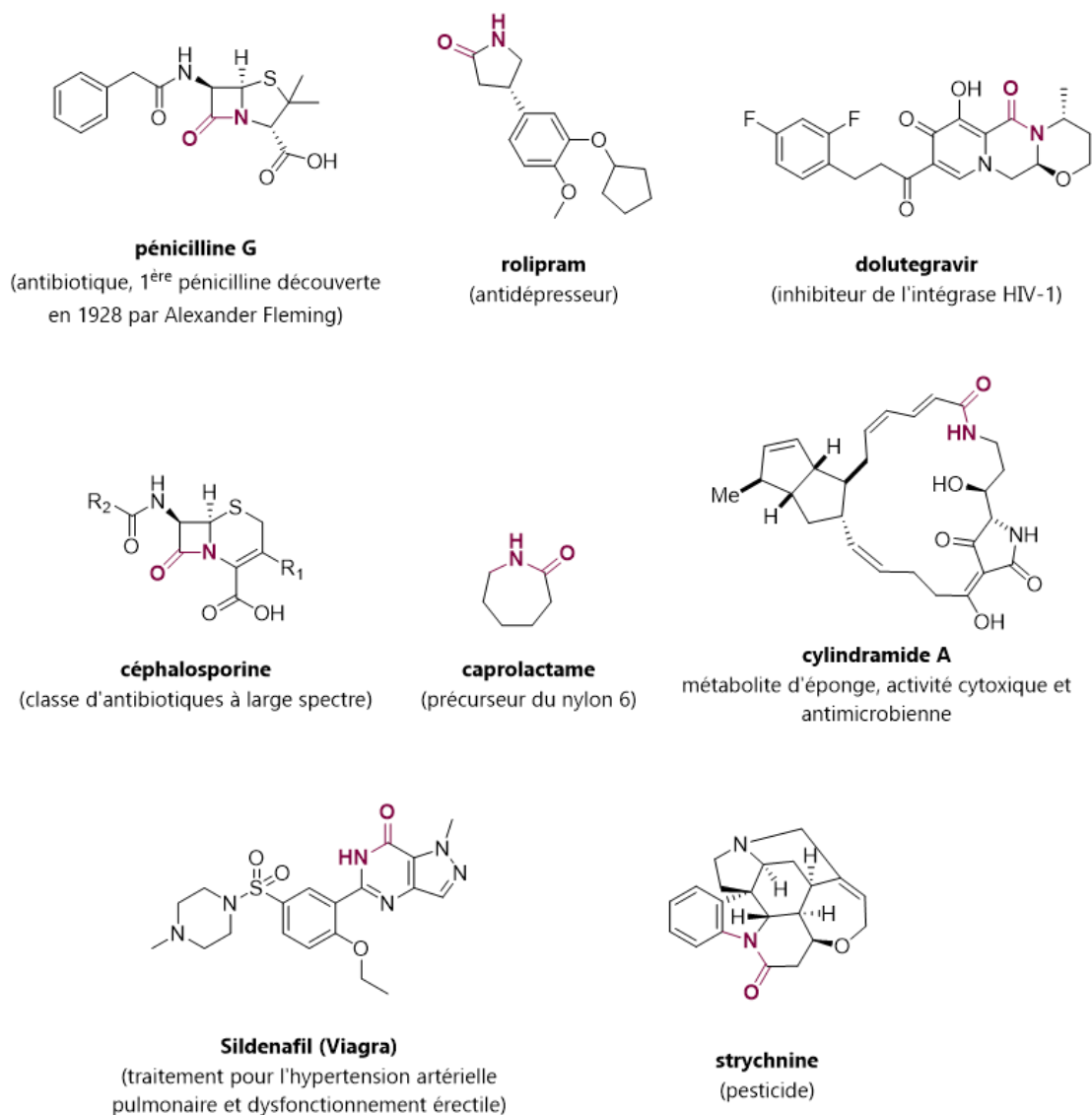


Figure 2. Exemples de composés naturels et synthétiques avec le motif lactame.

La principale voie de synthèse d'amides nécessite l'activation de la fonction carboxylique en utilisant des agents de couplage en quantité stœchiométrique, ces procédés sont donc à faible économie d'atome et génèrent des sous-produits indésirables (Arnold, Davies et al. 2006, Valeur and Bradley 2009, Sabatini, T. Boulton et al. 2019). De plus, cette réaction peut nécessiter l'utilisation de solvants dangereux pour la santé et l'environnement (Sabatini, T. Boulton et al. 2019).

Pour ces raisons, la recherche de méthodes de synthèse durable pour la formation de liaison amide fait partie des principaux axes de recherche en chimie verte définis lors des tables rondes de l'ACS Green Chemistry Institute® et des sociétés pharmaceutiques mondiales (*ACS Green Chemistry Institute® Pharmaceutical Roundtable*, abrégé en GCIPR) (Constable, Dunn et al. 2007, Bryan, Dunn et al. 2018). La biocatalyse est donc l'une des voies à explorer pour une formation de liaison amide plus respectueuse des principes de la chimie actuelle (Anastas, Kirchhoff et al. 2001, Anastas and Eghbali 2010, Sheldon and Woodley 2018). Si nous nous référons au guide de GSK (GlaxoSmithKline) pour la formation de liaison amide, l'utilisation d'enzymes ne présente presque aucun problème avec un score de « Environnement, Santé et Sécurité » (EHS pour *Environment, Health and Safety* en anglais) égal à 10/10. Certains de leurs APIs contenant le motif amide sont synthétisés par voie enzymatique (Adams, Alder et al. 2013, Adams, Brown et al. 2019).

La biocatalyse est définie comme l'utilisation d'enzymes pour catalyser des transformations chimiques à des fins de synthèse. Ces dernières décennies, plusieurs procédés biocatalytiques industriels ont été développés pour la synthèse de produits d'intérêt : en chimie des matériaux (acrylamide), chimie pharmaceutique (islatravir, traitement VIH produit par Merck), les détergents, les parfums (Ambrox, l'ambre grise) ou encore en chimie fine (vanilline, exemple d'ingénierie protéique) (Li and Rosazza 2000, Huffman, Fryszkowska et al. 2019, Bell, Finnigan et al. 2021, Eichhorn and Schroeder 2023). A travers ces exemples, il est clair que l'intérêt de la biocatalyse en synthèse organique et en biotechnologie n'est plus à démontrer. Les efforts se concentrent désormais sur l'intensification de la recherche de biocatalyseurs et leur optimisation pour étendre le champ d'applications à l'échelle industrielle (Bell, Finnigan et al. 2021, Yi, Bayer et al. 2021, Buller, Lutz et al. 2023). Ces efforts s'appuient sur de nouvelles méthodes de criblage à haut-débit, par exemple par micro-fluidique en gouttelettes ou en triage de cellules activées par fluorescence (FACS, de l'anglais *fluorescence-activated cell sorting*) (Sheludko and Fessner 2020, Jain, Stavakis et al. 2024). Ces technologies sont souvent utilisées lors de travaux d'évolution dirigée d'enzymes car elles permettent une miniaturisation du criblage, et donc l'étude de grandes bibliothèques de  $10^4$  à  $10^5$  mutants (Hecko, Schiefer et al. 2023). En amont des aspects expérimentaux, grâce à l'implémentation du *machine learning* (ML) et à l'intelligence artificielle (IAs), de nouveaux outils computationnels automatisés ont vu le jour comme AlphaFold qui permet de prédire la structure des protéines d'après leur séquence primaire, la dernière version AlphaFold3, permettant une meilleure prédiction des interactions biomoléculaires, est sortie très récemment (Abramson, Adler et al. 2024). Ces outils de prédictions, associés aux bases de données (Protein Data Bank, UniProt...) et aux logiciels existants (AutoDock, PyMOL, Rosetta), nous permettent d'explorer la structure et la dynamique moléculaire des enzymes et ainsi, d'orienter leur ingénierie pour les applications souhaitées. D'autres outils basés sur du *machine learning* ont été développés : on peut citer RetroBioCat, qui génère *in silico* des voies de rétrosynthèse enzymatique d'un produit cible (Finnigan, Hepworth et al. 2021, Finnigan, Lubberink et al. 2023). En aval, un travail important est mené sur la standardisation des données scientifiques dans le domaine de la biocatalyse et de l'enzymologie. On peut citer l'initiative EnzymeML, un format de transfert de données respectant les principes FAIR pour de meilleures reproductibilité, traçabilité, interopérabilité et stockage des résultats au sein de la communauté (Lauterbach, Dienhart et al. 2023).

Dans l'équipe Biocatalyse du Laboratoire Biodégradation, Biocatalyse et Métabolisme Synthétique (L2BMS), nous nous sommes intéressés à la formation de liaison amide par voie biocatalytique. Alignés avec les stratégies de l'unité Génomique Métabolique, nos travaux de recherche sont axés sur la recherche d'enzymes pour accéder à des produits d'intérêt par des voies synthétiques efficaces et respectueuses de l'environnement. Dans la littérature, des familles d'enzymes telles que les lipases, ou les nitrile hydratases sont décrites pour la synthèse d'amide (Petchey and Grogan 2019, Lubberink,

Finnigan et al. 2023). Il existe également des enzymes dépendantes de l'ATP, les Carboxylic Acid Reductases (CARs), les Amide Bond Synthetases (ABS) (Petchey, Rowlinson et al. 2020, Basri, Rahman et al. 2023, Lubberink, Finnigan et al. 2023). Parmi ces enzymes dépendantes de l'ATP, notre équipe s'est intéressée à la famille des enzymes adénylantes et plus précisément aux CoA ligases ou CoA synthétases. Nous reviendrons plus en détail sur ces différentes enzymes dans le chapitre 1 qui porte sur l'état de l'art de la synthèse enzymatique d'amides et de lactames.

Ce projet a été initié en 2017 avec la thèse de Chloé Lelièvre (2017-2020) qui a permis d'identifier des CoA ligases thermophiles capables de synthétiser des amides, en supprimant l'étape de thioesterification et en couplant cette synthèse à un système de régénération de l'ATP. Cette synthèse a ensuite été étendue à celle des lactames, qui est au cœur de mon projet de thèse. Ce projet est associé au projet ANR SMALA « From **S**mall to **MA**cro**L**Actam, an enzymatic approach » (2022-2026) qui consiste à chercher une approche globale pour accéder à des lactames de tailles et fonctionnalités diverses.

Afin d'identifier des CoA ligases pour la synthèse de lactames, deux approches ont été utilisées : l'exploration de la biodiversité et l'ingénierie protéique. Pour la partie exploration de la diversité, une collection de CoA ligases thermophiles a été criblée avec des substrats acides  $\omega$ -aminés simples et fonctionnalisés. Ce travail a été mené en collaboration interne avec la plateforme de criblage des activités biologiques FunDivEx (**F**unctional **D**iversity **E**xploration) du Laboratoire de Génomique et Biochimie du Métabolisme (LGBM) au sein de l'unité. Pour l'ingénierie protéique, une collaboration avec l'équipe des Pr. Dick Janssen et Pr Andy-Mark Thunnissen de l'université de Groningue (Pays-Bas) a été initiée durant la thèse de Dr Chloé Lelièvre (doctorante de 2017 à 2020). Cette collaboration a eu pour but de cristalliser une CoA ligase en vue de caractériser les déterminants structuraux responsables de la spécificité de substrats et modifier cette spécificité par ingénierie protéique rationnelle. Ces deux approches seront détaillées dans les chapitres 2 et 3 de ce manuscrit.

Sauf si indication contraire, j'ai participé activement à tous les travaux exposés dans ce manuscrit. Toute autre personne impliquée dans ces travaux est mentionnée dans le manuscrit. Quant à la partie expérimentale, les matériels utilisés et les protocoles sont rédigés en anglais.

## REFERENCES BIBLIOGRAPHIQUES

Abramson, J., J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Židek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper (2024). "Accurate structure prediction of biomolecular interactions with AlphaFold 3." Nature.

Adams, J. P., C. M. Alder, I. Andrews, A. M. Bullion, M. Campbell-Crawford, M. G. Darcy, J. D. Hayler, R. K. Henderson, C. A. Oare, I. Pendrak, A. M. Redman, L. E. Shuster, H. F. Sneddon and M. D. Walker (2013). "Development of GSK's reagent guides – embedding sustainability into reagent selection." Green Chemistry **15**(6): 1542-1549.

Adams, J. P., M. J. B. Brown, A. Diaz-Rodriguez, R. C. Lloyd and G.-D. Roiban (2019). "Biocatalysis: A Pharma Perspective." Advanced Synthesis & Catalysis **361**(11): 2421-2432.

Anastas, P. and N. Eghbali (2010). "Green Chemistry: Principles and Practice." Chemical Society Reviews **39**(1): 301-312.

Anastas, P. T., M. M. Kirchhoff and T. C. Williamson (2001). "Catalysis as a foundational pillar of green chemistry." Applied Catalysis A: General **221**(1): 3-13.

Arnold, K., B. Davies, R. L. Giles, C. Grosjean, G. E. Smith and A. Whiting (2006). "To Catalyze or not to Catalyze? Insight into Direct Amide Bond Formation from Amines and Carboxylic Acids under Thermal and Catalyzed Conditions." Advanced Synthesis & Catalysis **348**(7-8): 813-820.

Basri, R. S., R. N. Z. R. A. Rahman, N. H. A. Kamarudin and M. S. M. Ali (2023). "Carboxylic acid reductases: Structure, catalytic requirements, and applications in biotechnology." International Journal of Biological Macromolecules: 124526.

Bell, E. L., W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niiikura, S. Osuna, E. Romero, K. S. Ryan, N. J. Turner and S. L. Flitsch (2021). "Biocatalysis." Nature Reviews Methods Primers **1**(1): 46.

Boström, J., D. G. Brown, R. J. Young and G. M. Keserü (2018). "Expanding the medicinal chemistry synthetic toolbox." Nature Reviews Drug Discovery **17**(10): 709-727.

Brown, D. G. and J. Boström (2016). "Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have All the New Reactions Gone?" Journal of Medicinal Chemistry **59**(10): 4443-4458.

Bryan, M. C., P. J. Dunn, D. Entwistle, F. Gallou, S. G. Koenig, J. D. Hayler, M. R. Hickey, S. Hughes, M. E. Kopach, G. Moine, P. Richardson, F. Roschangar, A. Steven and F. J. Weiberth (2018). "Key Green Chemistry research areas from a pharmaceutical manufacturers' perspective revisited." Green Chemistry **20**(22): 5082-5103.

Buller, R., S. Lutz, R. J. Kazlauskas, R. Snajdrova, J. C. Moore and U. T. Bornscheuer (2023). "From nature to industry: Harnessing enzymes for biocatalysis." Science **382**(6673): eadh8615.

Caruano, J., G. G. Muccioli and R. Robiette (2016). "Biologically active  $\gamma$ -lactams: synthesis and natural sources." Organic & Biomolecular Chemistry **14**(43): 10134-10156.

Clayden, J., N. Greeves and S. Warren (2012). Organic Chemistry. Second Edition, Oxford University Press.

Constable, D. J. C., P. J. Dunn, J. D. Hayler, G. R. Humphrey, J. J. L. Leazer, R. J. Linderman, K. Lorenz, J. Manley, B. A. Pearlman, A. Wells, A. Zaks and T. Y. Zhang (2007). "Key green chemistry research areas—a perspective from pharmaceutical manufacturers." Green Chemistry **9**(5): 411-420.

Eichhorn, E. and F. Schroeder (2023). "From Ambergris to (–)-Ambrox: Chemistry Meets Biocatalysis for Sustainable (–)-Ambrox Production." Journal of Agricultural and Food Chemistry **71**(13): 5042-5052.

Finnigan, W., L. J. Hepworth, S. L. Flitsch and N. J. Turner (2021). "RetroBioCat as a computer-aided synthesis planning tool for biocatalytic reactions and cascades." Nature Catalysis **4**(2): 98-104.

Finnigan, W., M. Lubberink, L. J. Hepworth, J. Citoler, A. P. Matthey, G. J. Ford, J. Sangster, S. C. Cosgrove, B. Z. da Costa, R. S. Heath, T. W. Thorpe, Y. Yu, S. L. Flitsch and N. J. Turner (2023). "RetroBioCat Database: A Platform for Collaborative Curation and Automated Meta-Analysis of Biocatalysis Data." ACS Catalysis **13**(17): 11771-11780.

Hecko, S., A. Schiefer, C. P. S. Badenhorst, M. J. Fink, M. D. Mihovilovic, U. T. Bornscheuer and F. Rudroff (2023). "Enlightening the Path to Protein Engineering: Chemoselective Turn-On Probes for High-Throughput Screening of Enzymatic Activity." Chemical Reviews.

Huffman, M. A., A. Fryszkowska, O. Alvizo, M. Borra-Garske, K. R. Campos, K. A. Canada, P. N. Devine, D. Duan, J. H. Forstater, S. T. Grosser, H. M. Halsey, G. J. Hughes, J. Jo, L. A. Joyce, J. N. Kolev, J. Liang, K. M. Maloney, B. F. Mann, N. M. Marshall, M. McLaughlin, J. C. Moore, G. S. Murphy, C. C. Nawrat, J. Nazor, S. Novick, N. R. Patel, A. Rodriguez-Granillo, S. A. Robaire, E. C. Sherer, M. D. Truppo, A. M. Whittaker, D. Verma, L. Xiao, Y. Xu and H. Yang (2019). "Design of an in vitro biocatalytic cascade for the manufacture of islatravir." Science **366**(6470): 1255-1259.

Jain, A., S. Stavakis and A. deMello (2024). "Droplet-based microfluidics and enzyme evolution." Current Opinion in Biotechnology **87**: 103097.

Kumari, S., A. V. Carmona, A. K. Tiwari and P. C. Trippier (2020). "Amide Bond Bioisosteres: Strategies, Synthesis, and Successes." Journal of Medicinal Chemistry **63**(21): 12290-12358.

Lauterbach, S., H. Dienhart, J. Range, S. Malzacher, J.-D. Spöring, D. Rother, M. F. Pinto, P. Martins, C. E. Lagerman, A. S. Bommaris, A. V. Høst, J. M. Woodley, S. Ngubane, T. Kudanga, F. T. Bergmann, J. M. Rohwer, D. Iglezakis, A. Weidemann, U. Wittig, C. Kettner, N. Swainston, S. Schnell and J. Pleiss (2023). "EnzymeML: seamless data flow and modeling of enzymatic data." Nature Methods.

Li, T. and J. P. N. Rosazza (2000). "Biocatalytic Synthesis of Vanillin." Applied and Environmental Microbiology **66**(2): 684-687.

Lubberink, M., W. Finnigan and S. L. Flitsch (2023). "Biocatalytic amide bond formation." Green Chemistry.

Petchey, M. R. and G. Grogan (2019). "Enzyme-Catalysed Synthesis of Secondary and Tertiary Amides." Advanced Synthesis & Catalysis **361**(17): 3895-3914.

Petchey, M. R., B. Rowlinson, R. C. Lloyd, I. J. S. Fairlamb and G. Grogan (2020). "Biocatalytic Synthesis of Moclobemide Using the Amide Bond Synthetase McbA Coupled with an ATP Recycling System." ACS Catalysis **10**(8): 4659-4663.

Sabatini, M., L. T. Boulton, H. Sneddon and T. Sheppard (2019). A green chemistry perspective on catalytic amide bond formation.

Saldívar-González, F. I., E. Lenci, A. Trabocchi and J. L. Medina-Franco (2019). "Exploring the chemical space and the bioactivity profile of lactams: a chemoinformatic study." RSC Advances **9**(46): 27105-27116.

Sheldon, R. A. and J. M. Woodley (2018). "Role of Biocatalysis in Sustainable Chemistry." Chem. Rev. **118**(2): 801-838.

Sheludko, Y. V. and W.-D. Fessner (2020). "Winning the numbers game in enzyme evolution – fast screening methods for improved biotechnology proteins." Current Opinion in Structural Biology **63**: 123-133.

Valeur, E. and M. Bradley (2009). "Amide bond formation: beyond the myth of coupling reagents." Chemical Society Reviews **38**(2): 606-631.

Yi, D., T. Bayer, C. P. S. Badenhorst, S. Wu, M. Doerr, M. Höhne and U. T. Bornscheuer (2021). "Recent trends in biocatalysis." Chemical Society Reviews.

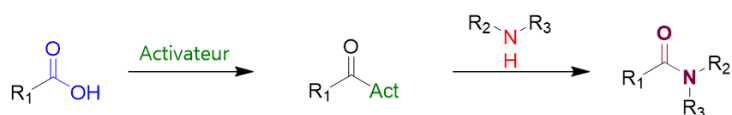




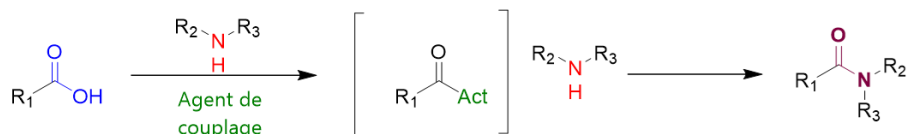
# CHAPITRE I : ETAT DE L'ART SUR LA SYNTHÈSE ENZYMATIQUE D'AMIDES ET DE LACTAMES

La formation de liaison amide est l'une des réactions chimiques les plus utilisées avec différentes méthodes de synthèse pour y accéder. La plus courante consiste en une amidation dite stœchiométrique d'un acide carboxylique, soit en utilisant un intermédiaire activé, soit avec la formation *in situ* de l'intermédiaire activé en utilisant un agent de couplage (Schéma 2) (Montalbetti and Falque 2005, Valeur and Bradley 2009, El-Faham and Albericio 2011, Dunetz, Magano et al. 2016). L'activation de l'acide carboxylique consiste à remplacer le groupe OH par un meilleur groupe partant. Cela permet d'obtenir un intermédiaire plus électrophile et ainsi de favoriser l'attaque nucléophile de l'amine, aussi appelée aminolyse. L'attaque nucléophile de l'amine passe par un intermédiaire tétraédrique instable qui va induire le départ de l'activateur. Cette réaction a généralement lieu en présence d'une base, comme la pyridine par exemple, afin de déprotoner l'amine (Schéma 2).

**A** Amidation via un intermédiaire en deux étapes



Amidation via un agent de couplage en « one-pot »



**B**

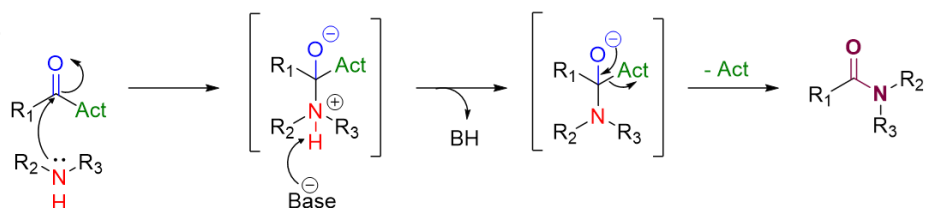


Schéma 2. **(A)** Schéma des deux stratégies d'amidation stœchiométrique et **(B)** son mécanisme réactionnel avec l'attaque nucléophile de l'amine.

En 2017-2018, la synthèse d'amide via les agents de couplage la plus employée est celle via des chlorures d'acyle, ou à partir de leur précurseurs les acyles de thionyle, représentaient respectivement 45% et 16% des réactions d'amidation reportées dans les articles dans la base de recherche Reaxys (Schéma 3) (Sabatini, T. Boulton et al. 2019).

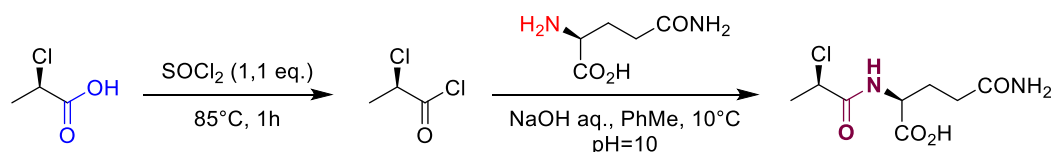


Schéma 3. Exemple d'amidation stœchiométrique à partir de chlorure de thionyle.

Des méthodes catalytiques sont de plus en plus employées. L'amidation peut être catalysée avec des organocatalyseurs, des catalyseurs à base de bore ou de métaux de transitions (Schéma 4) (Massolo, Pirola et al. 2020, Todorovic and Perrin 2020, Upadhyay, Rana et al. 2021).

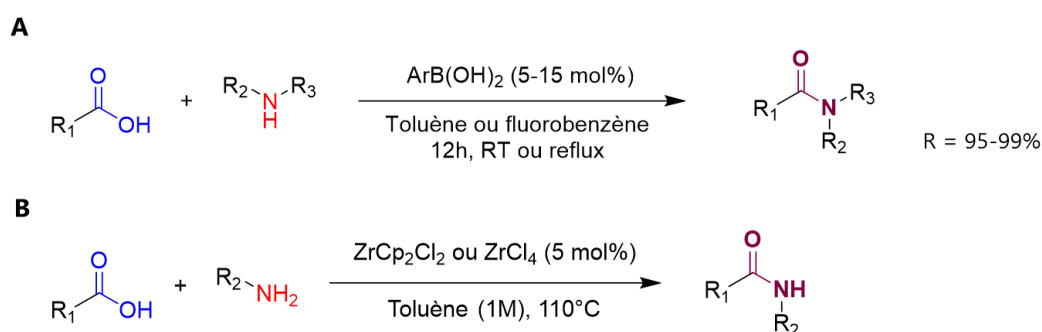


Schéma 4. Exemples d'amidation catalytique par un organocatalyseur à base de bore (**A**) et avec un métal de transition, le zirconium (**B**).

Outre ces méthodes catalytiques en chimie conventionnelle, de nombreuses méthodes biocatalytiques sont décrites. Les mécanismes catalytiques restent basés sur la même nécessité d'activation de l'acide. Les méthodes d'amidation en chimie conventionnelle ne seront pas plus détaillées et ce manuscrit se concentrera principalement sur la synthèse enzymatique d'amides et de lactames.

## I.1. ETAT DE L'ART SUR LA SYNTHÈSE ENZYMATIQUE DE LIAISON AMIDE

### I.1.1. FORMATION DE LIAISON AMIDE DANS LE VIVANT

La liaison amide la plus répandue dans le vivant est la liaison peptidique présente dans les peptides et les protéines. Dans cette première partie sont décrits les différentes enzymes et ensembles de modules enzymatiques qui catalysent la formation de liaison peptidique dans les réactions métaboliques (Goswami and Van Lanen 2015, Pitzer and Steiner 2016).

#### I.1.1.1. SYNTHÈSE PEPTIDIQUE RIBOSOMALE

Dans les organismes vivants, les ribosomes sont en charge de la synthèse peptidique. Ces complexes ribonucléoprotéiques sont impliqués dans la traduction des acides ribonucléiques messagers (ARNm) en protéines à l'aide d'acides ribonucléiques de transfert (ARNt). Lors de la lecture de l'ARNm sous forme de codons (séquence de 3 nucléotides), l'interaction avec les anticodons de ARNt va permettre l'assignation de l'acide aminé correspondant au codon lu. Les ribosomes sont composés de deux sous-unités : une petite sous-unité va permettre la fixation de l'ARNm au ribosome principalement et une large sous-unité du ribosome porte l'activité catalytique sous la forme de trois domaines nommés A, P et E (Figure 3).

Tout d'abord, les aminoacyl-ARNt synthétases (EC 6.1.1.-), des ligases dépendantes de l'ATP, catalysent la réaction entre un ARNt et un acide aminé générant un intermédiaire activé aminoacyl-ARNt (aa-ARNt) dans le domaine A (Figure 3B). Pour chaque acide aminé protéinogène, il existe une aminoacyl-ARNt synthétase qui lui est spécifique. Une fois l'acide aminé activé, le groupement amine de l'aa-ARNt fait une attaque nucléophile sur le carbonyle de la chaîne peptidique (pept-ARNt) en cours

dans le domaine P, formant une liaison peptidique. Cette aminolyse passe par un intermédiaire tétraédrique instable qui va entraîner le départ de l'ARNt qui est désacylé (Nissen, Hansen et al. 2000). Cette réaction est catalysée par le centre peptidyltransférase (CPT, EC 2.3.2.12), qui joue un rôle clé pour l'élongation de la chaîne peptidique. Le CPT catalyse aussi l'hydrolyse de peptidyl-ARNt qui va ensuite sur le domaine E qui va permettre la sortie du peptide hors du ribosome (Rodnina 2013). Ces mécanismes sont à la base de la synthèse des protéines. Le polypeptide formé peut subir des modifications post-traductionnelles comme l'acétylation, la phosphorylation, l'alkylation ou la glycosylation.

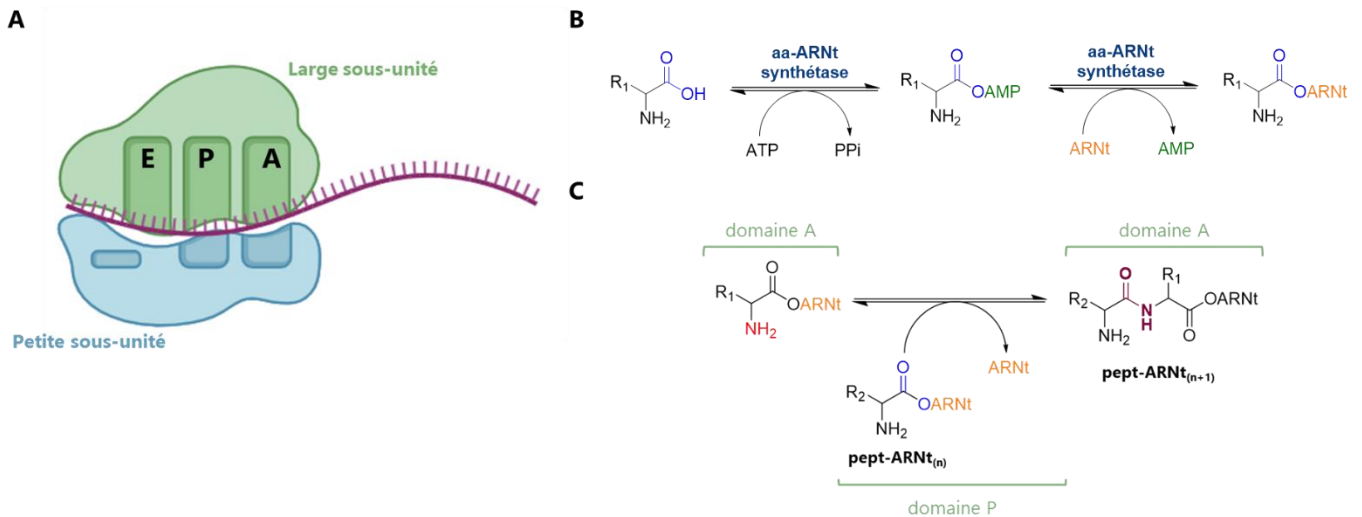


Figure 3. (A) Schéma de la structure d'un ribosome avec les domaines A, P et E. La chaîne violette correspond à la séquence peptidique en cours de synthèse par le ribosome. Créée avec BioRender.com. (B) Réaction d'activation de l'acide aminé par les aa-ARNt synthétases. (C) Formation de la liaison peptidique au sein du ribosome.

Tous les polypeptides ne deviennent pas des protéines, certains ont d'autres rôles avec des activités diverses, ce sont les « Ribosomally synthesized and post-translationally modified peptides » (RiPPs), appelés ainsi afin de les différencier des peptides non ribosomiques (NRP de l'anglais *Non-Ribosomal Peptides*) (Arnison, Bibb et al. 2013).

### I.1.1.2. PEPTIDES SYNTHETASES NON RIBOSOMIQUES (NRPS)

Chez les champignons et les bactéries, les peptides synthétases non ribosomiques (NRPSs, de l'anglais *Non-Ribosomal Peptide Synthetases*) permettent la synthèse d'un large spectre de peptides non ribosomiques qui sont des métabolites complexes. Parmi ces métabolites, on retrouve des composés présentant d'importantes propriétés biologiques, comme des antibiotiques, des anti tumoraux, des immunosuppresseurs ou des antidouleurs (Figure 4) (Süssmuth and Mainz 2017).

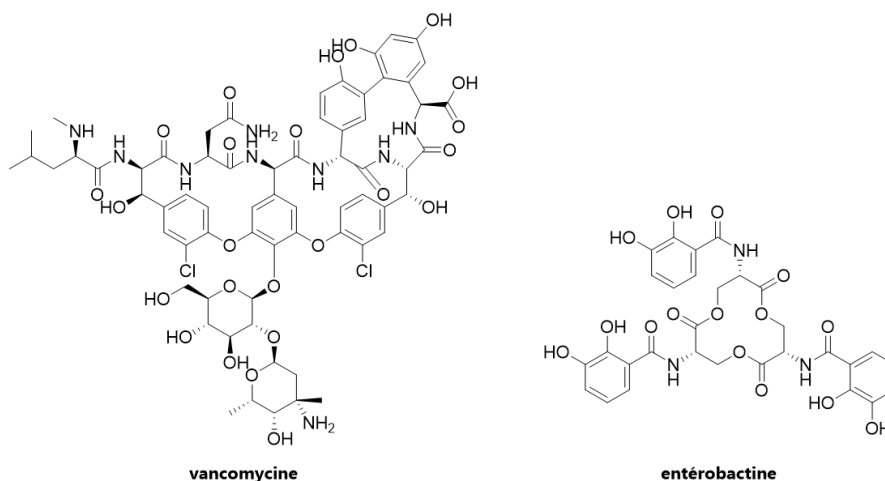


Figure 4. Exemples de peptides non ribosomiques biosynthésés par les NRPSs.

Les NRPS appartiennent à la famille des mégasynthases avec les polycétide synthases (PKS de l'anglais *polyketide synthase*) et les acide gras synthases (FAS de l'anglais *Fatty Acid Synthase*). Ce sont des enzymes à multi-domaines composés de 3 modules principaux responsables de la formation et de l'élongation du peptide : un domaine d'adénylation (A), un domaine de thiolation ou protéine porteuse de peptide (PCP de l'anglais *peptidyl carrier protein*), et un domaine de condensation (C). Les domaines d'adénylation appartiennent à la superfamille ANL, les enzymes adénylantes dépendantes de l'ATP comprenant aussi les luciférasés et les CoA ligases ou synthétases. Le domaine A reconnaît spécifiquement l'acide aminé et va l'activer en présence d'ATP pour générer un intermédiaire acyl-AMP. La formation de cet intermédiaire sera suivie d'une réaction de thioestérification avec le thiol du bras 4-phosphospantéthéine (PPan) rattaché par une sérine conservée dans le domaine PCP, produisant un intermédiaire acyl-thioester encore plus réactif. Cet intermédiaire thioester va ensuite s'additionner à la chaîne peptidique en cours par le domaine C, formant une liaison peptidique (Figure 5). Enfin pour la terminaison, un domaine de thioestérase (TE) est responsable de la libération du peptide, soit par hydrolyse de la liaison thioester ou par cyclisation, libérant alors un macrolactame ou une macrolactone (Challis and Naismith 2004, Marahiel 2009). D'autres modules optionnels peuvent être présents en amont du domaine TE pour réaliser des modifications sur la chaîne peptidique : méthylation, formylation, épimérisation, réduction, oxydation et cyclisation (Strieker, Tanović et al. 2010, Süßmuth and Mainz 2017).

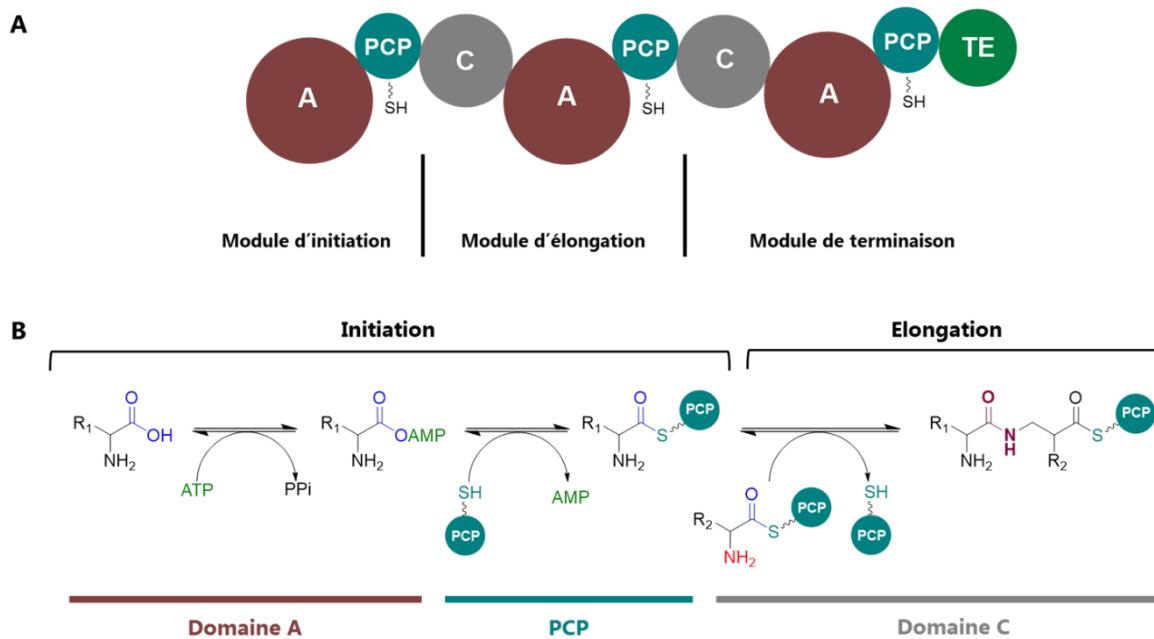


Figure 5. **(A)** Représentation schématique des modules constituant les NRPSs. Inspiré de l'article de Süssmuth et Mainz (2017). A correspond au domaine d'adénylation, PCP au peptidyl carrier protein, C au domaine de condensation et TE à la thioestérase. **(B)** Mécanisme de la formation de la liaison peptidique par les NRPSs.

Contrairement à la synthèse peptidique ribosomale, les NRPSs ne se limitent pas à l'intégration des acides aminés protéinogènes comme briques de construction des peptides, il peut également y avoir l'intégration de dérivés d'acides aminés. Ces dérivés peuvent provenir de voies biosynthétiques secondaires ou d'une modification par un module optionnel, après activation de l'acide aminé par un domaine autonome d'adénylation (SAAD de l'anglais *Stand-Alone Adenylation Domain*) (Süssmuth and Mainz 2017).

### I.1.1.3. AUTRES ENZYMES FORMANT DES LIAISONS AMIDES DANS LE VIVANT

#### I.1.1.3.1. Enzymes dépendantes de l'ATP

Les enzymes dépendantes d'ATP peuvent être classées en deux superfamilles : les enzymes capteuses d'ATP (*ATP-grasp enzymes* en anglais) et les enzymes adénylantes. Si ces deux superfamilles vont activer l'acide carboxylique, la principale différence se situe au niveau de l'intermédiaire formé en présence d'ATP. Les enzymes adénylantes vont générer un intermédiaire acyl-AMP comme il a déjà été mentionné précédemment avec les aa-ARNt synthétases ou les domaines d'adénylation des NRPS. Quant aux enzymes capteuses d'ATP, l'acide carboxylique sera activé par phosphorylation en produisant un intermédiaire acyl-phosphate très réactif (Schéma 5) (Healy, Mullins et al. 2000).

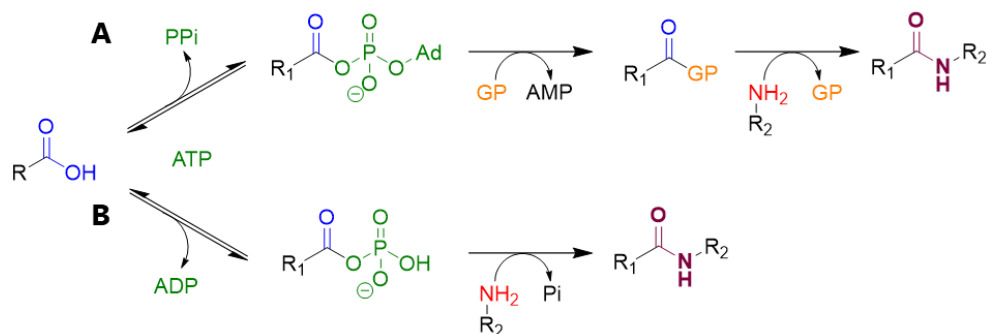


Schéma 5. Mécanisme d'activation de la fonction acide carboxylique par les enzymes ATP-dépendantes : les enzymes capteurs de l'ATP (A) et les enzymes adénylantes (B). Schéma inspiré de l'article Goswami *et al.* (2015). Légende : GP pour groupe partant et Ad pour adénosine.

#### I.1.1.3.1.1. Enzymes capteuses d'ATP

Cette superfamille, aussi connue sous le nom de ligase carboxylate-amine, compte actuellement 21 enzymes impliquées dans diverses voies de biosynthèse. Les premières à avoir été identifiées et caractérisées sont les biotine carboxylase, D-Alanine-D-Alanine ligase et glutathione synthétase (Schéma 6). Les autres membres de cette famille interviennent aussi bien dans le métabolisme central que dans le métabolisme spécialisé : le cycle de Krebs, la synthèse de purines, le catabolisme d'acides gras, la macrocyclisation de peptides ou encore la modification post-traductionnelle de RiPPs (Fawaz, Topper et al. 2011, Ogasawara and Dairi 2017).

##### A. Biotine carboxylase



CP : biotin carboxyl carrier protein

##### B. D-Ala-D-Ala ligase



##### C. Glutathione synthétase



Schéma 6. Exemple de formation de liaison amide catalysées par des enzymes capteuses d'ATP.

Toutes ces enzymes possèdent trois domaines conservés : le domaine N-ter, le domaine central et le domaine C-ter aussi appelés les domaines A, B et C respectivement. Elles tiennent leur nom de capteuses d'ATP dû à leur site de liaison à l'ATP particulier : il est formé par deux domaines  $\alpha + \beta$  qui vont saisir l'ATP, situés entre les domaines B et C. Les cations divalents les plus souvent impliqués sont  $Mg^{2+}$  et  $Ca^{2+}$ . Dans une première étape, ces enzymes vont donc former l'intermédiaire acyl-phosphate et dans une seconde étape, elles vont catalyser la formation de l'amide par addition nucléophile d'une amine (Schéma 5). Cette seconde étape nécessiterait également la présence d'un résidu basique dans le site actif afin de récupérer le proton libéré, mais ce dernier n'a toujours pas été identifié expérimentalement.

Une autre hypothèse est que le phosphate inorganique libéré jouerait ce rôle de base. Par ailleurs, en fonction de la source d'acide carboxylique, l'intermédiaire pourrait se décomposer en CO<sub>2</sub>, s'accompagnant d'un relargage de phosphate inorganique, avant l'attaque nucléophile de l'amine pour former une liaison amide (Fawaz, Topper et al. 2011, Goswami and Van Lanen 2015).

#### I.1.1.3.1.2. Enzymes adénylantes

La superfamille des enzymes adénylantes se divise en 4 classes d'enzymes. La classe I, appelée la superfamille **ANL** (de l'anglais *ANL superfamily*), regroupe les acyl/aryl-CoA ligases ou synthétases (**A**), les luciférases (**L**) et les domaines d'adénylation des NRPS (**N**). La classe II se réfère aux aminoacyl-ANRt synthétases, qui sont impliquées dans la synthèse peptidique ribosomale, la classe III correspond aux enzymes de synthèse des sidérophores indépendants des NRPS (NIS, de l'anglais *NPRS-independent siderophores*). Enfin, une quatrième classe s'est ajoutée dernièrement, les pimeloyl-CoA synthétases représentées par l'enzyme BioW, impliquée dans la biosynthèse de la biotine, qui possède les mêmes propriétés catalytiques communes aux enzymes adénylantes mais avec une structure tridimensionnelle différente (Estrada, Manandhar et al. 2017).

Toutes les enzymes adénylantes permettent l'activation de la fonction acide carboxylique en formant un intermédiaire acyl-AMP, l'addition nucléophile est catalysée par un autre domaine au sein des enzymes ou par une autre enzyme ou module enzymatique, comme c'est le cas pour les domaines d'adénylation des NRPS. Cependant, il existe une sous-classe d'enzymes de la famille ANL, les « Amide bond Synthetases » (ABS), qui catalysent aussi la formation de la liaison amide à partir de l'intermédiaire acyl-AMP. Elles ont été identifiées dans différentes voies de biosynthèse de métabolites secondaires, pour former des aminocoumarines comme la courmermycine A1 (Coul), la novobiocine (NovL) ou la clorobiocine (CloL), utilisées comme antibiotiques par exemple (Schéma 7) (Steffensky, Li et al. 2000, Eustáquio, Gust et al. 2003, Schmutz, Steffensky et al. 2003). On peut également citer d'autres ABS comme SimL, responsable de la formation de la liaison amide dans la biosynthèse de simocyclinone D (Pacholec, Freel Meyers et al. 2005), et McbA impliquée dans la biosynthèse d'alcaloïdes marinacarboline (Schéma 7) (Chen, Ji et al. 2013, Petchey, Cuetos et al. 2018). Ces ABSs sont également décrites dans le paragraphe I.1.2.3.1.2 car elles sont utilisées en biocatalyse.



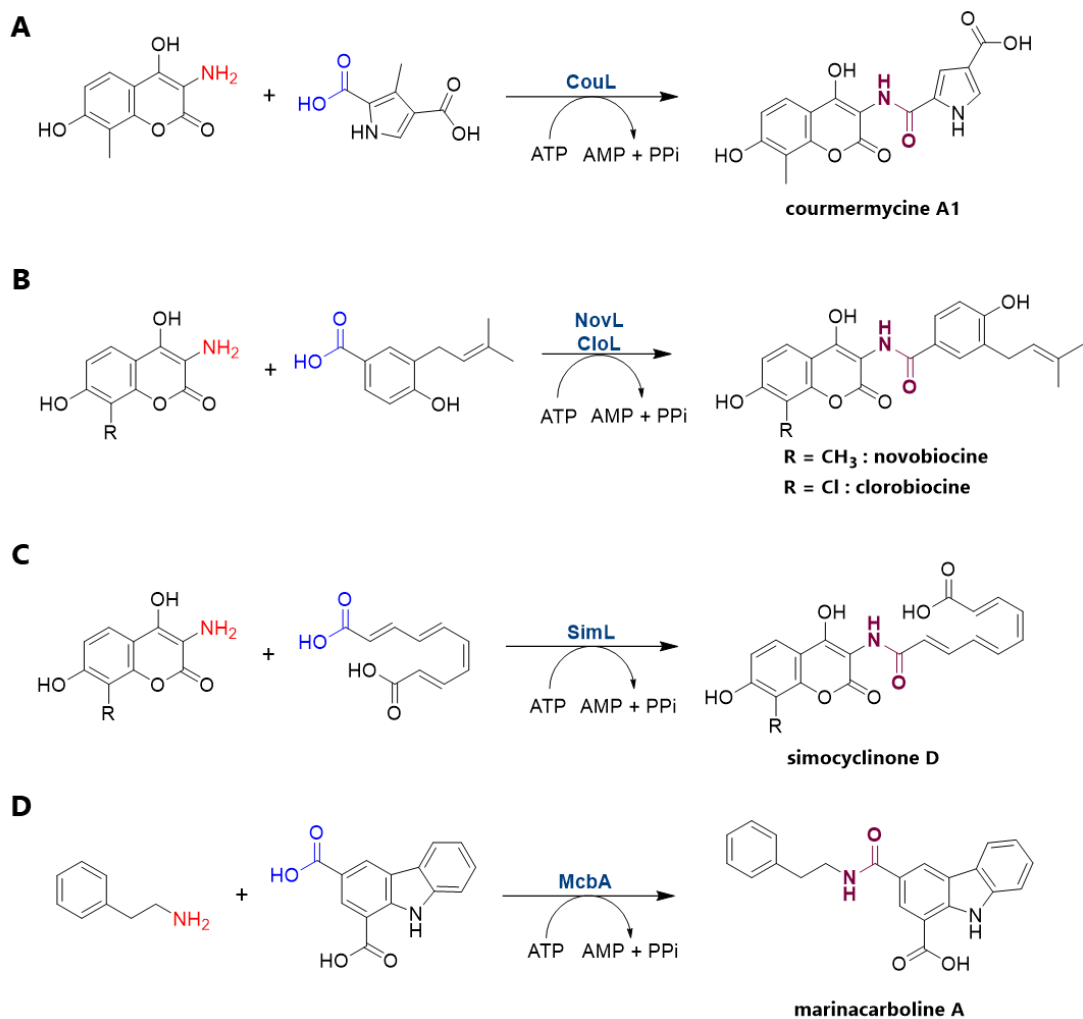
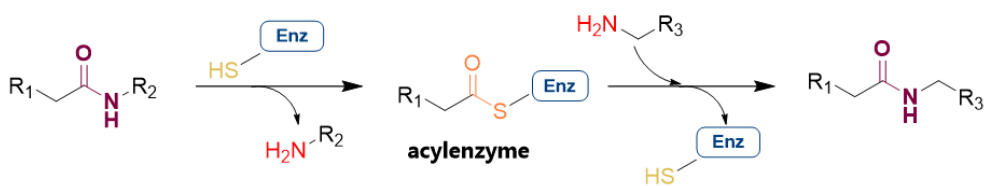


Schéma 7. Exemples de réactions métaboliques catalysées par des Amide Bond Synthétases (ABSs). (A) CouL, (B) NovL et CloL, (C) SimL et (D) McbA.

### I.1.1.3.2. Transpeptidases

Les transpeptidases catalysent le transfert d'acides aminés ou de peptides entre deux chaînes peptidiques. La formation de liaison peptidique se déroule donc entre un peptide et un autre substrat peptidique ou non. Elles appartiennent à la famille des hydrolases (EC 3.4-) ou des transférases (EC 2.3.2-) et catalysent la formation d'un intermédiaire acyl-enzyme. Cette catalyse covalente implique un résidu avec une fonction thiol ou alcool, selon qu'il s'agit d'une transpeptidase à cystéine ou à sérine respectivement. Ensuite, cet intermédiaire subit une attaque nucléophile de l'amine de la seconde chaîne peptidique pour mener à la formation d'une nouvelle liaison amide (Schéma 8).

### A. Transpeptidase à cystéine



### B. Transpeptidase à sérine

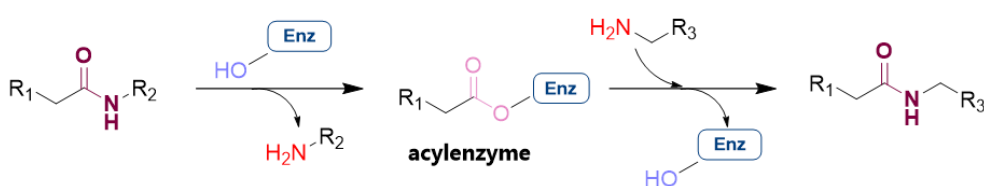


Schéma 8. Mécanisme réactionnel de la formation de la liaison peptidique par (A) des transpeptidases à cystéine ou (B) à sérine.

Un exemple métabolique de transpeptidase est la D-Ala-D-Ala transpeptidase à sérine (EC 3.4.16.4), une enzyme bactérienne qui permet le cross-linking des peptides dans le peptidoglycane, un composant de la membrane cellulaire des bactéries (Pratt 2008). Une autre transpeptidase impliquée dans du cross-linking de peptides est la transglutaminase, qui catalyse le transfert de résidus lysine avec des résidus glutamine et possède une cystéine dans sa triade catalytique à la place de la sérine (Ramos, Rollin et al. 2001). D'autres transpeptidases, comme le  $\gamma$ -glutamyltranspeptidase (EC 2.3.2.2) catalyse le transfert de glutathion et d'autres composés  $\gamma$ -glutamyl sur des acides aminés ou peptides chez tous les organismes vivants et ont des applications en biotechnologies ou pour la formation de peptides à visée thérapeutique (Suzuki, Kato et al. 2004).

#### I.1.1.3.3. N-acyltransférases

Les N-acyltransférases (NATs, EC 2.3.1.5) permettent d'accéder à des amides en catalysant le transfert de groupement acétyl d'un acyl-CoA vers une amine. Ces enzymes se retrouvent principalement dans les voies de biosynthèse de métabolites secondaires impliqués dans des processus de détoxification et qui possèdent des activités antibiotiques (Dunn, Wever et al. 2015, O'Flynn, Suarez et al. 2018). La réaction passe par un intermédiaire acyl-enzyme avant l'attaque nucléophile de l'amine pour donner l'amide (Schéma 9) (Sim, Walters et al. 2008).

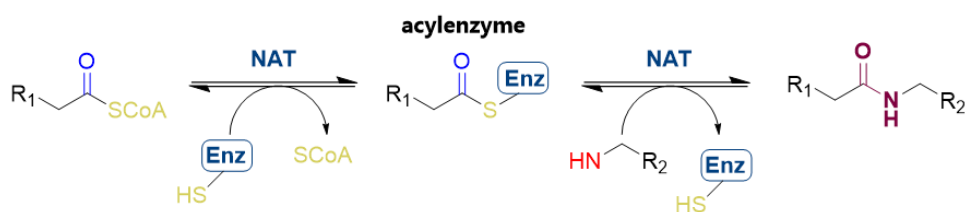


Schéma 9. Mécanisme réactionnel de la formation de liaison amide par les NATs.

CapW, auparavant présumée et annotée comme une  $\beta$ -lactamase de classe C, appartient à la famille des NATs et est impliquée dans une voie de biosynthèse d'antibiotiques dérivés de la capuramycine (Funabashi, Yang et al. 2010). Son substrat donneur est un ester de méthyle, issu de

l'activation de l'acide carboxylique catalysée par CapS, une carboxyle méthyltransférase dépendant du cofacteur S-adenosylméthionine (SAM). Ensuite, CapW catalyse la N-acylation avec l'amine d'un aminocaprolactame pour former l'antibiotique (Schéma 10).

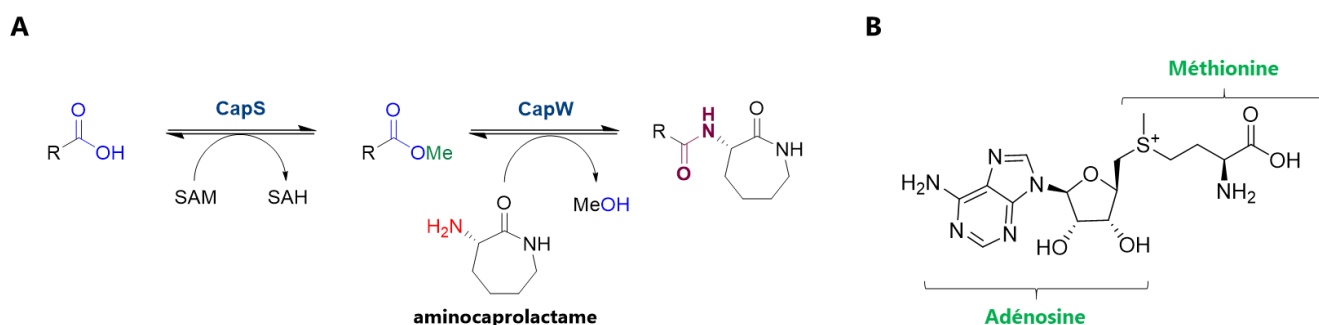


Schéma 10. (A) Mécanisme réactionnel de la formation de la liaison amide avec la NAT CapW. (B) Structure du cofacteur SAM.

#### I.1.1.3.4. Nitriles hydratases

Les nitrile hydratases (NHases, EC 4.2.1.84) sont des lyases qui vont catalyser l'hydratation de nitriles en l'amide primaire correspondant, dans les voies de dégradation des nitriles. Leur site actif contient soit un cation  $Fe^{3+}$  non hémique soit un cation  $Co^{3+}$  non corrinoïde. Leur mécanisme catalytique n'est pas encore tout à fait élucidé. Comme pour les transpeptidases, la formation d'amide primaire passerait d'abord par un intermédiaire acyl-enzyme qui nécessiterait une base dans le site actif. Les autres hypothèses de mécanismes catalytiques feraient également intervenir le cation dans l'intermédiaire acyl-enzyme (Schéma 11) (Prasad and Bhalla 2010, Cheng, Xia et al. 2020).

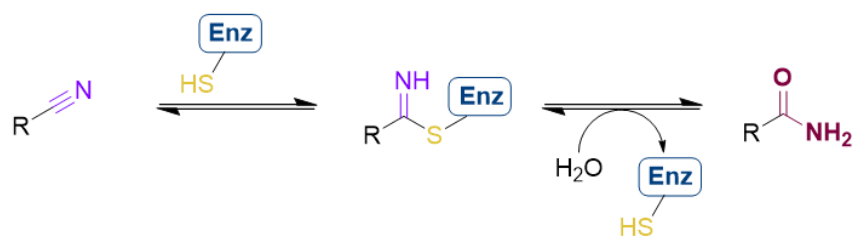


Schéma 11. Mécanisme réactionnel de formation de l'amide primaire par les NHases.

### I.1.2. SYNTHÈSE BIOCATALYTIQUE DE LIAISON AMIDE

Dans la partie précédente, nous avons vu des exemples d'enzymes catalysant la formation de liaison amide dans le vivant et leur mécanisme associé. La majorité de ces réactions métaboliques amènent à la formation de liaisons peptidiques, ce qui peut limiter leur application pour la synthèse biocatalytique de liaisons amides non peptidiques. Pour des applications en chimie, les enzymes utilisées sont contraintes par la facilité de mise en œuvre (résistance à l'oxygène, stabilité, disponibilité commerciale ou synthétique du cofacteur, facilité de production, etc.) et leur promiscuité de substrats. En effet, selon l'application visée, les substrats à transformer ont souvent des structures éloignées des substrats métaboliques. Les familles d'enzymes décrites dans le paragraphe précédent n'ont pas toutes des applications pour la synthèse biocatalytique d'amides et d'autres familles peuvent avoir un intérêt pour ces objectifs synthétiques.

Les différentes enzymes pour la synthèse biocatalytique d'amide sont détaillées dans la suite de ce manuscrit sur la base de revues et articles (Petchey and Grogan 2019, Winn, Richardson et al. 2020, Lubberink, Finnigan et al. 2023).

### I.1.2.1. HYDROLASES

Les hydrolases (EC 3.-) sont des enzymes qui catalysent la rupture d'une liaison covalente par action d'une molécule d'eau et sont classées en fonction du type de liaison hydrolysée. C'est la classe d'enzymes la plus utilisée à l'échelle industrielle pour la synthèse d'amides, mais également la plus étudiée avec de nombreuses structures cristallisées. Ces enzymes possèdent une bonne stabilité dans les solvants aqueux et organiques et ainsi qu'un large spectre d'activité sur divers substrats. De plus, beaucoup sont disponibles dans le commerce. La majorité des hydrolases possèdent la triade catalytique His-Asp-Ser et passent par la formation d'un intermédiaire acyl-enzyme. Au lieu d'être attaqué par l'eau ou l'anion hydroxyde, ce dernier peut subir l'attaque nucléophile de l'amine, comme il a été observé pour les transpeptidases (Schéma 12). Il est alors nécessaire pour cette aminolyse de se placer dans un solvant non aqueux. Enfin, cette famille d'enzyme est avantageuse pour faire du dédoublement cinétique et obtenir des amides énantiopurs.

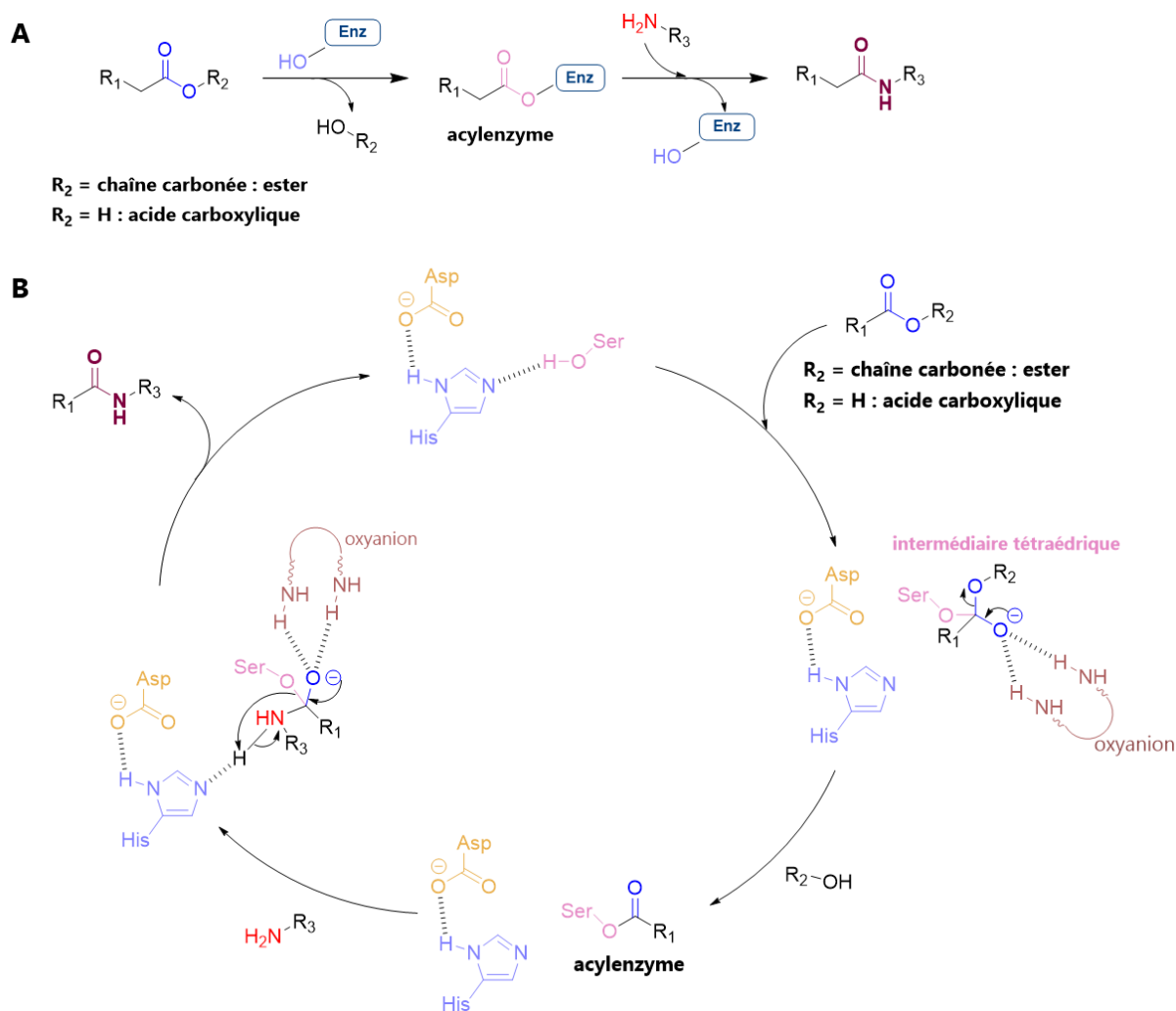


Schéma 12. Mécanisme général pour la formation d'amides avec les hydrolases (A) et le mécanisme réactionnel détaillé avec la triade catalytique His-Asp-Ser (B). Reproduction du schéma de l'article de Lima *et al.* (Lima, dos Anjos et al. 2019).

### I.1.2.1.1. Lipases

Dans le métabolisme, les lipases (EC 3.1.1.3) sont des hydrolases à sérine qui catalysent l'hydrolyse de liaisons esters de triglycérides. En chimie, ce sont des enzymes polyvalentes, très promiscuitaires, aussi capables d'hydrolyser des liaisons amides, de catalyser la formation de liaison C-C par addition de Michael par exemple (Busto, Gotor-Fernández et al. 2010). Les lipases permettent aussi bien d'accéder aux amides à partir d'esters que d'acides carboxyliques.

#### I.1.2.1.1.1. Aminolyse d'esters

Les premières lipases utilisées pour la synthèse d'amides ont été isolées de pancréas de porc (PPL, de l'anglais *pancreas porcine lipase*) et de *Candida cylindracea* (CCL, de l'anglais *Candida cylindracea lipase*). Ces aminolyses ont été réalisées dans un milieu organique anhydre pour éviter la réaction d'hydrolyse (Zaks and Klibanov 1985). Les lipases sont actives sur différents types d'esters et d'amines, pouvant être substitués avec des groupements hydrophobes ou halogénés par exemple (Schéma 13). Les lipases sont souvent utilisées pour du dédoublement cinétique d'amines, mais cela ne sera plus développé dans ce manuscrit car l'amine est la cible et non l'amide (Brieva, Rebolledo et al. 1990, Gotor, Brieva et al. 1991). La lipase de *Candida antarctica B* (CAL B) est la plus souvent employée, sous forme libre ou immobilisée. Elle présente en effet l'avantage d'accéder à des amides énantiopurs principalement (Lima, dos Anjos et al. 2019).

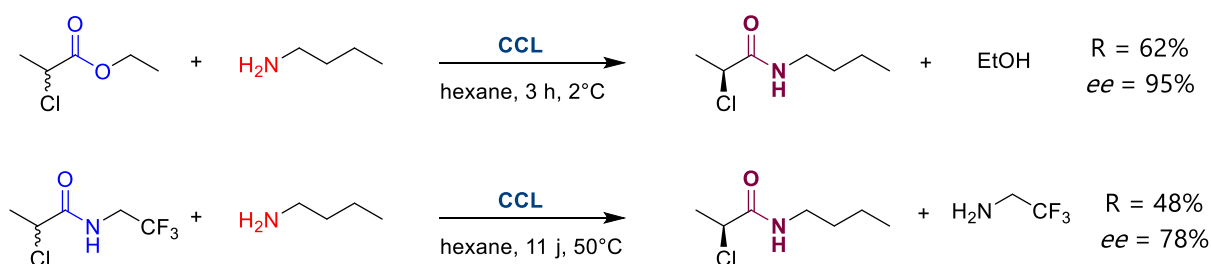
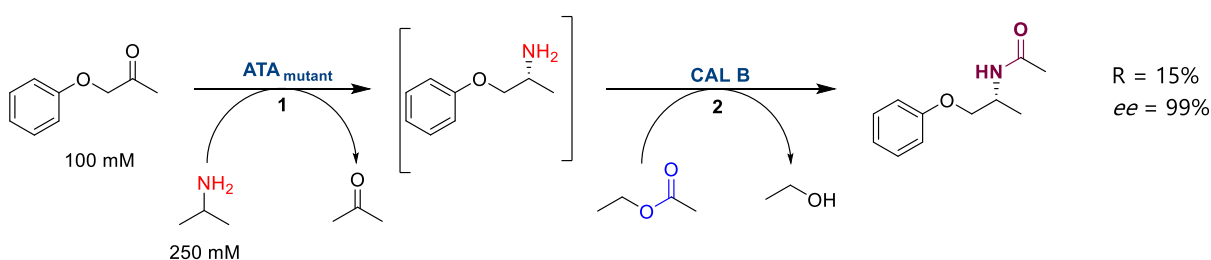


Schéma 13. Formation de liaison amide avec la CCL dans un solvant organique.

Les lipases se retrouvent aussi dans des cascades enzymatiques pour la synthèse d'amides. Une cascade en flux continu a par exemple été développée pour la synthèse stéréosélective d'amides en combinant CAL B immobilisée avec une transaminase, elle aussi immobilisée qui fournit l'amine nucléophile (*R*)-1-phenoxypropan-2-amine, pour l'aminolyse de l'acétate d'éthyle (Schéma 14) (Pintor, Matthey et al. 2023).



1) 1%v/v H<sub>2</sub>O dans EtOAc, 18h

2) EtOAc, 20 h, 37°C, 1200 rpm

Schéma 14. Formation d'un amide chirale en cascade avec une transaminase mutante (ATA) et CAL B immobilisés.

### I.1.2.1.1.2. Amidation directe d'acides carboxyliques

Les lipases catalysent aussi la synthèse d'amides en partant directement d'acides carboxyliques mais qui nécessitent toujours d'être activés pour que l'aminolyse ait lieu. Cependant, l'amidation avec des acides carboxyliques génère de l'eau comme sous-produit, ce qui peut entraîner l'hydrolyse de la liaison amide tout juste formée (Irimescu and Kato 2004).

Pour contourner ce problème, CAL B a été employée pour la synthèse d'un amide en présence d'éthanol comme solvant et pour activer l'acide carboxylique sous forme d'ester, qui subit alors une attaque nucléophile d'une amine primaire ou secondaire (Baldessari and Mangone 2001) (Schéma 15A). Par ailleurs, l'amidation d'acides carboxyliques a aussi été explorée dans des solvants ioniques ou dans un milieu sans solvant avec un retrait actif et continu de l'eau formée (Irimescu and Kato 2004) (Schéma 15B).

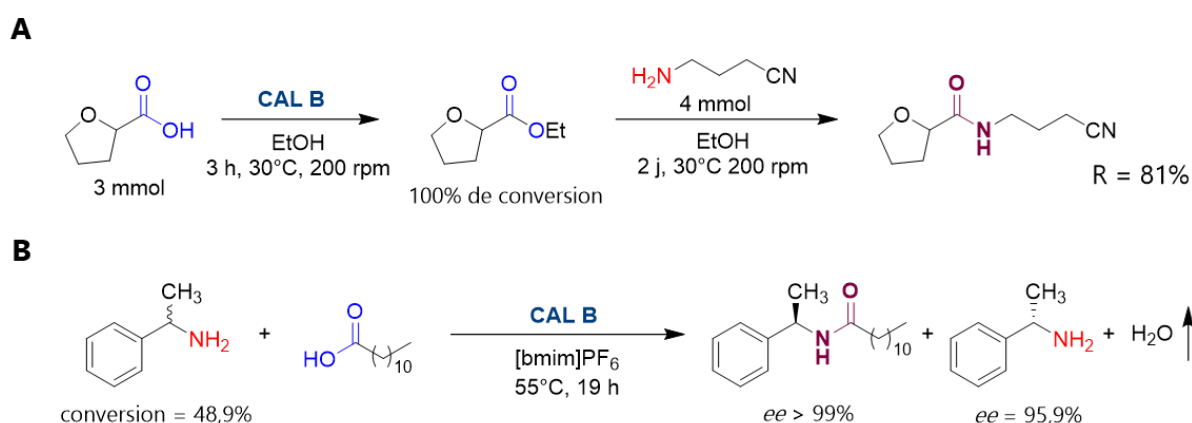


Schéma 15. Exemples d'amidation directe en partant d'acide carboxylique avec CAL B. (A) Avec utilisation d'éthanol comme solvant et réactif. (B) Dans un solvant ionique ou sans solvant avec retrait de l'eau formée sous atmosphère réduite.

Par ailleurs, la lipase de *Sphingomonas sp. HXN-200* (SpL) est aussi bien active avec des esters que des acides carboxyliques pour la formation d'amides (Zeng, Liu et al. 2018). Identifiée en criblant des organismes dégradant le n-octane, cette lipase produite par surexpression hétérologue chez *E. coli*, a montré des vitesses de réaction 1,76 fois plus rapides sur l'acide que l'ester, en présence d'eau dans le milieu organique. La synthèse du N-benzyl-3-phénylpropanamide dans du n-hexane avec 16% d'eau en volume a été réalisée avec 2 équivalent d'amine avec 99% de rendement en partant d'acide 3-phénylpropanoïque (Schéma 16). Une structure (PDB : 8OIM) récemment obtenue devrait permettre de comprendre la tolérance de cette enzyme à l'eau et donc de valoriser cette caractéristique qui la distingue de CAL B.

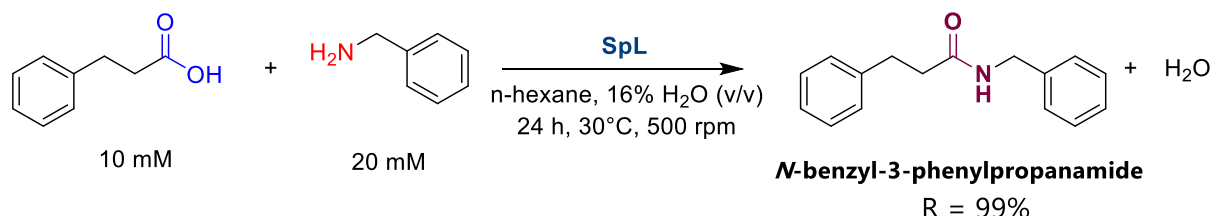


Schéma 16. Synthèse de N-benzyl-3-phénylpropanamide par la lipase SpL en présence d'eau dans le n-hexane.

Dernièrement, CAL B a été employée en chimie médicinale pour la synthèse de *N*-picolineamides qui servent d'agents anti proliférateurs dans le traitement de la maladie de Chagas, en partant de dérivés d'aminométhylpyridines et d'acides gras (Schéma 17) (García, Musikant et al. 2023).

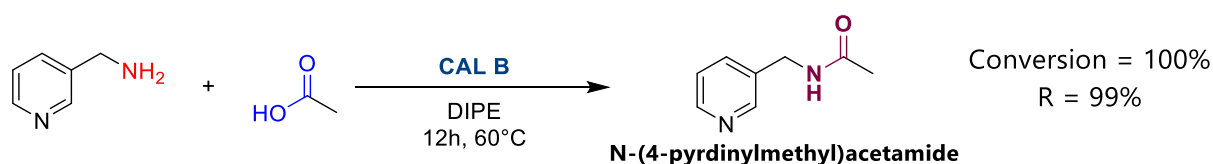


Schéma 17. Synthèse de la N-(4-pyridinylméthyl)acétamide (dérivé de *N*-picolineamide) par CAL B.

Les synthèses d'amides catalysées par les lipases sont d'un grand intérêt mais elles ont lieu dans des solvants organiques afin d'avoir la réaction d'aminolyse ou pour activer l'acide carboxylique. Mais certains de ces solvants présentent un danger environnemental et sanitaire (Lubberink, Finnigan et al. 2023). La synthèse d'amides avec des lipases dans des milieux aqueux ou dans des solvants plus durables serait une voie d'amélioration. Les études sur la lipase SpL ont montré la possibilité de synthèse en présence d'eau (Zeng, Liu et al. 2018). La compréhension structurale de cette propriété permettra peut-être de créer des lipases pour l'amidation fonctionnant avec encore plus d'eau que les 16% en volume de SpL. Plus récemment, l'équipe de Li a utilisé un milieu réactionnel glycérol/eau pour la synthèse de *N*-acylglycine en partant d'acides gras et de glycine comme source d'amine. Ainsi, la lipase RML active la fonction acide avec le glycérol, générant un intermédiaire ester-glycéryl avant l'aminolyse par la fonction amine de la glycine. Des travaux d'ingénierie protéique ont permis d'élargir le spectre d'activité de l'enzyme aux acides gras (jusqu'au C18) et d'améliorer son efficacité catalytique (Schéma 18) (Kua, Nguyen et al. 2023).

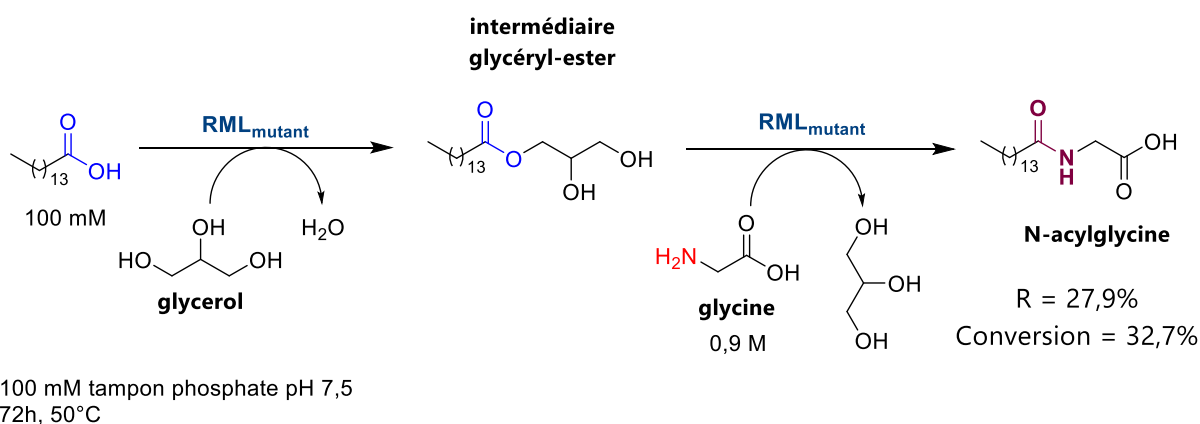


Schéma 18. Exemple de synthèse de *N*-acylglycine avec la lipase mutée de *Rhizomucor miehei* (RML) dans un milieu réactionnel eau-glycérol.

### I.1.2.1.2. Pénicilline acylases

Les pénicilline acylases (PAs, EC 3.5.1.11) sont des enzymes que l'on retrouve dans les microorganismes. Elles catalysent l'hydrolyse de la liaison amide *exo* de pénicillines, libérant l'acide 6-aminopénicillanique (6-APA), l'unité de base des pénicillines et l'acide phénylacétique, aussi appelé acide phénylglycine, activé sous forme d'amide (D-phénylglycine amide, D-PGA) (Grulich, Štěpánek et al. 2013). Cette réaction étant réversible *in vitro*, ces enzymes présentent donc un intérêt en synthèse pour la formation de nouvelles liaisons amide et elles sont souvent employées à l'échelle industrielle pour accéder à des pénicillines semi-synthétiques comme l'amoxicilline et le céfalexine par exemple (Schéma 19) (Jager, Jekel et al. 2007). Dans ces cas, la réaction d'amidation a lieu en tampon phosphate avec l'acide activé sous forme d'ester ou d'amide. Du fait de leur large application industrielle, les PAs font parties des enzymes les plus étudiées et caractérisées structurellement.

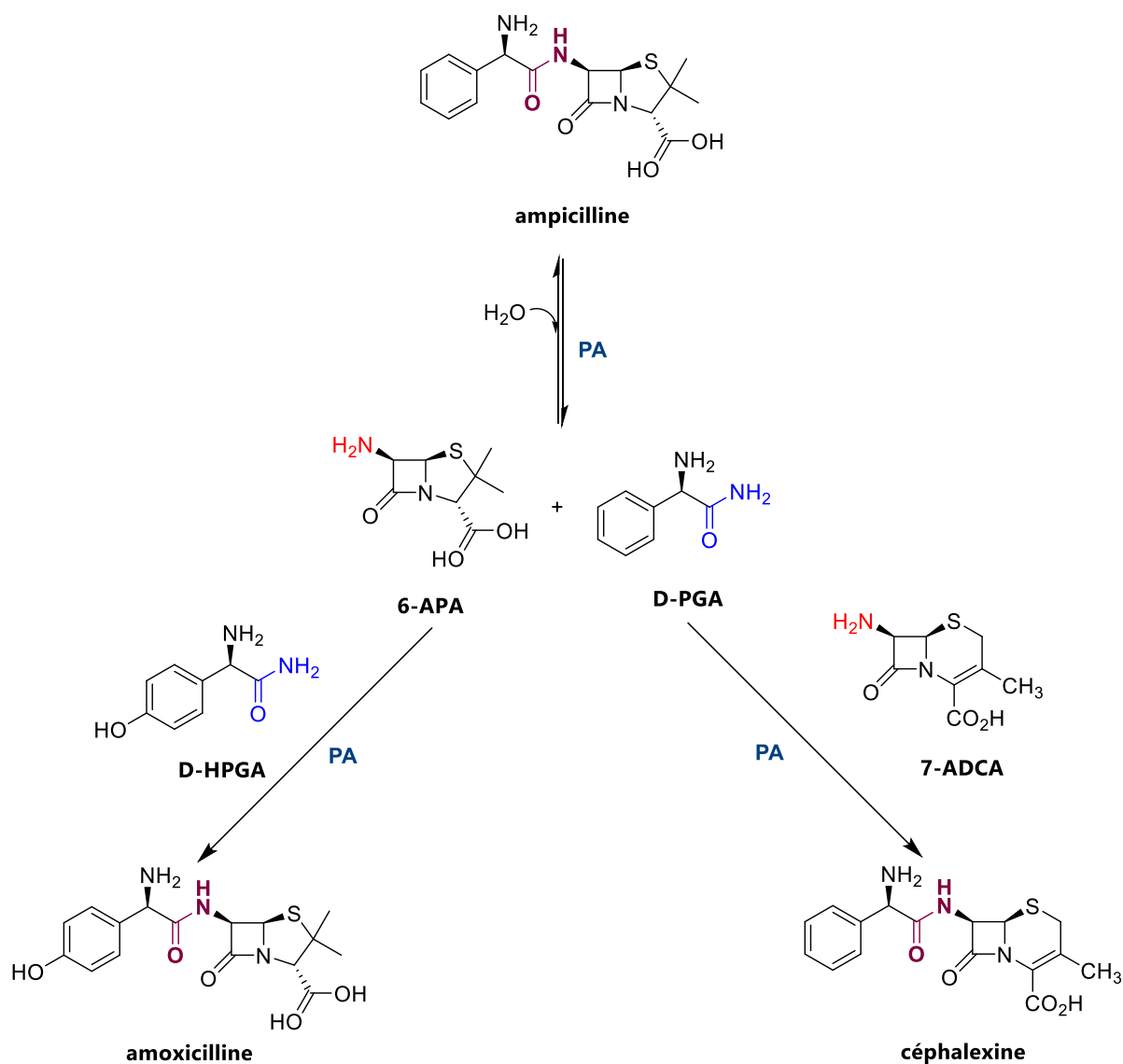


Schéma 19. Exemple de synthèse biocatalytique de pénicillines semi-synthétiques par des pénicillines acylases (PAs) en partant de l'hydrolyse de l'ampicilline. Légende : **6-APA** : acide 6-aminopénicillanique ; **D-PGA** : D-phénylglycine amide ; **D-HPGA** : D-*p*-hydroxyphénylglycine amide ; **7-ADCA** : acide 7-aminodésacétoxycéphalosporanique.

Cependant, la formation de liaison amide avec des PAs est en compétition avec deux réactions d'hydrolyse : l'hydrolyse de l'acide activé et l'hydrolyse de la liaison amide nouvellement formée. Ces réactions indésirables diminuent le rendement en antibiotique. Leur mécanisme étant bien connu, il est possible d'accéder à la formation de la liaison amide par contrôle cinétique de la réaction avec les PAs. Pour pallier ces problèmes, des travaux d'ingénierie protéique ont aussi été menés en complément d'études de pH, de solvants et de techniques de retrait du produit *in situ*. Ces expériences ont été réalisées avec la pénicilline G acylase (PGA) provenant d'*Achromobacter xylosoxidans* PX02, et ont permis d'atteindre respectivement 93,5% et 94,6% de rendement en ampicilline et amoxicilline avec une haute pureté (99%) (Pan, Li et al. 2020).

### 1.1.2.1.3. Protéases à serine

Les protéases (aussi appelées peptidases, EC 3.4.21-) catalysent l'hydrolyse des liaisons peptidiques au sein des protéines et sont généralement employées en synthèse peptidique car *in vitro*



cette réaction d'hydrolyse est réversible. Les protéases à sérine peuvent également catalyser la transamidation en présence d'amine comme déjà vu avec l'exemple des transpeptidases dans le métabolisme. Comme d'autres hydrolases, il existe une compétition entre la formation de la liaison peptidique et son hydrolyse. Les mêmes stratégies sont donc employées en partant de donneurs d'acyles activés sous forme d'ester ou d'amides.

L'utilisation des protéases permet de pallier les principaux inconvénients de la synthèse peptidique non enzymatique, à savoir la possible racémisation lors de l'activation de l'acide carboxylique par le donneur d'acyle et la purification du produit final dans un mélange de peptides isomères. Ces protéases sont stéréo- et régiospécifiques, elles ne nécessitent pas de cofacteurs coûteux et fonctionnent dans des conditions douces et sont plutôt stables (Bordusa 2002). Le principal inconvénient qui limite l'utilisation des protéases est que ces dernières possèdent une forte spécificité de substrat, d'où l'utilisation préférentielle d'autre hydrolases comme les lipases.

### I.1.2.2. ACYLTRANSFERASES

Les acyltransférases (ATase, EC 2.3.-) catalysent le transfert d'acyle d'un acide carboxylique activé vers une autre molécule. Elles sont impliquées dans des voies métaboliques centrales comme la biosynthèse des lipides ou la modification de protéines. La formation d'un intermédiaire acyl-enzyme permet de les utiliser en synthèse pour la transamidation. Elles peuvent être spécifiques de certains substrats (acides aminés ou acides gras), qui apparaîtront en suffixe dans le nom de l'enzyme comme la glycéronéphosphate O-acyltransférase par exemple. La sérotonine hydroxycinnamoyltransférase de *Capsicum annum* (66CaaT) est utilisée dans une cascade biocatalytique avec des CoA ligases qui vont générer l'intermédiaire thioester acyl-CoA, substrat de 66CaaT pour l'amidation (Philpott, Thomas et al. 2018). Cette cascade a été mise en application en cellule entière pour la synthèse biocatalytique d'un intermédiaire clé pour la synthèse du Losmapimod, un inhibiteur de tyrosine kinase (Schéma 20).

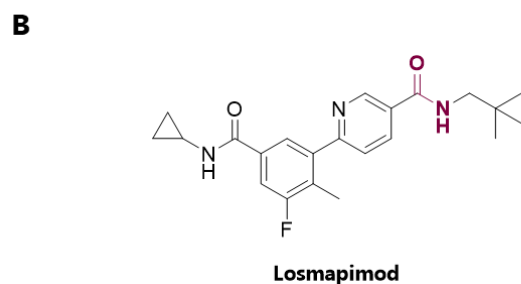
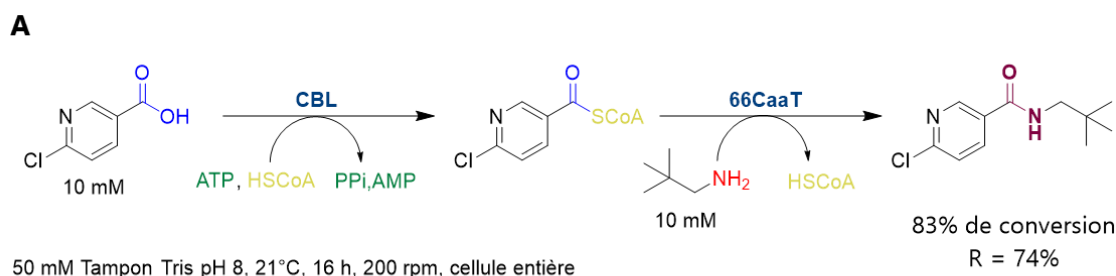


Schéma 20. (A) Synthèse d'un amide intermédiaire dans la synthèse du Losmapimod par 66CaaT. (B) Structure du Losmapimod.

Dans cette même famille, d'autres *N*-acyltransférases ont également été utilisées pour la synthèse d'amide comme celle de *Mycobacterium smegmatis* (MsAct) par exemple. Elles sont utilisées en biotransformation pour la production d'aminooxyacides comme l'acide suberanilique (SA de l'anglais

*suberanic acid*) et des dérivés, un intermédiaire de l'acide suberoylanique hydroxamique (SAHA de l'anglais *suberoylanilic hydroxamic acid*) qui est un potentiel inhibiteur d'histone désacétylases (Schéma 21A) (Christodoulou, Contente et al. 2022). Le SA et ses dérivés sont obtenus en partant d'anhydrides et d'anilines introduits à haute molarité (1M), avec de hauts rendements compris entre 68 et 94%. MsAcT a aussi été utilisée dans une cascade avec une transaminase (*Sp*-ATA) pour accéder à des amides en partant de cétones ou d'aldéhydes en solution aqueuse (Schéma 21B) (Land, Hendil-Forssell et al. 2016). Un des inconvénient de la MsAcT est qu'en milieu aqueux, elle catalyse l'hydrolyse de l'acide activé et du produit final comme tous les autres hydrolases présentées dans cette partie (Cannazza, Donzella et al. 2022).

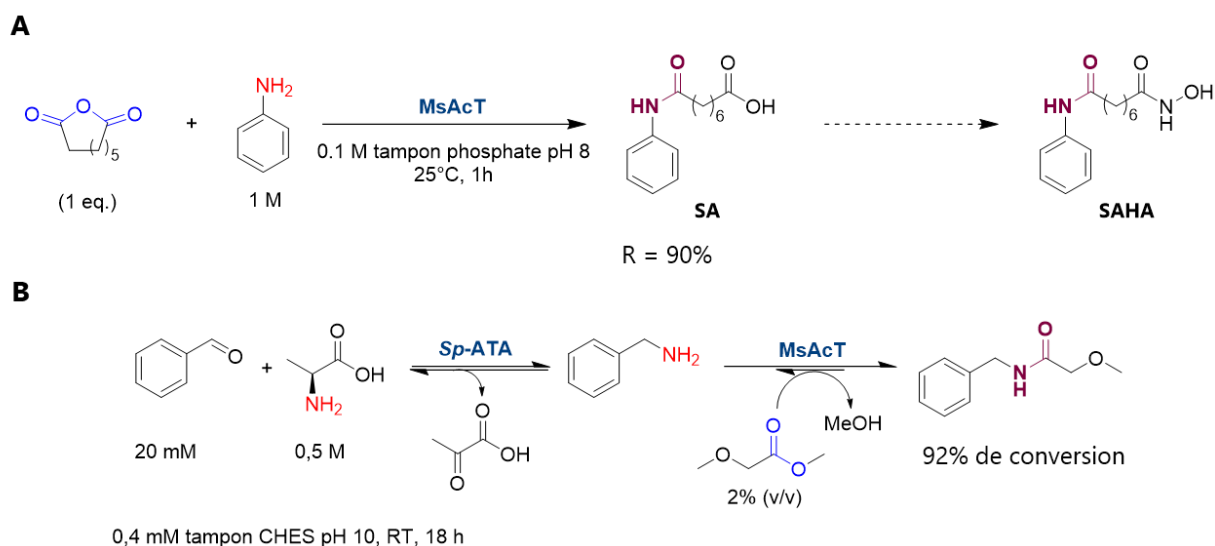


Schéma 21. Synthèse d'amides avec MsAcT. **(A)** Synthèse de l'acide suberalinique (SA), un intermédiaire du SAHA et **(B)** synthèse d'amides en cascade avec une transaminase en partant de benzaldéhyde. Légende : **CHES** : acide N-cyclohexyl-2-aminoéthanesulfonique, **Sp-ATA** : transaminase de *Silicibacter pomeroyi*.

Les transglutaminases (TGM, EC 2.3.2.13) appartenant à la sous-classe des aminoacyltransférases, sont largement employées dans les industries agroalimentaires et pharmaceutiques (Dorr and Fuerst 2018). Ces enzymes catalysent naturellement le transfert de glutamine sur le groupement amine d'un résidu lysine. La réaction s'accompagne d'une libération d'ammoniac (Schéma 22A).

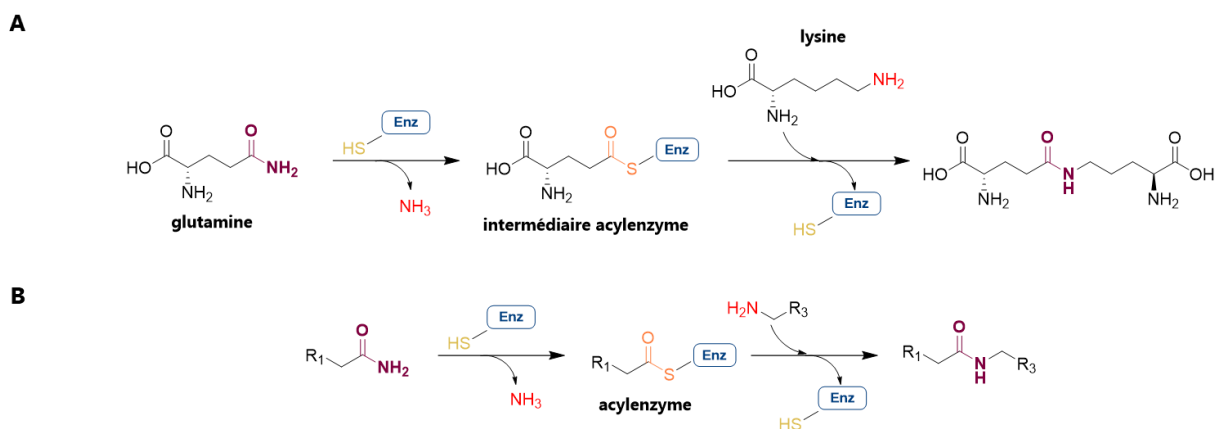


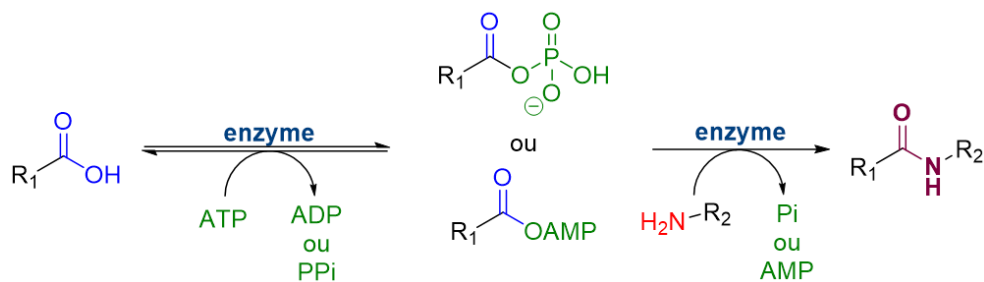
Schéma 22. **(A)** Réaction naturellement catalysée par la transglutaminase et **(B)** celle utilisée pour des applications biocatalytiques.

Le cross-linking (c'est-à-dire la modification de protéines par exemple par condensation de protéines) par les TGMs permet d'améliorer les propriétés de protéines alimentaires en agroalimentaire mais aussi d'ingrédients utilisés en formulation pour l'industrie cosmétique et biotechnologique (Bilawal, Gantumur et al. 2023, Vasić, Knez et al. 2023). Enfin, les TGMs permettent de faire de la bioconjugaison pour générer des conjugués de médicaments-anticorps à visée thérapeutique (Jeger, Zimmermann et al. 2010). La limitation majeure des transglutaminases reste leur spectre de substrat étroit car elles sont très spécifiques de la glutamine comme donneur d'acyle. Des réactions de désamidation du donneur d'acyle dans des solvants aqueux peuvent avoir lieu également, bien qu'elles soient minoritaires comparée au cross-linking.

### I.1.2.3. ENZYMES DEPENDANTES DE L'ATP

Avec les enzymes dépendantes de l'ATP, les synthèses de liaison amide peuvent être séparées en deux catégories : 1) les réactions d'amidation entre un donneur d'acyle et une amine comme c'est le cas avec les enzymes capteuses d'ATP et les Amide Bond Synthétases (ABS) ; 2) le détournement de l'activité naturelle de l'enzyme suite à l'étape d'activation de l'acide carboxylique (Figure 6). Ces cas sont détaillés dans les paragraphes suivants.

#### 1) Catalyse de l'amidation



#### 2) Détournement de l'activité de l'enzyme

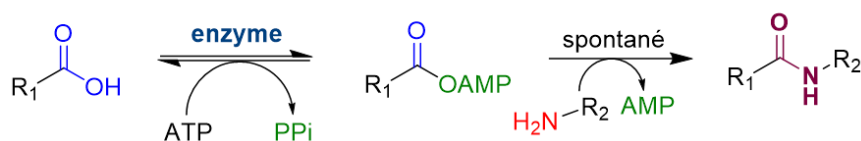


Figure 6. Stratégies employées pour la formation de liaisons amides par les enzymes dépendantes de l'ATP.

#### I.1.2.3.1. Formation de liaisons amides par catalyse de l'amidation

##### I.1.2.3.1.1. Enzymes capteuses d'ATP

Les enzymes capteuses d'ATP, notamment les L-amino acid ligases, catalysent la formation de liaisons peptidiques entre deux acides aminés via l'intermédiaire acyl-phosphate hautement réactif. La L-amino acid ligase ywfE de *Bacillus subtilis* a été testée pour la synthèse de dipeptides avec 21 acides aminés protéinogènes et sur les 231 combinaisons possibles, ywfE a catalysé l'amidation pour 111 d'entre elles (Tabata, Ikeda et al. 2005). Elle présente donc un large spectre de substrats avec une préférence pour des acides aminés sans encombrement comme l'alanine ou la sérine comme donneur d'acyle et des acides aminés aromatiques ou peu polaires comme donneur d'amine (Schéma 23A). Une

autre L-amino acid ligase, TabS de *Pseudomonas syringae*, présente un plus grand spectre de substrats avec les acides aminés protéinogènes, menant à la formation de dipeptides pour 136 combinaisons sur les 231 possibles et allant jusqu'à 96% de rendement pour la combinaison Gln-Thr (Schéma 23B) (Arai, Arimura et al. 2013). Ce dipeptide est un précurseur d'un autre dipeptide, Glu-Thr qui joue un rôle dans la salinité de produits alimentaires. Contrairement à ywfE, TabS a une préférence pour les acides aminés encombrés et polaires comme donneurs d'acyles et des acides aminés à chaîne latérale hydrophobe comme donneur d'amine. En outre, TabS est capable de synthétiser des peptides fonctionnalisés avec des acides aminés non protéinogènes comme la  $\beta$ -alanine, aussi bien utilisée comme donneur d'acyle ou donneur d'amine.

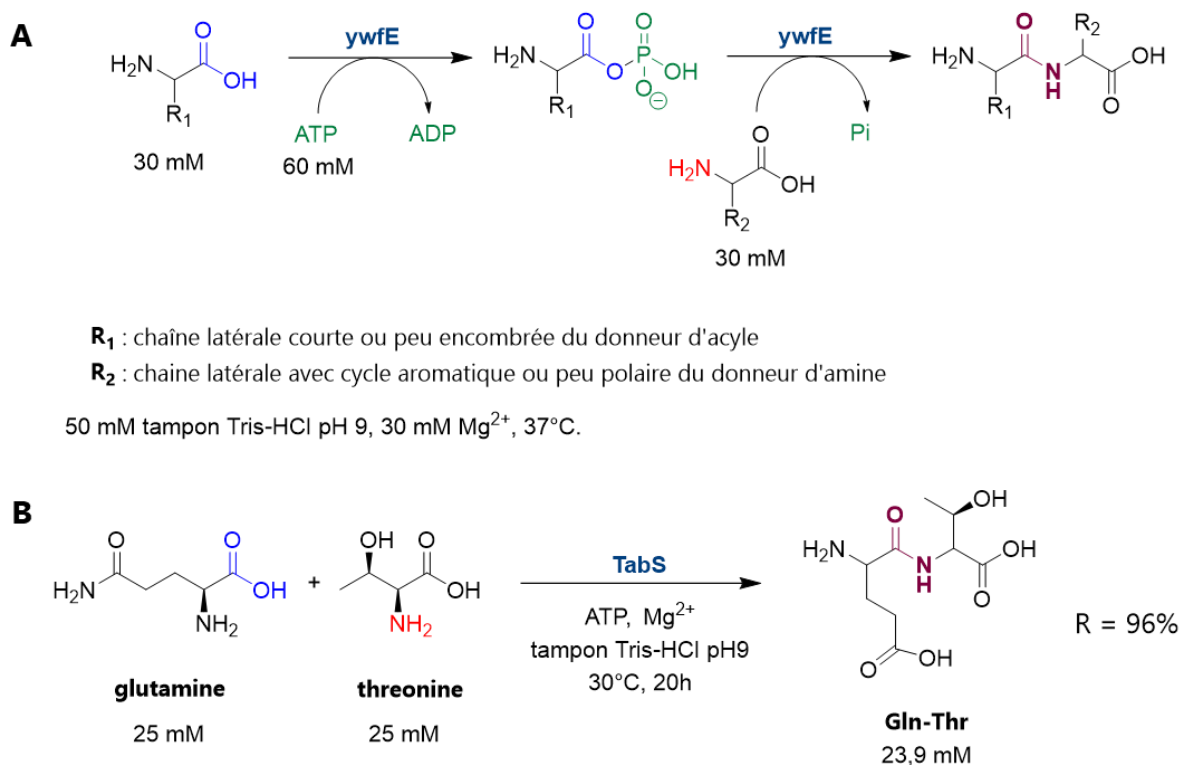
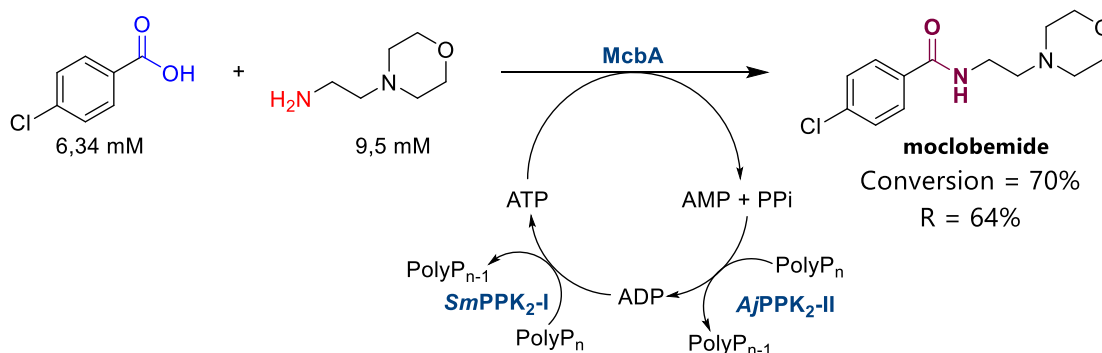


Schéma 23. Exemples de formation de liaison peptidiques par les L-amino acid ligases ywfE (A) et TabS (B).

#### I.1.2.3.1.2. Amide Bond Synthétases (ABS)

La première ABS utilisée pour la synthèse d'amide a été McbA de *Marinactinospora thermotolerans*, naturellement impliquée dans la biosynthèse d'alcaloïdes marinacarboline (voir paragraphe I.1.1.3.1.2). Comme il était déjà connu que McbA pouvait accepter des amines analogues ou dérivé de la 2-phényléthylamine, l'équipe du Pr. Grogan s'est intéressée à l'élargissement du spectre de substrats des acides carboxyliques avec d'autres dérivés de la  $\beta$ -carboline ou des substrats aromatiques portant un acide carboxylique (Petchey, Cuetos et al. 2018). Ces travaux ont mis en évidence que McbA peut réaliser le couplage d'une large gamme d'acides aromatiques, aussi bien le simple acide benzoïque que des acides avec deux cycles aromatiques ou contenant de la pyridine. Pour montrer son potentiel applicatif, McbA a été utilisée pour la synthèse biocatalytique de la monoclobémide, à l'échelle du milligramme, un antidépresseur qui a un rôle d'inhibiteur réversible de la monoamine-oxydase de type A (Petchey, Rowlinson et al. 2020). Cet inhibiteur est obtenu à partir d'acide 4-chlorobenzoïque et 4-(2-aminoethyl)morpholine. Un système de régénération de l'ATP, permettant d'utiliser l'ATP en quantité catalytique a été implémenté dans la synthèse (Schéma 24). McbA a également été cristallisée (PDB :

6SQ8) avec son substrat natif l'acide 1-acétyl-9H-β-carboline-3-carboxylique et de l'AMP, révélant une enzyme dimérique composée d'un large domaine N-ter et d'un petit domaine C-ter qui change de conformation entre l'adénylation et l'amidation. Son site actif qui se situe à l'interface des deux domaines est large et hydrophobe. Grâce à l'obtention de cette structure ainsi que des expériences de docking, il a été identifié qu'il y avait un tunnel d'accès pour l'amine dans le site actif et qu'un résidu aspartate (D201) jouerait un rôle dans l'activation de l'amine.



12,7 mM ATP, 50 mM tampon phosphate pH 7,5  
 0,2 mg/ml AjPPK<sub>2</sub>-II, 0,2 mg/ml SmPPK<sub>2</sub>-I  
 37°C, 250 rpm

Schéma 24. Synthèse du moclobémide par l'ABS McbA, couplé à un système enzymatique de régénération de l'ATP avec les polyphosphates kinases de *Sinorhizobium meliloti* (SmPPK<sub>2</sub>-I) et de *Acinetobacter johnsonii* (AjPPK<sub>2</sub>-II).

Plus récemment, une nouvelle ABS provenant de *Streptoalloteichus hindustanus* (ShABS) a été identifiée par homologie de séquences en utilisant la séquence de McbA comme modèle. Cette enzyme présente un plus large spectre d'activité que McbA sur des dérivés de la β-carboline et de diverses amines, dont certaines sont chirales, mais aussi de meilleurs rendements pour la synthèse des amides testés. La détermination de sa structure cristalline a permis une analyse fine du site actif et l'identification de résidus clés (Tang, Petchey et al. 2024). ShABS a été appliquée pour la synthèse d'APIs comme le cinépazide (vasodilatateur) avec 98% de rendement en 48h à 5 mM en substrat acide et 10 mM de substrat amine.

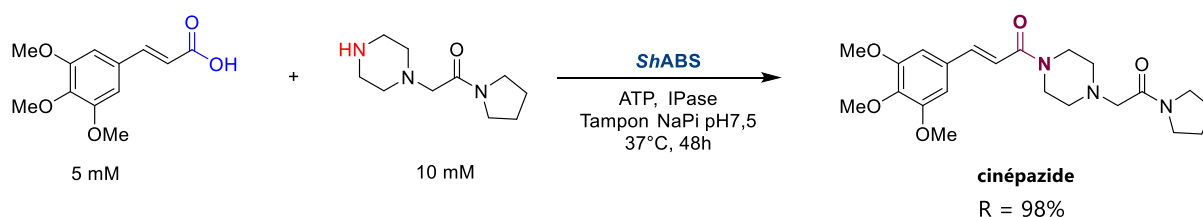


Schéma 25. Synthèse du cinépazide (vasodilatateur) catalysée par ShABS. Légende : IPase pour phosphatase inorganique.

Des travaux d'ingénierie protéique ont été effectués pour faire du dédoublement cinétique pour la synthèse d'amides en partant d'acides et d'amines racémiques et menant à de forts excès énantiomériques (*ee*) allant de 60% à 80%.

### I.1.2.3.2. Détournement de l'activité de l'enzyme

Dans cette partie, les enzymes étudiées ont toutes un point commun : en présence d'ATP, elles catalysent la formation de l'intermédiaire acyl-AMP. L'étape suivante est par contre différente de

l'activité métabolique : cette activation par adénylation catalysée est mise à profit pour permettre une seconde réaction non catalysée vers un amide (Figure 6). On peut donc parler de détournement de leur réaction naturelle. Concernant les enzymes adénylantes, nous nous focaliserons seulement sur la classe I des enzymes de la superfamille ANL et sur les Carboxylic Acid Reductases (CARs).

#### I.1.2.3.2.1. Domaines d'adénylation autonomes des NRPS

Les domaines d'adénylation autonomes des NRPS, appelés les SAADs en anglais pour *Stand-Alone Adenylation Domains*, servent à activer les substrats acides aminés en générant des intermédiaires aminoacyl-AMP avant leur transfert sur les autres modules. Les SAADs ont été utilisées pour la synthèse d'amides en jouant sur la promiscuité des substrats pour former une liaison amide. Par exemple, DhbE catalyse l'adénylation et l'amidation de l'acide 2,3-dihydroxybenzoïque (DHB), son substrat métabolique, avec de la cystéine pour produire du *N*-DHB-L-Cys, au lieu du transfert de l'acide activé sur le module PCP du NRPS (Abe, Hashimoto et al. 2017) (Schéma 26A). SfaB, une SAAD de *Streptomyces thioleus* impliquée dans la biosynthèse de SF2768 (chélateur de cuivre) via l'adénylation de l'acide (*R*)-3-isobutanoïque, a été testée avec différents acides aminés comme donneurs d'amine ainsi que différents acides carboxyliques. Suite à cette étude de spectre de substrats, les synthèses biocatalytiques d'amides comme le *N*-isovalérine-L-Lysine, l'heptanoyl-L-arginine ou le chloropentanoyl-L-arginine ont été réalisées (Schéma 26B) (Zhu, Wang et al. 2021).

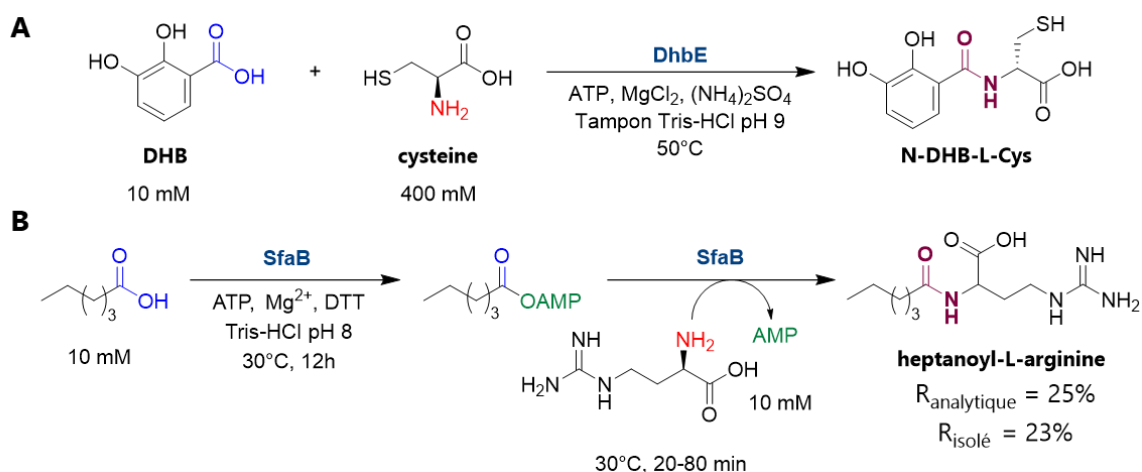


Schéma 26. (A) Formation de liaison amide avec DhE entre de l'acide 2,3-dihydroxybenzoïque (DHB) et de la cystéine. (B) Formation de liaison amide avec SfaB entre l'acide heptanoïque et l'arginine comme comme donneur d'amine.

Pendant, les SAADs se sont révélées très spécifiques pour l'acide aminé donc leur spectre d'activité peut être limité. La promiscuité en substrat est un peu plus grande pour les acides. DhbE possède une activité sur d'autres substrats ayant une structure très proche du DHB, comme les acides 2-hydroxybenzoïque et 3-hydroxybenzoïque. Quant à SfaB, son activité se limite à des acides gras courts de C4 à C7 mais ayant aussi une structure proche de l'acide (*R*)-3-isobutanoïque avec la présence d'un groupement substitué sur l'un des carbones terminaux, comme l'acide isovalérique ou l'acide 5-chlorovalérique pour lesquels elle présente les meilleures activités spécifiques (0,27  $\mu\text{mol/h/mg}$  pour ces deux substrats).

#### I.1.2.3.2.2. Carboxylic acid reductases

Les carboxylic acid reductases (CARs, EC 1.2.1.30), auparavant appelées aldéhydes oxydoréductases ou aryl aldéhyde déshydrogénases, catalysent la réduction ATP dépendante d'un acide

carboxylique en aldéhyde (Winkler 2018). Cette réduction est en trois étapes qui correspondent aux trois domaines qui composent les CARs nommés domaines A-T-R. La première étape qui se déroule dans le domaine A (en N-ter) est l'adénylation en présence d'ATP, réaction commune avec les enzymes adénylantes, qui génère un intermédiaire acyl-AMP. Ensuite, cet intermédiaire subit une réaction de thioestérification dans le domaine T avec la présence d'un groupe phosphopantéthéine pour obtenir un intermédiaire thioester acyl-CoA hautement réactif, qui est alors réduit en aldéhyde en présence de NADH dans le domaine R (en C-ter) (Figure 7) (Basri, Rahman et al. 2023). Les étapes d'adénylation et de thiolation sont similaires à ce qui est retrouvé chez les NRPS. Le principal avantage des CARs est leur large spécificité de substrats. Elles sont utilisées dans plusieurs procédés biocatalytiques, notamment en cascade pour la production de vanilline, d'alcane ou d'amines par exemple (Sheppard, Kunjapur et al. 2016, Hepworth, France et al. 2017, Horvat, Fiume et al. 2019, Basri, Rahman et al. 2023).

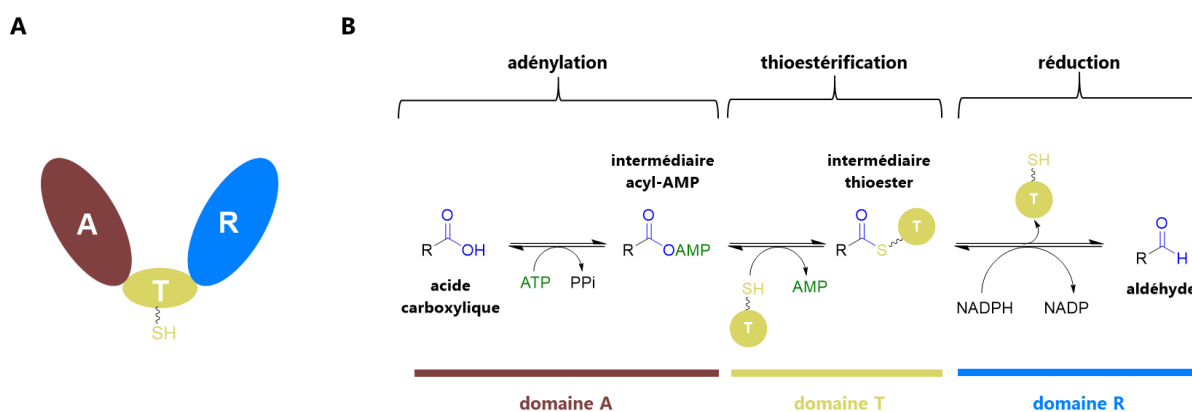


Figure 7. (A) Représentation schématique d'une CAR avec ses trois domaines : d'adénylation (A), de thiolation (T) et de réduction (R). (B) Mécanisme de la réaction catalysée par les CARs.

L'utilisation des CARs pour la synthèse biocatalytique d'amides a été particulièrement étudiée par l'équipe de Sabine Flitsch. Les CARs peuvent catalyser la formation de la liaison amide en détournant l'étape de réduction : le NADPH est remplacé par une amine qui s'additionne sur l'intermédiaire acyl-CoA et ainsi former l'amide correspondant (Wood, Weise et al. 2017). L'étape d'amidation n'est par contre pas catalysée par l'enzyme, ce qui nécessite l'emploi d'un large excès d'amine. Avec ce détournement de réaction, la synthèse biocatalytique d'un antiépileptique, l'ilepicimide, a été réalisée avec 96% de conversion en 72h avec une CAR provenant de *Mycobacterium marinum* (CAR<sub>mm</sub>). Dans ce même article, il a été montré que le domaine R n'est pas nécessaire à l'activité d'amidation en troquant ce domaine dans la CAR<sub>mm</sub>. Cette équipe est allée plus loin, en montrant que le domaine T n'est pas nécessaire non plus pour obtenir un domaine A fonctionnel et la formation de composés acyl-adénylés. Ainsi, la CAR<sub>mm</sub> tronquée ne conservant que le domaine d'adénylation (CAR<sub>mm</sub>-A) a montré une activité avec divers acides carboxyliques, en présence d'amines (Lubberink, Schnepel et al. 2020). De plus, cette réaction a été couplée à un système enzymatique de régénération de l'ATP utilisant une polyphosphate kinase (PolyP) de classe III de *Cytophaga hutchinsonii* (CHU) permettant d'utiliser de l'ATP en quantité catalytique (Schéma 27A). Avec cette CAR tronquée, plusieurs amides servant d'APIs ont pu être synthétisés comme le lazabémide avec 87% de conversion. Une montée en échelle à 100 mL a été réalisée pour la synthèse du cinépazide atteignant 18% de rendement. La CAR<sub>mm</sub>-A a été aussi employée pour la synthèse biocatalytique de tensioactifs non-ioniques stables et biodégradables, les N-alkanoyl-N-méthylglucamides (MEGAs) à partir de long acides gras et de N-méthylglucamine ou des dérivés (Schéma 27B) (Lubberink, Finnigan et al. 2022).

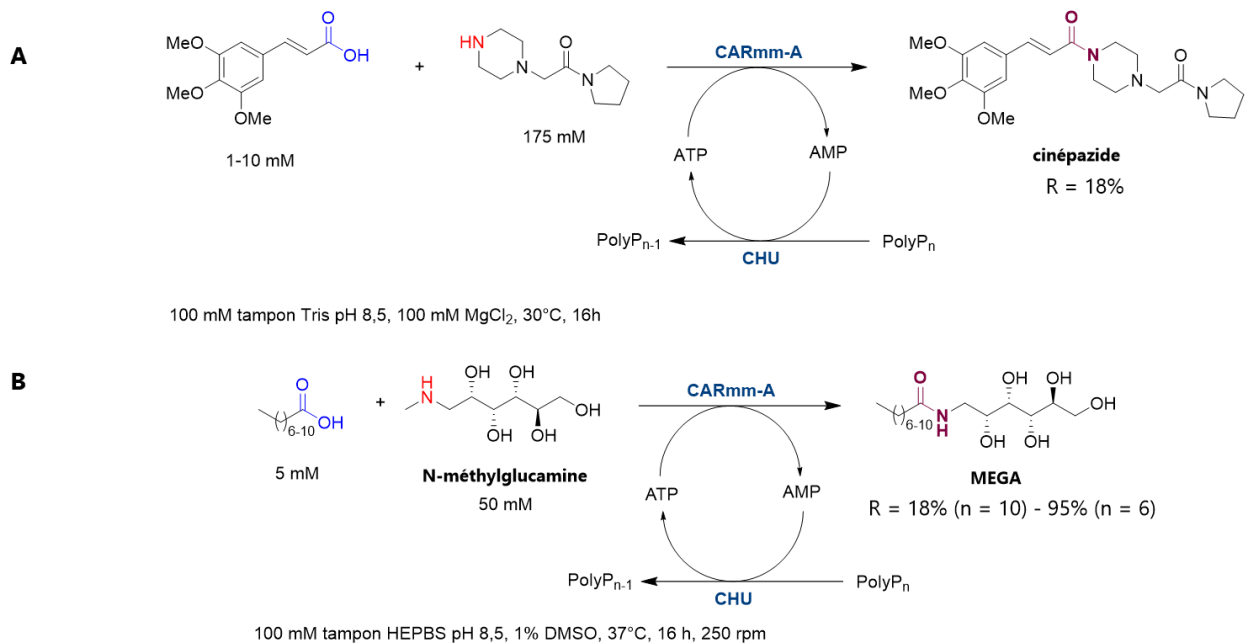


Schéma 27. Exemples de formation de liaison amide par le domaine d'adénylation de *CARmm* (*CARmm-A*) couplé à un système de régénération de l'ATP. **(A)** Synthèse de cinézapide (vasodilatateur) et **(B)** de tensioactifs MEGAs.

A travers ces applications, il apparaît que les CARs constituent une classe d'enzyme de choix pour la synthèse d'amide. En détournant la réaction originelle et en tronquant l'enzyme pour ne garder que le domaine d'adénylation, les CARs permettent d'accéder à des APIs. Enfin, l'utilisation de cofacteurs coûteux a pu être réduite en implémentant des systèmes de régénération enzymatique. L'avantage notable des CARs est la possibilité de formation sélective de liaisons amide sans la réaction en compétition d'estérification observée dans les réactions catalysées par les lipases.

#### I.1.2.3.2.3. CoA ligases

Les CoA ligases ou CoA synthétases (EC 6.2.1-) se retrouvent dans de nombreuses voies métaboliques centrales et spécialisées comme le cycle de Krebs, des voies de biosynthèse d'acides gras, de polycétides ou de terpènes par exemple. Les CoA ligases catalysent une étape d'adénylation suivie d'une étape de thioestérification en présence de HSCoA pour donner un intermédiaire thioester acyl-CoA (Schéma 28). Cet intermédiaire hautement réactif va être pris en charge par d'autres enzymes ou modules enzymatiques dans des voies biosynthétiques spécifiques.

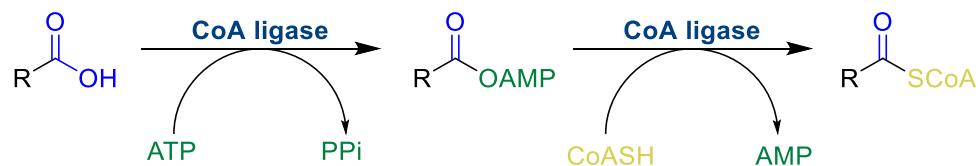


Schéma 28. Réaction naturelle catalysée par les CoA ligases.

Précédemment, nous avons pu voir un exemple où les CoA étaient employées en cascade avec des *N*-acyltransférases en leur fournissant le substrat thioester pour la formation d'amide (Philpott, Thomas et al. 2018). Il a été montré que les CoA ligases pouvaient être détournées pour la formation de liaison amide en substituant le CoA par une amine avec l'acyl-CoA ligase *AcsA* de *Pseudomonas Chlororaphis* (Abe, Hashimoto et al. 2008). Voulant étudier le spectre de substrats cette CoA ligase, le CoA avait été remplacé par de la L-cystéine pour suivre une réaction de thioestérification avec le groupement thiol de la cystéine. Cela a en fait résulté en la formation de la *N*-isobutyryl-L-cystéine,



mettant en évidence l'addition de l'amine sur l'intermédiaire acyl-adénylé (Schéma 29A). Plus récemment, la 4-coumarate CoA ligase de *Arabidopsis thaliana* (At4CL2) a été utilisée pour synthétiser des amides dérivés de l'acide 4-coumarique, le substrat natif de cette CoA ligase (Mori, Wanibuchi et al. 2021). Ces dérivés amide servent principalement d'inhibiteurs de croissance de bactéries comme *Bacillus subtilis* et *Staphylococcus aureus* ou de signal de quorum-sensing chez des espèces photosynthétiques, comme la *N*-(4-coumaroyl)-L-homosérine lactone (Schéma 29B). Le spectre de substrats de At4CL2 a été réalisé avec tous les acides aminés protéinogènes et d'autres non protéinogènes comme la L-homosérine lactone, avec l'obtention de meilleures activités avec les acides aminés protéinogènes. De plus, cette CoA ligase est également active avec d'autres substrats acides aromatiques proche de l'acide 4-coumarique comme l'acide caféique mais aussi des acides gras à chaîne moyenne comme l'acide hexanoïque, montrant une bonne versatilité de cette CoA ligase.

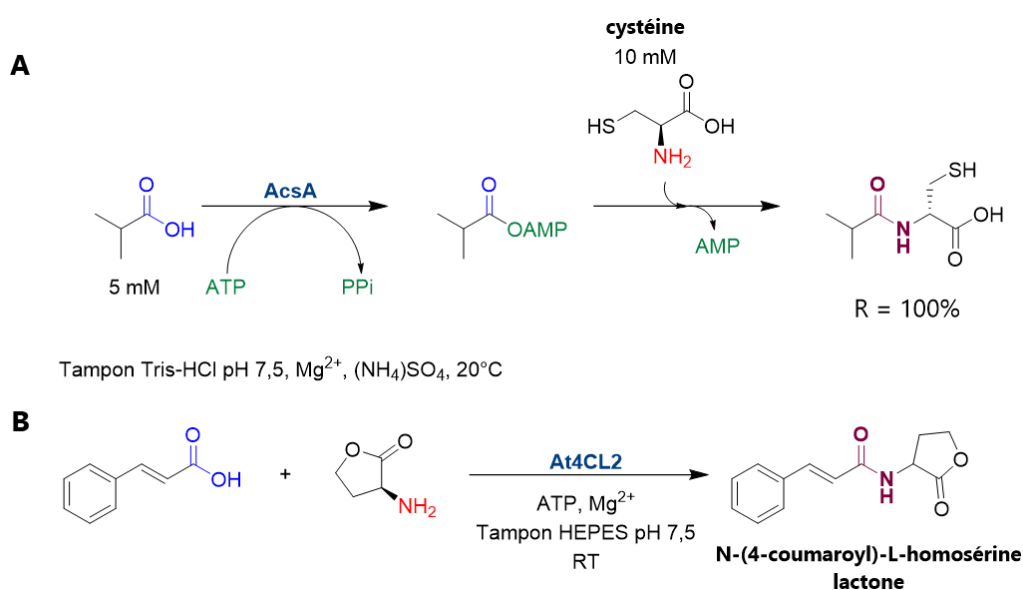


Schéma 29. Exemples de synthèse biocatalytique de liaison amide avec les CoA ligases AcsA (A) et At4CL2 (B).

Enfin, dans l'équipe de biocatalyse de l'unité Génomique Métabolique, la CoA ligase thermophile de *Metallosphaera sedulla* (code Uniprot : A4YDT1) a été étudiée pour la synthèse d'amides (Lelièvre, Balandras et al. 2020). Ces travaux constituent une partie de la thèse de Chloé Lelièvre effectuée entre 2017 et 2020. La thioestérification a été supprimée tout en promouvant l'attaque nucléophile de l'amine à haute température (60°C), ce qui justifie l'utilisation de CoA ligases thermoactives et thermostables. Cette réaction a été couplée à un système enzymatique de régénération de l'ATP permettant d'utiliser l'ATP en quantité catalytique (5% molaire). Comme cette réaction se déroule à haute température, ce système de régénération de l'ATP a nécessité également l'utilisation d'enzymes de recyclage thermoactives, une polyphosphate kinase de classe III et une pyrophosphatase. Cette réaction a ensuite été testée avec d'autres CoA ligases issues d'organismes thermophiles dont A4IBN3 de *Geobacillus thermodenitrificans* qui a permis une conversion totale de l'acide butyrique en *N*-méthylbutyramide, contre 88% de conversion obtenue avec A4YDT1 (Schéma 30). Une montée en échelle (10 mL, 10 mM d'acide butyrique) a été effectuée avec A4IBN3, menant à un rendement de 77% en *N*-méthylbutyramide isolé (95% de conversion).

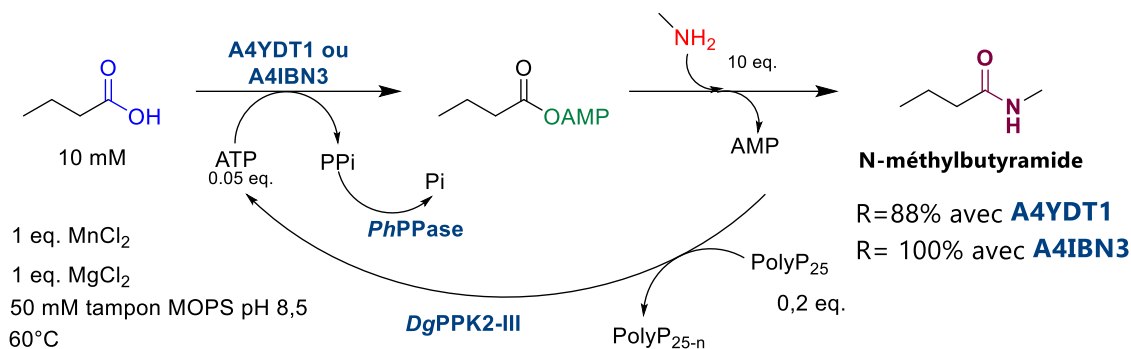


Schéma 30. Formation de liaison amide avec les CoA ligases thermophiles A4YDT1 ou A4IBN3, couplée à un système de régénération enzymatique de l'ATP avec la PPK2-III de *Deinococcus geothermalis* (DgPPK2-III) et la PPase de *Pyrococcus horikoshii* (PhPPase).

### I.1.3. CONCLUSION SUR L'ETAT DE L'ART POUR LA SYNTHÈSE ENZYMATIQUE DE LIAISON AMIDE

Tout d'abord, nous avons pu voir que dans le vivant, la majorité des enzymes catalysant la formation de liaison amide sont des enzymes impliquées dans les voies de biosynthèse des protéines et des peptides, catalysant la formation de liaisons peptidiques. Compte tenu de leur nature et de leur spécificité de substrats, ces enzymes du vivant sont peu applicables en synthèse biocatalytique. Cependant, d'autres classes d'enzymes ont pu être employées pour la synthèse de liaison amide comme les hydrolases qui ont un mécanisme réactionnel similaire aux transpeptidases, ou d'autres acyltransférases. Enfin, nous avons également vu les enzymes dépendantes de l'ATP, notamment les enzymes adénylantes qui représentent une classe de grand intérêt pour la synthèse de liaison amides et ont fait l'objet de travaux dans ces dernières décennies. Ces enzymes permettent d'accéder à la synthèse de la liaison amide soit par catalyse de l'amidation ou par détournement de leur activité. Récemment, la recherche et l'étude d'autres enzymes adénylantes pour la synthèse d'amides peut s'effectuer en partant d'études *in silico*. Ce type d'approche a été effectuée en utilisant l'ABS McbA comme point de départ pour remonter à l'enzyme ancestrale tout en prenant en compte la toxicité des substrats et des amides obtenus, dans un esprit d'une chimie toujours plus éco-responsable (Söderberg, Felicitas von Borries et al. 2023). Ces travaux ont permis d'identifier d'autres ABS mais aussi la synthèse divers amides durables pouvant être produits à partir de substrats acides et amines pas dangereux également. Cela met en exergue la possibilité de développer des voies de synthèses biocatalytiques durables tout en examinant l'aspect environnemental, éco-toxicologique de ces réactions.

### I.2. ETAT DE L'ART DE LA SYNTHÈSE ENZYMATIQUE DE LACTAMES

Les lactames étant des amides cycliques, la plupart des enzymes décrites précédemment pour la formation de liaison amide se retrouvent également pour la synthèse de lactames. Nous nous intéresserons principalement à la synthèse de lactames de petites taille (cycle de 5 à 7 chaînons). Les macrolactames n'ont pas été traités au cours de ces travaux mais font partie des cibles d'intérêt du projet ANR SMALA (ANR-21-CE07-0061) coordonné par le laboratoire.

## I.2.1. SYNTHÈSE DE LACTAMES DANS LE VIVANT

La synthèse de lactames dans le vivant se retrouve principalement dans la biosynthèse d'antibiotiques avec la classe des  $\beta$ -lactames. Par exemple, une  $\beta$ -lactame synthétase ( $\beta$ -LS) dépendante de l'ATP et de  $Mg^{2+}$  comme cofacteurs, est impliquée dans la biosynthèse de l'acide clavulanique, un inhibiteur de  $\beta$ -lactamases isolé de *Streptomyces clavuligerus* (Miller, Bachmann et al. 2001). Homologue des asparagine synthétases de classe B, cette  $\beta$ -lactamase catalyse l'adénylation en présence d'ATP pour activer l'acide carboxylique du N<sup>2</sup>-(carboxylethyl)-L-arginine (CEA), puis la  $\beta$ -lactamisation avec l'amine secondaire de la chaîne principale de la L-arginine pour conduire au lactame de l'acide déoxyguanidinoproclavaminique (DGPC), précurseur de l'acide clavulanique (Schéma 31A). Un autre exemple d'enzyme dépendante de l'ATP catalysant la  $\beta$ -lactamisation selon le même mécanisme, est la carbanepam synthase impliquée dans la biosynthèse de l'acide (5*R*)-carbapenem-2-em-3-carboxylique (Schéma 31B) (Gerratana, Stapon et al. 2003).

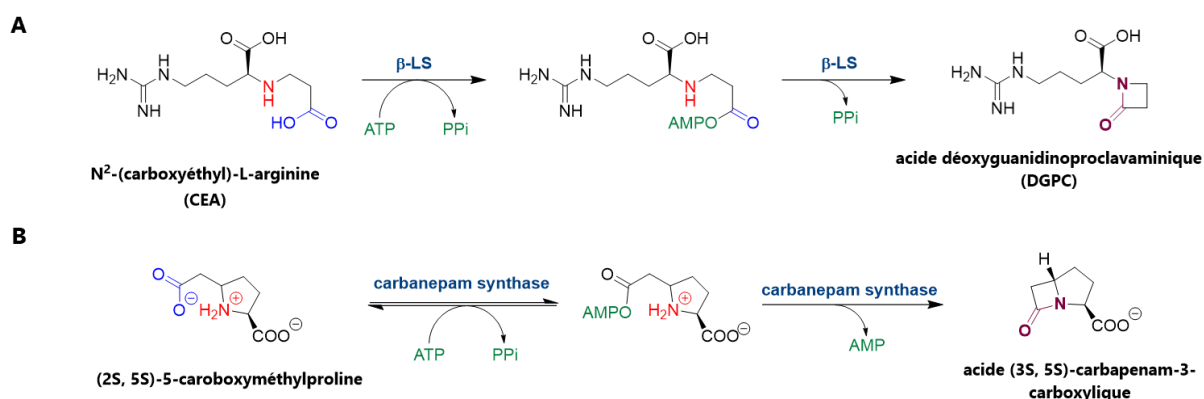


Schéma 31.  $\beta$ -lactamisation par  $\beta$ -lactame synthétase (A) et la cabanepam synthase (B).

L'isopénicilline N synthase (IPNS) impliquée dans la biosynthèse de l'isopénicilline N (IPN), est une autre enzyme catalysant la  $\beta$ -lactamisation (Chapman and Rutledge 2021). L'isopénicilline N appartient à la superfamille des oxygénases dépendantes du 2-oxoglutarate et possède un centre ferrique non hémique. Son mécanisme d'action pour la cyclisation passe par la coordination d'une molécule de dioxygène pour former un complexe oxo-ferrique hypervalent. Ce complexe hypervalent va réaliser l'oxydation de la  $\delta$ -(L- $\alpha$ -aminoadipoyl)-L-cysteinyl-D-valine (ACV) pour obtenir l'IPN (Schéma 32). Elle se différencie donc des deux autres enzymes car la liaison amide est déjà présente dans la structure de l'ACV et l'IPNS catalyse la fermeture du cycle en formant un liaison C-N.

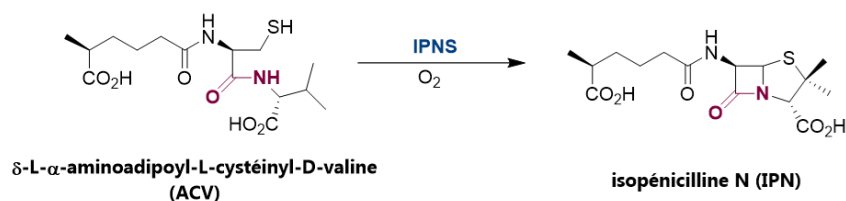
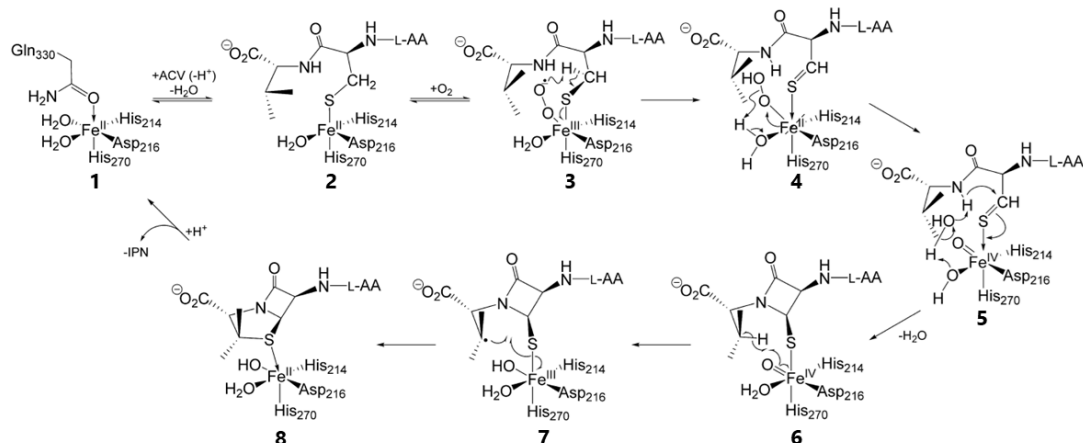
**A****B**

Schéma 32. Synthèse du noyau lactame via une oxydation par l'isopénicilline synthase (IPNS) et (B) son mécanisme réactionnel. Le schéma du mécanisme réactionnel est une image issue de l'article de Chapman (Chapman and Rutledge 2021).

Enfin, nous retrouvons les NRPS couplés à des modules de PKS pour la synthèse de lactames dans le vivant. Par exemple, chez le champignon *Pyricularia oryzae*, un système hybride de NRPS-PKS a été identifié pour la synthèse de métabolites contenant une structure de  $\gamma$ -butyrolactame (Motoyama, Nogawa et al. 2023). Ce système codé par le gène TAS1, ainsi que d'autres systèmes codés par des homologues de ce gène, catalysent la biosynthèse de l'acide tenuazonique qui est une mycotoxine. Ils sont également impliqués dans la biosynthèse des taslactames qui sont de potentiels inhibiteurs de protéasome, similaires à la lactacystine et à la salinsporamide A isolés d'actinomycètes, et qui présentent également des propriétés anti-cancéreuses. Le mécanisme proposé pour l'obtention du  $\gamma$ -lactame est une cyclisation résultant de la condensation entre une isoleucine et l'intermédiaire acétoacétyl-CoA qui forment un intermédiaire  $\beta$ -cétamide qui subit une cyclisation intramoléculaire de Dieckmann au niveau du domaine KS (Figure 8).

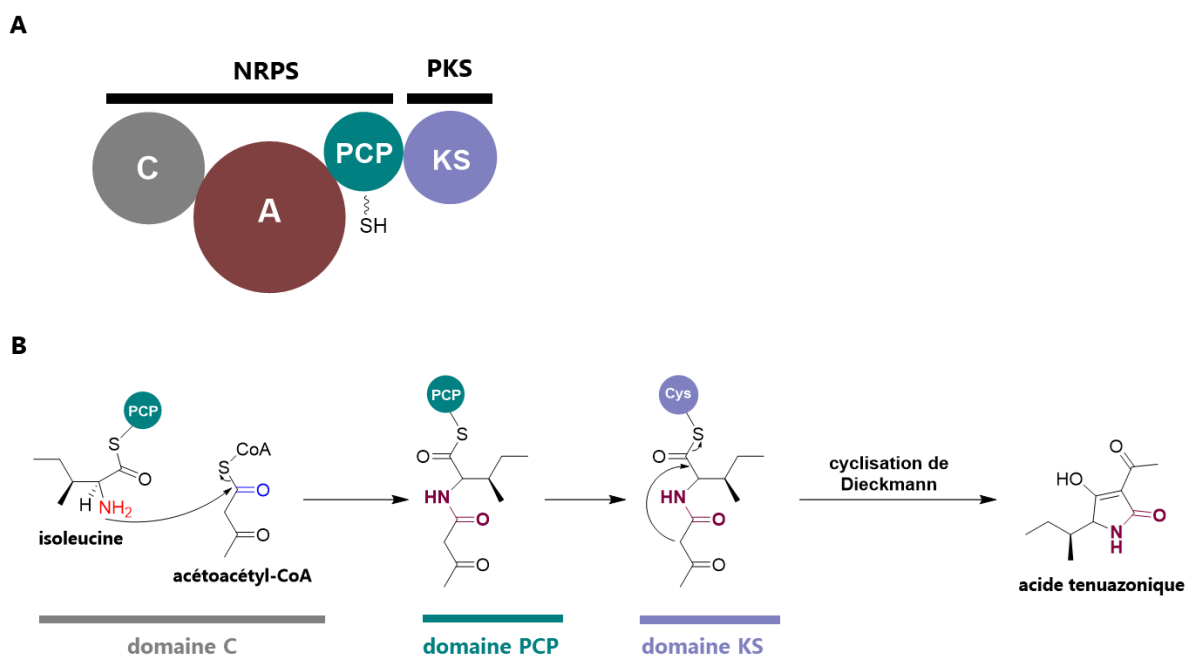


Figure 8. **(A)** Représentation des modules du gène TAS1 pour la biosynthèse de l'acide tenuazonique. **(B)** Mécanisme de la biosynthèse de l'acide tenuazonique par le système enzymatique codé par le gène TAS1 avec le domaine de condensation (C), le domaine peptidyl carrier protein (PCP) du NRPS, et le domaine cétosynthase (KS) du PKS.

A travers ces exemples, nous pouvons noter qu'il existe quelques voies pour la biosynthèse de lactames de petite taille, notamment pour la synthèse de  $\beta$ -lactames, et que ces dernières ne sont principalement présentes que chez des champignons. Les réactions de  $\beta$ -lactamisation sont majoritairement régies par l'activation de l'acide carboxylique puis catalyse de l'amidation intramoléculaire. Ces réactions se rapprochent donc de celles étudiées avec les enzymes adénylantes pour la synthèse d'amides. L'autre stratégie de formation de  $\beta$ -lactame consiste en la fermeture du cycle par une liaison C-C intramoléculaire, avec un substrat contenant déjà la liaison amide. Enfin les modules de NRPS peuvent se retrouver combinés avec des modules de PKS pour la synthèse de lactames supérieurs à 5 chaînons.

## I.2.2. SYNTHÈSE BIOCATALYTIQUE DE LACTAMES

Pour la synthèse de lactames, nous retrouvons quelques enzymes déjà exploitées pour la formation de liaisons amides, dont l'application a été étendue aux lactames. Les enzymes catalysant la  $\beta$ -lactamisation dans le vivant passent par l'activation de l'acide carboxylique en présence d'ATP, comme les enzymes dépendantes de l'ATP détaillées précédemment. C'est donc en toute logique que ce type d'enzymes est aussi exploité pour la synthèse de lactames. Toujours en s'inspirant des voies biosynthétiques, il est également possible d'obtenir des lactames par fermeture de cycle du substrat possédant déjà une liaison amide, par catalyse de réactions entraînant la formation de liaisons C-N ou C-C.

### I.2.2.1. SYNTHÈSE BIOCATALYTIQUE DE LACTAMES PAR FERMETURE DE CYCLE PAR LIAISON C-N ou C-C

Dans cette partie, nous allons détailler succinctement la synthèse de lactames avec des substrats possédant déjà une liaison amide, catalysées par des oxydoréductases (EC 1). Les oxydoréductases catalysent les réactions d'oxydoréduction et facilitent le transfert d'électrons d'une molécule à une autre.

Elles fonctionnent en présence d'un cofacteur ou coenzyme d'oxydoréduction comme le NAD(P)H. Ces enzymes sont impliquées dans de nombreux processus biologiques comme la photosynthèse, la respiration cellulaire ou pour la détoxification de radicaux libres. Pour la synthèse biocatalytique de lactames, les oxydoréductases vont principalement catalyser la formation intramoléculaire de liaison C-N ou C-H pour fermer le cycle du substrat.

### I.2.2.1.1. Galactose oxydases et alcool déshydrogénases

Comme leur nom l'indique, les galactoses oxydases (GOases, EC 1.1.3.9) sont impliquées dans l'oxydation du galactose chez les champignons en présence d'O<sub>2</sub> et elles nécessitent du Cu<sup>2+</sup> comme cofacteur. Dans le cadre de la synthèse biocatalytique de lactames, des mutants de GOases ont été employés pour leur activités sur des alcools primaires et secondaires, mais aussi sur des aminoalcools comparée à l'enzyme native qui est inactive sur ce type de substrat (Herter, McKenna et al. 2015). Une cascade a été développée pour la synthèse de lactames en partant donc d'aminoalcools qui sont oxydés par les GOase en leur aminoaldéhyde correspondant qui est ensuite cyclisé par une xanthine déshydrogénase. D'autre part, les aminoaldéhydes peuvent se tautomériser spontanément en imine. Pour la tautomérisation, il a été observé que cette réaction est dépendante du pH, dont l'optimum est situé à pH 8,5. L'imine formée est ensuite oxydée par une aldéhyde oxydase périplasmique de *E. coli* (PaoABC). Ainsi, la synthèse de  $\gamma$ -butyrolactames et de  $\delta$ -valérolactames a pu être réalisée en partant d'aminoaldéhydes ou de leur forme imine (Schéma 34).

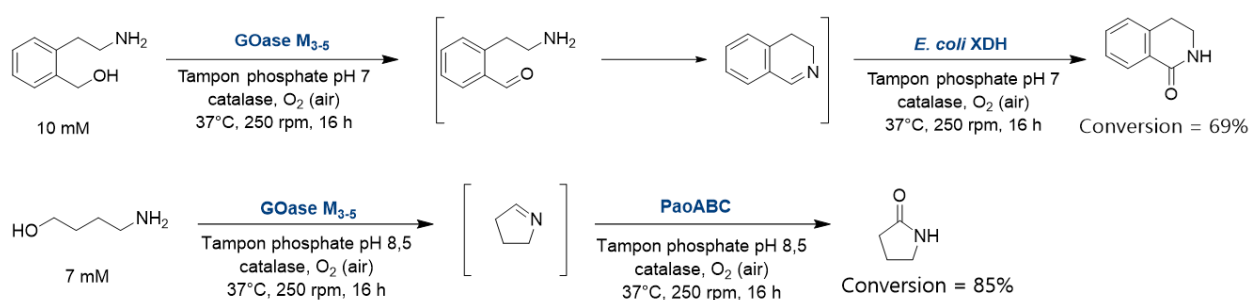


Schéma 33. Synthèse biocatalytique de lactames par une GOase en cascade avec une xanthine déshydrogénase (XDH) ou une aldéhyde oxydase (PaoABC).

La synthèse biocatalytique de lactames à partir d'aminoalcools a également été effectuée en utilisant des alcools déshydrogénases (ADH, EC 1.1.1.1) sur une réaction modèle en partant du 4-amino-1-butanol pour obtenir du  $\gamma$ -butyrolactame (Huang, Sayoga et al. 2018). Dans cet exemple, l'ADH a catalysé non seulement l'oxydation de l'aminoalcool en aminoaldéhyde mais aussi la forme tautomère hémicétabolique pour donner le lactame. Comme il a été vu avec les GOases, la formation de la forme hémicétabolique peut être favorisée en paramétrant le pH dont l'optimum se trouve à pH 8. De plus, cette réaction est couplée à un système de régénération enzymatique du NAD<sup>+</sup> par une NADH oxydase de *Streptococcus* mutants (*SmNOX*). Le  $\gamma$ -butyrolactame a pu être donc synthétisé avec un rendement analytique de 95% et la réaction a également été testée pour la synthèse de  $\delta$ -valérolactame et  $\epsilon$ -caprolactame, menant à des rendements plus faibles, de 38% et 14% respectivement (Schéma 34).

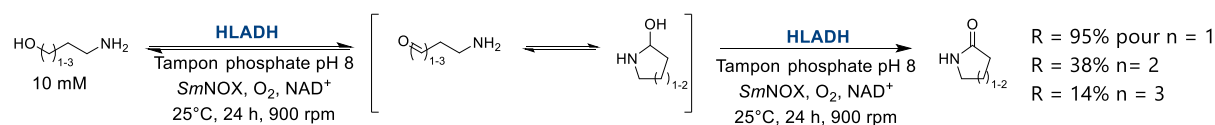


Schéma 34. Synthèse biocatalytique de lactames par une ADH de foie de cheval (HLADH pour *Horse Liver ADH* en anglais).

### I.2.2.1.2. Laccases et peroxydases

Les laccases (EC 1.10.3.2) sont des enzymes que l'on retrouve chez les plantes, champignons et bactéries et qui oxydent des composés phénoliques en utilisant le cuivre comme cofacteur. Elles sont impliquées dans la formation de la lignine mais aussi dans sa dégradation. Quant aux peroxydases, elles sont présentes chez tous les êtres vivants et elles catalysent l'oxydation de substrats en utilisant du peroxyde d'hydrogène  $H_2O_2$  comme accepteur d'électrons. Elles sont impliquées dans des processus de détoxification de radicaux libres ou dans des mécanismes de défense contre des agents pathogènes.

L'équipe de Jan Deska a travaillé sur la synthèse d'oxazolidinone (dérivé du  $\gamma$ -butyrolactame) et ses dérivés en employant la peroxydase de racine de Raifort, plus connue sous son nom anglophone Horseradish peroxidase (HRP), une enzyme couramment utilisée dans des applications biotechnologiques (Jäger, Gregori et al. 2023). Pour cela, une cascade enzymatique en aérobie a été mise en œuvre en partant de *N*-hydroxycarbamate. En présence de dioxygène, l'enzyme glucose oxydase (GOx) oxyde le D-glucose en gluconate et permet de régénérer *in situ* du  $H_2O_2$ . Ce dernier permet l'oxydation du *N*-hydroxycarbamate par la HRP, générant l'intermédiaire acyl-nitroso très réactif qui se cyclise spontanément par réaction type nitroso-ène intramolécule (Schéma 35).

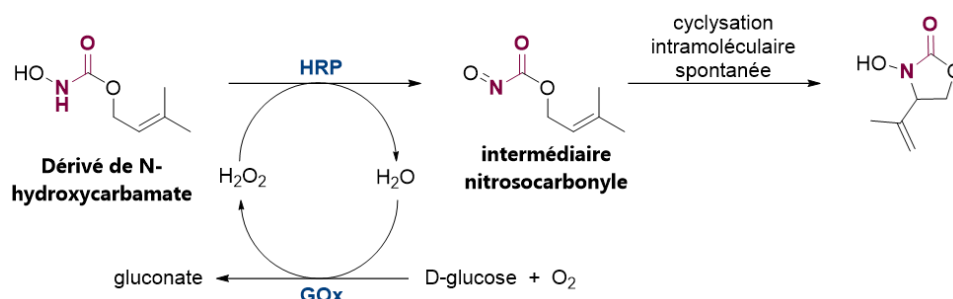


Schéma 35. Mécanisme général de formation de l'oxazolidinone par la peroxydase de racine de Raifort (HRP) couplée à un système de régénération enzymatique du peroxyde d'hydrogène par une glucose oxydase (GOx).

Grâce à ce système, l'oxazolidinone a été formée avec un rendement de 97% en 2 heures et à température ambiante. Par ailleurs, ce système permet un recyclage continu et *in situ* du  $H_2O_2$ . Cette même équipe a trouvé des laccases bactériennes et fongiques également capables de catalyser ces réactions, juste avec du  $O_2$  comme oxydant (Jäger, Haase et al. 2023). Les laccases et le système GOx/HRP ont été criblés sur différents dérivés de *N*-hydroxycarbamate. Le meilleur rendement avec les laccases a été obtenu avec la laccase de CotA de *Bacillus licheniformis* avec 83% de rendement en 26 heures d'incubation à 5°C sur le même substrat (Schéma 36). Ce système prometteur évite l'utilisation d'une deuxième enzyme tout en ayant des rendements très satisfaisants.

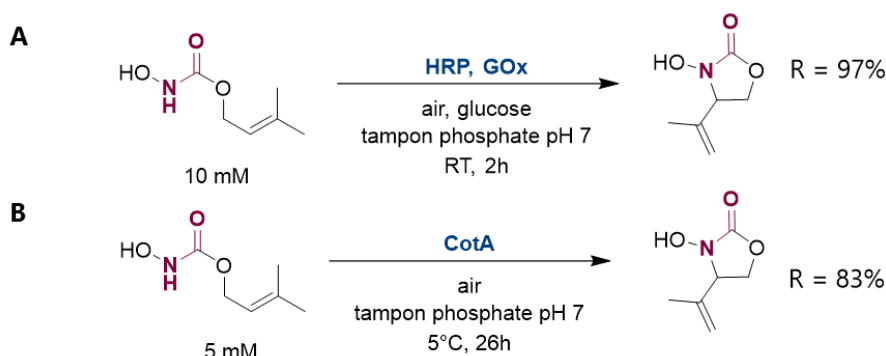


Schéma 36. Synthèse d'oxazolidinone par la peroxydase de racine de Raifort (HRP) couplée à un système de régénération du  $H_2O_2$  par une glucose oxydase (GOx), ou par la laccase CotA.

Enfin, le système HRP/GOx a été combiné à la lipase CALB pour la formation de dérivés de  $\gamma$ -butyrolactame en partant d'esters insaturés et d'hydroxylamine (Jäger, Nieger et al. 2023). Cette cascade multi-enzymatique a été réalisée en séquentiel, avec CALB qui active l'ester en présence de l'hydroxylamine dans du dioxane et à 50°C pour produire le dérivé N-hydroxycarbamate dans une première étape. Ce dérivé est ensuite utilisé par le système HRP/GOx pour synthétiser le dérivé  $\gamma$ -lactame comme vu précédemment. Avec cette cascade multi-enzymatique, des dérivés  $\gamma$ -lactames ont été synthétisés avec des rendements atteignant 98% avec une étape intermédiaire de purification du N-hydroxycarbamate (Schéma 37).

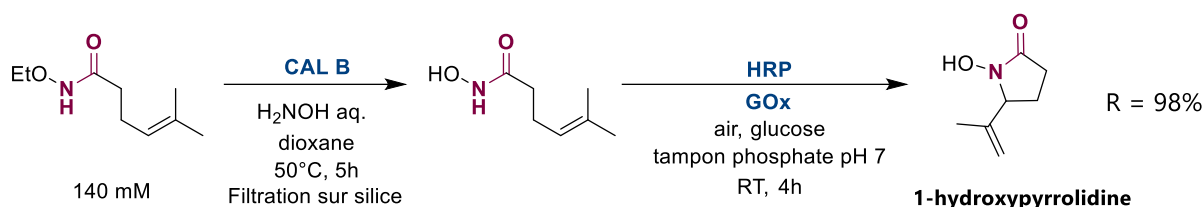


Schéma 37. Cascade enzymatique pour la synthèse de 1-hydroxypyrrolidine par la lipase CAL B et la peroxydase HRP.

Les laccases et les peroxydases représentent donc une classe d'intérêt pour la synthèse de lactames avec des réactions biocatalytiques réalisées en conditions douces. De hauts rendements sont atteints et ces enzymes peuvent être utilisées pour faire de synthèse de dérivés de lactames.

### I.2.2.1.3. Synthèse biocatalytique de lactames avec des réactions « new-to-nature »

Grâce à l'évolution dirigée et l'ingénierie protéique, il est possible de créer une fonctionnalité catalytique non naturelle à une enzyme, concept appelé « new-to-nature » (Renata, Wang et al. 2015). Dans ce cas, c'est vraiment le type de réaction qui est créé ou détourné et non juste la fonction native sur des substrats autres que le substrat métabolique. En partant d'une réaction désirée, une nouvelle enzyme peut être conçue en intégrant un site actif catalytique dans un échafaudage de protéine (synthèse *de novo*).

#### I.2.2.1.3.1. Ène-réductases

Les ène-réductases (EREDs) appartiennent à la famille des Old Yellow Enzymes (OYEs) et catalysent la réduction de double liaisons C=C. Ces enzymes sont dépendantes de la flavine mononucléotide (FMN) et du NAD(P). Leur mécanisme consiste en des transferts d'électrons initiés par la réduction du NAD(P) en NAD(P)H qui transfère un hydrure au FMN qui se réduit à son tour en FMNH<sub>2</sub>. Cela génère un nouveau transfert d'hydrure stéréosélectif sur le substrat afin de réduire la liaison C=C (Schéma 38) (Kumar Roy, Sreedharan et al. 2022). Pour avoir un recyclage du co-facteur NAD(P)H qui est coûteux, cette réaction est couplée soit à glucose déshydrogénase, une formate déshydrogénase ou une phosphite déshydrogénase.



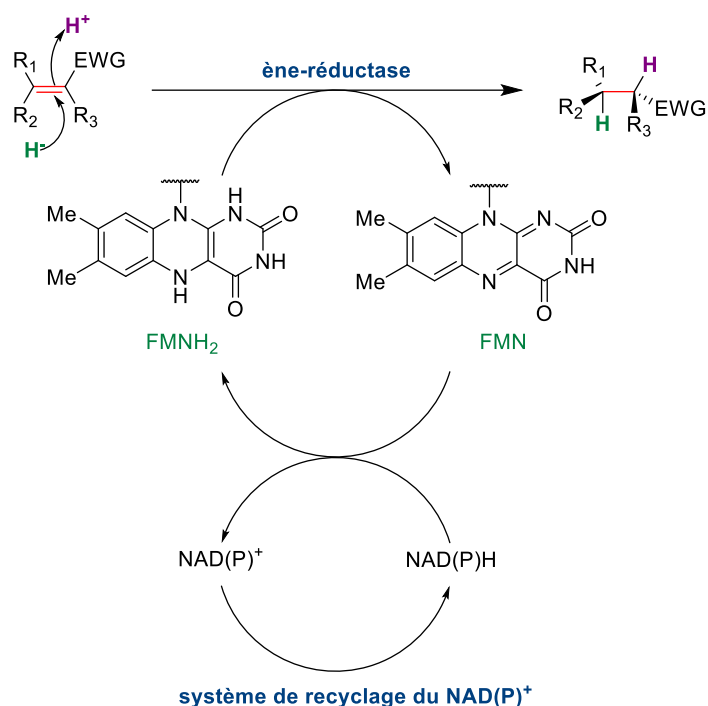


Schéma 38. Mécanisme réactionnel de réduction des alcènes par des EREDs.

Des EREDs modifiées ont été utilisées en photobiocatalyse pour une réaction différente de formation de lactames, à partir de substrats comportant déjà la liaison amide. Les mutations effectuées avaient pour but d'améliorer les rendements et la spécificité de substrats pour accommoder de plus larges substrats. Ces EREDs catalysent des cyclisations radicalaires intramoléculaires à partir de dérivés  $\alpha$ -chloroamides qui entraînent la fermeture du cycle par la formation d'une liaison C-H, sous effet de la lumière (Emmanuel, Bender et al. 2023). Le mécanisme est basé sur un clivage mésolytique au niveau du chlore en  $\alpha$ , générant un intermédiaire radicalaire qui réagit sur la double liaison C=C (Schéma 39) (Biegasiewicz, Cooper et al. 2019).

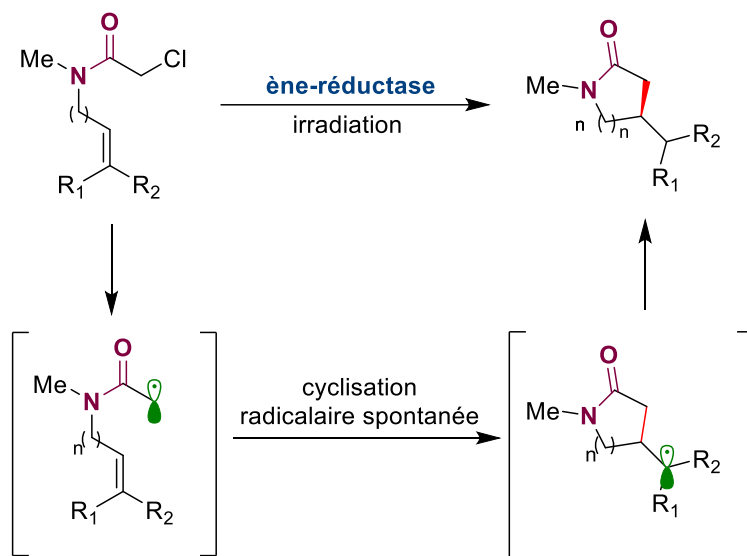


Schéma 39. Réaction "new-to-nature" pour la synthèse de lactames à partir de de  $\alpha$ -chloroamides par les EREDs.

Le mutant T36A de la ERED de *Gluconobacter oxydans* (GluER) a été appliqué pour la synthèse de différents dérivés de  $\gamma$ - et  $\epsilon$ -lactames sous une irradiation de LEDs cyans à 497 nm (Biegasiewicz,

Cooper et al. 2019), notamment avec des groupements encombrants comme la O-benzyloxime (OBn) (Gao, Turek-Herman et al. 2021). Divers dérivés *N*-méthylbutyrolactame, et dérivés de 2-azocanone, lactame à 8 chaînons, ont été obtenus avec des EREDs modifiées avec 4 mutations ciblées, avec des rendements allant jusqu'à 83% en 18h de réaction sous irradiation (Schéma 40). L'étude de l'énantiosélectivité de GluER et ses mutants a été réalisée (Capone, Dell'Orletta et al. 2023). On peut aussi noter que les ène-réductases ont été employées pour de la cyclisation intramoléculaire en formant une liaison C-N par une réaction d'hydroamination radicalaire (Ye, Cao et al. 2023).

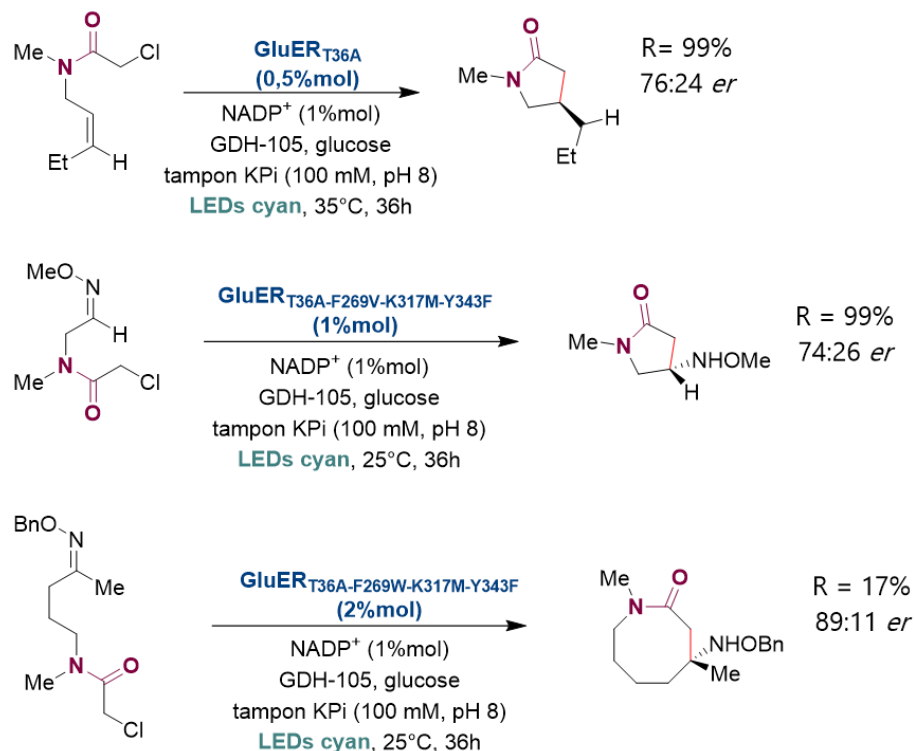


Schéma 40. Synthèse de  $\gamma$ -butyrolactames et 2-azocanones avec les mutants de ERED de *Gluconobacter oxydans* (GluER).

Les ène-réductases sont donc une classe d'enzyme d'intérêt pour la synthèse de lactames, et plus particulièrement de dérivés lactames avec différents types de substituants, simples comme encombrants avec des cycles aromatiques par exemple. Les réactions photobiocatalytiques se déroulent en conditions douces avec un recyclage du cofacteur mais sont pour le moment limitées au substrat  $\alpha$ -chloroamides ou ester hydroxamique avec la présence d'un groupement alcène.

#### I.2.2.1.3.2. Metalloenzymes naturelles et artificielles

Les metalloenzymes sont des enzymes qui possèdent un ou plusieurs ions métalliques dans leur site actif. Ces ions sont liés dans le site actif grâce à des interactions ioniques avec les chaînes latérales des résidus ou par coordination avec un ligand présent dans le site actif. Les metalloenzymes naturelles se retrouvent dans divers processus biologiques comme la photosynthèse ou la chaîne respiratoire. Une majorité des metalloenzymes appartient à la famille des oxydoréductases. Des metalloenzymes artificielles ont aussi été construites en introduisant un centre actif métallique ou cofacteur synthétique au sein d'un échafaudage d'une protéine ou en modifiant une metalloprotéine. L'avantage de ces constructions artificielles est de combiner l'efficacité de la catalyse métallique « synthétique » apportée par le catalyseur métallique, avec l'efficacité de la catalyse enzymatique grâce à la présence de la structure protéique. Cela permet de concevoir des catalyseurs plus efficaces et sélectifs pour une large gamme de réactions chimiques.

### I.2.2.1.3.2.1. Cytochromes P450

Les cytochromes P450 (CYP450) sont des hémoprotéines que l'on retrouve chez presque tous les êtres vivants (bactéries, champignons, humains). Ils sont impliqués dans le métabolisme et biotransformations des médicaments, des hormones, de toxines et d'autres métabolites secondaires, où ils catalysent des réactions d'oxydoréduction. Beaucoup de travaux d'ingénierie protéique ont été entrepris avec les CYP450 pour la biocatalyse de réactions non naturelles.

L'équipe de Yang Yang a introduit une activité « new-to-nature » de cyclisation radicalaire par transfert d'atome stéréosélective (ATRC de l'anglais *atom-transfer radical cyclization*), au sein d'un mutant du CYP450 de *Bacillus megaterium* obtenu après plusieurs cycles itératifs de mutagenèse saturante (Zhou, Chin et al. 2021). Ainsi, le meilleur mutant à activité ATRC, nommé P450<sub>ATRC<sub>Case3</sub></sub>, a été employé pour la synthèse biocatalytique d'un *N*-benzyl lactame en partant d'un dérivé d'amide portant un atome de brome en position β (Schéma 41). Cette réaction, réalisée avec des cellules entières de *E. coli* surexprimant cette enzyme, a mené à la formation du lactame fonctionnalisé avec un groupement halogénoalcane avec un rendement de 97% et un très bon excès énantiomérique. Ce lactame est une plateforme pour introduire d'autres groupements à la place du brome tout en gardant l'énantiosélectivité obtenue précédemment.

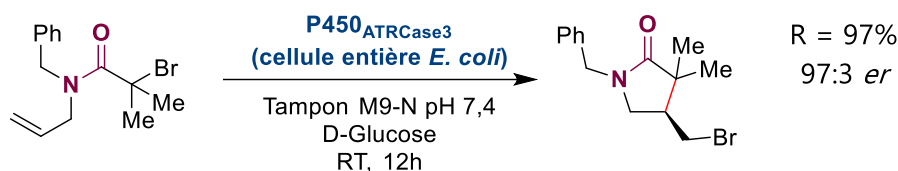


Schéma 41. Synthèse de dérivé de  $\gamma$ -butyrolactame par un CYP450 avec une activité « new-to-nature » de cyclisation radicalaire par transfert d'atomes (P450<sub>ATRC<sub>Case3</sub></sub>).

Cette même équipe a utilisé l'évolution dirigée sur des mutants de CYP450 pour la formation de lactames accolés à un cycle aromatique grâce à une activité de cyclase de radicaux aromatiques (P450<sub>arc</sub>) en partant d'un  $\alpha$ -bromo- $\beta$ -amidoester ou d'un  $\alpha$ -chloro- $\beta$ -amidoester lié un cycle aromatique via l'atome d'azote de l'amide (Schéma 42) (Fu, Neris et al. 2023).

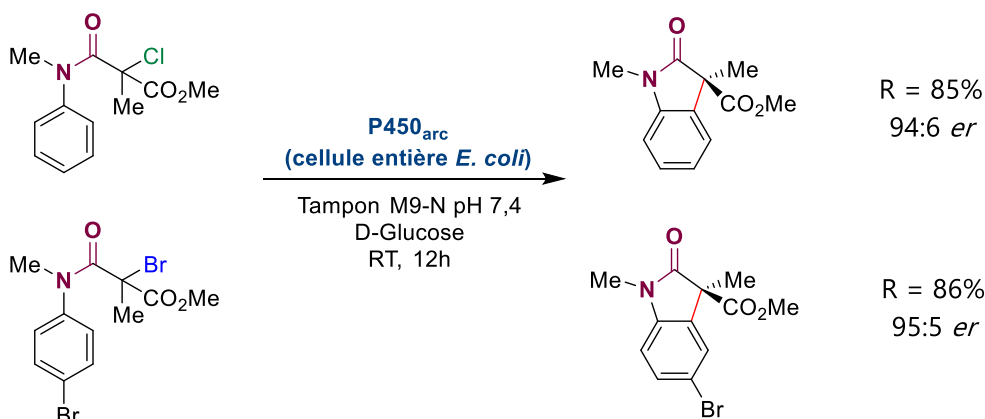


Schéma 42. Synthèse de dérivés de  $\gamma$ -butyrolactames avec une CYP450 par une activité "new-to-nature" de cyclase de radicaux aromatiques (P450<sub>arc</sub>).

### I.2.2.1.3.2.2. Métalloenzymes artificielles (ArMs)

La formation de lactames avec des ArMs repose sur l'utilisation du complexe biotin-streptavidine employé pour diverses applications biotechnologiques et chimiques avec des métaux de transition (Wilson and Whitesides 1978, Heinisch and Ward 2016). Le complexe de rhodénium [Cp\*RhX<sub>2</sub>]<sub>2</sub> introduit

dans le site actif, est inspiré du complexe organométallique  $[\text{Cp}^*\text{RhCl}_2]_2$ , un catalyseur couramment utilisé pour les réactions d'activation d'aromatiques électrophile (Hyster, Knörr et al. 2012). Ce complexe est coordonné dans la structure tétramérique de streptavidine (Sav) qui contient un site de liaison à la biotine. La synthèse de biocatalytique de dérivés de dihydroisoquinolone est réalisée à partir d'un acide benzhydroxamique et d'acrylate de méthyle avec de hauts rendements allant jusqu'à 95% en 72h de réaction en solution (Schéma 43).

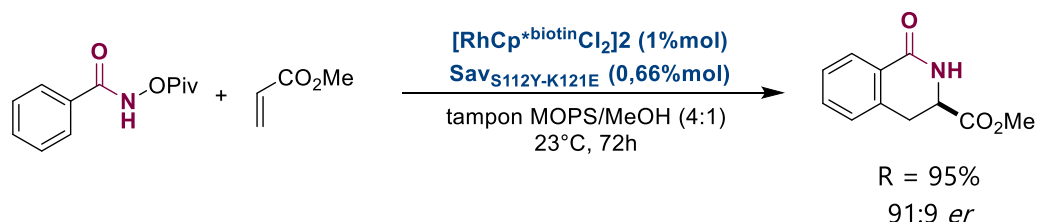


Schéma 43. Synthèse d'un dérivé de  $\delta$ -valérolactame par une ArM à base complexe de rhodium dans la structure biotine-streptavidine.

Toujours avec l'approche biotine-streptavidine, nous pouvons citer d'autres exemples, comme une ArM utilisant qu'un seul monomère de streptavidine (mSav) (Hassan, Ta et al. 2019). Cette nouvelle AM catalyse le couplage énantiosélectif entre différents esters d'acrylamide hydroxamates et des dérivés de styrène, menant à la synthèse de dérivés de lactames pouvant aller jusqu'à 99% de rendement. D'autres complexes organométalliques avec d'autres métaux de transition ont été employés comme Cu(I) hétéroscoopionate qui permet la synthèse de  $\beta$ - $\gamma$ -lactames et de spirolactames à partir de diazoacétamides après amélioration par évolution dirigée (Rumo, Stein et al. 2022). Dans l'exemple le plus récent, un complexe d'iridium  $[\text{Cp}^*\text{Ir}(\text{Boc-AQ-biot})\text{Cl}]$  avec une activité de nitrène insertase artificielle est employé pour la synthèse de  $\gamma$ -lactames à partir de dérivés de dioxazolone (Schéma 44) (Yu, Zou et al. 2023). Ce type de substrat présente l'avantage d'être facilement synthétisable et plus stable que des azotures qui peuvent être employés pour faire des transferts de nitrène également.

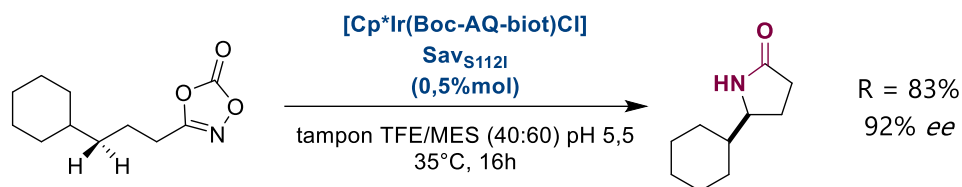


Schéma 44. Synthèse d'un dérivé de  $\gamma$ -butyrolactame à partir de dioxazolone par une ArM à base de complexe d'iridium dans la structure biotine-streptavidine.

L'équipe de Rudy Fasan a également produit des ArMs en utilisant la myoglobine (Mb) comme échafaudage, par mutation de résidus ligands de l'hème, pour la synthèse de  $\gamma$ -lactames fusionnés avec un cyclopropane à partir de diazoacétamide (Ren, Chandgude et al. 2020). Suite à des travaux d'ingénierie de la myoglobine modifiant la forme de la poche de l'hème pour améliorer son activité et énantiosélectivité, un quadruple mutant de Mb a été utilisé pour la synthèse de ces lactames fusionnés (Schéma 45A). La myoglobine modifiée a également permis la synthèse de  $\beta$ - $\gamma$ - et  $\delta$ -lactames par une amination C-H intramolécule en partant de dioxazolones, des précurseurs de nitrènes (Roy, Vargas et al. 2023). Le mutant de Mb nommé Mb\*, permet d'accéder à un panel de  $\beta$ - $\gamma$ - et  $\delta$ -lactames en modifiant la longueur de la chaîne ou ajoutant des substituants sur la partie arylalkyle du dérivé dioxazolone. Tous les lactames sont obtenus en majorité avec 99% d'*ee* en 3h de réaction à température ambiante et en anaérobiose en utilisant Mb\* à seulement 0,3% molaire (Schéma 45B).

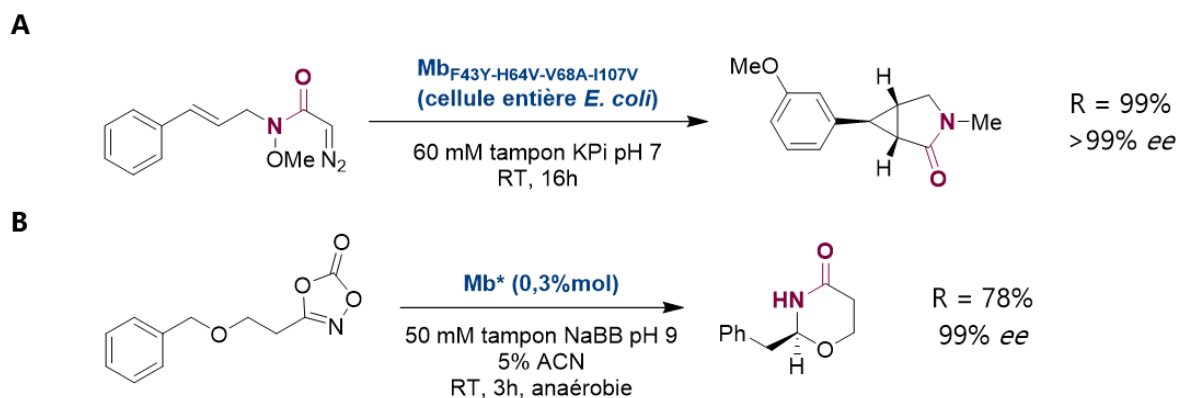


Schéma 45. Synthèse d'un spirolactame (A) et d'un dérivé de  $\delta$ -valérolactame (B) avec une myoglobine mutée pour catalyser l'amidation.

Les enzymes décrites dans cette partie permettent d'accéder à différents lactames par introduction d'activités non naturelles avec des complexes organométalliques ou par ingénierie protéique. A partir de substrats synthétiques faciles d'accès, ces métalloenzymes catalysent la synthèse stéréosélective de lactames, dans des conditions réactionnelles douces (température ambiante, pH neutre ou basique).

### I.2.2.2. SYNTHÈSE BIOCATALYTIQUE DE LACTAMES A PARTIR D'ACIDES $\omega$ -AMINÉS

De façon très générale, la synthèse de lactames à partir d'acides  $\omega$ -aminés nécessitent l'activation de l'acide carboxylique avant l'attaque nucléophile intramoléculaire de l'amine. Cette activation peut être soit sous forme d'ester ou d'amide, soit par l'intermédiaire acyl-adénylate comme nous l'avons pu le voir précédemment pour la synthèse d'amide.

#### I.2.2.2.1. Cyclisation spontanée d'esters d'acides $\omega$ -aminés

##### I.2.2.2.1.1. Transaminases

Les transaminases (TA) ou aminotransférases (ATA, EC 2.6.1-) catalysent le transfert d'un groupement amine d'un acide aminé (donneur d'amine) sur un acide  $\alpha$ -cétonique (accepteur), menant à la formation d'acides aminés chiraux. La transamidation consiste en 2 étapes simultanées : la déamination de l'acide aminé libérant le donneur d'amine, suivi de l'amination de l'acide  $\alpha$ -cétonique et nécessite du pyridoxal 5'-phosphate (PLP) comme cofacteur, qui est un dérivé de la vitamine B<sub>6</sub> (Schéma 46) (Guo and Berglund 2017). Le PLP est ancré de manière covalente à la lysine catalytique dans le site actif des TA et va servir d'hôte transitoire pour le groupement amine avant qu'il ne soit transféré sur l'acide  $\alpha$ -cétonique. Cela révèle une autre propriété importante des TA qui favorisent leur application en biocatalyse : leur stéréosélectivité. Les sites de liaisons de chaque substrat sont spécifiques et de tailles différents, mais en plus leur mode de liaison est contraint par la présence du PLP au sein du site actif, ce qui amène à des produits chiraux avec des bons excès énantiomériques (Guo and Berglund 2017). Pour former des amides, les TAs sont donc généralement employées en cascade avec une autre enzyme en utilisant des amines chirales.

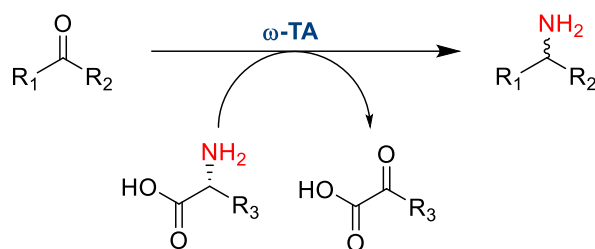


Schéma 46. Réaction naturellement catalysée par les ω-transaminases (TAs).

Les TAs sont utilisées pour catalyser la synthèse d'un ester ω-aminé en présence d'un donneur d'amine, puis cyclisation spontanée pour former le lactame correspondant. Les TAs ont été employées pour la première fois par dédoublement cinétique dynamique pour la synthèse de (*R*)-4-phénylpyrrolidin-2-one (Koszelewski, Clay et al. 2009). Le (*R*)-4-phénylpyrrolidin-2-one est obtenu en partant d'un mélange racémique de l'acide éthyle ester 4-oxo-3-phénylbutyrique et de la D-alanine comme donneur d'amine avec des ω-TAs de *Vibrio fluvialis* commerciales, menant à un rendement à 91% en 24 heures (Schéma 47).

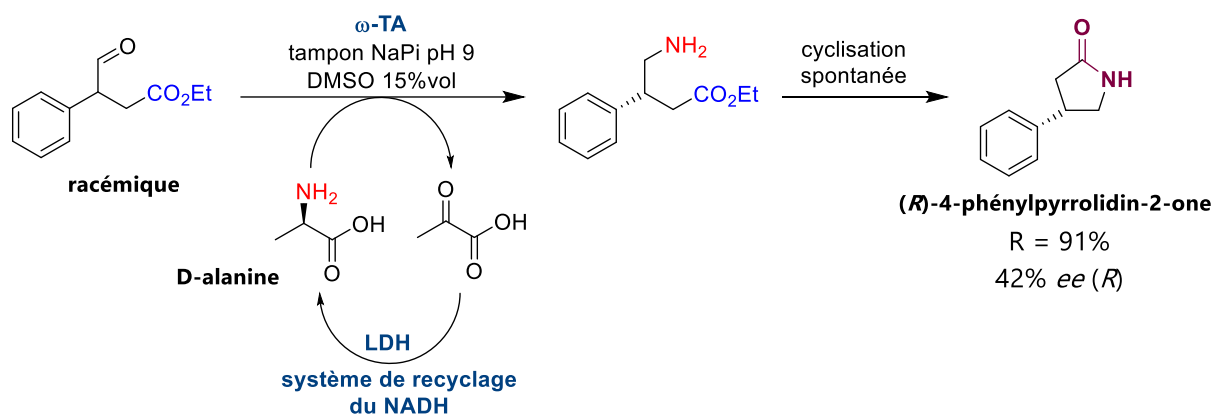


Schéma 47. Synthèse du (*R*)-4-phénylpyrrolidin-2-one par formation de l'acide ω-aminé par la ω-TA de *Vibrio fluvialis*.

Selon la même méthodologie, Mourelle et al. ont criblé des TAs commerciales avec différents cétoesters pour la synthèse de γ- et δ-lactames (Mourelle-Insua, Zampieri et al. 2018). Grâce à ce criblage, plusieurs TAs ont été utilisées pour la synthèse de lactames fonctionnalisés avec des groupements alkyles ou aryles avec de hautes conversions et des ee >99% aussi bien en faveur du (*R*) que du (*S*) (Schéma 48).

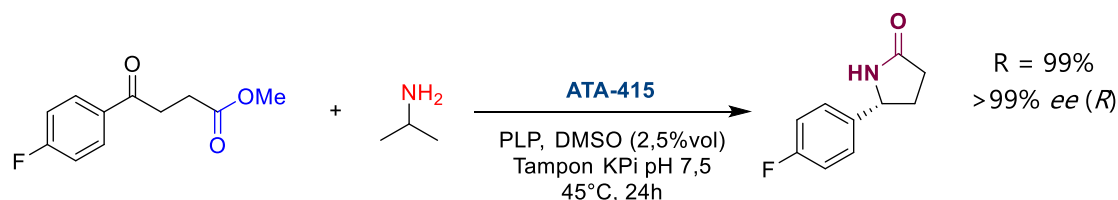


Schéma 48. Synthèse du (*R*)-4-(4-fluoro)phénylpyrrolidin-2-one par la transaminase ATA-115.

Plus récemment, de l'ingénierie protéique a été réalisée sur la ω-TA de *Bacillus megaterium* (ω-TAB<sub>m</sub>) pour agrandir la poche du site actif pour accéder à des lactames chiraux avec des donneurs d'amines substitués avec des groupements aromatiques (Jiang, Du et al. 2023). Cela permet donc d'obtenir un plus large panel de γ- et δ-lactames chiraux avec conversion totale du substrat et 99% d'ee avec les mutants de ω-TAB<sub>m</sub> (Schéma 49).

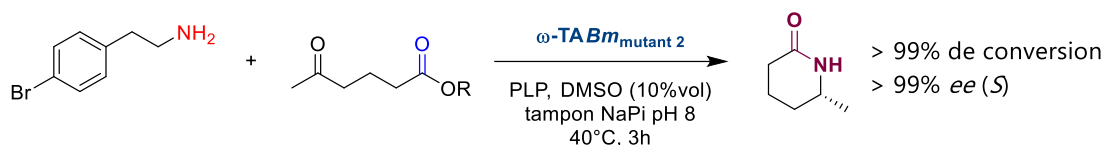


Schéma 49. Synthèse du 3-(S)-éthylpipéridin-2-one par  $\omega$ -TABm mutante.

Les TAs ont été aussi employées dans des cascades biocatalytiques pour de la bioconversion de substrats peu chers, issus de la biomasse et facilement valorisables comme des cycloalkylamines ou du cyclohexanol pour la synthèse de lactames (Sarak, Khobragade et al. 2021, Sarak, Sung et al. 2021). Les transaminases servent pour produire l'acide  $\omega$ -aminé mais sa cyclisation est catalysée par une cyclase (Schéma 50).

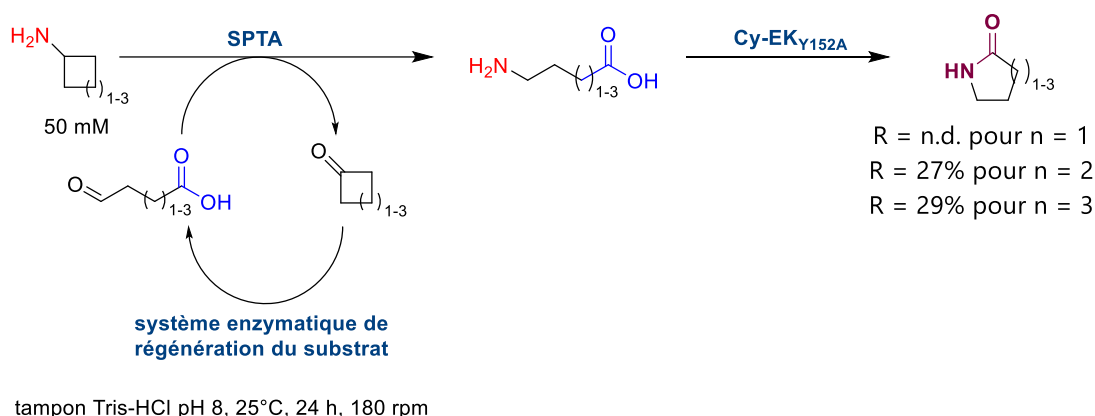


Schéma 50. Synthèse *in vivo* de lactames en partant de cycloalkylamines par une cascade enzymatique avec la transaminase de *Phaebacter porticola* (SPTA) et la cyclase mutée de *Enterobacter kobei* (Cy-EK<sub>Y152A</sub>). Légende : n.d. pour non déterminé.

Les transaminases constitue une classe d'enzymes adaptées pour la synthèse de lactames chiraux à partir d'esters avec de bons rendements et de forts excès énantiomériques dus à leur excellente stéréosélectivité.

#### I.2.2.2.1.2. Hydrolases

Comme pour la synthèse d'amides, les hydrolases (EC 3) ont été les premières enzymes utilisées pour la synthèse biocatalytique de lactames. La lactamisation a été illustrée pour la première fois avec la lipase pancréatique porcine (PPL) en partant d'aminoesters (Gutman, Meyer et al. 1992). La PPL est active sur différents aminoesters simples ou fonctionnalisés pour la formation de lactames de 5 à 7 chaînons, pouvant atteindre jusqu'à 80% de conversion (Schéma 51).

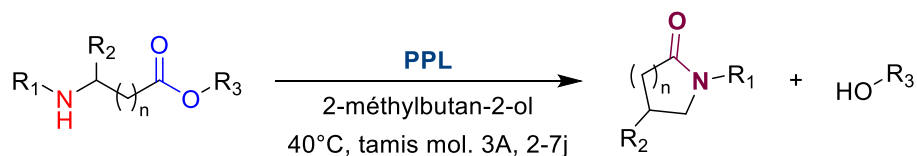


Schéma 51. Synthèses de lactames par la PPL.

La  $\omega$ -laurolactame hydrolase est la seule hydrolase trouvée pour la synthèse de lactames de taille moyenne (Ladkau, Hermann et al. 2011). Cette enzyme catalyse la cyclisation intramoléculaire de l'acide aminododécanoïque ou de l'acide activé sous forme d'aminoester (acide aminolaurique) atteignant 6,3% de rendement en lysat et 13% de rendement en cellule entière pour une réaction de bioconversion, au bout d'une heure et 30 minutes respectivement (Schéma 52).

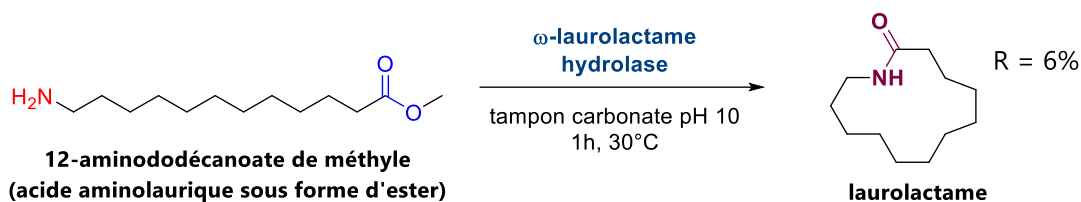


Schéma 52. Synthèse de lauro lactame par la ω-lauro lactame hydrolase.

La lipase CAL B a été également employée pour la cyclisation intramoléculaire d'acides ω-aminés pour la synthèse biocatalytique de lactames de 5 à 8 chaînons (Stavila and Loos 2013). CAL B utilisée sous forme immobilisée a permis d'atteindre jusqu'à 70% et 83% de conversion de l'acide 5-aminovalérique et de l'acide 6-aminohexanoïque au bout de 5 jours à 90°C (Schéma 53A). Cette réaction est tout de même réalisée sous atmosphère inerte et avec un système de séchage pour réduire la présence d'eau formée au cours de la réaction. Enfin, CALB a été exploitée pour la synthèse de fluorolactames qui sont des ingrédients actifs pharmaceutique (Willis, Fisher et al. 2016). La synthèse de fluorolactames est réalisée à partir d'un précurseur activé sous forme d'ester et obtenu avec un haut excès énantiomérique (98% *ee*), à pH neutre (pH = 7,3) et température ambiante (Schéma 53B).

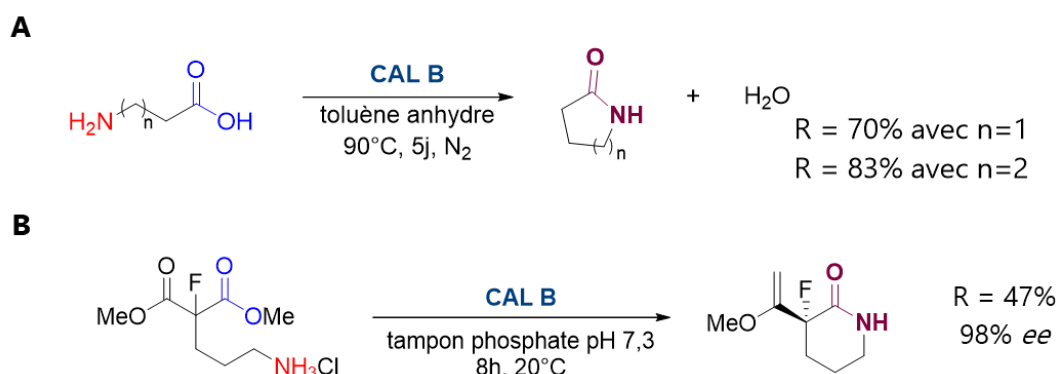


Schéma 53. Synthèse de lactames et d'un dérivé de fluorolactame avec CAL B.

Même si les hydrolases sont actives pour la synthèse de lactames, les réactions peuvent nécessiter un temps long pour atteindre de hauts rendements. De plus, ces synthèses sont limitées à des substrats hydrophobes et des acides activés sous forme d'esters.

#### I.2.2.2. Les enzymes dépendantes de l'ATP

Déjà identifiées pour la synthèse d'amides acycliques, les enzymes dépendantes de l'ATP ont été appliquées à la synthèse de lactames à partir d'acides ω-aminés, l'intermédiaire aminoacyl-AMP se cyclisant spontanément (Schéma 54).

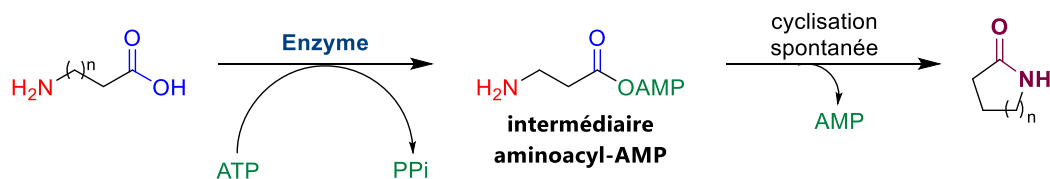


Schéma 54. Synthèse biocatalytique de lactames avec les enzymes dépendantes de l'ATP.

La synthèse de lactames avec les enzymes dépendantes de l'ATP se fait par réactions en cascade *in vitro* ou *in vivo* (Gordillo Sierra and Alper 2020).



La première enzyme à avoir été utilisée est ORF27, une acyl-CoA ligase de *Streptomyces aizunensis*, pour la synthèse de  $\gamma$ -butyrolactame dans *Escherichia coli* par l'équipe de Jay Keasling (Zhang, Kao et al. 2015). A partir du glutamate, en combinaison avec une glutamate décarboxylase qui permet de produire le GABA in situ, ORF27 a permis la synthèse  $\gamma$ -butyrolactame avec un rendement de 25% avec une conversion de 85,5% en 31h de réaction en partant d'une solution mère de glutamate à 9 g/L (Schéma 55A). L'équipe de Keasling a réutilisé cette même enzyme pour la formation de  $\delta$ - et  $\epsilon$ -lactames aussi en partant de leur précurseur acide  $\omega$ -aminé correspondant (Zhang, Barajas et al. 2017). En partant directement d'une solution de 2 g/L des précurseurs acides  $\omega$ -aminés, 152,7 mg/L et 2,02 mg/L de  $\delta$ -valérolactame et  $\epsilon$ -caprolactame ont été synthétisés respectivement en 72h. La formation de  $\delta$ -valérolactame est initiée à partir de la lysine transformée en son précurseur, l'acide 5-aminovalérique (5-AVA) par une lysine monooxygénase et une 5-aminovaléramide amidohydrolase (Schéma 55B). Cette voie de biosynthèse a permis d'obtenir 7,1 mmol soit 705 mg/L de  $\delta$ -valérolactame en partant de 68,4 mmol (10 g/L) de lysine totalement consommée au bout de 72 h, ce qui indiquerait que la cyclisation du 5-AVA en  $\delta$ -valérolactame avec la CoA ligase serait l'étape limitante. Pour pallier cette étape limitante, un système ascendant de régulation dynamique basé sur un biosenseur de  $\delta$ -valérolactame est utilisé pour surexprimer la CoA ligase pour convertir plus de 5-AVA en  $\delta$ -valérolactame (Zhao, Wu et al. 2023). Grâce à ce système, la production de valérolactame est passée à 11.88 g/L en 24h.

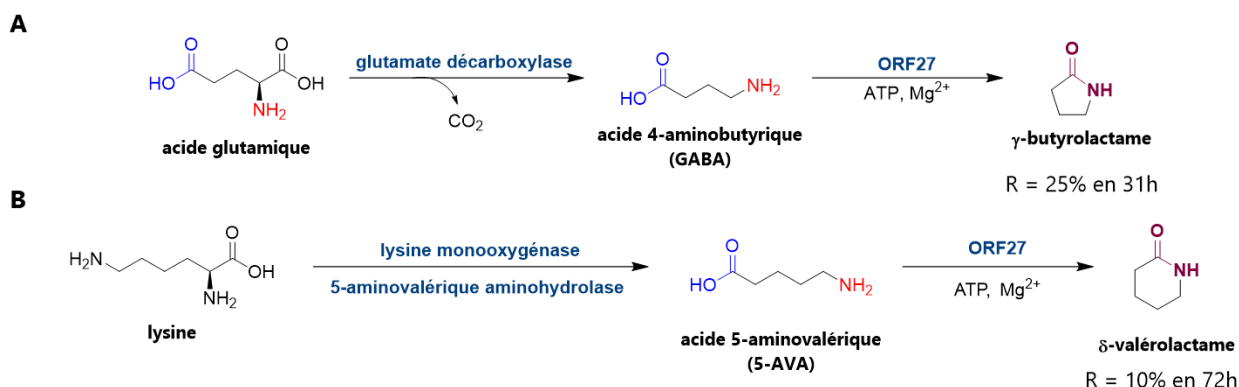


Schéma 55. Synthèse *in vivo* de  $\gamma$ -butyrolactame et de  $\delta$ -valérolactame avec la CoA ligase ORF27 en partant de l'acide glutamique (A) ou de la lysine (B) respectivement.

Très récemment, une CoA ligase de *Metallosphaera sedulla* (*msed*), de la même famille que celle utilisée au sein de l'équipe de biocatalyse pour la synthèse d'amides, a été utilisée pour la synthèse de lactames *in vivo* (Shen, Jiang et al. 2024). Cette CoA ligase est employée pour la biosynthèse de  $\gamma$ -butyrolactame en partant du glucose comme source de carbone. Le  $\gamma$ -butyrolactame est synthétisé à partir du glutamate qui se retrouve décarboxylé pour former le GABA qui sera pris en charge par la CoA ligase *msed* pour donner le lactame correspondant. L'utilisation de *msed* a été comparée avec d'autres CoA ligases qui sont employées pour la biosynthèse de lactames comme ORF26. *Msed* permet d'accéder à plus large production de  $\gamma$ -butyrolactame de 920,8 mg/L contre seulement 301,5 mg/L pour ORF26, en 84 h de fermentation (Schéma 56). Une montée en échelle dans un fermenteur de 3 litres a permis de produire jusqu'à 45,2 g/L de  $\gamma$ -butyrolactame en 84 h.

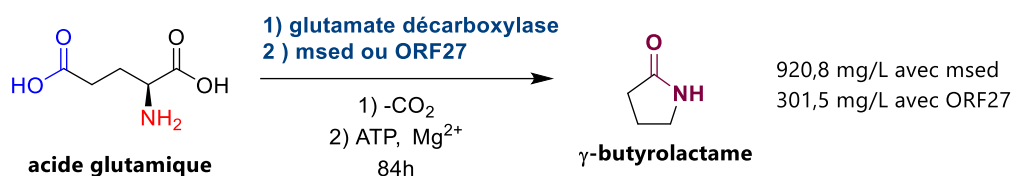


Schéma 56. Synthèse *in vivo* pour la synthèse de  $\gamma$ -butyrolactame en employant la CoA ligase msed ou ORF27 pour l'activation de l'acide  $\omega$ -aminé.

Basée sur les travaux de l'équipe de Flitsch principalement, les CARs ont été employés pour la première fois pour la synthèse de lactames simples et chiraux en prenant deux CARs modèles : *MmCAR* et *SrCAR* (Qin, Zhang et al. 2022). Dans le cadre des CARs, il a été démontré que la cyclisation spontanée pouvait avoir lieu suite à l'étape d'adénylation ou suite à la thiolation et que ces deux mécanismes étaient en compétition. La mutation d'un résidu impliqué dans l'étape de thioestéification a permis d'augmenter la synthèse de lactames. Des mutations ciblées ont permis d'obtenir des mutants avec des activités améliorées allant jusqu'à une conversion totale du substrat (Schéma 57). Des lactames chiraux ont été synthétisés par dédoublement cinétique mais ces résultats restent en dessous de ce qu'il a été possible d'obtenir avec des transaminases notamment.

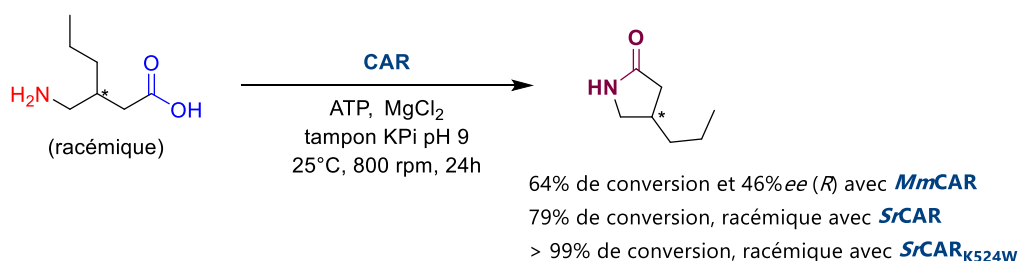


Schéma 57. Synthèse de lactames chiraux par dédoublement cinétique avec des CARs.

### 1.3. CONCLUSION SUR L'ETAT DE L'ART POUR LA SYNTHÈSE ENZYMATIQUE DE LIAISON AMIDE ET DE LACTAMES

Dans le vivant, la formation de liaison amide est majoritairement de la formation de liaison peptidiques, il n'y a peu d'enzymes catalysant la formation de liaison amide non peptidique acyclique ou cyclique. De plus, ces enzymes étant généralement spécifiques de leur substrat les applications en biocatalyse sur des substrats non métaboliques sont limitées.

En biocatalyse, si on excepte les réactions de type « new-to-nature », le plus souvent la formation de la liaison amide se fait par réaction entre un acide carboxylique ou un dérivé (ester, amide) et une amine. On rencontre deux grands types de mécanisme, par catalyse covalente avec la formation d'un acyl-enzyme, comme avec les hydrolases, et par formation d'un (amino)acyl-phosphate ou (amino)acyl-adénylate, avec les enzymes dépendantes de l'ATP, comme par exemple avec les CARs et les CoA ligases, ces dernières étant les enzymes utilisées dans nos travaux. La formation de la liaison amide se fait en deux étapes, et le plus souvent la seconde étape, l'addition nucléophile de l'amine, est spontanée.

S'il existe des exemples assez variés de synthèse biocatalytique d'amides acycliques, les exemples de synthèse biocatalytique de lactames sont limités. On peut citer les travaux de Keasling et de son équipe, qui par ingénierie métabolique de *E. coli* avec une cascade incluant une CoA ligase, ont obtenu la production de lactames de 5 à 6 chaînons. Plus récemment, Qin et al. ont obtenu des résultats très

intéressants avec des mutants de CARs pour la synthèse de lactames substitués par des groupes alkyles. Ce domaine d'études est encore récent et demande à être exploré, en particulier au niveau du spectre de substrats pour permettre d'étendre la synthèse biocatalytique à des lactames fonctionnalisés.

## RÉFÉRENCES BIBLIOGRAPHIQUES

Abe, T., Y. Hashimoto, H. Hosaka, K. Tomita-Yokotani and M. Kobayashi (2008). "Discovery of Amide (Peptide) Bond Synthetic Activity in Acyl-CoA Synthetase." Journal of Biological Chemistry **283**(17): 11312-11321.

Abe, T., Y. Hashimoto, S. Sugimoto, K. Kobayashi, T. Kumano and M. Kobayashi (2017). "Amide compound synthesis by adenylation domain of bacillibactin synthetase." The Journal of Antibiotics **70**(4): 435-442.

Arai, T., Y. Arimura, S. Ishikura and K. Kino (2013). "l-Amino Acid Ligase from *Pseudomonas syringae* Producing Tabtoxin Can Be Used for Enzymatic Synthesis of Various Functional Peptides." Applied and Environmental Microbiology **79**(16): 5023-5029.

Arnison, P. G., M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K.-D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H.-G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süßmuth, J. R. Tagg, G.-L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey and W. A. van der Donk (2013). "Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature." Natural Product Reports **30**(1): 108-160.

Baldessari, A. and C. P. Mangone (2001). "One-pot biocatalyzed preparation of substituted amides as intermediates of pharmaceuticals." Journal of Molecular Catalysis B: Enzymatic **11**(4): 335-341.

Basri, R. S., R. N. Z. R. A. Rahman, N. H. A. Kamarudin and M. S. M. Ali (2023). "Carboxylic acid reductases: Structure, catalytic requirements, and applications in biotechnology." International Journal of Biological Macromolecules: 124526.

Biegasiewicz, K. F., S. J. Cooper, X. Gao, D. G. Oblinsky, J. H. Kim, S. E. Garfinkle, L. A. Joyce, B. A. Sandoval, G. D. Scholes and T. K. Hyster (2019). "Photoexcitation of flavoenzymes enables a stereoselective radical cyclization." Science **364**(6446): 1166-1169.

Bilawal, A., M.-A. Gantumur, Y. Huang, A. Qayum, M. Ishfaq and Z. Jiang (2023). "Effect of transglutaminase cross-linking and emulsification on the stability and digestion of limonene emulsions." Process Biochemistry **131**: 210-216.

Bordusa, F. (2002). "Proteases in Organic Synthesis." Chemical Reviews **102**(12): 4817-4868.

Brieva, R., F. Rebolledo and V. Gotor (1990). "Enzymatic synthesis of amides with two chiral centres." Journal of the Chemical Society, Chemical Communications(20): 1386-1387.

Busto, E., V. Gotor-Fernández and V. Gotor (2010). "Hydrolases: catalytically promiscuous enzymes for non-conventional reactions in organic synthesis." Chemical Society Reviews **39**(11): 4504-4523.

Cannazza, P., S. Donzella, A. Pellis and M. L. Contente (2022). "Mycobacterium smegmatis acyltransferase: The big new player in biocatalysis." Biotechnology Advances **59**: 107985.

Capone, M., G. Dell'Orletta, B. T. Nicholls, G. D. Scholes, T. K. Hyster, M. Aschi and I. Daidone (2023). "Evidence of a Distinctive Enantioselective Binding Mode for the Photoinduced Radical Cyclization of  $\alpha$ -Chloroamides in Ene-Reductases." ACS Catalysis: 15310-15321.

Challis, G. L. and J. H. Naismith (2004). "Structural aspects of non-ribosomal peptide biosynthesis." Current Opinion in Structural Biology **14**(6): 748-756.

Chapman, N. C. and P. J. Rutledge (2021). "Isopenicillin N Synthase: Crystallographic Studies." ChemBioChem **22**(10): 1687-1705.

Chen, Q., C. Ji, Y. Song, H. Huang, J. Ma, X. Tian and J. Ju (2013). "Discovery of McbB, an Enzyme Catalyzing the  $\beta$ -Carboline Skeleton Construction in the Marinacarboline Biosynthetic Pathway." Angewandte Chemie International Edition **52**(38): 9980-9984.

Cheng, Z., Y. Xia and Z. Zhou (2020). "Recent Advances and Promises in Nitrile Hydratase: From Mechanism to Industrial Applications." Frontiers in Bioengineering and Biotechnology **8**.

Christodoulou, M. S., M. L. Contente, S. Dallavalle and A. Pinto (2022). "Enzymatic amide bond formation: synthesis of amino-oxo-acids through a Mycobacterium smegmatis acyltransferase." Green Chemistry **24**(11): 4432-4436.

Dorr, B. M. and D. E. Fuerst (2018). "Enzymatic amidation for industrial applications." Curr Opin Chem Biol **43**: 127-133.

Dunetz, J. R., J. Magano and G. A. Weisenburger (2016). "Large-Scale Applications of Amide Coupling Reagents for the Synthesis of Pharmaceuticals." Organic Process Research & Development **20**(2): 140-177.

Dunn, Z. D., W. J. Wever, N. J. Economou, A. A. Bowers and B. Li (2015). "Enzymatic Basis of "Hybridity" in Thiomarinol Biosynthesis." Angewandte Chemie International Edition **54**(17): 5137-5141.

El-Faham, A. and F. Albericio (2011). "Peptide Coupling Reagents, More than a Letter Soup." Chemical Reviews **111**(11): 6557-6602.

Emmanuel, M. A., S. G. Bender, C. Bilodeau, J. M. Carceller, J. S. DeHovitz, H. Fu, Y. Liu, B. T. Nicholls, Y. Ouyang, C. G. Page, T. Qiao, F. C. Raps, D. R. Sorigué, S.-Z. Sun, J. Turek-Herman, Y. Ye, A. Rivas-Souchet, J. Cao and T. K. Hyster (2023). "Photobiocatalytic Strategies for Organic Synthesis." Chemical Reviews.

Estrada, P., M. Manandhar, S.-H. Dong, J. Deveryshetty, V. Agarwal, J. E. Cronan and S. K. Nair (2017). "The pimeloyl-CoA synthetase BioW defines a new fold for adenylate-forming enzymes." Nature Chemical Biology **13**(6): 668-674.

Eustáquio, A. S., B. Gust, T. Luft, S.-M. Li, K. F. Chater and L. Heide (2003). "Clorobiocin Biosynthesis in Streptomyces: Identification of the Halogenase and Generation of Structural Analogs." Chemistry & Biology **10**(3): 279-288.

Fawaz, M. V., M. E. Topper and S. M. Firestine (2011). "The ATP-grasp enzymes." Bioorganic Chemistry **39**(5): 185-191.

Fu, W., N. M. Neris, Y. Fu, Y. Zhao, B. Krohn-Hansen, P. Liu and Y. Yang (2023). "Enzyme-controlled stereoselective radical cyclization to arenes enabled by metalloredox biocatalysis." Nature Catalysis.

Funabashi, M., Z. Yang, K. Nonaka, M. Hosobuchi, Y. Fujita, T. Shibata, X. Chi and S. G. Van Lanen (2010). "An ATP-independent strategy for amide bond formation in antibiotic biosynthesis." Nature Chemical Biology **6**(8): 581-586.

Gao, X., J. R. Turek-Herman, Y. J. Choi, R. D. Cohen and T. K. Hyster (2021). "Photoenzymatic Synthesis of  $\alpha$ -Tertiary Amines by Engineered Flavin-Dependent "Ene"-Reductases." Journal of the American Chemical Society **143**(47): 19643-19647.

García, F. F., D. Musikant, J. L. Escalona, M. M. Edreira and G. G. Liñares (2023). "Lipase-Catalyzed Synthesis and Biological Evaluation of N-Picolineamides as Trypanosoma cruzi Antiproliferative Agents." ACS Medicinal Chemistry Letters **14**(1): 59-65.

Gerratana, B., A. Stapon and C. A. Townsend (2003). "Inhibition and Alternate Substrate Studies on the Mechanism of Carbapenam Synthetase from Erwinia carotovora." Biochemistry **42**(25): 7836-7847.

Gordillo Sierra, A. R. and H. S. Alper (2020). "Progress in the metabolic engineering of bio-based lactams and their  $\omega$ -amino acids precursors." Biotechnology Advances **43**: 107587.

Goswami, A. and S. G. Van Lanen (2015). "Enzymatic strategies and biocatalysts for amide bond formation: tricks of the trade outside of the ribosome." Molecular BioSystems **11**(2): 338-353.

Gotor, V., R. Brieva, C. González and F. Rebolledo (1991). "Enzymatic aminolysis and transamidation reactions." Tetrahedron **47**(44): 9207-9214.

Grulich, M., V. Štěpánek and P. Kyslík (2013). "Perspectives and industrial potential of PGA selectivity and promiscuity." Biotechnology Advances **31**(8): 1458-1472.

Guo, F. and P. Berglund (2017). "Transaminase biocatalysis: optimization and application." Green Chemistry **19**(2): 333-360.

Gutman, A. L., E. Meyer, X. Yue and C. Abell (1992). "Enzymatic formation of lactams in organic solvents." Tetrahedron Letters **33**(27): 3943-3946.

Hassan, I. S., A. N. Ta, M. W. Danneman, N. Semakul, M. Burns, C. H. Basch, V. N. Dippon, B. R. McNaughton and T. Rovis (2019). "Asymmetric  $\delta$ -Lactam Synthesis with a Monomeric Streptavidin Artificial Metalloenzyme." Journal of the American Chemical Society **141**(12): 4815-4819.

Healy, V. L., L. S. Mullins, X. Li, S. E. Hall, F. M. Raushel and C. T. Walsh (2000). "d-Ala-d-X ligases: evaluation of d-alanyl phosphate intermediate by MIX, PIX and rapid quench studies." Chemistry & Biology **7**(7): 505-514.

Heinisch, T. and T. R. Ward (2016). "Artificial Metalloenzymes Based on the Biotin–Streptavidin Technology: Challenges and Opportunities." Accounts of Chemical Research **49**(9): 1711-1721.

Hepworth, L. J., S. P. France, S. Hussain, P. Both, N. J. Turner and S. L. Flitsch (2017). "Enzyme Cascades in Whole Cells for the Synthesis of Chiral Cyclic Amines." ACS Catalysis **7**(4): 2920-2925.

Herter, S., S. M. McKenna, A. R. Frazer, S. Leimkühler, A. J. Carnell and N. J. Turner (2015). "Galactose Oxidase Variants for the Oxidation of Amino Alcohols in Enzyme Cascade Synthesis." ChemCatChem **7**(15): 2313-2317.

Horvat, M., G. Fiume, S. Fritsche and M. Winkler (2019). "Discovery of carboxylic acid reductase (CAR) from *Thermothelomyces thermophila* and its evaluation for vanillin synthesis." Journal of Biotechnology **304**: 44-51.

Huang, L., G. V. Sayoga, F. Hollmann and S. Kara (2018). "Horse Liver Alcohol Dehydrogenase-Catalyzed Oxidative Lactamization of Amino Alcohols." ACS Catalysis: 8680-8684.

Hyster, T. K., L. Knörr, T. R. Ward and T. Rovis (2012). "Biotinylated Rh(III) Complexes in Engineered Streptavidin for Accelerated Asymmetric C–H Activation." Science **338**(6106): 500-503.

Irimescu, R. and K. Kato (2004). "Lipase-catalyzed enantioselective reaction of amines with carboxylic acids under reduced pressure in non-solvent system and in ionic liquids." Tetrahedron Letters **45**(3): 523-525.

Jäger, C., B. J. Gregori, J. A. S. Aho, M. Hallamaa and J. Deska (2023). "Peroxidase-induced C–N bond formation via nitroso ene and Diels–Alder reactions." Green Chemistry.

Jäger, C., M. Haase, K. Koschorreck, V. B. Urlacher and J. Deska (2023). "Aerobic C–N Bond Formation through Enzymatic Nitroso-Ene-Type Reactions\*\*." Angewandte Chemie International Edition **62**(7): e202213671.

Jäger, C., M. Nieger, K. Rissanen and J. Deska (2023). "Multienzymatic Synthesis of  $\gamma$ -Lactam Building Blocks from Unsaturated Esters and Hydroxylamine." European Journal of Organic Chemistry **26**(39): e202300288.

Jager, S. A. W., P. A. Jekel and D. B. Janssen (2007). "Hybrid penicillin acylases with improved properties for synthesis of  $\beta$ -lactam antibiotics." Enzyme and Microbial Technology **40**(5): 1335-1344.

Jeger, S., K. Zimmermann, A. Blanc, J. Grünberg, M. Honer, P. Hunziker, H. Struthers and R. Schibli (2010). "Site-Specific and Stoichiometric Modification of Antibodies by Bacterial Transglutaminase." Angewandte Chemie International Edition **49**(51): 9995-9997.

Jiang, N., X. Du and L. Zheng (2023). "Highly efficient synthesis of chiral lactams by using a  $\omega$ -transaminase from *Bacillus megaterium* and its mutant enzymes." Molecular Catalysis **547**: 113364.

Koszelewski, D., D. Clay, K. Faber and W. Kroutil (2009). "Synthesis of 4-phenylpyrrolidin-2-one via dynamic kinetic resolution catalyzed by  $\omega$ -transaminases." Journal of Molecular Catalysis B: Enzymatic **60**(3): 191-194.

Kua, G. K. B., G. K. T. Nguyen and Z. Li (2023). "Enzyme Engineering for High-Yielding Amide Formation: Lipase-Catalyzed Synthesis of N-Acyl Glycines in Aqueous Media." Angewandte Chemie International Edition **62**(14): e202217878.

Kumar Roy, T., R. Sreedharan, P. Ghosh, T. Gandhi and D. Maiti (2022). "Ene-Reductase: A Multifaceted Biocatalyst in Organic Synthesis." Chemistry – A European Journal **28**(21): e202103949.

Ladkau, N., I. Hermann, B. Bühler and A. Schmid (2011). "Enzyme-Catalyzed Laurolactam Synthesis via Intramolecular Amide Bond Formation in Aqueous Solution." Advanced Synthesis & Catalysis **353**(13): 2501-2510.

Land, H., P. Hendil-Forsell, M. Martinelle and P. Berglund (2016). "One-pot biocatalytic amine transaminase/acyl transferase cascade for aqueous formation of amides from aldehydes or ketones." Catalysis Science & Technology **6**(9): 2897-2900.

Lelièvre, C. M., M. Balandras, J.-L. Petit, C. Vergne-Vaxelaire and A. Zaparucha (2020). "ATP Regeneration System in Chemoenzymatic Amide Bond Formation with Thermophilic CoA Ligase." ChemCatChem **12**(4): 1184-1189.

Lima, R. N., C. S. dos Anjos, E. V. M. Orozco and A. L. M. Porto (2019). "Versatility of *Candida antarctica* lipase in the amide bond formation applied in organic synthesis and biotechnological processes." Molecular Catalysis **466**: 75-105.

Lubberink, M., W. Finnigan and S. L. Flitsch (2023). "Biocatalytic amide bond formation." Green Chemistry.

Lubberink, M., W. Finnigan, C. Schnepel, C. R. Baldwin, N. J. Turner and S. L. Flitsch (2022). "One-Step Biocatalytic Synthesis of Sustainable Surfactants by Selective Amide Bond Formation\*\*." Angewandte Chemie International Edition **61**(30): e202205054.

Lubberink, M., C. Schnepel, J. Citoler, S. R. Derrington, W. Finnigan, M. A. Hayes, N. J. Turner and S. L. Flitsch (2020). "Biocatalytic Monoacylation of Symmetrical Diamines and Its Application to the Synthesis of Pharmaceutically Relevant Amides." ACS Catalysis **10**(17): 10005-10009.

Marahiel, M. A. (2009). "Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis." Journal of Peptide Science **15**(12): 799-807.

Massolo, E., M. Pirola and M. Benaglia (2020). "Amide Bond Formation Strategies: Latest Advances on a Dateless Transformation." European Journal of Organic Chemistry **2020**(30): 4641-4651.

Miller, M. T., B. O. Bachmann, C. A. Townsend and A. C. Rosenzweig (2001). "Structure of  $\beta$ -lactam synthetase reveals how to synthesize antibiotics instead of asparagine." Nature Structural Biology **8**(8): 684-689.

Montalbetti, C. A. G. N. and V. Falque (2005). "Amide bond formation and peptide coupling." Tetrahedron **61**(46): 10827-10852.

Mori, T., K. Wanibuchi, H. Morita and I. Abe (2021). "Amide Bond Formation Using 4-Coumarate: CoA Ligase from *Arabidopsis thaliana*." Chemical and Pharmaceutical Bulletin **69**(8): 717-720.

Motoyama, T., T. Nogawa, T. Shimizu, M. Kawatani, T. Kashiwa, C.-S. Yun, D. Hashizume and H. Osada (2023). "Fungal NRPS-PKS Hybrid Enzymes Biosynthesize New  $\gamma$ -Lactam Compounds, Taslactams A-D, Analogous to Actinomycete Proteasome Inhibitors." ACS Chemical Biology **18**(2): 396-403.

Mourelle-Insua, Á., L. A. Zampieri, I. Lavandera and V. Gotor-Fernández (2018). "Conversion of  $\gamma$ - and  $\delta$ -Keto Esters into Optically Active Lactams. Transaminases in Cascade Processes." Advanced Synthesis & Catalysis **360**(4): 686-695.

Nissen, P., J. Hansen, N. Ban, P. B. Moore and T. A. Steitz (2000). "The Structural Basis of Ribosome Activity in Peptide Bond Synthesis." Science **289**(5481): 920-930.

O'Flynn, B. G., G. Suarez, A. J. Hawley and D. J. Merkler (2018). "Insect Arylalkylamine N-Acyltransferases: Mechanism and Role in Fatty Acid Amide Biosynthesis." Frontiers in Molecular Biosciences **5**.

Ogasawara, Y. and T. Dairi (2017). "Biosynthesis of Oligopeptides Using ATP-Grasp Enzymes." Chemistry – A European Journal **23**(45): 10714-10724.



Pacholec, M., C. L. Freel Meyers, M. Oberthür, D. Kahne and C. T. Walsh (2005). "Characterization of the Aminocoumarin Ligase SimL from the Simocyclinone Pathway and Tandem Incubation with NovM,P,N from the Novobiocin Pathway." Biochemistry **44**(12): 4949-4956.

Pan, X., A. Li, Z. Peng, X. Ji, J. Chu and B. He (2020). "Efficient synthesis of  $\beta$ -lactam antibiotics with in situ product removal by a newly isolated penicillin G acylase." Bioorganic Chemistry **99**: 103765.

Petchey, M., A. Cuetos, B. Rowlinson, S. Dannevald, A. Frese, P. W. Sutton, S. Lovelock, R. C. Lloyd, I. J. S. Fairlamb and G. Grogan (2018). "The Broad Aryl Acid Specificity of the Amide Bond Synthetase McbA Suggests Potential for the Biocatalytic Synthesis of Amides." Angewandte Chemie International Edition **57**(36): 11584-11588.

Petchey, M. R. and G. Grogan (2019). "Enzyme-Catalysed Synthesis of Secondary and Tertiary Amides." Advanced Synthesis & Catalysis **361**(17): 3895-3914.

Petchey, M. R., B. Rowlinson, R. C. Lloyd, I. J. S. Fairlamb and G. Grogan (2020). "Biocatalytic Synthesis of Moclobemide Using the Amide Bond Synthetase McbA Coupled with an ATP Recycling System." ACS Catalysis **10**(8): 4659-4663.

Philpott, H. K., P. J. Thomas, D. Tew, D. E. Fuerst and S. L. Lovelock (2018). "A versatile biosynthetic approach to amide bond formation." Green Chemistry **20**(15): 3426-3431.

Pintor, A., A. P. Matthey, I. Lavandera, V. Gotor-Fernández and A. Volkov (2023). "Development of an amine transaminase-lipase cascade for chiral amide synthesis under flow conditions." Green Chemistry **25**(15): 6041-6050.

Pitzer, J. and K. Steiner (2016). "Amides in Nature and Biocatalysis." Journal of Biotechnology **235**: 32-46.

Prasad, S. and T. C. Bhalla (2010). "Nitrile hydratases (NHases): At the interface of academia and industry." Biotechnology Advances **28**(6): 725-741.

Pratt, R. F. (2008). "Substrate specificity of bacterial DD-peptidases (penicillin-binding proteins)." Cellular and Molecular Life Sciences **65**(14): 2138-2155.

Qin, Z., X. Zhang, X. Sang, W. Zhang, G. Qu and Z. Sun (2022). "Carboxylic acid reductases enable intramolecular lactamization reactions." Green Synthesis and Catalysis.

Ramos, D., P. Rollin and W. Klaffke (2001). "Chemoenzymatic synthesis of neoglycopeptides: Application to an  $\alpha$ -gal-terminated neoglycopeptide." Journal of Organic Chemistry **66**(9): 2948-2956.

Ren, X., A. L. Chandgude and R. Fasan (2020). "Highly Stereoselective Synthesis of Fused Cyclopropane- $\gamma$ -Lactams via Biocatalytic Iron-Catalyzed Intramolecular Cyclopropanation." ACS Catalysis **10**(3): 2308-2313.

Renata, H., Z. J. Wang and F. H. Arnold (2015). "Expanding the Enzyme Universe: Accessing Non-Natural Reactions by Mechanism-Guided Directed Evolution." Angewandte Chemie International Edition **54**(11): 3351-3367.

Rodnina, M. V. (2013). "The ribosome as a versatile catalyst: reactions at the peptidyl transferase center." Current Opinion in Structural Biology **23**(4): 595-602.

Roy, S., D. A. Vargas, P. Ma, A. Sengupta, L. Zhu, K. N. Houk and R. Fasan (2023). "Stereoselective construction of  $\beta$ -,  $\gamma$ - and  $\delta$ -lactam rings via enzymatic C–H amidation." Nature Catalysis.

Rumo, C., A. Stein, J. Klehr, R. Tachibana, A. Prescimone, D. Häussinger and T. R. Ward (2022). "An Artificial Metalloenzyme Based on a Copper Heteroscorpionate Enables sp<sup>3</sup> C–H Functionalization via Intramolecular Carbene Insertion." Journal of the American Chemical Society **144**(26): 11676-11684.

Sabatini, M., L. T. Boulton, H. Sneddon and T. Sheppard (2019). A green chemistry perspective on catalytic amide bond formation.

Sarak, S., T. P. Khobragade, H. Jeon, A. D. Pagar, P. Giri, S. Lee and H. Yun (2021). "One-pot biocatalytic synthesis of nylon monomers from cyclohexanol using Escherichia coli-based concurrent cascade consortia." Green Chemistry.

Sarak, S., S. Sung, H. Jeon, M. D. Patil, T. P. Khobragade, A. D. Pagar, P. E. Dawson and H. Yun (2021). "An Integrated Cofactor/Co-Product Recycling Cascade for the Biosynthesis of Nylon Monomers from Cycloalkylamines." Angewandte Chemie International Edition **60**(7): 3481-3486.

Shen, X., X. Jiang, X. Sun, Q. Yuan and J. Wang (2024). "Identifying and charactering a 4-aminobutyryl-CoA ligase for the production of butyrolactam." Journal of Biotechnology **382**: 21-27.

Sheppard, M. J., A. M. Kunjapur and K. L. J. Prather (2016). "Modular and selective biosynthesis of gasoline-range alkanes." Metabolic Engineering **33**: 28-40.

Sim, E., K. Walters and S. Boukouvala (2008). "Arylamine N-acetyltransferases: From Structure to Function." Drug Metabolism Reviews **40**(3): 479-510.

Söderberg, E. S., K. Felicitas von Borries, U. Norinder, M. Petchey, S. Chavan, H. Holmquist, M. Johansson, I. Cotgreave, M. Hayes, P. Fantke and P.-O. Syrén (2023). "A safer and more sustainable by design perspective in biocatalytic amide-bond coupling of in silico controlled chemical space." ChemRxiv.

Stavila, E. and K. Loos (2013). "Synthesis of lactams using enzyme-catalyzed aminolysis." Tetrahedron Letters **54**(5): 370-372.

Steffensky, M., S.-M. Li and L. Heide (2000). "Cloning, Overexpression, and Purification of Novobiocin Acid Synthetase from *Streptomyces spheroides* NCIMB 11891 \*." Journal of Biological Chemistry **275**(28): 21754-21760.

Strieker, M., A. Tanović and M. A. Marahiel (2010). "Nonribosomal peptide synthetases: structures and dynamics." Current Opinion in Structural Biology **20**(2): 234-240.

Süssmuth, R. D. and A. Mainz (2017). "Nonribosomal Peptide Synthesis—Principles and Prospects." Angewandte Chemie International Edition **56**(14): 3770-3821.

Suzuki, H., K. Kato and H. Kumagai (2004). "Development of an efficient enzymatic production of  $\gamma$ -d-glutamyl-L-tryptophan (SCV-07), a prospective medicine for tuberculosis, with bacterial  $\gamma$ -glutamyltranspeptidase." Journal of Biotechnology **111**(3): 291-295.

Tabata, K., H. Ikeda and S.-i. Hashimoto (2005). "ywfE in *Bacillus subtilis* Codes for a Novel Enzyme, L-Amino Acid Ligase." Journal of Bacteriology **187**(15): 5195-5202.

Todorovic, M. and D. M. Perrin (2020). "Recent developments in catalytic amide bond formation." Peptide Science **112**(6): e24210.

Upadhyay, R., R. Rana and S. K. Maurya (2021). "Organocatalyzed C–N bond-forming Reactions for the Synthesis of Amines and Amides." ChemCatChem **13**(8): 1867-1897.

Valeur, E. and M. Bradley (2009). "Amide bond formation: beyond the myth of coupling reagents." Chemical Society Reviews **38**(2): 606-631.

Vasić, K., Ž. Knez and M. Leitgeb (2023). "Transglutaminase in Foods and Biotechnology." International Journal of Molecular Sciences **24**(15): 12402.

Willis, N. J., C. A. Fisher, C. M. Alder, A. Harsanyi, L. Shukla, J. P. Adams and G. Sandford (2016). "Sustainable synthesis of enantiopure fluorolactam derivatives by a selective direct fluorination – amidase strategy." Green Chemistry **18**(5): 1313-1318.

Wilson, M. E. and G. M. Whitesides (1978). "Conversion of a protein to a homogeneous asymmetric hydrogenation catalyst by site-specific modification with a diphosphinerhodium(I) moiety." Journal of the American Chemical Society **100**(1): 306-307.

Winkler, M. (2018). "Carboxylic acid reductase enzymes (CARs)." Current Opinion in Chemical Biology **43**: 23-29.

Winn, M., S. M. Richardson, D. J. Campopiano and J. Micklefield (2020). "Harnessing and engineering amide bond forming ligases for the synthesis of amides." Current Opinion in Chemical Biology **55**: 77-85.

Wood, A. J. L., N. J. Weise, J. D. Frampton, M. S. Dunstan, M. A. Hollas, S. R. Derrington, R. C. Lloyd, D. Quaglia, F. Parmeggiani, D. Leys, N. J. Turner and S. L. Flitsch (2017). "Adenylation Activity of Carboxylic Acid Reductases Enables the Synthesis of Amides." Angewandte Chemie International Edition **56**(46): 14498-14501.

Ye, Y., J. Cao, D. G. Oblinsky, D. Verma, C. K. Prier, G. D. Scholes and T. K. Hyster (2023). "Using enzymes to tame nitrogen-centred radicals for enantioselective hydroamination." Nature Chemistry **15**(2): 206-212.

Yu, K., Z. Zou, N. V. Igareta, R. Tachibana, J. Bechter, V. Köhler, D. Chen and T. R. Ward (2023). "Artificial Metalloenzyme-Catalyzed Enantioselective Amidation via Nitrene Insertion in Unactivated C(sp<sup>3</sup>)–H Bonds." Journal of the American Chemical Society.

Zaks, A. and A. M. Klibanov (1985). "Enzyme-catalyzed processes in organic solvents." Proceedings of the National Academy of Sciences **82**(10): 3192-3196.

Zeng, S., J. Liu, S. Anankanbil, M. Chen, Z. Guo, J. P. Adams, R. Snajdrova and Z. Li (2018). "Amide Synthesis via Aminolysis of Ester or Acid with an Intracellular Lipase." ACS Catalysis **8**(9): 8856-8865.

Zhang, J., J. F. Barajas, M. Burdu, G. Wang, E. E. Baidoo and J. D. Keasling (2017). "Application of an Acyl-CoA Ligase from *Streptomyces aizunensis* for Lactam Biosynthesis." ACS Synth Biol **6**(5): 884-890.

Zhang, J., E. Kao, G. Wang, E. E. K. Baidoo, M. Chen and J. D. Keasling (2015). "Metabolic engineering of *Escherichia coli* for the biosynthesis of 2-pyrrolidone." Metabolic engineering communications **3**: 1-7.

Zhao, X., Y. Wu, T. Feng, J. Shen, H. Lu, Y. Zhang, H. H. Chou, X. Luo and J. D. Keasling (2023). "Dynamic upregulation of the rate-limiting enzyme for valerolactam biosynthesis in *Corynebacterium glutamicum*." Metabolic Engineering.

Zhou, Q., M. Chin, Y. Fu, P. Liu and Y. Yang (2021). "Stereodivergent atom-transfer radical cyclization by engineered cytochromes P450." Science **374**(6575): 1612-1616.

Zhu, M., L. Wang and J. He (2021). "Repurposing the 3-Isocyanobutanoic Acid Adenylation Enzyme SfaB for Versatile Amidation and Thioesterification." Angewandte Chemie International Edition **60**(4): 2030-2035.



## CONTEXTE DE LA THESE ET OBJECTIFS

Comme nous avons pu le voir dans l'état de l'art portant sur la synthèse biocatalytique d'amides, et surtout de lactames, les enzymes dépendantes de l'ATP ouvrent des perspectives intéressantes. Parmi celles-ci, les CoA ligases ont conduit à des résultats prometteurs obtenus en bioconversion.

La synthèse d'amides par des CoA ligases thermophiles étudiée lors de la thèse de Dr Chloé Lelièvre de 2017 à 2020, présente plusieurs avantages. Notamment, la diversité des amides obtenue peut être, *a priori*, très vaste car limitée par la seule réactivité des amines, l'étape suivant l'adénylation n'étant pas catalysée par l'enzyme. Il a été également montré que l'étape de thioestérification peut être supprimée, l'attaque de l'amine pouvant intervenir directement sur l'intermédiaire acyl-AMP, évitant ainsi l'utilisation de CoA qui est un cofacteur coûteux. Pour limiter l'utilisation d'ATP, un autre co-facteur coûteux, un système de régénération de l'ATP a été mis en place (Lelièvre, Balandras et al. 2020).

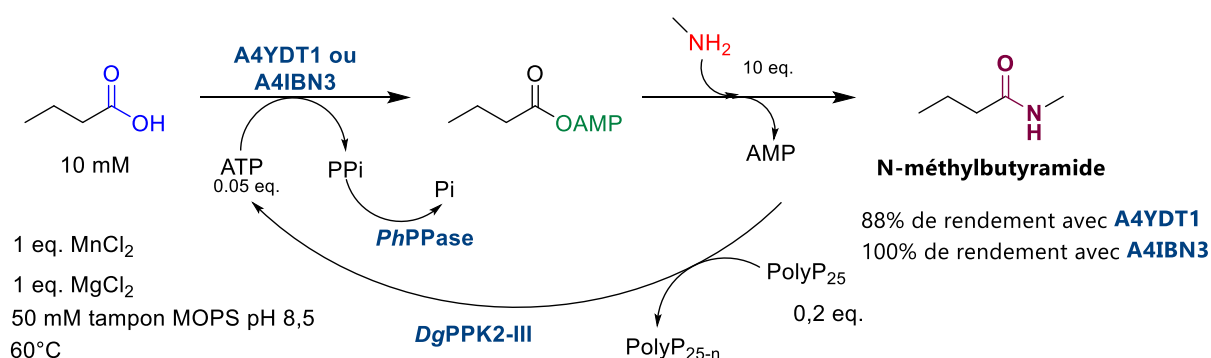


Schéma 58. Synthèse de N-méthylbutyramide avec les CoA ligases A4YDT1 ou A4IBN3 couplée à un système de régénération enzymatique de l'ATP par une PPK-III de *Deinococcus geothermalis* (DgPPK2-III) et la PPase de *Pyrococcus horikoshii* (PhPPase).

Des résultats préliminaires prometteurs ont été obtenus pour l'application de cette voie enzymatique aux lactames. Des acides  $\omega$ -aminés simples se sont révélés substrats de CoA ligases identifiées dans la littérature ou par criblage de la biodiversité, en particulier A4YDT1, une CoA ligase de *Metallosphaera sedulla*, active sur l'acide 4-aminobutyrique, précurseur du  $\gamma$ -butyrolactame (lactame à 5 chaînons) (Hawkins, Han et al. 2013). La collaboration entreprise avec des chercheurs de l'Université de Groningen (Dr Nikolas Capra, Pr Andy-Mark Thunnissen et Pr Dick Janssen) a permis une analyse de la structure de cette enzyme et de proposer des résidus cibles à modifier pour accéder à des activités sur des acides  $\omega$ -aminés plus encombrants, notamment l'acide 6-amino hexanoïque, précurseur du caprolactame, un composé d'intérêt car précurseur du nylon 6,6.

Ce projet de thèse s'est inscrit dans la continuité de ces résultats préliminaires. L'objectif global était d'identifier, caractériser et appliquer en réaction des CoA ligases pour la formation de lactames à 5-7 chaînons simples ou fonctionnalisés. Ceci permettant, par la suite, de proposer une voie enzymatique pour la synthèse de lactames d'intérêts, complémentaire aux voies pré-existantes. Pour cela, plusieurs objectifs ont été fixés, basés sur deux stratégies parallèles : 1) la recherche d'enzymes par approche génomique au sein de la biodiversité et 2) l'ingénierie protéique de la CoA ligase de *Metallosphaera sedulla*. Le premier objectif a été d'identifier les déterminants structuraux responsables de sa spécificité de substrats afin d'obtenir des mutants compétents pour l'obtention de lactames variés. Ce travail a été réalisé toujours en collaboration avec l'Université de Groningen (Schéma 1). Ces travaux et résultats font l'objet du chapitre 2 de ce manuscrit. Le deuxième objectif était d'identifier des CoA ligases au sein de

la biodiversité répertoriée en effectuant une exploration avec deux itérations minimum. Pour cela, la stratégie était d'utiliser des séquences de CoA ligases identifiées actives sur des acides  $\omega$ -aminés, idéalement fonctionnalisés, pour bâtir une nouvelle collection rassemblant des homologues plus actifs sur les substrats cibles. Ce travail a été effectué en collaboration avec la plateforme FunDivEx de l'unité. Ces travaux sont détaillés dans le chapitre 3 de ce manuscrit.

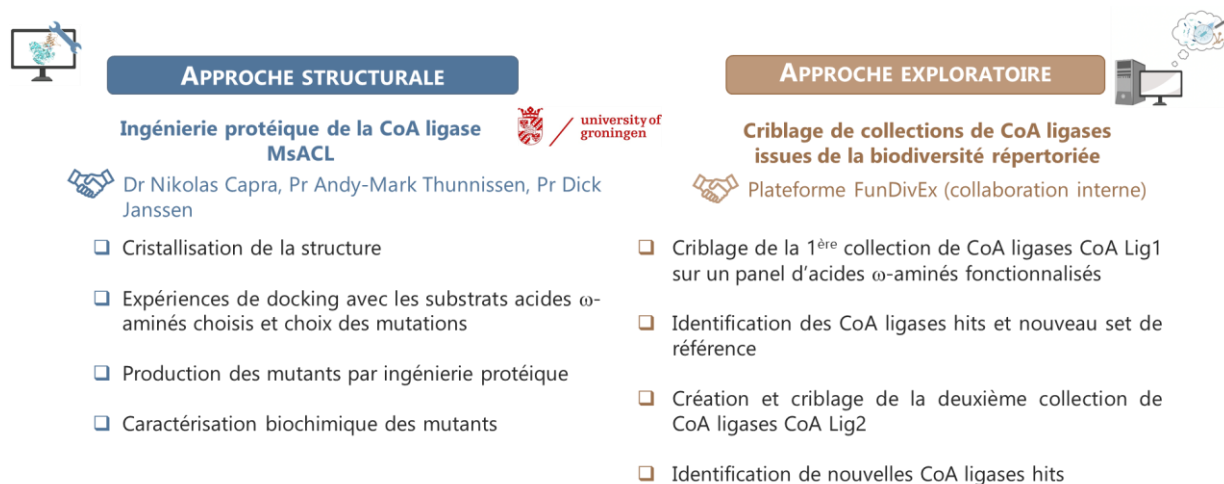


Figure 9. Objectifs de thèse pour la recherche CoA ligases pour la synthèse de lactames.

Les deux approches sont complémentaires. En connaissant les facteurs structuraux responsables de l'activité de la CoA ligase A4YDT1, nous avons modifié spécifiquement son site de liaison du substrat et dirigé son activité pour la synthèse de lactames d'intérêt. Par l'approche exploratoire, nous avons pu identifier de nouvelles CoA ligases actives afin d'établir une nouvelle collection plus ciblée sur notre panel de substrats acides  $\omega$ -aminés par itération et continuer notre exploration de CoA ligases plus diverses pour la synthèse de lactames. De plus, nous nous sommes appuyés sur les travaux d'étude structurale menés sur A4YDT1 pour une analyse plus poussée de nos résultats, notamment pour expliquer l'activité des CoA ligases que nous avons identifiées. A travers ces deux approches, nous avons donc voulu disposer de plus de CoA ligases pour la synthèse de divers lactames.

## RÉFÉRENCES BIBLIOGRAPHIQUES

Hawkins, A. S., Y. Han, R. K. Bennett, M. W. W. Adams and R. M. Kelly (2013). "Role of 4-Hydroxybutyrate-CoA Synthetase in the CO<sub>2</sub> Fixation Cycle in Thermoacidophilic Archaea." Journal of Biological Chemistry **288**(6): 4012-4022.

Lelièvre, C. M., M. Balandras, J.-L. Petit, C. Vergne-Vaxelaire and A. Zapparucha (2020). "ATP Regeneration System in Chemoenzymatic Amide Bond Formation with Thermophilic CoA Ligase." ChemCatChem **12**(4): 1184-1189.





## CHAPITRE II : INGENIERIE PROTEIQUE D'UNE COA LIGASE POUR LA SYNTHESE DE LACTAMES : CAS DE L'ENZYME A4YDT1

L'ingénierie protéique fait partie des approches les plus courantes afin d'étendre et améliorer les propriétés des enzymes pour diverses applications biotechnologiques. Cela permet de lever des limitations liées à leur stabilité, leur tolérance à de fortes concentrations et substrat et/ou de solvant, leur activité catalytique ou encore leur sélectivité (Paul, Hanefeld et al. 2024).

Deux stratégies sont principalement employées en ingénierie protéique : l'évolution dirigée et le design rationnel (Penning and Jez 2001). L'évolution dirigée est basée sur le concept de l'évolution darwiniste et repose sur la mutagenèse aléatoire portée par des techniques comme « l'error-prone PCR » (epPCR). L'epPCR consiste à réduire la fidélité de copie de la Taq polymérase lors de l'amplification de l'ADN, en jouant sur les conditions réactionnelles pour augmenter son taux d'erreur (Cadwell and Joyce 1992). Cette technique permet de générer de larges bibliothèques de mutants qui sont ensuite criblées pour une propriété cible, ce qui, par conséquent, nécessite d'utiliser une méthode de criblage adaptée (Porter, Rusli et al. 2016). La pionnière de l'utilisation de l'évolution dirigée est Frances Arnold dont les travaux pour l'ingénierie protéique ont été récompensés du prix Nobel de chimie en 2018 (Chen and Arnold 1993). Le design rationnel se base sur la connaissance de la structure de l'enzyme étudiée, de son site actif, des interactions entre les résidus du site actif et du substrat et de son mécanisme réactionnel. Par mutation des résidus clés, le plus souvent ceux du site actif (ou de toute autre partie de l'enzyme, le tunnel d'accès au site actif par exemple), il est possible d'adapter l'enzyme pour permettre la catalyse sur des substrats structurellement différents du substrat natif ou de d'accéder à une stéréosélectivité différente (Penning and Jez 2001, Paul, Hanefeld et al. 2024). On parle souvent alors de « redesign de site actif » (Toscano, Woycechowsky et al. 2007). Ces deux approches, aléatoire et rationnelle, sont souvent combinées en une approche semi-rationnelle avec la création de bibliothèques de mutants ciblés suite à l'analyse structurale de l'enzyme. Les travaux d'ingénierie protéique sont facilités grâce à l'impressionnant développement d'outils bio-informatiques, comme les outils de prédiction de variants (VEP) ou AlphaFold par exemple (Jumper, Evans et al. 2021, Horne and Shukla 2022). Ainsi, il est possible, pour des applications ciblées, d'obtenir un modèle de structure d'enzyme à partir de la séquence protéique, d'explorer toute la dynamique moléculaire des enzymes, ou encore de prédire l'effet de mutations (Bell, Finnigan et al. 2021).

Le projet d'ingénierie protéique d'une CoA ligase a débuté durant la thèse de Dr Chloé Lelièvre (2017-2020). Ces travaux sont le fruit de la collaboration avec Pr Andy-Mark Thunnissen et Pr Dick Janssen ainsi que le Dr Nikolas Capra, encore doctorant à ce moment, de l'Université de Groningen (Pays-Bas). Les travaux d'ingénierie ont été effectués sur la CoA ligase *MsACL* ou A4YDT1 (identifiant UniProt) de *Metallosphaera sedulla*. Cette enzyme fait partie des CoA ligases décrites dans la littérature utilisées pour constituer le set de référence de la première collection de CoA ligases issues de la biodiversité (Hawkins, Han et al. 2013, Hawkins, Adams et al. 2014) Elle a été choisie car son activité est limitée sur les substrats cibles (acides  $\omega$ -aminés), et les déterminants structuraux responsables de la spécificité de substrat sont mal connus. C'était donc la protéine candidate idéale pour cette étude visant à étendre l'activité des CoA ligases pour la synthèse de lactames variés. Au début de ma thèse, la structure d'A4YDT1 avait été résolue par Dr Nikolas Capra. Des travaux de modélisation moléculaire avaient permis de proposer des résidus cibles pour mutation et les mutants correspondants avaient été produits et testés par Dr Chloé Lelièvre permettant l'obtention de résultats préliminaires. La

consolidation de ces premiers résultats et la suite de ces travaux font l'objet de ce chapitre. Pour permettre une bonne compréhension des travaux que j'ai mené, je rappelle dans les paragraphes II.1 et II.2 les résultats obtenus par Nikolas Capra et Chloé Lelièvre. Ce travail a été publié dans *Frontiers in Catalysis* en 2024 (Capra, Lelièvre et al. 2024).

## II.1. DESCRIPTION STRUCTURALE DE LA COA LIGASE DES COA LIGASES ET RESULTATS PRELIMINAIRES POUR LA FORMATION DE LACTAMES AVEC A4YDT1

### II.1.1. STRUCTURE GLOBALE DES COA LIGASES ET LEUR MECANISME

Les CoA ligases appartiennent à la superfamille ANL qui comprend également les domaines d'adénylation des NRPS et les luciféras, dont la caractéristique commune est la réalisation d'une étape d'adénylation en présence d'ATP générant un intermédiaire acyl-AMP. Ce domaine d'adénylation est bien conservé parmi tous les membres de cette superfamille. Une autre caractéristique commune aux CoA ligases est leur structure bien conservée avec un large domaine N-terminal et un petit domaine C-terminal, avec le site actif à l'interface de ces deux domaines (Figure 10). Le site de liaison du substrat se trouve sur la partie N-ter et le site de liaison de l'AMP est en grande partie sur le côté N-ter. Le site de liaison du HSCoA est dans la partie C-ter avec un tunnel dans lequel le motif panthéteine du HSCoA va se retrouver enfoui et qui traverse ces deux domaines.

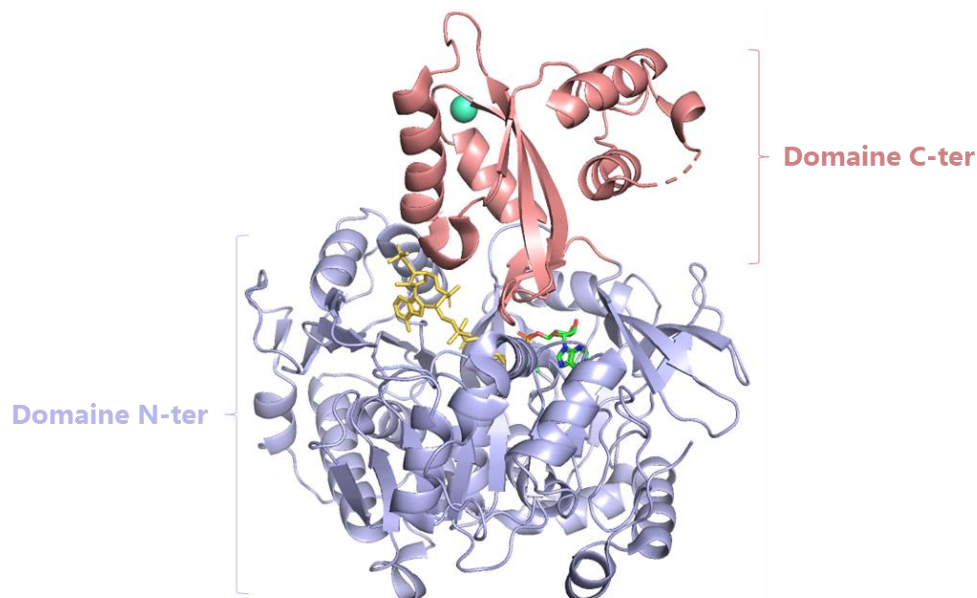


Figure 10. Structure tridimensionnelle de l'acétyl-CoA ligase Acs de *Salmonella enterica* en conformation de thioestérification (PDB : 1PG4) (Gulick *et al.*, 2003). Le domaine N-terminal est représenté en bleu-violet et le domaine C-terminal en rose clair. Structure co-cristallisée avec l'intermédiaire adénylé propyl-AMP représenté en vert, le HSCoA en jaune-orange et le cation  $Mg^{2+}$  représenté par une sphère cyan.

Les CoA ligases ont des séquences dans lesquelles on trouve des motifs conservés. Ces motifs, numérotés de A1 à A10, correspondent à des motifs aussi identifiés chez les NRPS (Clark, Leatherby et al. 2018). Les motifs A1, A2 et A6 ont des rôles structuraux dans le maintien de la structure secondaire des CoA ligases (Gulick 2009). Le motif A3 est une boucle riche en glycine et en sérine qui forment la

« P-loop ». Cette P-loop constitue un site de liaison et d'activation du phosphate de l'ATP permettant la coupure de la liaison phosphoester lors la formation de l'intermédiaire acyl-AMP, libérant ainsi du pyrophosphate. Le motif A10 situé dans le domaine C-ter, contient une lysine bien conservée dont la chaîne latérale est en interaction avec le substrat et l'AMP, permettant ainsi leur bon alignement dans le site actif (Conti, Stachelhaus et al. 1997). Après des travaux d'études structurales de la superfamille ANL, il a été établi que les motifs A3 et A10 jouent des rôles cruciaux dans l'étape d'adénylation (Figure 11). Le motif A8, localisé dans la partie C-ter, contient deux résidus bien conservés : un résidu aspartate charnière et une glycine (Figure 11). Le résidu acide aspartique se situe dans la partie charnière entre les domaines C-ter et N-ter et va adopter des angles de torsion sur sa chaîne principale phi et psi différents en fonction de la conformation de la protéine (Shah, Ingram-Smith et al. 2009). Par exemple, dans la 4-chlorobenzoate CoA ligase, le résidu Asp402 va passer d'un angle phi  $-66^\circ$  à  $-89^\circ$  et d'un angle psi  $-29^\circ$  à  $-164^\circ$  en passant de la conformation adénylation à la conformation thioestérification (Wu, Reger et al. 2009). Quant au résidu glycine, il est le seul résidu conservé dans le tunnel d'interaction avec la partie pantéthéine du CoA et plus précisément avec le motif  $\beta$ -alanine. Des mutations de ce résidu par une leucine ou une alanine chez l'acétyl-CoA ligase Acs de *Salmonella enterica* ont résulté en une diminution drastique de l'efficacité catalytique de l'enzyme pour le HSCoA ou une suppression de l'étape de thioestérification (Gulick, Starai et al. 2003, Reger, Wu et al. 2008).

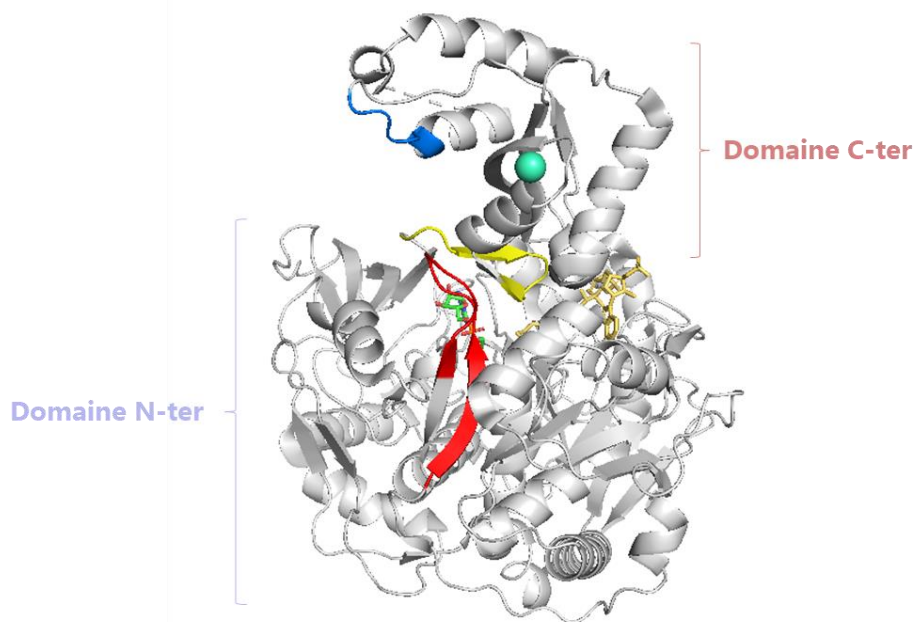


Figure 11. Structure tridimensionnelle de la CoA ligase Acs cristallisée en conformation de thioestérification (PDB : 1PG4). Les motifs A3 et A10 impliqués dans l'adénylation sont indiqués en rouge et bleu marine respectivement. Le motif A8 contenant la région charnière à l'interface des deux domaines est représenté en jaune. L'intermédiaire propyl-AMP est représenté en vert tandis que le HSCoA est représenté en jaune-orange. Le cation  $Mg^{2+}$  est représenté par une sphère cyan.

Les CoA ligases catalysent la formation de thioesters de CoA à partir d'acides carboxyliques. La réaction se fait en deux étapes avec une étape d'adénylation en présence d'ATP puis l'étape de thioestérification avec du CoA. Pour ces deux réactions, les CoA ligases adoptent des deux conformations différentes dites ouverte et fermée, qui s'accompagnent d'un réarrangement du site actif (Figure 12). Ces changements de conformation sont définis comme une alternance de domaines (Wu, Cao et al. 2008). La thioestérification provoque un changement de conformation avec un angle de  $140^\circ$  qui peut être variable selon la CoA ligase. Lors du passage de la conformation d'adénylation à la

conformation de thioestérification, des résidus catalytiques impliqués dans l'adénylation comme la lysine du motif A10, vont être éloignés du site actif tandis que d'autres résidus du même domaine C-ter se retrouveront positionnés dans le site actif mais aussi permettre l'ouverture du tunnel d'interaction avec la partie pantéthéine du HSCoA, comme le résidu glycine du motif A8.

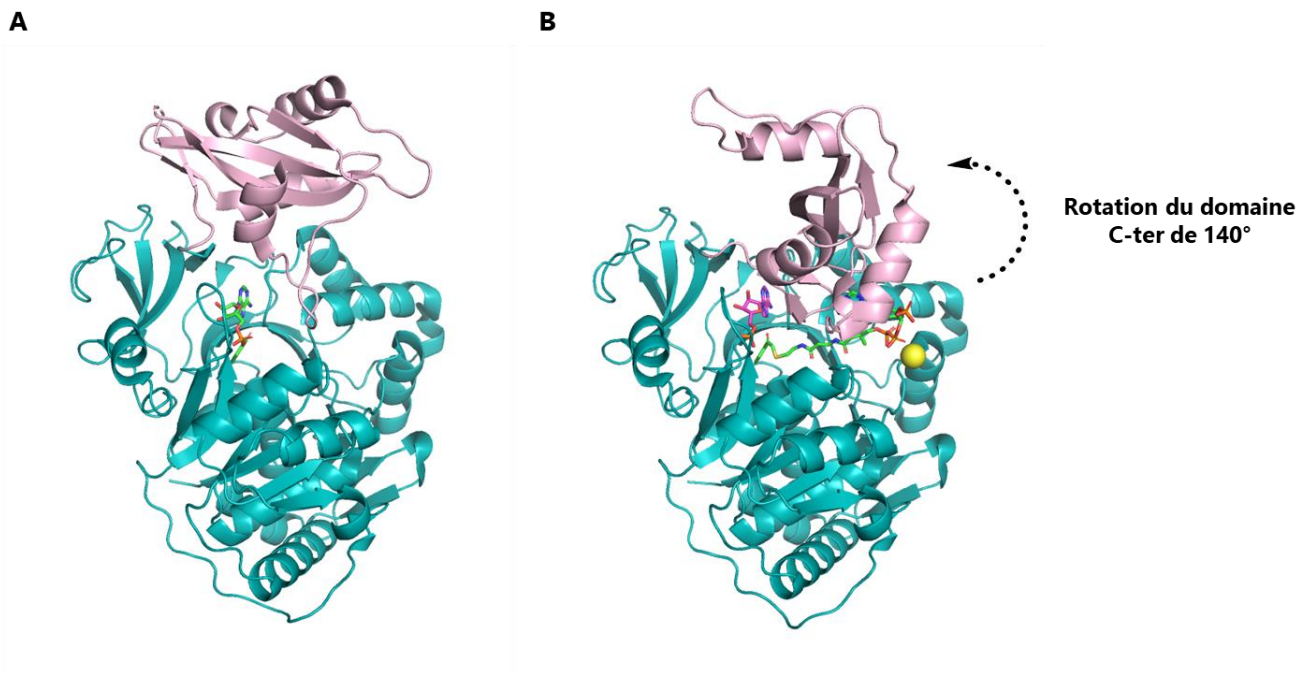


Figure 12. Représentation tridimensionnelle des deux conformations d'une 4-chlorobenzoate CoA ligase (4-CBL) de *Alcaligenes sp. AL3007* (Wu *et al.*, 2008 ; Reger *et al.*, 2008). Le domaine N-ter est représenté en cyan et le domaine C-ter en rose. **(A)** Conformation d'adénylation avec l'intermédiaire adénylé du substrat, 4-chlorobenzoate-AMP, représenté en vert (PDB : 3CW8). **(B)** Conformation de thioestérification avec le produit thioester 4-chlorobenzoate-CoA représenté en vert, l'AMP représenté en violet et  $Mg^{2+}$  représenté par une boule jaune (PDB : 3CW9).

Le mécanisme catalytique des CoA ligases est un mécanisme ping-pong « Bi Uni Uni Bi », c'est-à-dire que les deux demi-réactions composent un système à deux substrats formant un seul produit puis ce dernier va être libéré avant qu'un autre substrat se fixe dans le site actif pour former deux produits. Ce mécanisme a été mis en évidence avec la 4-chlorobenzoate CoA ligase de *Alcaligenes sp. AL3007* et la CoA ligase pour les acides gras à longue chaîne de *Thermus Thermophilus* grâce à l'obtention des structures RX de ces CoA ligases dans chaque conformation (Hisanaga, Ago et al. 2004, Wu, Cao et al. 2008).

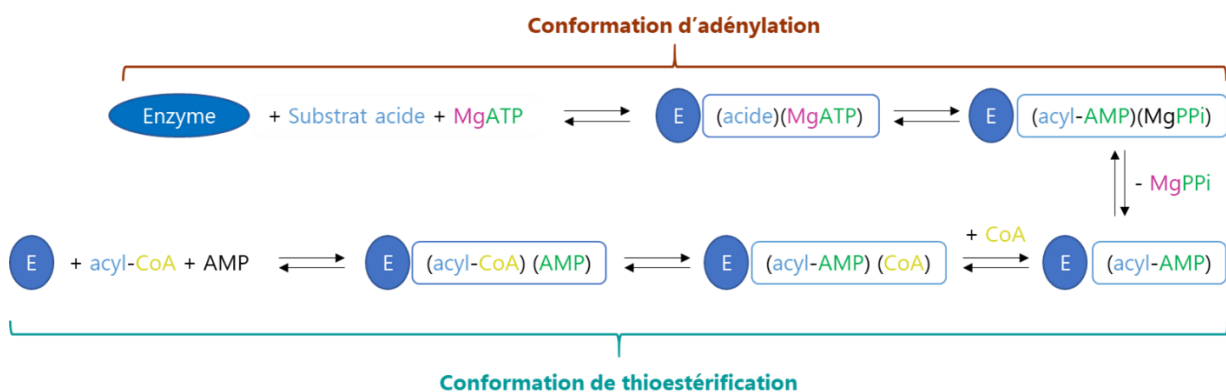


Schéma 59. Mécanisme enzymatique simplifié des deux réactions catalysées par les CoA ligases. Figure inspirée de l'article de Wu *et al.* (Wu *et al.*, 2008).

Lors de l'adénylation, le carbonyle du groupe COOH du substrat établit des liaisons hydrogène avec une histidine du N-ter ce qui permet de bien aligner le substrat mais aussi d'obstruer le site de liaison du HSCoA. L'ATP va être positionné dans le site grâce au « P-loop » (motif A3) qui contient des résidus thréonine qui vont former une liaison hydrogène avec l'un des oxygènes des phosphates  $\alpha$  et  $\gamma$  de l'ATP. La lysine du motif A10 contribue également au bon alignement de l'ATP en établissant des liaisons hydrogène avec les groupements hydroxyle 2' et 3' de la partie adénosine. Cette lysine conservée contribue également à l'activité de l'enzyme durant l'adénylation par une liaison ionique avec le carboxyle du substrat jusqu'à la formation de l'intermédiaire acyl-AMP. Cette demi-réaction nécessite un cation divalent  $Mg^{2+}$  qui active l'ATP par coordination aux phosphates  $\beta$  et  $\gamma$ , ce qui correspond au 1<sup>er</sup> « Bi » du mécanisme : MgATP. Le phosphate  $\alpha$  de l'ATP puis de l'AMP est quant à lui en interaction avec des résidus thréonine/sérine par une liaison hydrogène avec l'un des oxygènes du groupement. L'adénylation va donc produire l'intermédiaire acyl-AMP et va libérer le produit MgPPi qui est le 1<sup>er</sup> produit « Uni ». La dissociation du MgPPi provoque l'alternance de domaine avec la rotation du domaine C-ter. Ce changement de conformation éloigne la lysine catalytique du site actif mais aussi l'histidine, ce qui permet d'ouvrir le tunnel et laisser entrer le HSCoA dans le site actif sans qu'il y ait de gêne stérique. En effet, due à la rotation du domaine C-ter, le résidu glutamate conservé est orienté vers le site actif et va être en interaction avec l'histidine catalytique du N-ter par une liaison hydrogène, ce qui permet d'ouvrir le tunnel pantéthéine. Le HSCoA qui représente le 2<sup>ème</sup> « Uni », est bien aligné dans le site actif via son noyau pantéthéine par des liaisons hydrogènes dont une avec l'oxygène de la chaîne principale du résidu glycine conservé appartenant au motif A8. Par ailleurs, des résidus aromatiques du domaine C-ter, aussi orientés vers le site actif, vont stabiliser le noyau adénine du HSCoA. Le HSCoA « Uni » est donc bien aligné dans le site actif pour attaquer l'intermédiaire acyl-AMP et générer les deux produits finaux « Bi » résultants de la thioestérification : l'intermédiaire thioester acyl-CoA et l'AMP.

## II.1.2. PRESENTATION DE L'ENZYME A4YDT1 ET RESULTATS PRELIMINAIRES POUR LA SYNTHÈSE DE LACTAMES

La CoA ligase A4YDT1 a été identifiée chez l'archée thermoacidophile *Metallosphaera sedulla*, où elle est impliquée dans la voie de biosynthèse du 3-hydroxypropionate (3-HP)/4-hydroxybutyrate (4-HB) dans le cycle de fixation du carbone pour obtenir de l'acétyl-CoA (Hawkins, Adams et al. 2014). Si son substrat naturel est l'acide 4-hydroxybutyrique, il a été montré qu'elle était également active sur d'autres acides carboxyliques comme l'acide butyrique ou l'acide valérique pour lequel elle présente une meilleure affinité (Schéma 60) (Hawkins, Han et al. 2013). De plus, l'acide 4-hydroxybutyrique possède une structure similaire à l'acide 4-aminobutanoïque (ou GABA, de l'anglais  *$\gamma$ -aminobutanoic acid*), qui est l'un des acides  $\omega$ -aminés testés.

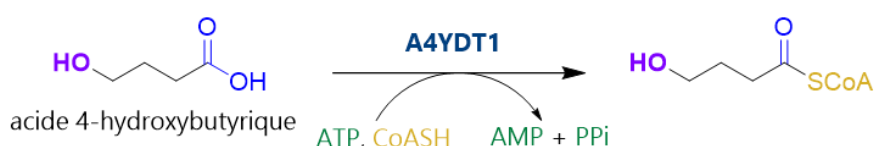


Schéma 60. Réaction naturelle catalysée par la CoA ligase A4YDT1 de *Metallosphaera sedulla*.

Compte tenu de ses activités, cette enzyme a été choisie comme CoA ligase de référence, pour être testée pour la formation de lactames de 5 à 7 chaînons, et comme point de départ pour chercher de nouvelles CoA ligases. A4YDT1 a été testée sur trois acides  $\omega$ -aminés : le GABA, l'acide 5-

aminopentanoïque (5-APA) et l'acide 6-aminohexanoïque (6-AHA) et n'a montré de l'activité qu'avec le GABA, atteignant 63% de rendement en  $\gamma$ -butyrolactame en 24h.

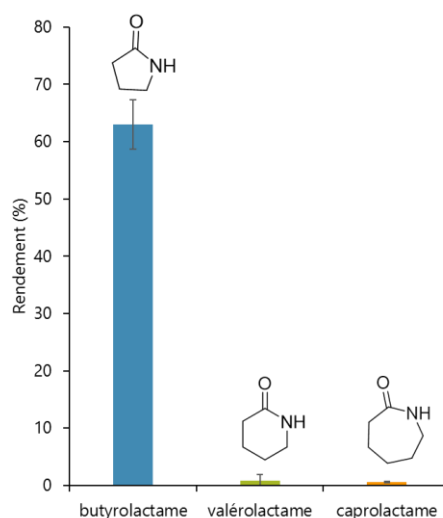


Figure 13. Rendement en  $\gamma$ -butyrolactame,  $\delta$ -valérolactame et  $\epsilon$ -caprolactame avec la CoA ligase A4YDT1. Conditions réactionnelles : 5 mM substrat  $\omega$ -acide aminé, 5 mM ATP, 5 mM  $MnCl_2$ , 50 mM tampon MOPS pH 8.5, 2 mg/mL A4YDT1 purifié. 24h, 60°C. Analyse par UHPLC-MS (APCI).

Ce résultat a été à l'origine du projet d'ingénierie protéique entrepris avec nos collaborateurs de l'Université de Groningen. L'étude s'est focalisée sur la modification du site actif d'A4YDT1 afin d'accommoder des substrats les acides  $\omega$ -aminés plus grands pour étendre son applicabilité dans la synthèse des lactames. Pour cela, la structure tridimensionnelle d'A4YDT1 a été résolue par cristallographie aux rayons X par nos collaborateurs. En comparant ces structures avec celles de CoA ligases apparentées et en effectuant des dockings avec des intermédiaires de substrats adénylés, des informations sur les résidus clés composant le site de fixation du substrat et déterminant la spécificité de l'enzyme ont été obtenus. Ces résultats ont guidé la prédiction rationnelle de mutations permettant d'élargir le champ des substrats d'A4YDT1.

## II.2. ETUDE STRUCTURALE DE LA COA LIGASE A4YDT1

La cristallisation et la résolution de la structure de A4YDT1 ont été réalisées par Dr Nikolas Capra. Grâce à la structure obtenue et en s'aidant de la littérature portant sur d'autres CoA ligases cristallisées, les caractéristiques structurales communes à cette famille ont été identifiées chez A4YDT1. Des expériences de docking avec les substrats adénylés ont ensuite permis de proposer des résidus clés comme responsables de la spécificité de substrat de l'enzyme. Des mutations ont alors été proposées suite à l'identification de ces déterminants structuraux. Les résultats sont présentés ci-dessous.

### II.2.1. CRISTALLISATION ET DETERMINATION DE LA STRUCTURE D'A4YDT1

#### II.2.1.1. CRISTALLISATION D'A4YDT1

Tout d'abord, des expériences de co-cristallisation de A4YDT1 ont été faites en présence de différents ligands (AMP, ATP et CoA), mais aussi des combinaisons de ces ligands (ATP+ $MgCl_2$ , ATP+ $MgCl_2$ +CoA, ATP+ $MgCl_2$ +acide butyrique, AMP+CoA, AMPPNP+ $MgCl_2$ ) à une concentration de 5



mM pour chaque ligand. Le but de cette co-cristallisation est de pouvoir accéder à différentes conformations en présence des ligands. Des cristaux ont été formés pour toutes les conditions après 48h avec des conditions réactionnelles différentes dans une solution de 0,1 M de tampon acétique de pH 5.2, 200 mM de NaCl et 8% de PEG3350. Au final, deux cristaux ont été obtenus avec une résolution de 2,8 Å pour la combinaison de ligands de 5 mM d'ATP et 5 mM MgCl<sub>2</sub>, et une résolution de 3,1 Å la combinaison à 5 mM d'ATP, 5 mM MgCl<sub>2</sub> et 5 mM CoA.

Pour reconstruire la structure de A4YDT1, un remplacement moléculaire a été effectué en utilisant la structure cristallisée d'une autre CoA ligase homologue, la CoA ligase humaine ACSM2A (PDB : 3B7W) avec laquelle A4YDT1 partage 33% d'identité (Kochan, Pilka et al. 2009). Ce remplacement a été effectué en deux étapes, une pour chaque domaine, en cherchant leurs bonnes positions et rotations afin de prendre en compte la flexibilité de l'enzyme. Un modèle AlphaFold2 a été employé afin d'améliorer la reconstitution dans les zones avec peu de densité électronique. Par ailleurs, l'étude de la carte de densité électronique a révélé la présence d'une molécule acétyl-AMP dans le site actif indiquant que A4YDT1 a réagi avec une molécule de tampon acétate comme substrat pour générer cet intermédiaire avant qu'il y ait cristallisation. En comparant avec la structure de ACSM2A obtenue dans les conformations d'adénylation et de thioestérification, il a été montré que A4YDT1 a été cristallisée en conformation de thioestérification (Kochan, Pilka et al. 2009).

### **II.2.1.2. STRUCTURE GLOBALE D'A4YDT1**

La structure de A4YDT1 cristallisée en conformation de thioestérification est référencée dans la base de données de la Protein Data Bank (PDB) sous le code 8BIT pour l'enzyme co-cristallisée avec l'AMP et 8BIQ la co-cristallisation avec AMP et le CoA. A4YDT1 possède une structure quaternaire bien conservée dans la famille des CoA ligases et autres membres de la superfamille ANL avec les résidus 16 à 449 qui forment le large domaine N-ter et les 120 derniers résidus (456 à 570) qui forment le plus petit domaine C-ter (Figure 14). Ces deux domaines sont connectés par le domaine charnière constitué de 5 résidus de 450 à 455. La co-cristallisation de l'acétyl-AMP en présence du HSCoA a permis l'identification du tunnel en interaction avec la partie panthéine du CoA ainsi que de la poche de liaison du substrat par étude de la densité électronique de ces ligands.



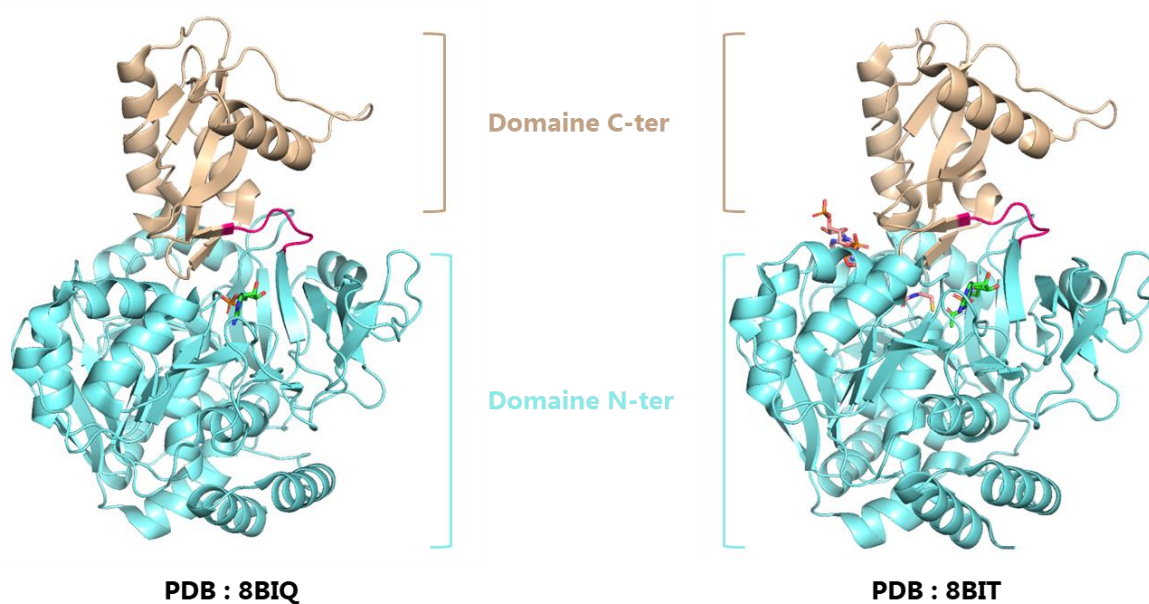


Figure 14. Structures cristallisées de la CoA ligase A4YDT1 en conformation de thioestérification, en présence d'AMP représenté en vert (PDB : 8BIQ) et en présence d'acétyl-AMP représenté en vert et du HSCoA représenté en rose clair (PDB : 8BIT). Les domaines N-ter et C-ter sont représentés en cyan et beige respectivement et la région charnière en rose.

L'identification des motifs et des résidus catalytiques chez A4YDT1 a été réalisée par superposition avec 5 CoA ligases dont les structures ont été cristallisées en conformation de thioestérification : l'acétyl-CoA ligase de *Salmonella enterica*, l'acyl-CoA ligase de *Methanosarcina acetivorans*, l'acyl-CoA ligase humaine ACSM2A, la benzoate CoA ligase de *Rhodopseudomas palustris* et l'antranilate-CoA ligase de *Stigmatella aurantiaca* (Tableau 1) (Gulick, Starai et al. 2003, Kochan, Pilka et al. 2009, Shah, Ingram-Smith et al. 2009, Thornburg, Wortas-Strom et al. 2015, Jackson, Tu et al. 2016). Les CoA ligases partagent une faible homologie avec A4YDT1, avec un pourcentage d'identité de séquence variant entre 30 et 38%.

Tableau 1. CoA ligases homologues à A4YDT1 dont la structure a été cristallisée en conformation de thioestérification.

Code PDB	Code UniProt	Nomenclature dans le texte	Annotation	Organisme	Spectre de substrats acides	Pourcentage d'identité de séquence avec A4YDT1 (%)
1PG4	Q8ZKF6	<i>Se</i> ACL	Acétyl-CoA synthétase	<i>Salmonella enterica</i>	acides à chaîne courte	32
3ETC	Q8TLW1	<i>Ma</i> ACL	Acyl-adénylate CoA ligase	<i>Methanosarcina acetivorans</i>	acides à chaîne courte et moyennes ramifiées	38
2WD9	Q08AH3	<i>Hs</i> ACL	Acyl-CoA synthétase ACSM2A	<i>Homo sapiens</i>	acides à chaîne moyenne et acides aromatiques	33
4RM3	Q93TK0	<i>Rp</i> ACL	Benzoate-CoA ligase	<i>Rhodopseudomas palustris</i>	acides aromatiques	30
4WV3	F3Y661	<i>Sa</i> ACL	Antranilate-CoA ligase	<i>Stigmatella aurantiaca</i>	acides aromatiques	31

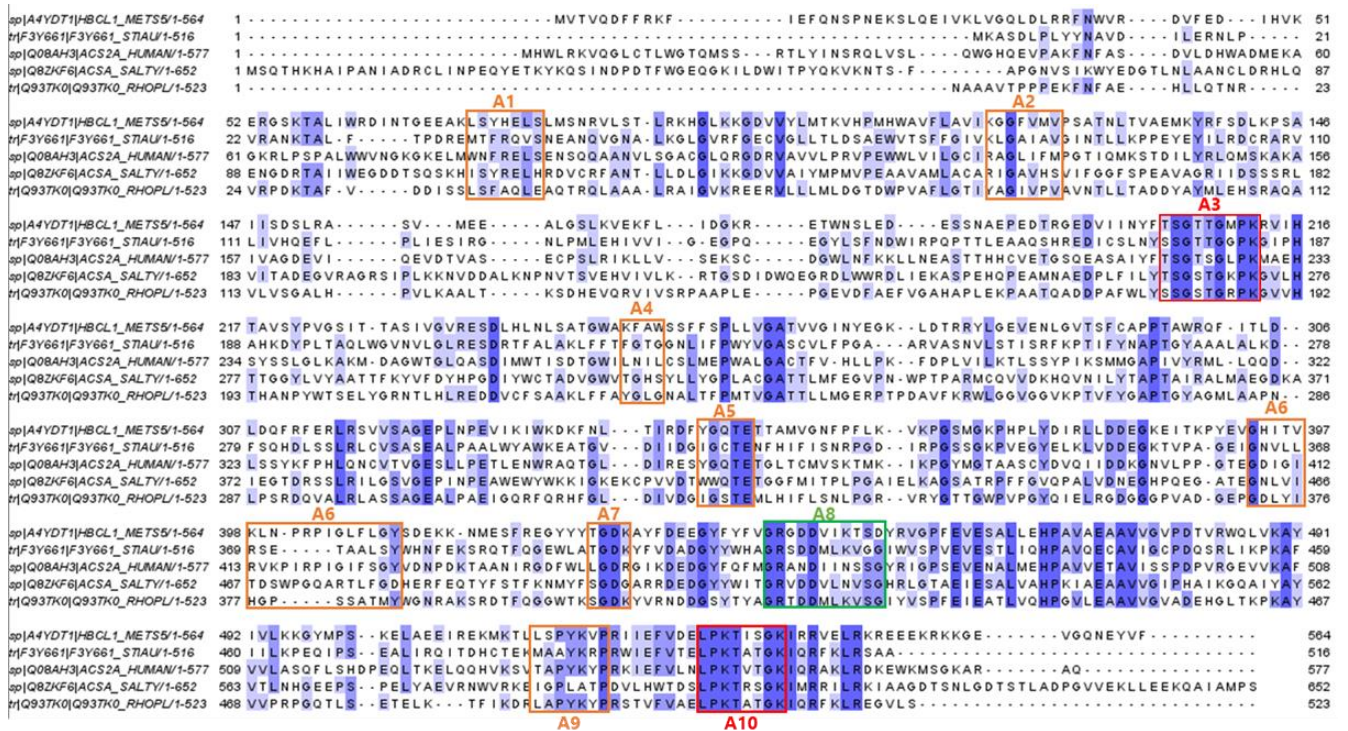


Figure 15. Alignement des séquences protéiques ClustalOmega d'A4YDT1 et des CoA ligases servant pour la reconstruction structurale de A4YDT1. Les motifs structuraux sont encadrés et indiqués sur la figure.

On retrouve les motifs communs A1 à A10 aux CoA ligases et membres de la superfamille ANL, notamment les motifs catalytiques impliqués dans l'étape d'adénylation avec le motif A3 riche en glycine, sérine et thréonine et le motif A10 avec la lysine, et le motif A8 avec le résidu aspartate charnière conservée par toutes les CoA ligases.

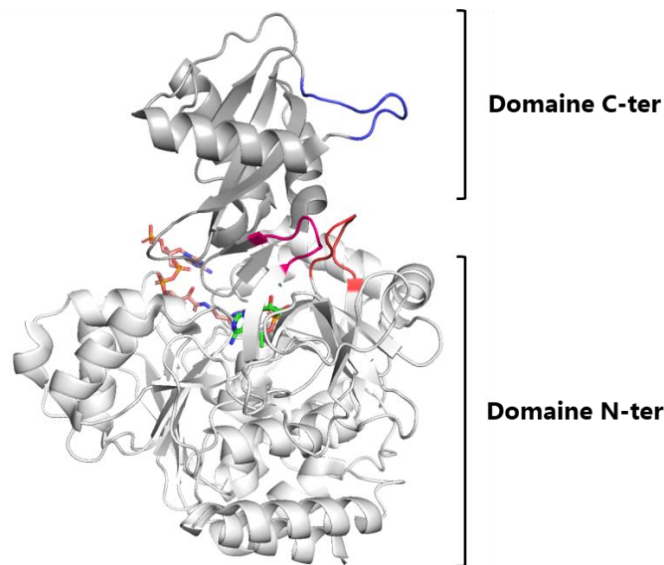


Figure 16. Structure tridimensionnelle de A4YDT1 en conformation de thioestérification avec les motifs A3 et A10 impliqués dans l'adénylation, représentés en rouge et bleu marine respectivement. Le motif A8 correspondant à la région charnière est coloré en rose foncé. L'acétyl-AMP et le HSCoA sont colorés en vert et en rose pâle respectivement.

### II.2.1.2.1. Site de liaison de l'AMP du site actif de A4YDT1

Comme décrit précédemment, le site de liaison de l'AMP est situé à l'interface des deux domaines, avec une majorité des résidus catalytiques dans la partie N-ter. Les groupements amine de la partie adénine de l'AMP forment des liaisons hydrogène avec la chaîne latérale du résidu D349 et la chaîne principale du résidu F350. L'un des oxygènes du phosphate est en interaction avec le résidu Q353 de la « P-loop » (motif A3) ainsi que le résidu R461 située dans le domaine C-ter. Nous retrouvons également la lysine 456 du motif A10 qui établit également une liaison hydrogène avec l'un des oxygènes du groupement phosphate ainsi que l'oxygène du carbonyle de la liaison phosphoester de l'AMP. Cette lysine intervient donc dans le bon positionnement de la partie AMP pour l'attaque nucléophile du groupement thiol du HSCoA dans l'étape de thioestérification (Figure 17).

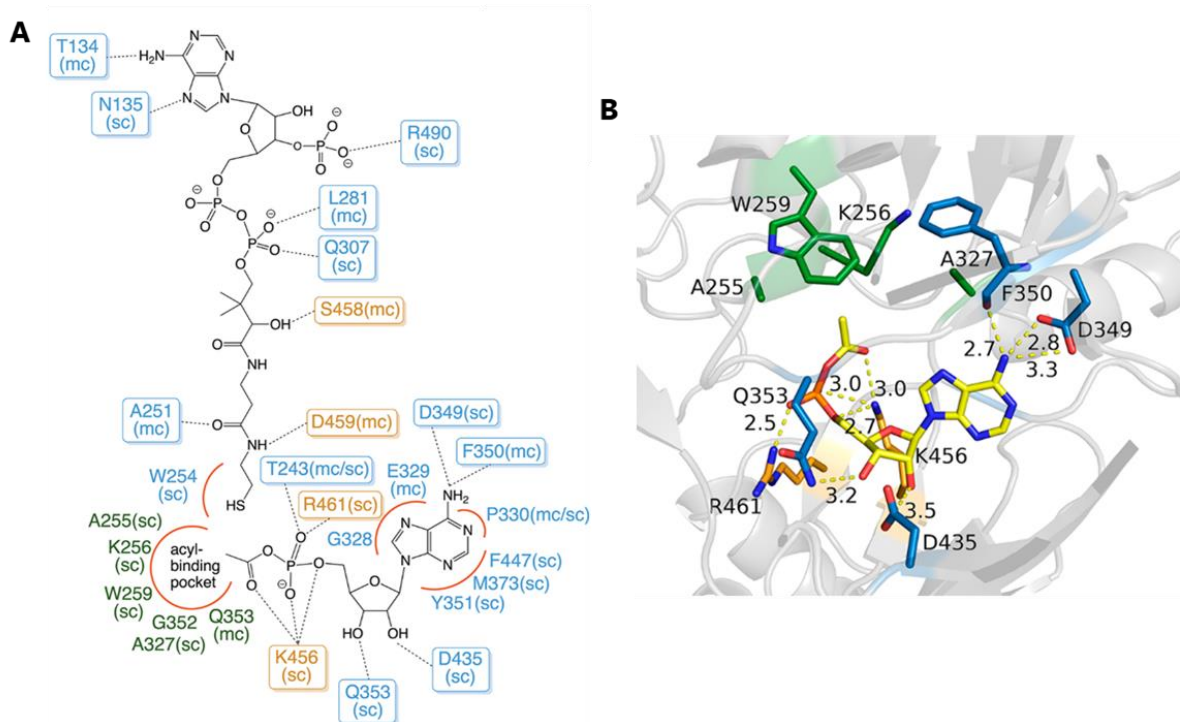


Figure 17. (A) Diagramme schématisé montrant les interactions entre le HSCoA, l'AMP et le substrat avec les résidus présents dans le site actif de A4YDT1, basé sur l'analyse de sa structure. Les résidus appartenant au domaine N-ter sont indiqués en bleu et vert et les résidus du domaine C-ter sont indiqués en orange. (B) Zoom sur les résidus en interaction avec l'AMP (en jaune) dans la structure 3D de A4YDT1. Figures faites par Dr Nikolas Capra.

### II.2.1.2.2. Site de liaison du substrat dans le site actif de A4YDT1

Le site de liaison de liaison du substrat est situé dans la partie N-ter. Il est constitué par les chaînes latérales des résidus A255, K256, G352, Q353 et W259 (Figure 18).



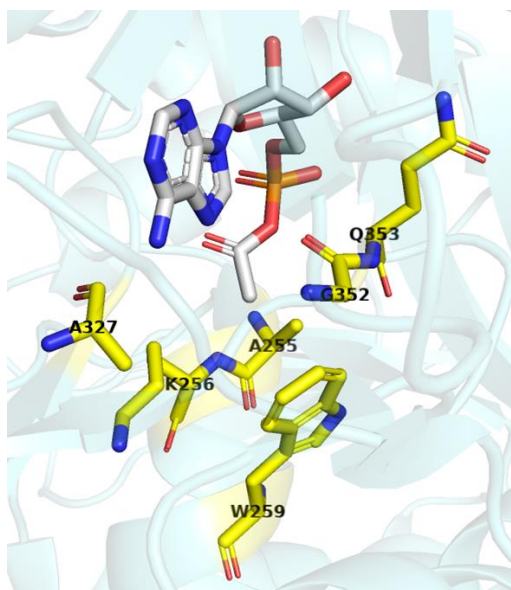


Figure 18. Zoom sur les résidus constituant le site de liaison du substrat dans la partie N-ter de A4YDT1 (PDB : 8BIT), colorés en jaune. L'intermédiaire acétyl-AMP est coloré en blanc.

Cette poche est petite et majoritairement hydrophobe avec le résidu W259 présent dans le fond. L'alignement structural et la superposition des sites actifs avec les autres CoA ligases montrent qu'elles possèdent également toutes une poche majoritairement hydrophobe. Toutefois, il existe une différence dans la profondeur et la forme de la poche, qui est spécifique du type de substrats sur lesquels ces CoA ligases sont actives.

Tableau 2. Comparaison des résidus constituant le site de liaison de substrat de A4YDT1 avec les autres CoA ligases homologues, identifiés par superposition des structures dans le logiciel PyMOL.

Résidus dans la poche de A4YDT1	Rôle dans la poche	<i>SeACL</i> (acides courts)	<i>MaACL</i> (acides moyens, ramifiés)	<i>HsACL</i> (acides moyens et aromatiques)	<i>RpACL</i> (acides aromatiques)	<i>SaACL</i> (acides aromatiques)
A255	Constitutif de la poche, pas conservé	V310	G255	I266	A227	T221
K256	Constitutif de la poche, conservé chez certaines CoA ligases à acides moyens	T311	K256	L267	Y228	F222
W259	Fond de poche, conservé chez certaines CoA ligases à acides moyens	S314	W259	L270	G231	G225
A327	Constitutif de la poche, conservé	V386	A326	V337	A302	A293
G352	Constitutif de la poche, très conservé chez les acyl-CoA ligases sauf ceux à chaîne courtes	W414	G351	G362	G327	G318

Nous avons comparé A4YDT1 avec la CoA ligase *MaACL* (PDB : 3ETC), celle avec laquelle elle partage la plus forte homologie de séquences (38%), nous avons observé presque les mêmes résidus pour ces deux CoA ligases, notamment avec ce tryptophane en fond de poche. Expérimentalement, on observe bien des activités semblables sur des substrats acides de taille moyenne pour les deux enzymes, A4YDT1 et *MaACL* (Shah, Ingram-Smith et al. 2009). La CoA ligase *SeACL* dont le spectre de substrat est limité aux acides courts, possède aussi un résidu tryptophane plancher, situé à la position équivalente

de G352 chez A4YDT1 (Gulick, Starai et al. 2003). Cette spécificité se retrouve chez les CoA ligases actives sur des substrats courts. A contrario, les CoA ligases *HsACL* (PDB : 2WD9), *RpACL* (PDB : 4RM3) et *SaACL* (PDB : 4WV3) spécifiquement actives sur des acides aromatiques, ne possèdent pas de résidu tryptophane plancher en position équivalente à W259. Elles ont soit un résidu glycine soit leucine qui sont plus flexibles et moins encombrants, et témoignent donc d'une poche de liaison du substrat plus profonde.

Des expériences de docking moléculaire ont été réalisées avec les intermédiaires aminoacyl-AMP.

## II.2.2. MODIFICATION DU SITE DE LIAISON DU SUBSTRAT DANS LE SITE ACTIF DE A4YDT1

### II.2.2.1. DOCKING DES INTERMÉDIAIRES ACYL-AMP DANS LE SITE ACTIF ET CHOIX DES RÉSIDUS A MUTER

Les intermédiaires aminoacyl-AMP des substrats acides aminés (GABA, 5-APA et 6-AHA) ont été générés en utilisant le logiciel gratuit JSME et les expériences de docking sont menées sur le logiciel AutoDock Vina. En présence de ces intermédiaires aminoacyl-AMP, les groupements amine terminaux du 5-aminopentanoyl-AMP et du 6-aminohexanoyl-AMP sont en clash stérique avec le résidu W259 en fond de poche. W259 était donc un résidu clé, la mutation W259G a été proposée pour pouvoir agrandir le site de liaison du substrat pour la prise en charge de plus larges acides aminés (Figure 19).

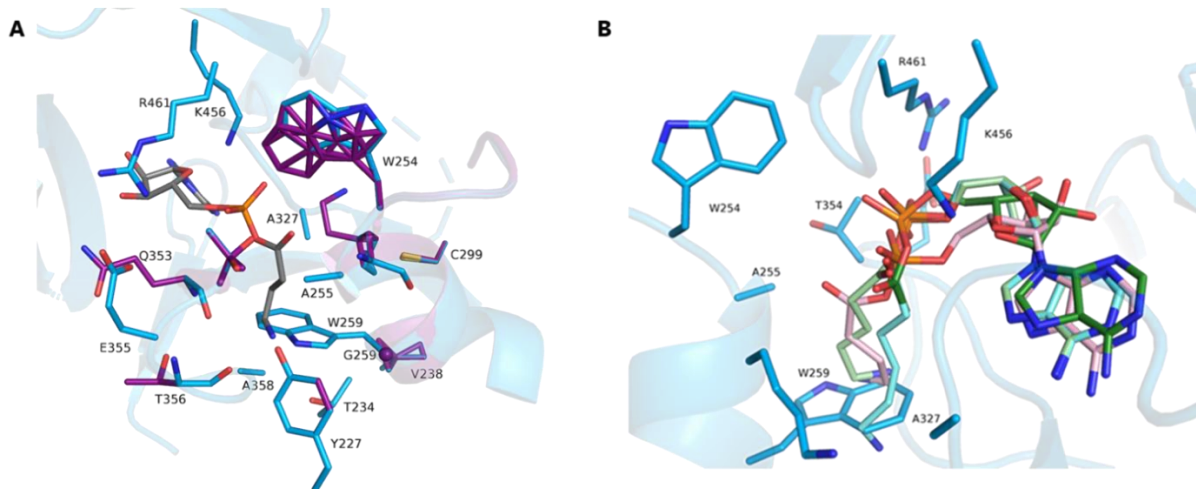


Figure 19. (A) Docking de l'intermédiaire 4-aminobutyryl-AMP (en gris) dans le site de liaison du substrat de A4YDT1. (B) Docking des intermédiaires 4-aminobutyryl-AMP (en rose), 5-aminopentanoyl-AMP (vert pomme) et 6-aminohexanoyl-AMP (cyan) dans le site de liaison de substrat de A4YDT1. Les résidus constituant le site de liaison du substrat sont colorés en bleu.

Cette modification a déjà été réalisée chez des CoA ligases acceptant des substrats à chaîne courte comme les acétyl-CoA ligases de *Methanothermobacter thermautotrophicus* Z245 (MT-ACS) et de *Arabidopsis thaliana* (atACS) (Ingram-Smith, Woods et al. 2006, Sofeo, Hart et al. 2019). Ces dernières sont passées d'une activité limitée à l'acétate à une activité sur de plus longs acides carboxyliques jusqu'à l'octanoate, mais également sur des chaînes ramifiées dans le cas de MT-ACS. Cette mutation W259G permet d'accéder à de nouveaux résidus qui étaient auparavant masqués par le résidu tryptophane et qui pourraient intervenir dans l'ancrage du groupement amine terminal dans le site actif par introduction de groupements polaires.

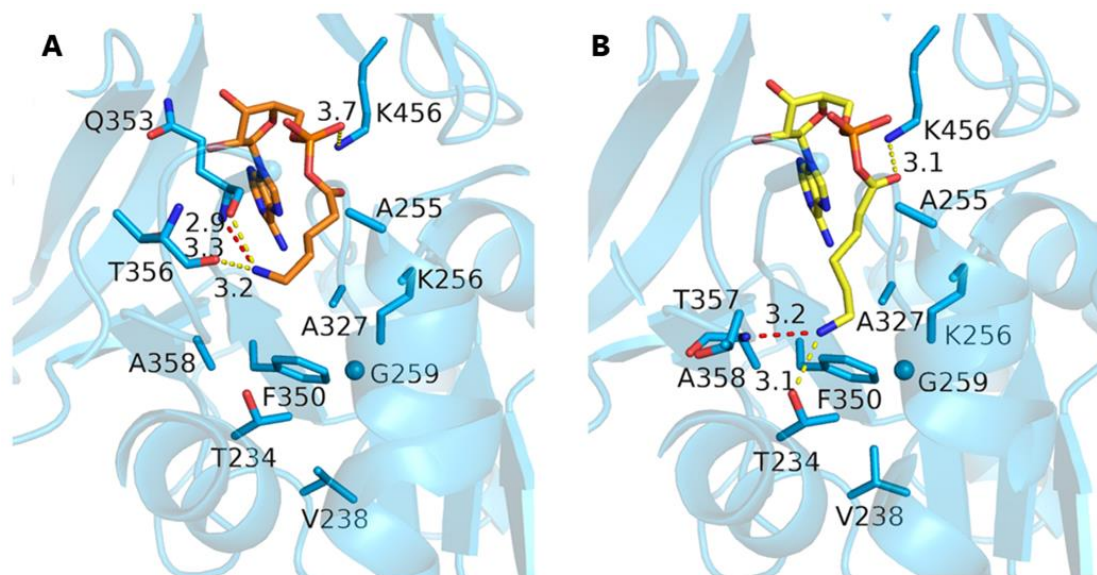


Figure 20. Docking des intermédiaires 5-aminopentanoyl-AMP en orange (A) et 6-aminohexanoyl-AMP (B) en jaune en présence de la mutation W259G représentée par une sphère bleue.

Pour identifier d'autres résidus cibles à muter accessibles au substrat dans le mutant W259G, les intermédiaires aminoacyl-AMP ont été dockés (Figure 20). Les résidus V238 et F350 ont ainsi été ciblés et les mutations proposées ont été V238T et F350Y, afin de conserver les effets stériques tout en introduisant une fonction polaire pour interagir avec le groupement amine des intermédiaires aminoacyl-AMP. Les doubles mutants W259G/V238T et W259G/F350Y ont également été proposés pour étudier un éventuel effet synergique des mutations.

### II.2.2.2. MUTAGENESE DIRIGEE DU SITE DE LIAISON DU SUBSTRAT ET SUREXPRESSON DES MUTANTS

Les simples mutants W259G, V238T et F350Y et les doubles mutants W259G/V238T et W259G/F350Y ont été produits par mutagenèse directe. Pour cela, les kits d'Agilent « QuikChange® II XL Site-Directed Mutagenesis Kit » et « QuikChange® Multi Site-Directed Mutagenesis Kit » ont été utilisés. Au préalable, les amorces contenant chaque mutation ont été conçues et optimisées avec l'aide de Jean-Louis Petit de la plateforme FunDivEx de l'unité. Le protocole fourni par le fournisseur a été utilisé, avec une 1<sup>ère</sup> étape consistant en une amplification PCR du gène contenant la mutation en appliquant des cycles thermiques. Un cycle thermique se découpe en trois étapes : la dénaturation de l'ADN puis hybridation des amorces et élongation des primers par l'ADN polymérase. Ensuite, le plasmide template est détruit par digestion par une enzyme de restriction DpnI avant de transformer celui contenant la mutation dans des cellules compétentes XL Gold fournies avec le kit et de les étaler sur une boîte agar-LB (Schéma 61).

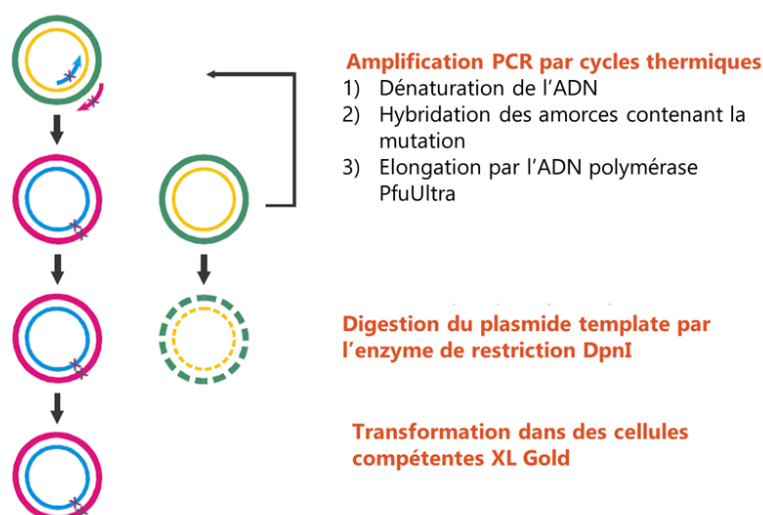


Schéma 61. Protocole de la méthode « QuikChange<sup>®</sup> II site-directed mutagenesis » (Agilent) appliquée à la CoA ligase A4YDT1. Schéma issu du manuel d'instruction du kit.

Suite à la transformation, huit colonies isolées ont été sélectionnées pour chaque mutant et mises en culture afin d'extraire l'ADN pour les analyser en séquençage et vérifier que ces colonies ont bien intégré le plasmide avec le gène mutant. Les colonies positives ont été cultivées et surexprimées selon le protocole utilisé en routine par la plateforme FunDivEx, décrit dans la partie expérimentale du manuscrit. Enfin, les enzymes ont été purifiées par mini-colonne Ni-NTA manuellement pour une petite production ou par chromatographie IMAC réalisée par Peggy Sirvain de l'équipe LGBM (Laboratoire de Génomique et Biochimie du Métabolisme) de l'unité. Les gels SDS-PAGE des mutants sont présentés en annexe de ce manuscrit.

## II.3. SYNTHÈSE DE LACTAMES ET D'AMIDES AVEC LES MUTANTS DE A4YDT1

### II.3.1. ACTIVITÉ DES MUTANTS POUR LA SYNTHÈSE DE LACTAMES

A4YDT1 et chaque mutant ont été testés pour la formation de lactames en 24 h à 60°C dans des tubes Eppendorf contenant une solution aqueuse de 5 mM de substrat acide  $\omega$ -aminé, 5 mM d'ATP, 5 mM  $MnCl_2$ , 50 mM tampon MOPS pH 8,5 et 2 mg/ml d'enzyme purifiée. Ces conditions réactionnelles ont été mises au point durant la thèse de C. Lelièvre et les réactions ont été faites en duplicat. Le rendement analytique pour chaque lactame a été déterminé à partir d'une gamme de calibration avec les standards lactames commerciaux, par UHPLC-MS en mode APCI. Cette méthode de détection a également été développée durant la thèse C. Lelièvre et le gradient appliqué est détaillé dans la partie expérimentale. Les rendements après 24 h de réaction sont montrés dans l'histogramme ci-dessous (Figure 21).

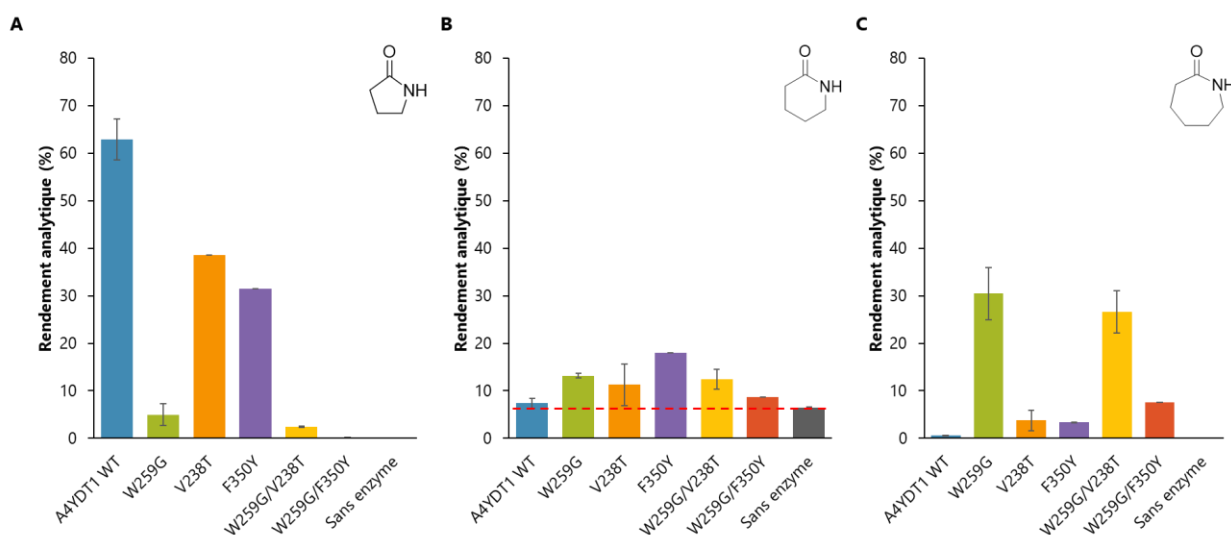


Figure 21. Rendements en  $\gamma$ -butyrolactame,  $\delta$ -valérolactame et  $\epsilon$ -caprolactame. Conditions réactionnelles : 5 mM substrat acide  $\omega$ -aminé, 5 mM ATP, 5 mM  $\text{MnCl}_2$ , 50 mM tampon MOPS pH 8.5 ; 2 mg/ml d'enzyme purifiée. 24h, 60°C, 400 rpm. Analyse par UHPLC-MS, mode APCI.

Pour la synthèse de  $\gamma$ -butyrolactame, le meilleur rendement a été obtenu avec l'enzyme *wild-type* A4YDT1 avec 63% de rendement, suivi par les simples mutants V238T et F350Y avec  $\approx 38\%$  et  $\approx 31\%$  de rendement respectivement. Ce résultat n'est pas surprenant compte tenu de l'activité d'A4YDT1 sur les substrats acides carboxyliques courts (Hawkins, Han et al. 2013, Hawkins, Adams et al. 2014). Pour le simple mutant W259G et les doubles mutants W259G/V238T et W259G/F350Y, le rendement en  $\gamma$ -butyrolactame diminue drastiquement passant à 5% et  $\approx 2\%$  avec W259G et W259G/V238T respectivement, et aucun rendement n'a été obtenu avec le double mutant W259G/F350Y. Le point commun entre ces trois mutants étant la mutation W259G, nous en avons donc déduit que l'élargissement de poche liée à cette mutation est défavorable pour la liaison de l'acide 4-aminobutyrique dans le site actif de A4YDT1. Il y a eu une perte d'interaction et d'affinité pour ce substrat dû à l'agrandissement de la poche du substrat. Ce même résultat a déjà été observé chez la CoA ligase *atACS* où la mutation du tryptophane plancher en glycine (W427G) a diminué l'activité native de l'enzyme limitée aux acides à chaîne courtes (jusqu'à C3) au profit d'acides de plus longues tailles et ramifiés (Sofeo, Hart et al. 2019). Avec l'acide 5-aminopentanoïque qui conduit à la formation de  $\delta$ -valérolactame, une importante cyclisation spontanée a été observée, représentant 6,5% de rendement. A4YDT1 WT n'a aucune activité sur l'acide 5-aminopentanoïque, tandis que les mutants sont légèrement actifs. Le meilleur résultat a été obtenu avec le simple mutant F350Y avec un modeste rendement de 18%, et  $\approx 13\%$  et  $\approx 11\%$  pour les mutants W259G et V238T respectivement. Les doubles mutants ont également donné des rendements similaires ou moindre avec  $\approx 12\%$  et  $\approx 9\%$  de rendement, il n'y a donc pas d'effet coopératif des mutations pour améliorer le rendement en  $\delta$ -valérolactame. Enfin, les mutants sont également actifs avec l'acide 6-aminohexanoïque pour la synthèse de caprolactame. Le rendement atteint jusqu'à  $\approx 30\%$  et  $\approx 27\%$  pour le simple mutant W259G et le double mutant W259G/V238T respectivement, ce qui montre que l'élargissement de la poche est nécessaire pour accepter des acides  $\omega$ -aminés de plus longue taille, jusqu'à C6 dans notre cas. Les simples mutants V238T et F350Y n'ont donné que de faibles rendements en caprolactame, avec  $\approx 4\%$  et  $\approx 3\%$  respectivement, et il n'y a aucune activité avec A4YDT1 WT, ce qui était attendu. Les doubles mutations n'ont pas apporté pas l'effet escompté. En effet nous n'observons aucun effet bénéfique de la mutation V238T dans le double mutant W259G/V238T qui permet la formation du caprolactame avec un rendement  $\approx 27\%$ , donc avec un rendement à peine inférieur à celui obtenu avec le simple mutant ( $\approx 30\%$ ). Quant à la mutation F350Y son effet est même délétère puisque le rendement en lactame n'est que de  $\approx 8\%$  dans le double mutant



W259G/F350Y. Notre hypothèse que ces résidus permettraient d'ancrer le groupement amine terminal du substrat acide  $\omega$ -aminé et d'avoir une meilleure activité ne se vérifie donc pas.

### II.3.2. SYNTHÈSE DE LACTAMES AVEC LA 2<sup>ÈME</sup> SÉRIE DE MUTANTS DE A4YDT1

Les premiers résultats n'ayant pas montré d'effet bénéfique des doubles mutations, nous avons voulu tester de nouvelles mutations. La mutation V238T étant quasi sans effet tandis que la F350Y a un effet négatif, nous avons décidé de combiner le double mutant W259G/V238T avec de nouvelles mutations en position de F350. Puisque la présence d'un résidu polaire d'encombrement stérique très similaire au le résidu du WT est délétère à l'activité lorsque ce résidu est accessible par la mutation W259G, nous avons pensé introduire des résidus de plus petite taille, apolaire ou polaire. Trois nouvelles mutations ont donc été proposées : F350A pour l'agrandissement de la poche, F350V idem avec un effet moins prononcé que l'alanine, enfin F350T afin d'introduire un groupement polaire pouvant entrer en interaction avec le substrat tout maintenant l'agrandissement de la poche. Au final, les trois simples mutants F350X et les trois triples mutants W259G/V238T/F350X avec X = A, V, T ont été produits en utilisant le kit QuikChange® Site-Directed Mutagenesis en partant de l'ADN natif d'A4YDT1 ou celui du mutant W259G/V238T (Annexe 1).

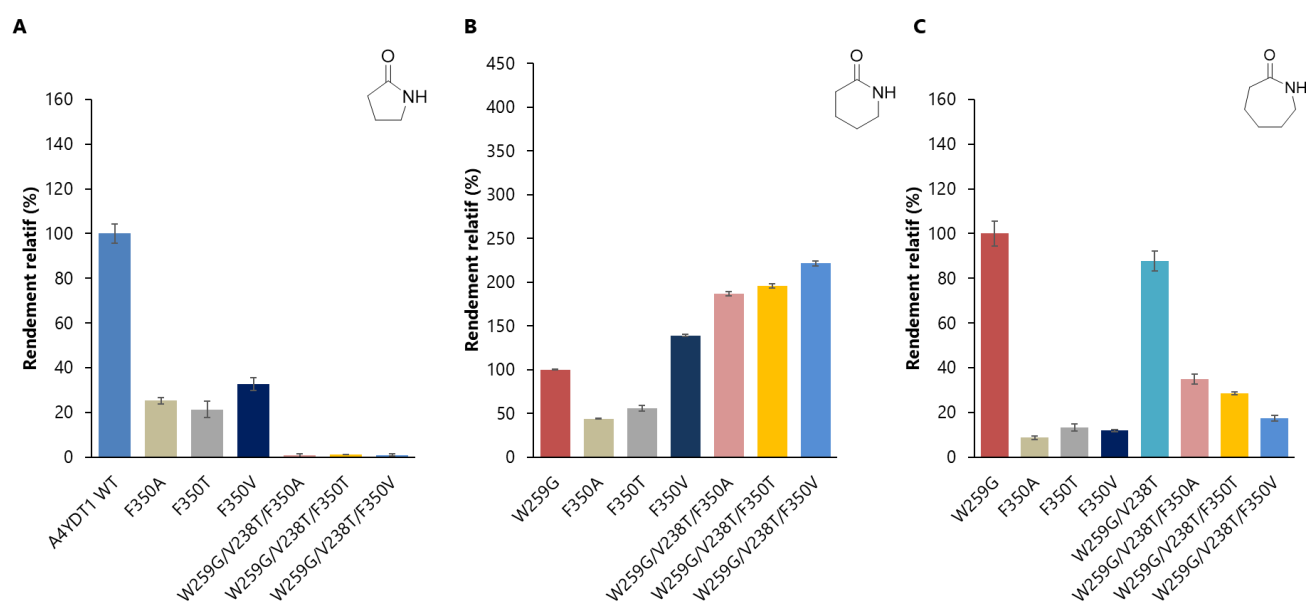


Figure 22. Rendements de la 2<sup>ème</sup> série de mutants en relatif par rapport à A4YDT1 WT pour la synthèse de  $\gamma$ -butyrolactame (62% de rendement), et W259G pour la synthèse de  $\delta$ -valérolactame (13% de rendement) et de  $\epsilon$ -caprolactame (30% de rendement). Conditions réactionnelles : 5 mM substrat acide  $\omega$ -aminé, 5 mM ATP, 5 mM  $MnCl_2$ , 50 mM tampon MOPS pH 8.5, 2 mg/ml d'enzyme purifiée. 60°C, 24h, 400 rpm. Analyse par UHPLC-MS (APCI).

La figure 22 présente les rendements obtenus avec les simples mutants F350X et les triples mutants W259G/V238T/F350X en relatif par rapport au meilleur rendement obtenu avec A4YDT1 pour la synthèse du  $\gamma$ -butyrolactame ; et par rapport à W259G pour la synthèse de  $\delta$ -valérolactame et  $\epsilon$ -caprolactame. Nous avons choisi de comparer les rendements obtenus pour la synthèse de  $\delta$ -valérolactame et  $\epsilon$ -caprolactame avec W259G comme il s'agissait du meilleur mutant pour la synthèse de  $\epsilon$ -caprolactame. Comme attendu, aucun des mutants n'a montré d'activité supérieure à celle de l'enzyme native pour la formation de  $\gamma$ -butyrolactame. Aucune activité n'a été observée avec les triples mutants W259G/V238T/F350X dû à l'élargissement de la poche par la mutation W259G, qui ne permet vraisemblablement pas l'ancrage du substrat, comme cela a été observé précédemment avec la première

série de mutants. Pour la formation de  $\delta$ -valérolactame avec les simples mutants, F350V a été le plus actif et nous avons pu observer une légère amélioration du rendement avec les triples mutants, notamment un gain allant de 10 à 20% avec le triple mutant W259G/V238T/F350V. Ces résultats étant relatifs par rapport à W259G, et puisque le rendement est seulement de 13% avec ce mutant, les rendements obtenus avec ces nouveaux mutants restent faibles comparés à ce qui a été obtenu avec le  $\epsilon$ -caprolactame (Annexe 7). Enfin, pour la formation de  $\epsilon$ -caprolactame, aucun simple mutant F350X n'a été plus actif que le simple mutant W259G, ce qui met de nouveau en évidence que l'élargissement de la poche est nécessaire pour accommoder de plus long acides  $\omega$ -aminés. Par ailleurs, les trois mutations combinées n'ont montré aucun effet combiné comme nous l'observons sur le 3<sup>ème</sup> histogramme. En effet, nous pouvons voir que l'ajout de la mutation F350X a été défavorable et a diminué drastiquement l'activité du double mutant W259G/V238T. Le résidu à cette position a certainement un rôle important, mais nous ne l'avons pas identifié à ce jour.

### II.3.3. SPECTRE D'ACTIVITE DES MUTANTS D'A4YDT1 SUR DES SUBSTRATS ACIDES POUR LA SYNTHÈSE D'AMIDES

Durant sa thèse, Dr Chloé Lelièvre a étudié le spectre d'activité de A4YDT1 sur différents acides carboxyliques dans une réaction modèle en présence de méthylamine en excès (Schéma 62). Pour rappel, la CoA ligase est employée pour l'activation de l'acide et l'attaque nucléophile de la méthylamine est spontanée pour former le N-méthylbutyramide correspondant. Il avait été déterminé que l'activité de A4YDT1 était limitée à des substrats de taille moyenne avec des ramifications et des acides courts portant des groupements aromatiques comme l'acide benzoïque ou l'acide phénylpropanoïque. Ces résultats sont en accord avec les résultats obtenus précédemment (Hawkins, Han et al. 2013).

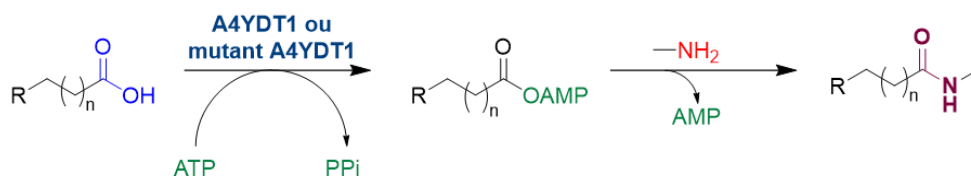


Schéma 62. Synthèse de N-méthylamides avec A4YDT1 WT ou les mutants A4YDT1.

Nous avons vu que les mutants W259G et W25G/V238T permettent une meilleure activité sur les acides  $\omega$ -aminés à 6 et 7 chaînons. L'agrandissement de la poche de liaison du substrat devrait permettre d'avoir une réactivité sur des substrats acides carboxyliques simple avec une chaîne longue. Nous avons donc testé des acides carboxyliques de taille variée, plus ou moins encombrés, toujours en présence de méthylamine (Figure 23).

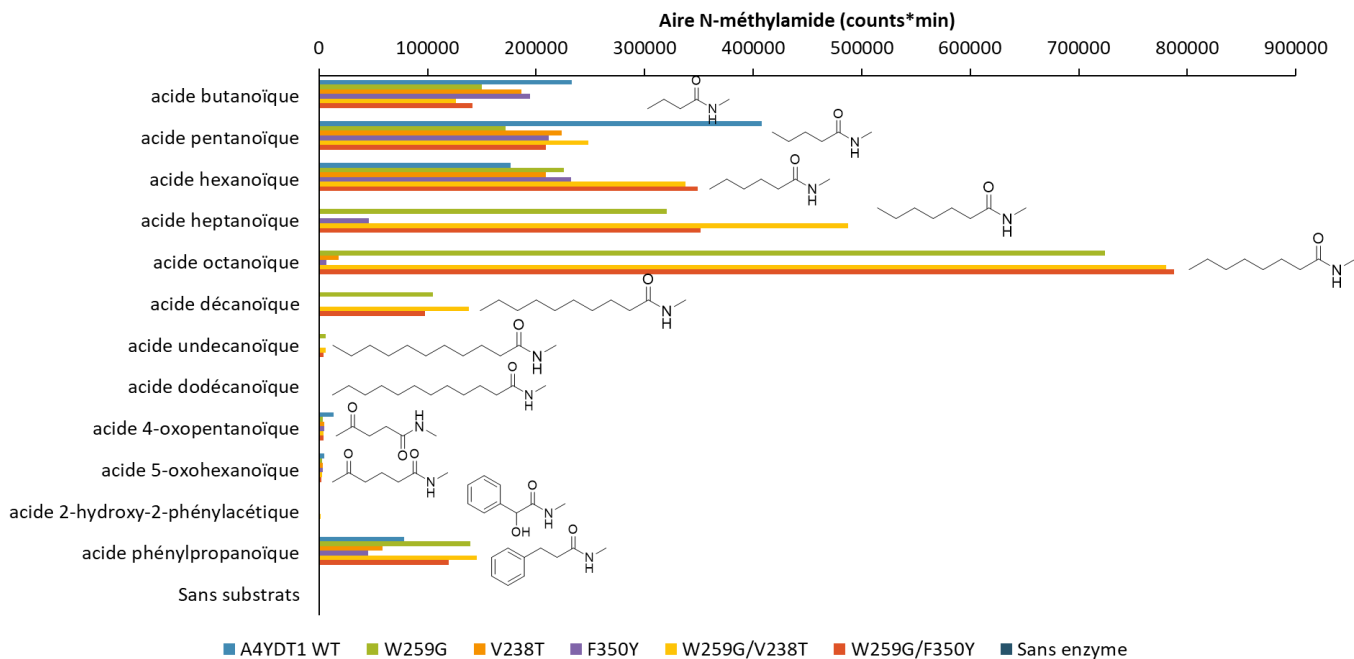


Figure 23. Comparaison du spectre d'activité de A4YDT1 et des mutants A4YDT1 avec divers substrats acides par détection du produit N-méthylamide correspondant. Conditions réactionnelles : 5 mM substrat acide, 5 mM ATP, 5 mM MgCl<sub>2</sub>, 50 mM méthylamine, 50 mM tampon phosphate pH 8, 0.2 mg/ml d'enzyme purifiée. 60°C, 24h, 400 rpm. Détection déduite d'injections de standards synthétisés au laboratoire du C4 au C8 et la présence d'une aire au m/z attendue pour chaque N-méthylamide par UPLC-MS (ESI).

Les résultats présentés ci-dessus correspondent à la réponse en aire obtenue pour chaque N-méthylamide. Des essais de quantifications ont été réalisés pour la formation des N-méthylamides obtenus avec les acides carboxyliques de C4 à C8. Les réponses en aire étant différentes d'un N-méthylamide à un autre, il n'était donc pas possible de déduire une droite de calibration moyenne pour comparer tous nos résultats. Pour l'analyse de ces résultats, nous pouvons donc comparer les réponses en aire entre les enzymes pour un substrat donné, mais pas les aires obtenues pour chaque substrat pour une enzyme donnée. Dans le cas de A4YDT1 WT, nous retrouvons le même spectre d'activité que précédemment, qui s'étend du C4, correspondant à environ 40% de rendement analytique, jusqu'au C6. Avec les mutants, le spectre d'activité s'étend jusqu'à l'acide undécanoïque (C11) avec le simple mutant W259G et les deux doubles mutants W259G/V238T et W259G/F350Y. L'élargissement du site de liaison du substrat portée par la mutation W259G est majoritairement responsable de l'accès à de plus grands amides comme nous avons pu l'observer précédemment pour la synthèse de lactames. De surcroît, les doubles mutants W259G/V238T et W259G/F350Y sont plus actifs que le simple mutant W259G, ce qui laisse à penser que ces mutations pourraient contribuer, avec W259G, à l'activité sur des substrats plus longs. Une très faible, voire aucune activité, n'a été détectée pour les acides 4-oxopentanoïque, 5-oxohexanoïque et l'acide 2-hydroxy-2-phénylacétique, aussi bien avec A4YDT1 WT qu'avec les mutants.

### II.3.4. PARAMETRES CINETIQUES DES MEILLEURS MUTANTS

Les paramètres cinétiques ont été déterminés pour A4YDT1 WT avec l'acide 4-aminobutyrique ainsi que pour les deux meilleurs mutants pour la synthèse de  $\epsilon$ -caprolactame : le simple mutant W259G et le double mutant W259G/V238T.

Il existe de nombreuses méthodes de suivi d'activité des CoA ligases afin de déterminer leurs paramètres cinétiques dont la plus commune consiste à suivre l'étape de thioestérification par

spectrophotométrie en suivant la formation de la liaison thioester ou la consommation du HSCoA. Comme les CoA ligases sont justes employées pour l'étape d'adénylation pour la synthèse de lactames et d'amides, nous devons disposer d'un test qui reflète leur activité. Pour rappel, cette étape permet d'activer l'acide carboxylique en présence d'ATP pour former l'intermédiaire (amino)acyl-AMP, libérant du pyrophosphate (PPi). L'addition de l'amine est, quant à elle, spontanée et ne fait donc pas intervenir la CoA ligase.

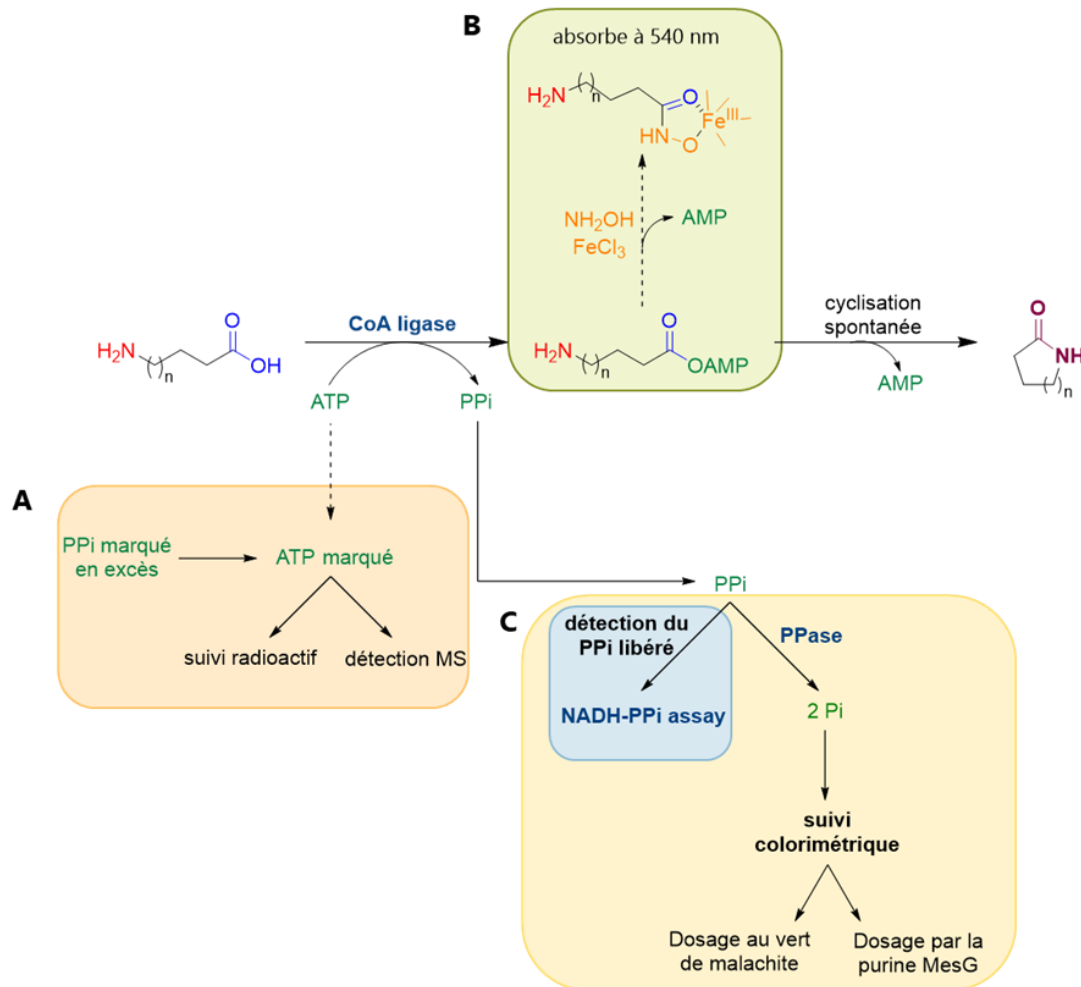


Schéma 63. Méthodes de suivi cinétique de l'étape d'adénylation des CoA ligases et des enzymes adénylantes : par marquage de l'ATP (A), détection de l'intermédiaire aminoacyl-AMP (B) et par détection du pyrophosphate libéré (C). Figure inspirée de l'article de Kittila (Kittila *et al.*, 2016).

Différentes méthodes de suivi de l'adénylation ont été développées comme la méthode d'échange ATP-PPi, basée sur la réversibilité de cette étape pour quantifier l'ATP radiomarquée (Otten, Schaffer et al. 2007). Ce principe nécessite de travailler avec des éléments radioactifs [<sup>32</sup>P]PPi, et sa mise en place peut être laborieuse (Schéma 63A). Les autres méthodes sont des méthodes indirectes par dosage des produits de la réaction d'adénylation : l'intermédiaire (amino)acyl-AMP et le PPI. L'intermédiaire (amino)acyl-AMP peut être dosé à l'aide d'hydroxylamine (NH<sub>2</sub>OH) qui fait une attaque nucléophile sur l'intermédiaire, entraîne le départ de l'AMP et forme un dérivé d'acide acétohydroxamique. Cet acide va ensuite se complexer avec l'ion ferrique Fe<sup>3+</sup> pour produire un complexe ferrique tris-hydroxamate détectable spectrophotométriquement à 540 nm (Schéma 62B). Le dosage du PPI peut être réalisé de façon indirecte par le dosage spectrophotométrique de phosphate inorganique (Pi) libéré par l'hydrolyse du PPI (Schéma 63C). Le Pi peut, alors, être complexé avec du vert de malachite, initialement de couleur orange en solution acide, qui va alors changer de couleur et

devenir vert en présence de phosphate inorganique dû à un changement de pH (absorbance à 623 nm). Basé sur ce même principe, le dosage par le 2-amino-6-mercapto-7-méthylpurineribonucleoside ou MesG implique une complexation avec le Pi et, sous l'action d'une purine nucléoside phosphorylase, avec une libération du complexe ribose-phosphate d'une part et la partie purine libre d'autre part, détectée spectrophométriquement à 360 nm. Ces deux méthodes sont commerciales sous forme de kit, le « Malachite Green Phosphate Assay Kit® » et le « EnzCheck® Phosphate Assay Kit » et sont un peu onéreuses. Nous avons décidé d'utiliser un 3<sup>ème</sup> type de test, le NADH-PPi Assay, dans lequel la formation de pyrophosphate (PPi) est couplée à la consommation de NADH par quatre étapes enzymatiques permettant un suivi spectrophotométrique à 340 nm (Schéma 63C) (Kittilä, Schoppet et al. 2016).

### II.3.4.1. PRINCIPE DU SUIVI CINÉTIQUE DE L'ADÉNYLATION PAR LA MÉTHODE NADH-PPi

Le test NADH-PPi consiste en une cascade de 4 enzymes bien connues car ce sont des enzymes qui sont impliquées dans la glycolyse : la fructose-6-phosphate kinase dépendante du pyrophosphate (PPi-PFK), l'aldolase, la triosephosphate isomérase (TPI) et la glycérophosphate déshydrogénase (GPDH) (Schéma 64).

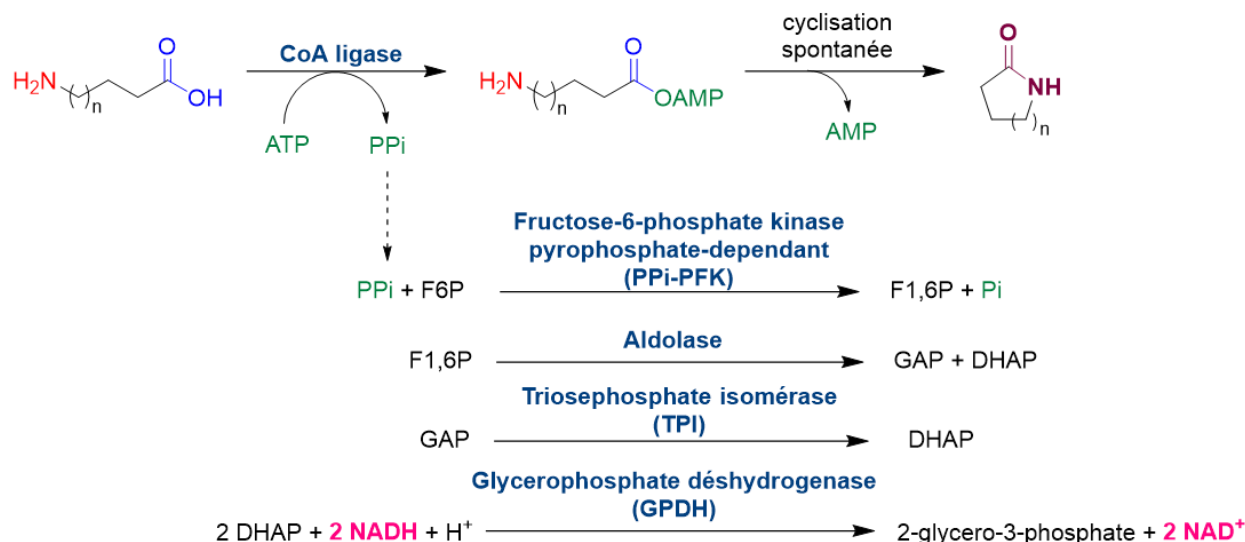


Schéma 64. Cascade enzymatique du test NADH-PPi pour le suivi cinétique de l'étape d'adénylation des CoA ligases.

Contrairement aux méthodes détaillées précédemment, la cascade est déclenchée directement à partir du pyrophosphate et nous suivons spectrophométriquement la consommation de NADH à 340 nm. Ces enzymes ayant un optimum d'activité à 37°C, elles ne sont donc pas compatibles pour un dosage en temps réel avec les CoA ligases thermophiles à 60°C. Nous nous sommes donc tournés vers une méthode en point final pour déterminer les paramètres cinétiques de A4YDT1 et ses mutants. Suite à l'arrêt de la vente unitaire de la PPi-PFK, nous avons utilisé le kit commercial Pyrophosphate Reagent® vendu par Sigma-Aldrich. Les constantes cinétiques ont été calculées à l'aide de la formule ci-dessous donnée par le fournisseur :

$$\frac{\Delta Abs - V_{total}}{6.22 \times 2 \times V_{\text{échantillon}}}$$

### II.3.4.2. PARAMETRES CINETIQUES DE A4YDT1 ET DES MUTANTS W259G ET W259G/V238T

Le tableau ci-dessous résume les données cinétiques ( $K_M$ ,  $k_{cat}$  et  $k_{cat}/K_M$ ) obtenues pour l'activité de A4YDT1 WT et des mutants les plus performants pour la synthèse de  $\epsilon$ -caprolactame : W259G et W259G/V238T. Pour l'enzyme native A4YDT1, les constantes cinétiques ont été déterminées pour son activité envers le GABA, tandis que pour les mutants W259G et W259G cela a été fait sur le 6-AHA.

Tableau 3. Paramètres cinétiques de l'enzyme native A4YDT1 avec l'acide 4-aminobutyrique (GABA) et des mutants W259G et W259/V328T avec l'acide hexanoïque (6-AHA). Détermination des données cinétiques en utilisant le kit Pyrophosphate Reagent® en méthode de point final. Expériences faites en duplicat.

	Substrat	$k_{cat}$ ( $s^{-1}$ )	$K_M$ (mM)	$k_{cat}/K_M$ ( $s^{-1}.M^{-1}$ )
<b>A4YDT1 WT</b>	ATP	$0,014 \pm 0,001$	$0,06 \pm 0,01$	$2,2.10^2$
	GABA	$0,021 \pm 0,001$	$0,87 \pm 0,20$	24,6
<b>W259G</b>	ATP	$3,6 \pm 0,2$	$0,12 \pm 0,06$	$3,0.10^4$
	6-AHA	$3,8 \pm 0,5$	$279 \pm 85$	13,4
<b>W259G/V238T</b>	ATP	$5,6 \pm 0,5$	$0,15 \pm 0,05$	$3,7.10^4$
	6-AHA	$0,80 \pm 0,04$	$220 \pm 121$	3,6

Tableau 4. Données cinétiques de A4YDT1 d'après l'article de Hawkins *et al.* (Hawkins, Han et al. 2013) et de quelques CoA ligases décrites dans la littérature (Wu, Cao et al. 2008, Zhang, Barajas et al. 2017, Sofeo, Hart et al. 2019).

	Substrat	$k_{cat}$ ( $s^{-1}$ )	$K_M$ (mM)	$k_{cat}/K_M$ ( $s^{-1}.M^{-1}$ )
<b>CBL (RpACL)</b> 4-chlorobenzoate-CoA ligase de <i>Rhodopseudomonas palustris</i>	ATP	$13,6 \pm 0,2$	$1,2.10^{-1} \pm 6.10^{-3}$	$1,0.10^5$
	4-chlorobenzoate	$14,2 \pm 0,3$	$1,4.10^{-3} \pm 0,1.10^{-3}$	$1,0.10^7$
<b>MT-ACS</b> Acétyl-CoA ligase d' <i>Arabidopsis thaliana</i>	ATP	$3,56 \pm 0,12$	$0,20 \pm 0,03$	$1,8.10^4$
	acétate	$2,41 \pm 0,03$	$0,27 \pm 0,001$	$8,9.10^4$
	propionate	$0,66 \pm 0,07$	$4,0 \pm 1,2$	$1,7.10^2$
<b>ORF26</b> Acyl-CoA ligase de <i>Streptomyces aizunensis</i>	GABA	$0,15 \pm 5,8.10^{-3}$	$4,96 \pm 0,56$	30,2
	5-APA	$0,13 \pm 2.10^{-3}$	$0,08 \pm 0,01$	$1,5.10^3$
	6-AHA	$0,098 \pm 2,8.10^{-3}$	$0,16 \pm 0,03$	$6,1.10^2$
<b>A4YDT1</b> Résultats issus de l'article de Hawkins <i>et al.</i>	4-hydroxybutyrate	1,8	2,0	$9,1.10^2$
	butyrate	8,4	0,32	$2,6.10^4$
	valérate	5,6	0,74	$7,5.10^3$

Pour l'enzyme native A4YDT1, le  $K_M$ (ATP) est de 0,06 mM (Tableau 3). Cette valeur se retrouve dans la gamme de valeurs de  $K_M$  que l'on peut retrouver pour d'autres CoA ligases comme MT-ACS ou CBL présentes dans le tableau 4 ou encore d'autres CoA ligases dont les données cinétiques sont répertoriées dans la base de données BRENDA (Tableau 4). Nous restons également dans cette même gamme de valeur de  $K_M$ (ATP) avec les mutants, même s'ils ont légèrement moins affins avec 0,12 et 0,15 mM respectivement. Cela montre que l'affinité de la CoA ligase pour l'ATP n'a pas été affectée par les mutations. De manière plus globale, nous avons noté que les mutants possèdent une faible activité et affinité, mais cela n'est pas au détriment de leur potentielle application en biocatalyse. En effet, il a déjà été démontré que des enzymes avec de faibles affinités pouvaient être avantageuses en synthèse biocatalytique en présence de fortes concentration de substrats sans qu'il n'y ait d'effet d'inhibition (Mayol, David et al. 2016).

Si nous comparons les données cinétiques d'A4YDT1 pour le GABA avec celles de ORF26, l'une des rares CoA ligases décrites pour la synthèse de lactames, nous pouvons noter que A4YDT1 WT possède une meilleure affinité pour ce substrat que ORF26. Cependant, son  $k_{cat}$  est plus faible et finalement nous nous retrouvons des efficacités catalytiques similaires entre ces deux CoA ligases.

Avec le 6-AHA les deux mutants ont une affinité similaire pour ce substrat, 279 mM et 220 mM pour W259G et W259/V238T respectivement, mais bien moindre que celle d'ORF26 à 0,16 mM. Les constantes catalytiques sont inférieures avec les mutants. Cela amène donc à une faible efficacité catalytique  $k_{cat}/K_M$ , inférieure à ORF26 active sur ce même substrat.

Les activités spécifiques ont également été déterminées pour les mutants de la 1<sup>ère</sup> série W259G, W259G/F350Y et W259G/V238T. Les activités ont été mesurées sur les acides hexanoïque, heptanoïque et octanoïque et les résultats sont présentés dans le tableau ci-dessous (Tableau 5).

Tableau 5. Activité spécifique des mutants de A4YDT1, le simple mutant W259G et les doubles mutants W259G/V238T et W259G/F350Y. Expériences faites en duplicat avec le kit Pyrophosphate Reagent®.

	<b>Activité spécifique (mU/mg d'enzyme)</b>		
	<b>W259G</b>	<b>W259G/F350Y</b>	<b>W259G/V238T</b>
acide hexanoïque	101 ± 11	86 ± 3	71 ± 16
acide heptanoïque	157 ± 73	156 ± 73	131 ± 49
acide octanoïque	198 ± 22	166 ± 43	194 ± 22

Globalement, les résultats sont cohérents avec les conversions observées lors de l'analyse des N-méthylamides par UHPLC-MS. Pour un substrat donné, les activités spécifiques sont très similaires pour les trois mutants, indiquant que la deuxième mutation, F350Y ou V238T, n'a presque pas d'effet. Ces données nous permettent de conclure que les mutants sont, comme attendu, plus actifs vis-à-vis des acides carboxyliques en C8 qu'en C6.

## **II.4. CONCLUSION ET DISCUSSION SUR L'INGENIERIE PROTEIQUE DE L'ENZYME A4YDT1**

Grâce à la structure tridimensionnelle de l'enzyme A4YDT1 obtenue par nos collaborateurs, Dr Nikolas Capra, Pr Andy-Mark Thunnissen et Pr Dick Janssen et aux expériences de modélisation moléculaire qu'ils ont menées, nous avons pu utiliser une approche rationnelle d'ingénierie protéique afin d'étendre le spectre de l'enzyme. L'étude structurale a permis d'identifier les facteurs structuraux jouant un rôle dans la spécificité de substrats de A4YDT1, notamment le résidu tryptophane qui forme un plancher dans la poche de liaison du substrat dans le site actif de A4YDT1. Ce résidu a été alors muté en glycine afin d'agrandir la poche, stratégie qui avait déjà été utilisée pour l'étude d'autres CoA ligases (Ingram-Smith, Woods et al. 2006, Reger, Carney et al. 2007, Sofeo, Hart et al. 2019). Il en résulte que la CoA ligase native limitée à la synthèse de  $\gamma$ -butyrolactame, a vu son activité étendue à la synthèse de  $\delta$ -valérolactame et  $\epsilon$ -caprolactame avec 30% de rendement. Cela est d'autant plus intéressant que ce dernier lactame a un intérêt industriel car il est le précurseur du nylon-6,6, très utilisé dans les domaines des matériaux et du textile. Ce résultat permet d'allonger la courte liste des CoA ligases connues pour la synthèse de lactames comme ORF26 et msed issu du même organisme que A4YDT1 (Zhang, Barajas et al. 2017, Shen, Jiang et al. 2024). Cependant, les rendements obtenus restent modestes comparés à ceux

obtenus avec ces CoA ligases mais aussi avec les CARs utilisés pour la synthèse de  $\delta$ -valérolactames chiraux avec des rendements allant jusqu'à 85% (Qin, Zhang et al. 2022). Même si ces travaux avaient pour but de permettre la synthèse de lactames, les mutants ont également été testés sur des acides carboxyliques et nous avons montré que l'élargissement de la poche permet une activité sur les substrats longs. Ces résultats ont fait l'objet d'un article paru dans *Frontiers in Catalysis* (Capra, Lelièvre et al. 2024).

En outre, bien que les structures d'un certain nombre de CoA ligases ont été résolues et que les substrats de ces enzymes sont connus, il n'y en a pas suffisamment pour bien comprendre les spécificités de substrats des CoA ligases et encore moins pour pouvoir faire de la prédiction de spécificité de CoA ligase à partir de séquence protéique de manière universelle (Gulick, Starai et al. 2003, Reger, Wu et al. 2008, Wu, Cao et al. 2008, Shah, Ingram-Smith et al. 2009, Sofeo, Hart et al. 2019). Il existe un logiciel nommé AdenylPred (<https://srobinson.shinyapps.io/AdenylPred/>) qui permet de faire de la prédiction de la classe fonctionnelle et de la spécificité de substrats des enzymes de la superfamille ANL, dont des CoA ligases (Robinson, Terlouw et al. 2020). Cet outil basé sur le *machine learning*, a intégré plus de 1 700 séquences peptidiques d'enzymes de la superfamille ANL avec les données structurales et cinétiques ainsi que sur leur classe fonctionnelle et leur spécificité de substrats pour celles qui ont été décrites dans la littérature. Nous avons testé AdenylPred avec un set de CoA ligases : la bonne classe d'enzyme a été prédite avec un bon score mais la prédiction de spécificité de substrats était relativement peu fiable pour la majorité des enzymes. Cela met donc en avant l'intérêt et l'importance d'étudier la spécificité de substrats de CoA ligases de façon expérimentale afin que ces données puissent être incluses dans l'apprentissage machine du logiciel et améliorer leur prédiction.

Pour conclure, cette approche guidée par la structure de la CoA ligase pour étendre l'application en synthèse biocatalytique de lactames est complémentaire de la seconde approche utilisée durant la thèse. Dans ce chapitre, nous sommes partis de la structure de l'enzyme en ciblant un substrat cible sur lequel nous voulions que la CoA ligase soit active. Dans la partie exploratoire, nous avons criblé une collection de CoA ligases issues de la biodiversité répertoriée sur un panel d'acides  $\omega$ -aminés pour trouver de nouvelles enzymes avec de nouvelles spécificités de substrats. Cette approche est détaillée dans le chapitre suivant.



## REFERENCES BIBLIOGRAPHIQUES

Bell, E. L., W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niiikura, S. Osuna, E. Romero, K. S. Ryan, N. J. Turner and S. L. Flitsch (2021). "Biocatalysis." Nature Reviews Methods Primers **1**(1): 46.

Cadwell, R. C. and G. F. Joyce (1992). "Randomization of genes by PCR mutagenesis." Genome Research **2**(1): 28-33.

Capra, N., C. Lelièvre, O. Touré, A. Fossey-Jouenne, C. Vergne-Vaxelaire, D. B. Janssen, A.-M. W. H. Thunnissen and A. Zapparucha (2024). "Adapting an acyl CoA ligase from *Metallosphaera sedula* for lactam formation by structure-guided protein engineering." Frontiers in Catalysis **4**.

Chen, K. and F. H. Arnold (1993). "Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide." Proceedings of the National Academy of Sciences **90**(12): 5618-5622.

Clark, L., D. Leatherby, E. Krilich, A. J. Ropelewski and J. Perozich (2018). "In silico analysis of class I adenylate-forming enzymes reveals family and group-specific conservations." PLOS ONE **13**(9): e0203218.

Conti, E., T. Stachelhaus, M. A. Marahiel and P. Brick (1997). "Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S." The EMBO Journal **16**(14): 4174-4183.

Gulick, A. M. (2009). "Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase." ACS Chemical Biology **4**(10): 811-827.

Gulick, A. M., V. J. Starai, A. R. Horswill, K. M. Homick and J. C. Escalante-Semerena (2003). "The 1.75 Å Crystal Structure of Acetyl-CoA Synthetase Bound to Adenosine-5'-propylphosphate and Coenzyme A." Biochemistry **42**(10): 2866-2873.

Hawkins, A. B., M. W. W. Adams and R. M. Kelly (2014). "Conversion of 4-Hydroxybutyrate to Acetyl Coenzyme A and Its Anapleurosis in the *Metallosphaera sedula* 3-Hydroxypropionate/4-Hydroxybutyrate Carbon Fixation Pathway." Appl Environ Microbiol **80**(8): 2536-2545.

Hawkins, A. S., Y. Han, R. K. Bennett, M. W. W. Adams and R. M. Kelly (2013). "Role of 4-Hydroxybutyrate-CoA Synthetase in the CO<sub>2</sub> Fixation Cycle in Thermoacidophilic Archaea." Journal of Biological Chemistry **288**(6): 4012-4022.

Hisanaga, Y., H. Ago, N. Nakagawa, K. Hamada, K. Ida, M. Yamamoto, T. Hori, Y. Arai, M. Sugahara, S. Kuramitsu, S. Yokoyama and M. Miyano (2004). "Structural Basis of the Substrate-specific Two-step Catalysis of Long Chain Fatty Acyl-CoA Synthetase Dimer\*." Journal of Biological Chemistry **279**(30): 31717-31726.

Horne, J. and D. Shukla (2022). "Recent Advances in Machine Learning Variant Effect Prediction Tools for Protein Engineering." Industrial & Engineering Chemistry Research **61**(19): 6235-6245.

Ingram-Smith, C., B. I. Woods and K. S. Smith (2006). "Characterization of the Acyl Substrate Binding Pocket of Acetyl-CoA Synthetase." Biochemistry **45**(38): 11482-11490.

Jackson, D. R., S. S. Tu, M. Nguyen, J. F. Barajas, A. J. Schaub, D. Krug, D. Pistorius, R. Luo, R. Müller and S.-C. Tsai (2016). "Structural Insights into Anthranilate Priming during Type II Polyketide Biosynthesis." ACS Chemical Biology **11**(1): 95-103.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis (2021). "Highly accurate protein structure prediction with AlphaFold." Nature **596**(7873): 583-589.

Kittilä, T., M. Schoppet and M. J. Cryle (2016). "Online Pyrophosphate Assay for Analyzing Adenylation Domains of Nonribosomal Peptide Synthetases." Chembiochem **17**(7): 576-584.

Kochan, G., E. S. Pilka, F. von Delft, U. Oppermann and W. W. Yue (2009). "Structural Snapshots for the Conformation-dependent Catalysis by Human Medium-chain Acyl-coenzyme A Synthetase ACSM2A." Journal of Molecular Biology **388**(5): 997-1008.

Mayol, O., S. David, E. Darii, A. Debard, A. Mariage, V. Pellouin, J.-L. Petit, M. Salanoubat, V. de Berardinis, A. Zaparucha and C. Vergne-Vaxelaire (2016). "Asymmetric reductive amination by a wild-type amine dehydrogenase from the thermophilic bacteria *Petrotoga mobilis*." Catalysis Science & Technology **6**(20): 7421-7428.

Otten, L. G., M. L. Schaffer, B. R. M. Villiers, T. Stachelhaus and F. Hollfelder (2007). "An optimized ATP/PPi-exchange assay in 96-well format for screening of adenylation domains for applications in combinatorial biosynthesis." Biotechnology Journal **2**(2): 232-240.

Paul, C., U. Hanefeld, F. Hollmann, G. Qu, B. Yuan and Z. Sun (2024). "Enzyme engineering for biocatalysis." Molecular Catalysis **555**: 113874.

Penning, T. M. and J. M. Jez (2001). "Enzyme Redesign." Chemical Reviews **101**(10): 3027-3046.

Porter, J. L., R. A. Rusli and D. L. Ollis (2016). "Directed Evolution of Enzymes for Industrial Biocatalysis." ChemBioChem **17**(3): 197-203.

Qin, Z., X. Zhang, X. Sang, W. Zhang, G. Qu and Z. Sun (2022). "Carboxylic acid reductases enable intramolecular lactamization reactions." Green Synthesis and Catalysis.

Reger, A. S., J. M. Carney and A. M. Gulick (2007). "Biochemical and Crystallographic Analysis of Substrate Binding and Conformational Changes in Acetyl-CoA Synthetase." Biochemistry **46**(22): 6536-6546.

Reger, A. S., R. Wu, D. Dunaway-Mariano and A. M. Gulick (2008). "Structural Characterization of a 140° Domain Movement in the Two-Step Reaction Catalyzed by 4-Chlorobenzoate:CoA Ligase." Biochemistry **47**(31): 8016-8025.

Robinson, S. L., B. R. Terlouw, M. D. Smith, S. J. Pidot, T. P. Stinear, M. H. Medema and L. P. Wackett (2020). "Global analysis of adenylate-forming enzymes reveals  $\beta$ -lactone biosynthesis pathway in pathogenic *Nocardia*." Journal of Biological Chemistry **295**(44): 14826-14839.

Shah, M. B., C. Ingram-Smith, L. L. Cooper, J. Qu, Y. Meng, K. S. Smith and A. M. Gulick (2009). "The 2.1 Å crystal structure of an acyl-CoA synthetase from *Methanosarcina acetivorans* reveals an alternate

acyl-binding pocket for small branched acyl substrates." Proteins: Structure, Function, and Bioinformatics **77**(3): 685-698.

Shen, X., X. Jiang, X. Sun, Q. Yuan and J. Wang (2024). "Identifying and charactering a 4-aminobutyryl-CoA ligase for the production of butyrolactam." Journal of Biotechnology **382**: 21-27.

Sofoe, N., J. H. Hart, B. Butler, D. J. Oliver, M. D. Yandeau-Nelson and B. J. Nikolau (2019). "Altering the Substrate Specificity of Acetyl-CoA Synthetase by Rational Mutagenesis of the Carboxylate Binding Pocket." ACS Synthetic Biology **8**(6): 1325-1336.

Thornburg, C. K., S. Wortas-Strom, M. Nosrati, J. H. Geiger and K. D. Walker (2015). "Kinetically and Crystallographically Guided Mutations of a Benzoate CoA Ligase (BadA) Elucidate Mechanism and Expand Substrate Permissivity." Biochemistry **54**(40): 6230-6242.

Toscano, M. D., K. J. Woycechowsky and D. Hilvert (2007). "Minimalist Active-Site Redesign: Teaching Old Enzymes New Tricks." Angewandte Chemie International Edition **46**(18): 3212-3236.

Wu, R., J. Cao, X. Lu, A. S. Reger, A. M. Gulick and D. Dunaway-Mariano (2008). "Mechanism of 4-Chlorobenzoate:Coenzyme A Ligase Catalysis." Biochemistry **47**(31): 8026-8039.

Wu, R., A. S. Reger, X. Lu, A. M. Gulick and D. Dunaway-Mariano (2009). "The Mechanism of Domain Alternation in the Acyl-Adenylate Forming Ligase Superfamily Member 4-Chlorobenzoate: Coenzyme A Ligase." Biochemistry **48**(19): 4115-4125.

Zhang, J., J. F. Barajas, M. Burdu, G. Wang, E. E. Baidoo and J. D. Keasling (2017). "Application of an Acyl-CoA Ligase from *Streptomyces aizunensis* for Lactam Biosynthesis." ACS Synth Biol **6**(5): 884-890.

# CHAPITRE III : EXPLORATION DE LA BIODIVERSITE POUR LA RECHERCHE D'ENZYMES POUR LA SYNTHÈSE BIOCATALYTIQUE DE LACTAMES

Dans le chapitre précédent, nous avons modifié par ingénierie une CoA ligase (A4YDT1) pour la rendre active sur l'acide 6-aminohexanoïque, précurseur de  $\epsilon$ -caprolactame, utilisé pour obtenir du nylon-6,6. Dans ce second volet de la thèse, nous adoptons une approche exploratoire pour disposer de nouvelles CoA ligases actives sur divers acides  $\omega$ -aminés.

L'exploration de la biodiversité constitue une approche éprouvée pour découvrir de nouvelles enzymes capables de catalyser la synthèse de composés d'intérêt. Grâce entre autres aux campagnes de séquençage massives, le nombre de séquences protéiques d'enzymes disponibles ne cesse d'augmenter (Buller, Lutz et al. 2023). Ces enzymes, issues de divers environnements peuvent posséder des caractéristiques biochimiques uniques ou des spécificités de substrats particulières. En tirant à profit cette diversité enzymatique naturelle, il est possible d'élargir le répertoire d'outils disponibles pour la biocatalyse et d'améliorer des méthodes de synthèse chimique. Cela passe notamment par le criblage de cette diversité à la recherche de l'activité désirée avant d'optimiser les conditions réactionnelles avec les enzymes candidates.

Au laboratoire, une première collection de CoA ligases issues de la biodiversité, nommée CoA Lig1, a été constituée. Cette collection a été construite et criblée durant la thèse de Chloé Lelièvre, à la recherche d'enzymes actives pour la synthèse de lactames simples de 5 à 7 chaînons. Dans un second temps, nous avons criblé cette collection sur des acides  $\omega$ -aminés fonctionnalisés. Par itération, ces enzymes identifiées ont ensuite servi pour construire une deuxième collection, ciblée sur les substrats acides  $\omega$ -aminés fonctionnalisés. Ces travaux sont détaillés dans la suite de ce chapitre. Toutes les enzymes citées dans la suite de ce manuscrit sont nommées selon leur code d'identification UniProt.

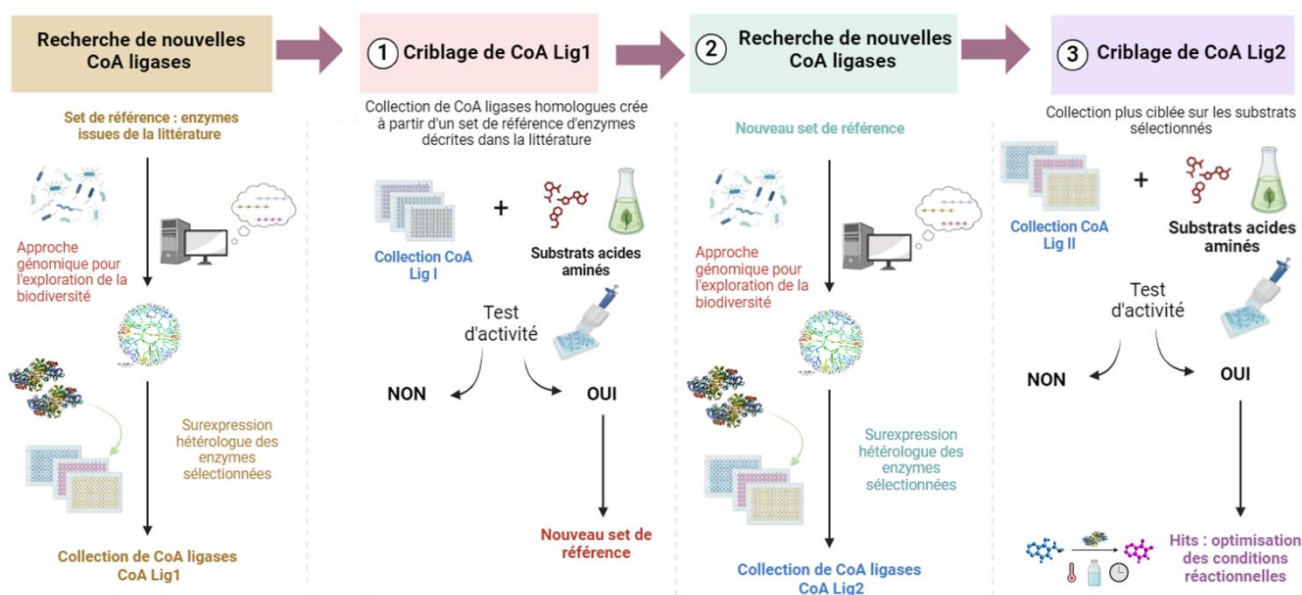


Schéma 65. Stratégie pour la recherche de nouvelles CoA ligases actives sur des divers acides  $\omega$ -aminés fonctionnalisés. Schéma inspiré du résumé de stage de Camille Baland (stagiaire M2 dans l'équipe). Fait avec BioRender.com.

### III.1. CRIBLAGE DE LA 1ERE COLLECTION DE COA LIGASES (COA LIG1)

Ces travaux ont été faits en collaboration avec la plateforme de clonage et de criblage des activités enzymatiques FunDivEx du Laboratoire de Génomique et Biochimie du Métabolisme (LGBM) au sein de l'unité. La recherche des CoA ligases dans la biodiversité a été effectuée par approche génomique selon un protocole bien établi pour la recherche de nouvelles enzymes au sein du laboratoire.

#### III.1.1. RECHERCHE DE NOUVELLES ENZYMES DANS LA BIODIVERSITE PAR APPROCHE GENOMIQUE

L'approche génomique pour la recherche de nouvelles enzymes peut être réalisée de différentes manières (Figure 24). Elle peut être guidée par le criblage d'une caractéristique phénotypique ou en se basant sur l'annotation fonctionnelle précise à l'aide des bases de données (Zaparucha, de Berardinis et al. 2018, Robinson, Piel et al. 2021). De nouvelles classes d'enzymes ont pu être aussi découvertes par la méthode PCR en mettant à profit cette technique d'amplification de séquences d'ADN en utilisant des amorces de gènes dites dégénérées. Ces amorces dégénérées contiennent des positions variables où différentes bases nucléotidiques peuvent être placées et donc créer de la diversité. Parmi les autres approches génomiques possibles, nous retrouvons également la recherche par identification de motifs conservés impliqués dans l'activité enzymatique, ou à partir de la structure 3D d'enzymes connues. En effet, en partant de la structure, il est possible de rechercher de nouvelles enzymes en définissant un ou des catalophores qui sont une carte contenant les résidus d'acides aminés nécessaires pour l'activité catalytique dans le site actif. La recherche s'effectue ensuite par criblage de ces catalophores dans la PDB. Cette dernière stratégie peut être déployée grâce aux outils de prédiction comme AlphaFold qui permet de générer des modèles de structures de protéines à partir de leurs séquences.

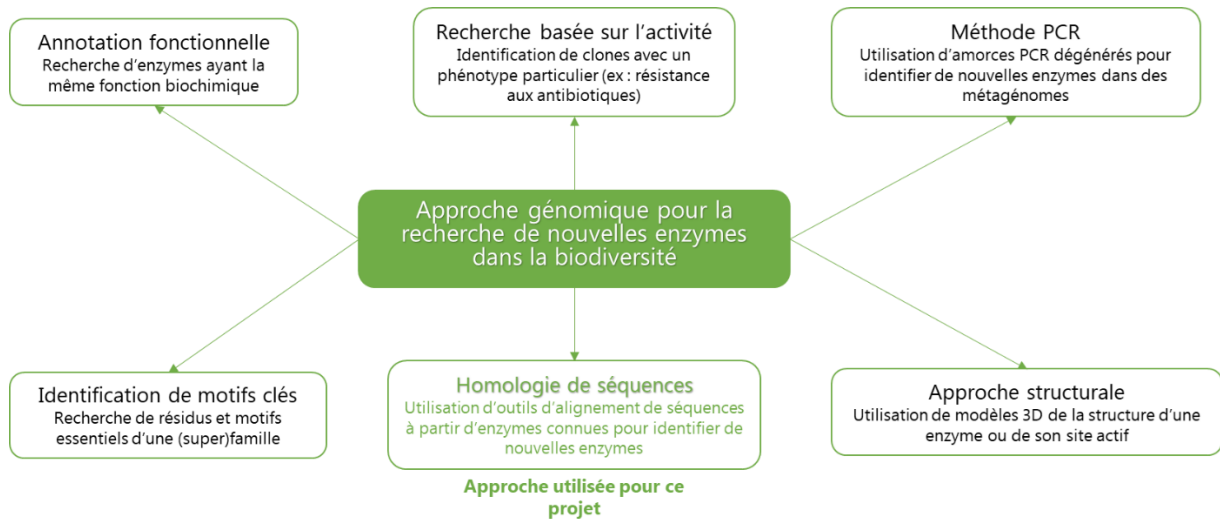


Figure 24. Récapitulatif des différentes stratégies possibles pour l'approche génomique pour la recherche de nouvelles enzymes.

La recherche d'enzymes peut s'effectuer également par homologie de séquences en partant d'enzymes connues et caractérisées. En se servant d'outils d'alignement de séquences comme BLASTP disponible depuis la base de données UniProt par exemple, il est possible de récupérer des enzymes homologues. Ensuite, différents critères biochimiques (pH, température de l'organisme duquel elles sont issues) ou de l'activité désirée peuvent être prises en compte pour réduire cette liste. En fonction du pourcentage d'identité de séquence, il est possible de répartir ces enzymes dans des clusters qui

regroupent les enzymes les plus proches en terme d'homologie de séquence et former des groupes d'enzymes dites isofonctionnels. Ces groupes isofonctionnels vont donc contenir des enzymes ayant la même activité et qui possèderaient des spécificités de substrats communes. A l'inverse, avoir des enzymes avec des homologies de séquences faibles peut permettre d'explorer et d'apporter la diversité. Les enzymes identifiées peuvent catalyser le même type de réaction mais avoir des spécificités de substrats différentes (Zaparucha, de Berardinis et al. 2018). Cette stratégie est celle qui a été utilisée pour construire les collections de CoA ligases.

### III.1.1.1. APPROCHE GENOMIQUE POUR LA RECHERCHE DE CoA LIGASES

Les collections d'enzymes faites au sein du laboratoire sont créées en collaboration avec la plateforme FunDivEx selon un protocole utilisé en routine. Tout d'abord, un set de référence a été constitué, regroupant des enzymes issues de la littérature dont l'activité type CoA ligase a été prouvée expérimentalement. Ce set de référence a servi de point de départ pour rechercher des enzymes homologues en utilisant l'outil BLAST pour explorer les séquences dans les bases de données UniProt ou des données métagénomiques provenant de projets internes du Genoscope. De façon très classique, un critère d'identité de séquence variant entre 30% à 80% a été choisi. La borne de 30% est choisie pour accéder à des enzymes relativement différentes, mais catalysant la même réaction, pour accéder à de nouvelles spécificités de substrats, la borne supérieure de 80% permet de disposer d'enzymes proches en terme de spécificité de substrats mais avec possiblement des différences de sélectivité et/ou d'efficacité catalytique (Zaparucha, de Berardinis et al. 2018). Si le nombre d'enzymes ainsi récupéré est élevé, les enzymes sont réparties en clusters en regroupant les enzymes ayant un fort pourcentage d'identité de séquence (>80%), considérées comme isofonctionnelles. Le choix des enzymes dans un cluster se fait en fonction du pourcentage GC des gènes et de la disponibilité de l'ADN dans la souchothèque du laboratoire. Pour ce projet comme nous recherchions des enzymes thermophiles, seules les séquences provenant d'organismes thermophiles ont été retenues. Les CoA ligases ont ensuite été produites en plaque 96 puits et criblées.



Figure 25. Approche génomique mis en place au laboratoire pour la recherche de nouvelles enzymes au sein de la biodiversité.

### III.1.1.2. SET DE REFERENCE DE CoA Lig1

Le set de référence utilisé pour construire la collection CoA Lig1 était composé de quatre enzymes : la CoA ligase thermophile A4YDT1 de *Metallosphaera sedulla*, les travaux d'ingénierie

réalisées sur cette enzyme sont présentés dans le chapitre précédent ; l'enzyme A4YDR9 qui provient du même organisme et qui a la même annotation que A4YDT1 ; les CoA ligases non thermophiles B3G250 et A0A0A8RCP0 de *Pseudomonas aeruginosa*. Depuis mai 2023, la référence UniProt de A0A0A8RCP0 est archivée car la séquence est obsolète donc elle a été supprimé des bases de données métagénomiques.

Tableau 6. Set de référence composé de CoA ligases issues de la littérature pour construire la collection CoA Lig1.

Code UniProt	Organisme	Annotation
A4YDT1	<i>Metallosphaera sedulla</i>	4-hydroxybutyrate--CoA ligase 1
A4YDR9	<i>Metallosphaera sedulla</i>	4-hydroxybutyrate--CoA ligase 2
B3G250	<i>Pseudomonas aeruginosa</i>	AMP-dependent synthetase and ligase family protein
A0A0A8RCP0	<i>Pseudomonas aeruginosa</i>	Acetate/butyrate CoA ligase AAE7, peroxisomal

A partir de ce set, la première collection de CoA ligases issue de la biodiversité, nommée CoA Lig1, a été construite et produite par la plateforme FunDivEx. Au final, 181 gènes ont été sélectionnés soit l'équivalent de deux plaques de 96 puits. Enfin, il y a eu une troisième plaque correspondant à des CoA ligases issues d'archées thermophiles avec 79 gènes. La majorité des CoA ligases sélectionnées ont pour homologue le plus proche soit A4YDT1 ou B3G250. Les CoA ligases du set de référence ont servi de témoins positifs pour la production et le criblage de la collection. Le taux de production des enzymes pour ce projet est relativement bas avec seulement 31% de réussite de production de CoA ligases.

### III.1.2. CRIBLAGE DE COA LIG1

Durant la thèse de Dr Chloé Lelièvre, la collection CoA Lig1 avait été criblée pour la recherche d'enzymes actives pour la synthèse de lactames simples de 5 à 7 chaînons. Cela a permis d'identifier et d'étudier 7 nouvelles CoA ligases. Au cours de ma thèse, cette même collection a été criblée à nouveau mais avec des substrats acides  $\omega$ -aminés fonctionnalisés, afin de trouver des CoA ligases actives pour la synthèse de lactames fonctionnalisés. Les hits identifiés ont été utilisés pour constituer un nouveau set de référence afin d'étendre l'exploration la biodiversité à la recherche de nouvelles CoA ligases présentant de nouvelles et/ou meilleures activités, toujours ciblées pour la synthèse de lactames fonctionnalisés.

Le criblage de la collection de CoA ligases avec des acides  $\omega$ -aminés fonctionnalisés est divisée en deux étapes. La première étape consistait à identifier les potentiels hits à l'aide d'un test de suivi de la réaction simple et robuste. Puis dans une seconde étape, les hits potentiels identifiés ont été vérifiés en UHPLC afin de ne pas inclure de faux positifs dans notre set de référence. Les résultats sont présentés ci-dessous.

#### III.1.2.1. CHOIX DES SUBSTRATS POUR LE CRIBLAGE

L'objectif de ce nouveau criblage était de trouver des enzymes actives pour la synthèse de lactames fonctionnalisés. Pour cela nous avons testé CoA Lig1 sur des acides  $\omega$ -aminés fonctionnalisés, toujours pour la formation de lactames de 5 à 7 chaînons. Nous souhaitons partir de substrats simples pouvant être biosourcés ou biosourçables. Nous avons également sélectionné des substrats portant des centres chiraux permettant accéder à des lactames chiraux comme il a été possible de le faire avec les CARs ou des oxydoréductases par exemple. Le choix des acides  $\omega$ -aminés a été fait en fonction de leur nature, source ou du lactame d'intérêt qu'il serait possible d'obtenir par cyclisation spontanée de l'acide

activé. Au final, 17 substrats répartis en pools ont été criblés avec cette collection (Figure 3). Ces pools ont été constitués en fonction des substitutions présentes sur la chaîne latérale des acides  $\omega$ -aminés afin de faciliter leur détection par UHPLC-MS ou UV en appliquant des méthodes analytiques communes pour chaque pool.

Lors de la thèse de Dr Chloé Lelièvre, CoA Lig1 a été criblée sur trois acides  $\omega$ -aminés simples, l'acide 4-aminobutyrique (**1a**, GABA), l'acide 5-aminopentanoïque (**2a**, 5-APA) et l'acide 6-aminohexanoïque (**3a**, 6-AHA) pour la formation des lactames correspondants, le  $\gamma$ -butyrolactame,  $\delta$ -valérolactame et le  $\varepsilon$ -caprolactame. Ce criblage a permis d'identifier de nouvelles CoA ligases comme Q3ABN2 de *Carboxythermus hydrogenoformans* ou encore K6Q029 de *Thermaerobacter subtanerreus* qui a été caractérisée biochimiquement. Ces acides  $\omega$ -aminés constituent notre premier pool de substrats. Cribler de nouveau CoA Lig1 avec ce pool a servi de contrôle positif.

Le deuxième pool est constitué d'acides  $\omega$ -aminés ayant un groupement amine primaire supplémentaire avec une taille de chaîne latérale variable : l'acide 2,4-diaminobutanoïque (**4a**), la L-ornithine (**5a**) et la L-lysine (**6a**). La lysine est une source de  $\varepsilon$ -caprolactame par désamination du 3-aminoazépen-2-one ou D,L-aminocaprolactame (**6b**), le lactame formé par cyclisation de lysine. Ce deuxième pool contenant également l'acide 2-hydroxy-4-aminobutanoïque (**7a**) qui diffère juste du substrat **4a** en position 2 par un groupement hydroxyle au lieu d'un groupement amine. Le troisième pool contient des acides  $\omega$ -aminés substitués avec un groupement acide carboxylique : l'acide glutamique (**8a**), l'acide 2-aminoadipique ou 2-aminohexanedoïque (**9a**) et l'acide aminopimélique ou l'acide 2-aminoheptanedoïque (**10a**). Le quatrième pool est composée du (2-aminoéthyl)glycine (**11a**) et de l'acide 3-(2-aminoéthoxy)propanoïque (**12a**) qui possèdent un hétéroatome dans leur chaîne latérale. Ces substrats amènent à des lactames d'intérêt pour la synthèse de médicaments : la piperazin-2-one (**11b**) qui appartient donc la famille des pipérazines, l'une des classes de N-hétérocycles les plus employées et approuvées par la FDA (Food and Drug Administration) ; et le 1,4-oxazépan-5-one (**12b**) qui appartient à la famille des morpholinones utilisées en chimie médicinale. Le cinquième pool contient aussi des substrats qui contiennent également des hétéroatomes mais plus complexes. Nous retrouvons la créatine (**13a**) qui est un dérivé d'acide aminé biosynthétisé à partir de glycine, arginine et méthionine et le lactame attendu est la créatinine (**13b**). Ce pool contient également la glycyglycine (**14a**) avec une liaison amine dont la cyclisation permettrait d'accéder à la 2,5-dicétopipérazine (**14b**). L'acide  $\delta$ -aminolévulinique (**15a**, ALA) appartient aussi à ce pool, ce substrat dérive de la glycine et de succinyl-CoA provenant du cycle de Krebs. Enfin, le sixième et dernier pool contient les acides 4-(méthylamino)- (**16a**) et 4-(éthylamino)butyrique (**17a**) permettant d'accéder à des  $\gamma$ -butyrolactames *N*-méthylé et *N*-éthylé correspondant (**16b** et **17b**).



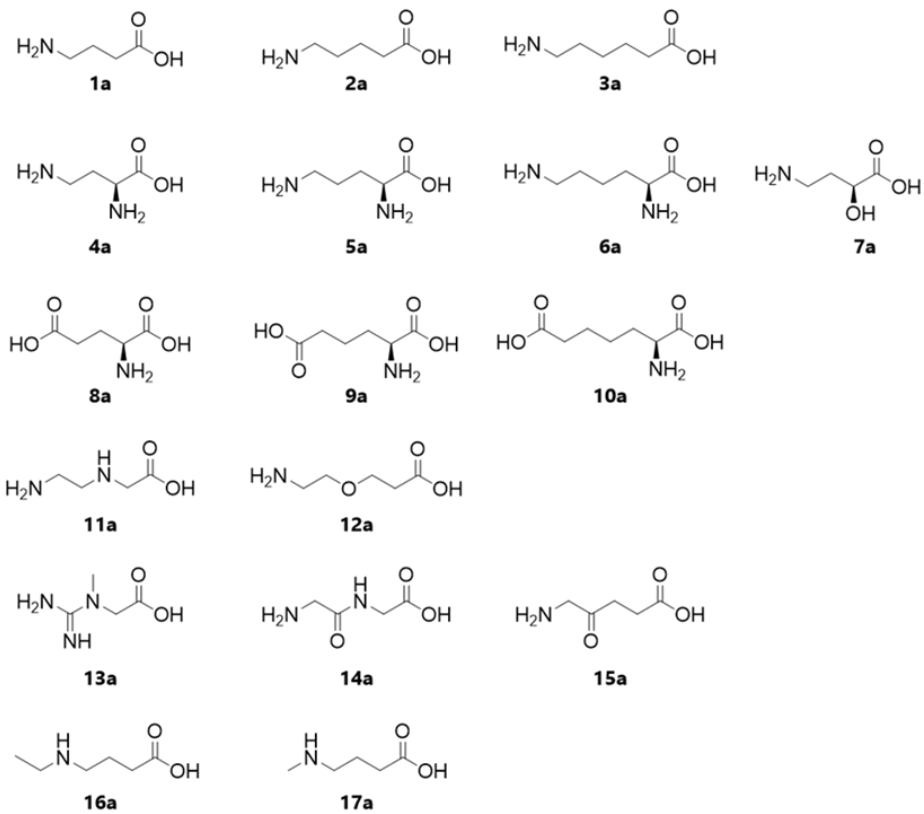
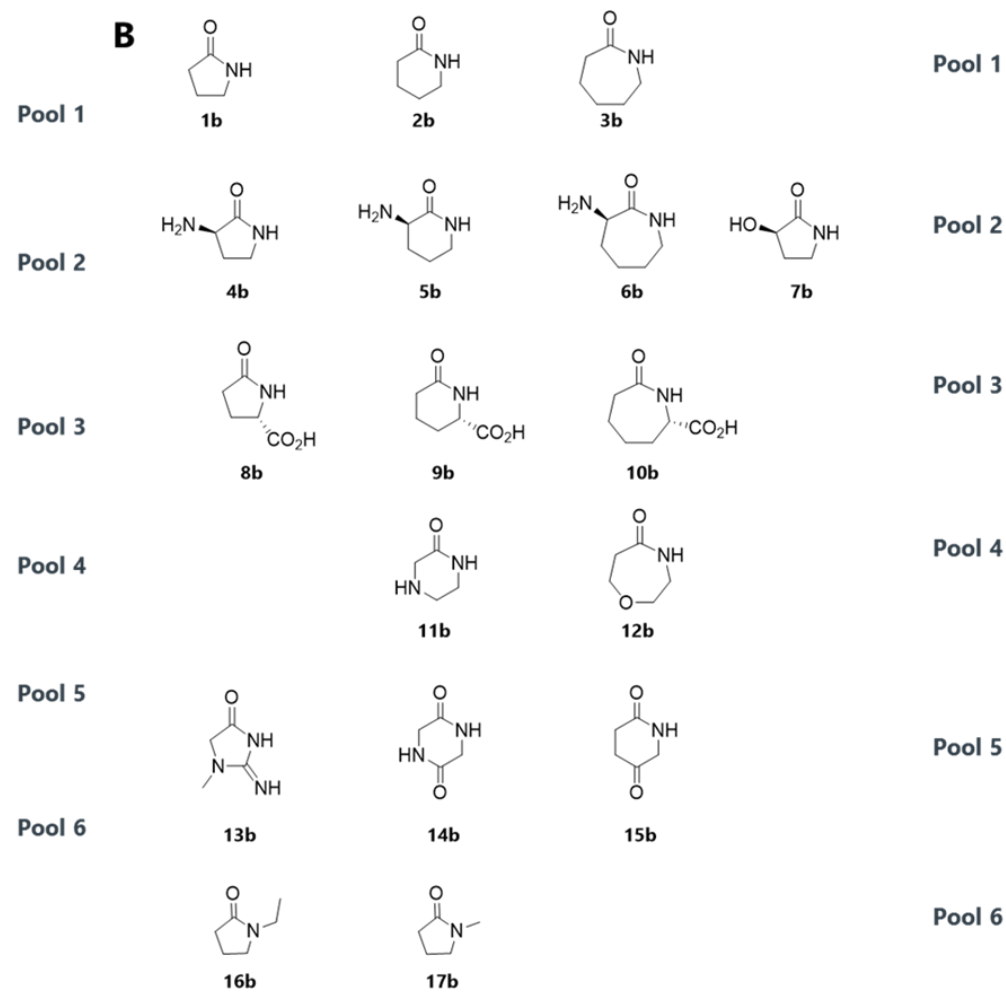
**A****B**

Figure 26. Liste des 17 acides ω-aminés répartis par pool pour le criblage de la collection de CoA ligases CoA Lig1.

### III.1.2.2. METHODES DE SUIVI DE L'ACTIVITE DES CoA LIGASES EN CONDITIONS DE CRIBLAGE

La réaction de formation de la liaison amide se fait en deux étapes : la 1<sup>ère</sup> catalysée par une CoA ligase conduit à la formation d'un intermédiaire (amino)acyl-AMP avec libération de PPi, la seconde est spontanée et conduit à l'amide (ou au lactame selon le substrat) avec libération d'AMP (Schéma 66).

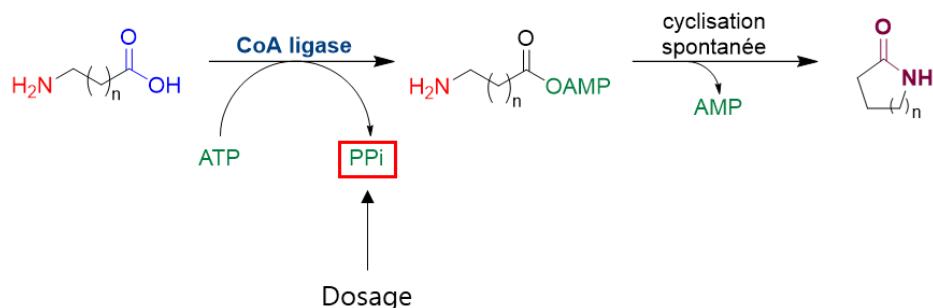


Schéma 66. Schéma simplifié de la formation de lactames avec des CoA ligases et dosage de la libération de pyrophosphate.

Lors du 1<sup>er</sup> criblage de la collection CoA Lig1, le test utilisé pour suivre l'activité était un dosage de l'AMP libéré par un test enzymatique couplé. Pour ce nouveau criblage, nous souhaitons disposer d'un test facile à mettre en œuvre, de préférence colorimétrique, et nous nous sommes tournés vers des dosages du PPi libéré lors de la première étape. Dans les deux cas on ne suit pas la même chose, le test de dosage de l'AMP permet de suivre les deux étapes de la cascade menant à la formation de lactame (ou d'amide), le test de dosage du PPi reflète uniquement l'activité de la CoA ligase. Dans la mesure où nous souhaitons d'abord, avec un crible rapide, trouver des enzymes actives sur nos pools de substrats, puis dans un second temps vérifier la formation de lactames, nous avons étudié la possibilité de mettre en place un test colorimétrique.

#### III.1.2.2.1. Méthode de suivi colorimétrique de l'adénylation : dosage au vert de malachite (MG assay)

Le dosage au vert de malachite permet de suivre l'adénylation grâce à la libération du pyrophosphate inorganique (PPi). En présence d'une pyrophosphatase dans le milieu réactionnel, le PPi est hydrolysé en phosphate inorganique (Pi) qui se complexe avec le vert de malachite (MG pour Malachite Green). Initialement, le vert de malachite est en solution acide de couleur orangée puis la complexation de Pi va induire un changement de couleur vers le vert (Schéma 67). Nous avons utilisé le protocole développé par Maria Vardakou *et al.* pour préparer notre solution de vert de malachite en solution aqueuse acide (dans de l'acide sulfurique) que nous nommerons la solution de colorant vert de malachite (Vardakou, Salmon et al. 2014). Au moment de son utilisation, cette solution est mélangée à du molybdate pour donner la solution dite de développement. Un inconvénient du dosage au vert de malachite est l'apparition d'un précipité au cours du temps dû à la complexation. Pour pallier ce problème, le polysorbate TWEEN®20 est ajouté comme tensioactif afin de stabiliser la solution, qui doit tout de même préparée juste avant utilisation.

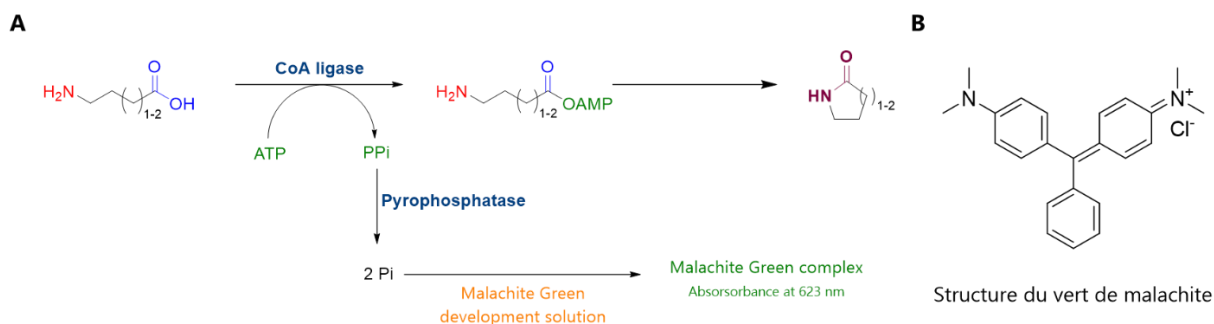


Schéma 67. (A) Dosage du PPI par le vert de malachite. (B) Structure du vert de malachite.

Nous nous sommes tournés vers ce test car il est colorimétrique, donc qu'il nous permettrait d'obtenir une réponse rapide du type OUI/NON sur l'activité d'une CoA ligase par observation de la couleur verte et de son intensité. De plus, le dosage au MG est une méthode couramment employée pour un suivi d'activité d'enzymes libérant du PPI en conditions de criblage à bas et haut-débit, mais aussi pour déterminer des paramètres cinétiques (McQuade, Shallop et al. 2009, Scott, Yang et al. 2010, Vardakou, Salmon et al. 2014, Kittilä, Schoppet et al. 2016).

Des premiers essais du MG assay avec du vecteur vide pET-22 sous forme de lysat nous ont montré que le dosage était sensible au milieu utilisé. En effet, de grandes variations ont été observées pour réaliser des gammes de calibration en présence de lysat. Il était donc nécessaire de travailler sur extraits cellulaires clarifiés ou enzymes purifiés. Pour enlever tout autre biais, les enzymes ont été produites dans du tampon MOPS (acide 3-morpholino-propanesulfonique), alors qu'habituellement elles sont produites en tampon phosphate. Nous avons choisi ce tampon car il a déjà été utilisé pour l'étude des CoA ligases A4YDT1 et K6Q029 lors la thèse de Dr C. Lelièvre (2017-2020). Dans le cadre du criblage, le test a été effectué en point final avec une pyrophosphatase thermophile pour qu'il y ait une libération « continue » de phosphate inorganique dans le milieu. Suite à la réaction enzymatique, la solution de développement de vert de malachite, préparée juste avant utilisation a été ajoutée dans le milieu réactionnel et la solution incubée pendant 15 minutes à température ambiante avant la lecture de la plaque au spectrophotomètre.

Finalement, après plusieurs tentatives avec des CoA ligases servant de contrôles positifs ou négatifs mais aussi avec un test dans les conditions de criblage avec le pool 1 qui sert également de contrôle, le test ne s'est pas révélé assez fiable pour être utilisé. Nous avons utilisé la méthode de suivi d'activité déjà utilisée pour le premier criblage : la cascade Pyruvate Kinase-Lactate Déshydrogénase (cascade PK-LDH).

### III.1.2.2.2. Cascade pyruvate kinase-lactate déshydrogénase (cascade PK-LDH)

La cascade enzymatique PK-LDH, utilisée pour le 1<sup>er</sup> criblage de la CoA Lig1 est un test souvent utilisé dans l'unité (Vergne-Vaxelaire, Mariage et al. 2018). Cette cascade est composée de 3 enzymes et permet de doser l'AMP libéré par suivi de la consommation de NADH au cours du temps (Schéma 68). Le criblage avec cette cascade a été faite en méthode de point final car les CoA ligases sont thermophiles contrairement aux enzymes de la réaction analytique dans la cascade. La réaction biocatalytique est chauffée 5h à 60°C puis stoppée par addition d'acide chlorhydrique 6M (HCl), le mélange est ensuite neutralisé par de la soude 5M (NaOH) avant analyse.

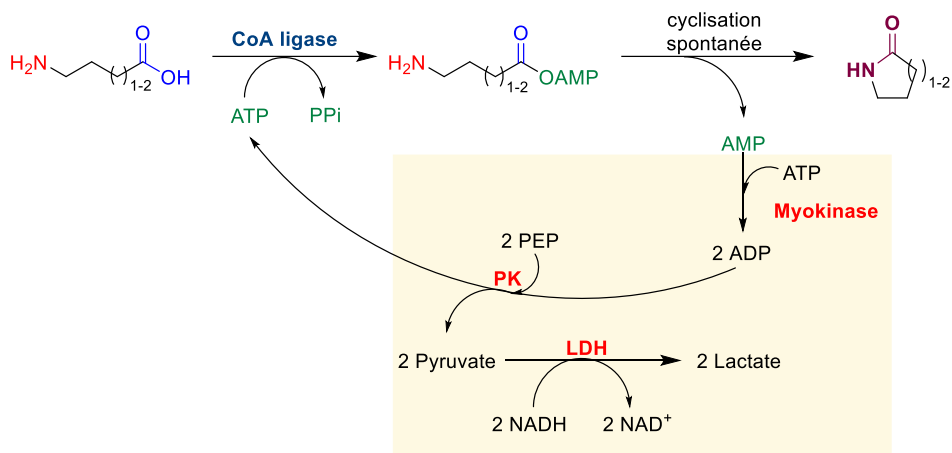


Schéma 68. Dosage de l'AMP libéré lors de la cyclisation spontanée par la cascade PK-LDH.

### III.1.3. RESULTATS DU CRIBLAGE DE CoA Lig1

Les gels SDS-PAGE, les concentrations des enzymes ainsi que toutes les données traitées du criblage sont en annexe de ce manuscrit (Annexe 13 et Annexe 14). Le criblage a été effectué en plaque de 96 puits à isovolume de CoA ligase purifiée dans chaque puit, 5h dans une étuve à 60°C sans agitation. Au total 88 CoA ligases ont été testées, les 83 de la collection CoA Lig1, en plus les 5 CoA ligases du set de référence qui servent de témoins.

#### III.1.3.1. IDENTIFICATION DES HITS

Le pool 1 correspond à celui des acides  $\omega$ -aminés simples déjà testés lors du premier criblage de la collection durant la thèse de Dr C. Lelièvre, il nous a donc servi de référence pour ce nouveau criblage (Tableau 7). Les listes des CoA ligases positives sont données par pool de substrats. A noter, pour être sûrs de ne pas écarter des enzymes positives nous avons été larges dans notre choix d'enzymes en prenant les enzymes positives à partir de 2% de rendement par rapport au seuil de détection.

Tableau 7. Liste des CoA ligases retenues pour une vérification de l'activité par UHPLC-MS pour le pool 1 suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les enzymes du set de référence de CoA Lig1 sont indiquées en gras, les enzymes déjà identifiées lors du 1<sup>er</sup> criblage de CoA Lig1 sont indiquées en bleu.

Position plaque	Code UniProt	Organisme	Annotation
CoA_1_C01	A4IMI8	<i>Geobacillus thermodenitrificans</i> (strain NG80-2)	Long-chain-fatty-acid-CoA ligase
CoA_1_C02	A4IMJ1	<i>Geobacillus thermodenitrificans</i> (strain NG80-2)	Medium-chain-fatty-acid-CoA ligase
CoA_1_C10	A5D1Z4	<i>Pelotomaculum thermopropionicum</i> (strain DSM 13744 / JCM 10971 / SI)	Acyl-coenzyme A synthetases/AMP-(Fatty acid ligases)
CoA_1_D07	F8A8M3	<i>Thermodesulfator indicus</i> (strain DSM 15286 / JCM 11887 / CIR29812)	Acetyl-coenzyme A synthetase (AcCoA synthetase) (Acs) (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-activating enzyme)
CoA_1_F05	Q3ABN2	<i>Carboxydotherrnus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	AMP-binding enzyme family protein

CoA_1_G12	B9L4W4	<i>Thermomicrobium roseum</i> (strain ATCC 27502 / DSM 5159 / P-2)	AMP-dependent synthetase and ligase
CoA_1_H03	D1A9C5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_1_H05	A4YDT1	<b><i>Metallosphaera sedula</i> (strain ATCC 51363 / DSM 5348 / JCM 9185 / NBRC 15509 / TH2)</b>	<b>4-hydroxybutyrate--CoA ligase 1</b>
CoA_2_C02	F2NLM4	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_2_D11	Q72L27	<i>Thermus thermophilus</i> (strain HB8 / ATCC 27634 / DSM 579)	Medium-chain-fatty-acid-CoA ligase
CoA_2_F06	D1AB87	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_G09	F2NPU2	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	O-succinylbenzoate--CoA ligase (EC 6.2.1.26)
CoA_2_H05	K6Q029	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_3_B09	C5A3C0	<i>Thermococcus gammatolerans</i> (strain DSM 15229 / JCM 11827 / EJ3)	Acetyl-coenzyme A synthetase (AcsA) (EC 6.2.1.1)

Pour le pool 2, nous avons décidé de retenir 13 CoA ligases, parmi elles deux enzymes du set de référence (A4YDT1 et B3G250) (Tableau 8). On retrouve également K6P2L2 une CoA ligase provenant du même organisme que K6Q029, identifiée sur le pool 1.

Tableau 8. Liste des CoA ligases retenues pour une vérification de l'activité par UHPLC-UV pour le pool 2 suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les enzymes du set de référence de CoA Lig1 sont indiquées en gras. Les CoA ligases retenues pour d'autres pools de substrats sont indiqués en vert.

Position plaque	Code UniProt	Organisme	Annotation
CoA_1_C02	A4IMJ1	<i>Geobacillus thermodenitrificans</i> (strain NG80-2)	Medium-chain-fatty-acid-CoA ligase
CoA_1_F04	Q3ABD3	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_1_H05	A4YDT1	<b><i>Metallosphaera sedula</i> (strain ATCC 51363 / DSM 5348 / JCM 9185 / NBRC 15509 / TH2)</b>	<b>4-hydroxybutyrate--CoA ligase 1</b>
CoA_1_H06	B3G250	<b><i>Pseudomonas aeruginosa</i></b>	<b>AMP-dependent synthetase and ligase family protein</b>
CoA_2_G09	F2NPU2	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	O-succinylbenzoate--CoA ligase (EC 6.2.1.26)
CoA_2_H02	K6P2L2	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_2_A12	D3PTT7	<i>Meiothermus ruber</i> (strain ATCC 35948 / DSM 1279 / VKM B-1258 / 21) ( <i>Thermus ruber</i> )	AMP-dependent synthetase and ligase
CoA_2_B03	D6Y433	<i>Thermobispora bispora</i> (strain ATCC 19993 / DSM 43833 / CBS 139.67 / JCM 10125 / NBRC 14880 / R51)	AMP-dependent synthetase and ligase

CoA_2_B07	D7BGN5	<i>Meiothermus silvanus</i> (strain ATCC 700542 / DSM 9946 / VI-R2) ( <i>Thermus silvanus</i> )	AMP-dependent synthetase and ligase
CoA_2_B08	D7BIL1	<i>Meiothermus silvanus</i> (strain ATCC 700542 / DSM 9946 / VI-R2) ( <i>Thermus silvanus</i> )	AMP-dependent synthetase and ligase
CoA_3_A06	Q97UW7	<i>Sulfolobus solfataricus</i> (strain ATCC 35092 / DSM 1617 / JCM 11322 / P2)	Medium-chain-fatty-acid--CoA ligase (AlkK-4) (EC 6.2.1.-)
CoA_3_D06	E1QT70	<i>Vulcanisaeta distributa</i> (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)	AMP-dependent synthetase and ligase
CoA_3_E06	O29881	<i>Archaeoglobus fulgidus</i> (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	Acetyl-CoA synthetase (Acs-2)

Pour le pool 3, 23 CoA ligases ont été retenues pour une vérification par UHPLC pour leur activité sur des acides  $\omega$ -aminés substitués avec des groupements carboxyliques (Tableau 9). Même si l'annotation n'est pas l'élément sur lequel nous nous focalisons pour choisir des CoA ligases, on remarque que certaines d'entre elles sont annotées comme des O-succinylbenzoate CoA ligases qui sont une classe de CoA ligases assez bien étudiée, ou comme des CoA ligases d'acides gras à chaîne moyenne ou longue. Nous retrouvons de nouveau K6Q029 et K6P2L2 déjà retenues pour leur activité sur d'autres pools.

Tableau 9. Liste des CoA ligases retenues pour le pool 3 pour une vérification en UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert.

Position plaque	Code UniProt	Organisme	Annotation
CoA_1_C02	A4IMJ1	<i>Geobacillus thermodenitrificans</i> (strain NG80-2)	Medium-chain-fatty-acid-CoA ligase
CoA_1_D03	D3SQH8	<i>Thermocrinis albus</i> (strain DSM 14484 / JCM 11386 / HI 11/12)	Acetate/CoA ligase
CoA_1_E03	L0R9S0	<i>Desulfovibrio hydrothermalis</i> AM13 = DSM 14728	Acetyl-coenzyme A synthetase (EC 6.2.1.1)
CoA_1_E04	L0RAD3	<i>Desulfovibrio hydrothermalis</i> AM13 = DSM 14728	Acetyl-coenzyme A synthetase (AcCoA synthetase) (Acs) (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-activating enzyme)
CoA_1_F04	Q3ABD3	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_1_G03	Q9HI39	<i>Thermoplasma acidophilum</i> (strain ATCC 25905 / DSM 1728 / JCM 9062 / NBRC 15155 / AMRC-C165)	Probable SA protein
CoA_1_H03	D1A9C5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_1_H04	D1ABC1	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_A02	D1C3U2	<i>Sphaerobacter thermophilus</i> (strain DSM 20745 / S 6022)	AMP-dependent synthetase and ligase
CoA_2_A03	D1C448	<i>Sphaerobacter thermophilus</i> (strain DSM 20745 / S 6022)	AMP-dependent synthetase and ligase
CoA_2_B03	D6Y433	<i>Thermobispora bispora</i> (strain ATCC 19993 / DSM 43833 / CBS 139.67 / JCM 10125 / NBRC 14880 / R51)	AMP-dependent synthetase and ligase
CoA_2_C02	F2NLM4	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)

CoA_2_D11	Q72L27	<i>Thermus thermophilus</i> (strain HB8 / ATCC 27634 / DSM 579)	Medium-chain-fatty-acid-CoA ligase
CoA_2_E02	S9S7W1	<i>Rubellimicrobium thermophilum</i> DSM 16684	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II (EC 6.2.1.-)
CoA_2_E03	D1A3M5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_E10	D1A5W6	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_G09	F2NPU2	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	O-succinylbenzoate--CoA ligase (EC 6.2.1.26)
CoA_2_H02	K6P2L2	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_2_H05	K6Q029	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_3_C09	D9PUA1	<i>Methanothermobacter marburgensis</i> (strain ATCC BAA-927 / DSM 2133 / JCM 14651 / NBRC 100331 / OCM 82 / Marburg) ( <i>Methanobacterium thermoautotrophicum</i> )	Predicted acetyl-coenzyme A synthetase (EC 6.2.1.1)
CoA_3_D07	F2KPD0	<i>Archaeoglobus veneficus</i> (strain DSM 11195 / SNP6)	Acetyl-coenzyme A synthetase (AcCoA synthetase) (Acs) (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-activating enzyme)
CoA_3_E06	O29881	<i>Archaeoglobus fulgidus</i> (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	Acetyl-CoA synthetase (Acs-2)
CoA_3_G07	B0R4T6	<i>Halobacterium salinarum</i> (strain ATCC 29341 / DSM 671 / R1)	O-succinylbenzoate--CoA ligase (EC 6.2.1.26)

Pour le pool 4, 15 CoA ligases ont montré une activité, à noter que pour la majorité d'entre elles il n'y a aucune activité sur les substrats du pool 1 (Tableau 10).

Tableau 10. Liste des CoA ligases retenues pour le pool 4 pour une vérification en UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert.

Position plaque	Code UniProt	Organsime	Annotation
CoA_1_B01	A0A150LLT2	<i>Anoxybacillus flavithermus</i>	2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58)
CoA_1_E04	L0RAD3	<i>Desulfovibrio hydrothermalis</i> AM13 = DSM 14728	Acetyl-coenzyme A synthetase (AcCoA synthetase) (Acs) (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-activating enzyme)
CoA_1_F04	Q3ABD3	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_1_F06	Q3ABP3	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Long-chain-fatty-acid--CoA ligase
CoA_1_F09	Q3AEI5	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Medium-chain-fatty-acid--CoA ligase (EC 6.2.1.-)
CoA_1_H03	D1A9C5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_1_H04	D1ABC1	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_C05	G2SH91	<i>Rhodothermus marinus</i> SG0.5JP17-172	4-coumarate--CoA ligase (EC 6.2.1.12)

CoA_2_C06	G2SHU3	<i>Rhodothermus marinus</i> SG0.5JP17-172	Acetate--CoA ligase (EC 6.2.1.1)
CoA_3_A06	Q97UW7	<i>Sulfolobus solfataricus</i> (strain ATCC 35092 / DSM 1617 / JCM 11322 / P2)	Medium-chain-fatty-acid--CoA ligase (AlkK-4) (EC 6.2.1.-)
CoA_3_C02	D3RYS2	<i>Ferroglobus placidus</i> (strain DSM 10642 / AEDII12DO)	AMP-dependent synthetase and ligase
CoA_3_C03	D3RZ03	<i>Ferroglobus placidus</i> (strain DSM 10642 / AEDII12DO)	AMP-dependent synthetase and ligase
CoA_3_D04	E1QQ97	<i>Vulcanisaeta distributa</i> (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)	AMP-dependent synthetase and ligase
CoA_3_G01	V9S912	<i>Sulfolobus acidocaldarius</i> SUSAZ	Acyl-CoA synthetase
CoA_3_G03	A8A939	<i>Ignicoccus hospitalis</i> (strain KIN4/I / DSM 18386 / JCM 14125)	Acyl-coenzyme A synthetase/AMP-(Fatty) acid ligase-like protein

Pour le pool 5, 17 CoA ligases ont été retenues dont Q3ABD3, Q3ABN2, D1A9C5, Q72L27 et K6Q029 déjà retenues pour d'autres pools précédemment (Tableau 11).

Tableau 11. Liste des CoA ligases retenues pour le pool 5 pour une vérification par UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert.

Position plaque	Code UniProt	Organisme	Annotation
CoA_1_C02	A4IMJ1	<i>Geobacillus thermodenitrificans</i> (strain NG80-2)	Medium-chain-fatty-acid-CoA ligase
CoA_1_D11	K8DYP0	<i>Desulfotomaculum hydrothermale</i> Lam5 = DSM 18033	Acetyl-coenzyme A synthetase (AcCoA synthetase) (Acs) (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-activating enzyme)
CoA_1_F04	Q3ABD3	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_1_F05	Q3ABN2	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	AMP-binding enzyme family protein
CoA_1_H03	D1A9C5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_D11	Q72L27	<i>Thermus thermophilus</i> (strain HB8 / ATCC 27634 / DSM 579)	Medium-chain-fatty-acid-CoA ligase
CoA_2_F05	D1AB70	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_H05	K6Q029	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_3_A08	Q97VT6	<i>Sulfolobus solfataricus</i> (strain ATCC 35092 / DSM 1617 / JCM 11322 / P2)	Long-chain-fatty-acid--CoA ligase (FadD-2) (EC 6.2.1.3)
CoA_3_B09	C5A3C0	<i>Thermococcus gammatolerans</i> (strain DSM 15229 / JCM 11827 / EJ3)	Acetyl-coenzyme A synthetase (AcsA) (EC 6.2.1.1)
CoA_3_C04	D3RZA8	<i>Ferroglobus placidus</i> (strain DSM 10642 / AEDII12DO)	AMP-dependent synthetase and ligase



CoA_3_D04	E1QQ97	<i>Vulcanisaeta distributa</i> (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)	AMP-dependent synthetase and ligase
CoA_3_D09	O28246	<i>Archaeoglobus fulgidus</i> (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	Medium-chain acyl-CoA ligase (AlkK-5)
CoA_3_E06	O29881	<i>Archaeoglobus fulgidus</i> (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	Acetyl-CoA synthetase (Acs-2)
CoA_3_F11	Q9HLD6	<i>Thermoplasma acidophilum</i> (strain ATCC 25905 / DSM 1728 / JCM 9062 / NBRC 15155 / AMRC-C165)	4-coumarate-CoA ligase related protein
CoA_3_G01	V9S912	<i>Sulfolobus acidocaldarius</i> SUSAZ	Acyl-CoA synthetase
CoA_3_G03	A8A939	<i>Ignicoccus hospitalis</i> (strain KIN4/I / DSM 18386 / JCM 14125)	Acyl-coenzyme A synthetase/AMP-(Fatty acid ligase-like protein

Enfin pour le pool 6, 11 potentielles CoA ligases dont certaines enzymes déjà identifiées comme actives sur les pools 1 et 5 (Tableau 12).

Tableau 12. Liste des CoA ligases retenues pour le pool 6 pour une confirmation en UHPLC-MS suite au test de criblage par dosage de l'AMP par le test PK-LDH. Les CoA ligases déjà retenues pour d'autres pools de substrats sont indiquées en vert.

Position plaque	Code UniProt	Organisme	Annotation
CoA_1_C02	A4IMJ1	<i>Geobacillus thermodenitrificans</i> (strain NG80-2)	Medium-chain-fatty-acid-CoA ligase
CoA_1_F04	Q3ABP3	<i>Carboxydotherrnus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_1_F05	Q3ABN2	<i>Carboxydotherrnus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	AMP-binding enzyme family protein
CoA_1_G03	Q9HI39	<i>Thermoplasma acidophilum</i> (strain ATCC 25905 / DSM 1728 / JCM 9062 / NBRC 15155 / AMRC-C165)	Probable SA protein
CoA_1_H03	D1A9C5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase
CoA_2_C03	F2NLM5	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)
CoA_2_D11	Q72L27	<i>Thermus thermophilus</i> (strain HB8 / ATCC 27634 / DSM 579)	Medium-chain-fatty-acid-CoA ligase
CoA_2_H02	K6P2L2	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_2_H05	K6Q029	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II
CoA_3_C01	D3RXJ3	<i>Ferroglobus placidus</i> (strain DSM 10642 / AEDII12DO)	AMP-dependent synthetase and ligase
CoA_3_D09	O28246	<i>Archaeoglobus fulgidus</i> (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	Medium-chain acyl-CoA ligase (AlkK-5)

Au total, nous avons retenu 55 CoA ligases dont certaines sont actives sur plusieurs pools de substrats (Annexe 15).

Suite à cette sélection de CoA ligases actives pour chaque pool, nous avons vérifié les activités en UHPLC. Le but du suivi UHPLC est de valider l'activité par suivi de la formation du lactame, pour rappel lors du 1<sup>er</sup> criblage nous avons sélectionné « large » en gardant des enzymes pour lesquelles le résultat était juste au-dessus du bruit de fond. Cela permet aussi d'éliminer les faux positifs. L'avantage du suivi par UHPLC est la possibilité de continuer à travailler en pools car au sein d'un même pool on s'attend à des temps de rétention différent pour chaque lactame. L'inconvénient est qu'il faut développer une méthode pour chaque pool avec le risque de ne pas réussir à mettre au point une méthode de suivi.

### III.1.3.2. VERIFICATION PAR UHPLC-MS OU UHPLC-UV

De nouvelles réactions ont été lancées avec les CoA ligases sélectionnées et dans les mêmes conditions que le criblage avec la cascade PK-LDH. A noter, les réactions ont été effectuées à iso volume d'enzyme et non iso concentration, ce qui a rendu plus difficile l'analyse des résultats. Ces analyses n'ont donc pu être que qualitatives par simple détection des aires correspondant aux lactames attendus pour chaque pool de substrat.

Nous avons mis au point des méthodes permettant la détection des lactames cibles en UHPLC-MS ou UV. Quand le standard commercial d'un lactame était disponible, nous avons également fait de la quantification afin de déterminer un rendement analytique avec les CoA ligases actives, à partir d'une gamme de calibration incubée dans les mêmes conditions que les réactions enzymatiques. Cette quantification a été effectuée pour les lactames correspondants aux substrats du pool 1 (GABA **1a**, 5-APA **2a** et 6-AHA **3a**), la lysine (**6a**) du pool 2 et le glutamate (**8a**) du pool 3. Les histogrammes avec les rendements analytiques ou les aires détectées sont présents en annexe de ce manuscrit, ainsi les droites de calibration des standards (Annexes 15 à 23). N'ayant pas trouvé de méthodes analytiques en UHPLC pour détecter les lactames résultant des pools 4 et 5 et l'acide 2-hydroxy-4-aminobutyrique (**7a**) du pool 2, nous n'avons pas été en mesure de vérifier les activités par UHPLC avec les CoA ligases sélectionnées.

Pour le pool 1, la détection des lactames simples ( $\gamma$ -butyrolactame **1b**,  $\delta$ -valérolactame **2b**,  $\varepsilon$ -caprolactame **3b**) avait déjà été mise au point par Chloé Lelièvre (Figure 27). Des rendements avec les CoA ligases ont pu être calculés à partir d'une droite de calibration des standards commerciaux des lactames (Annexe 18). Pour le  $\delta$ -valérolactame, la cyclisation spontanée de l'acide 5-aminopentanoïque a été déduite des rendements obtenus pour ce lactame.

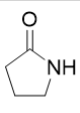
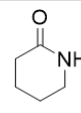
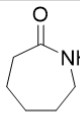
	 <b>1b</b>	 <b>2b</b>	 <b>3b</b>
A4IM18	2%	8%	
A4IMJ1			
A5D1Z4			
F8A8M3	3%		
Q3ABN2	62%		
B9L4W4	19%		
D1A9C5	3%		
A4YDT1	8%	2%	
F2NLM4			
Q72L27		3%	
D1AB87	2%	5%	
F2NPU2	2%		
K6Q029	12%	10%	
C5A3C0		3%	

Figure 27. Heat map des activités des CoA ligases sélectionnées pour le pool 1 suite à la détection des lactames attendus (**1b-3b**) par UHPLC-MS (APCI). Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl<sub>2</sub>, 1,5 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Analyse par UHPLC-MS (APCI). Rendements déduits d'une gamme de calibration avec les standards commerciaux. Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge.

Après 5h de réaction, les rendements obtenus pour la synthèse du  $\gamma$ -butyrolactame (**1b**) et/ou  $\delta$ -valérolactame (**2b**) sont très modestes dans l'ensemble (Figure 27). L'activité a été confirmée pour 11 CoA ligases sur les 14 sélectionnées. Deux enzymes sortent du lot, Q3ABN2 de *Carboxydotherrus hydrogenoformans* et B9L4W4 de *Thermomicrobium roseum*, avec environ 62% et 19% de rendement en  $\gamma$ -butyrolactame respectivement. Ces mêmes CoA ligases étaient ressorties comme hits lors de la confirmation des hits pour le 1<sup>er</sup> criblage de cette collection par Chloé Lelièvre. Nous avons aussi retrouvé l'activité de K6Q029 qui avait déjà été identifiée, mais les résultats ne sont pas cohérents avec ce qui a été obtenu précédemment. Nous attendons une activité croissante en fonction de la taille du substrat de départ et donc un rendement en caprolactame supérieur à 10%.

Les lactames du 2<sup>ème</sup> pool sont substitués par un groupement amine primaire et ils sont détectés par dérivation de cette amine avec du chlorure de benzoyle (BzCl) (Figure 28). Pour ce pool de substrats, nous ne disposons que du standard lactame du 2-aminocaprolactame (**6b**) qui résulte de la cyclisation de la lysine (**6a**). Nous avons donc d'un côté vérifié l'activité des enzymes retenues sans quantification possible pour les lactames **4b** et **5b**, de l'autre suivi l'activité en déterminant le rendement analytique de la formation du lactame **6b** (voir Annexe 18 et 19). Globalement, une activité n'a été confirmée que pour 6 des 13 enzymes retenues après le 1<sup>er</sup> criblage et pour chacun des substrats il y a au moins une enzyme active. Dans l'ensemble, les aires obtenues pour les lactames **4b** et **5b** sont faibles, à noter que comme précédemment, de la cyclisation spontanée en lactame à 6 chaînons a été observée (2-aminovalérolactame, **5b**), ici à partir de l'ornithine (**5a**) (Annexe 19).

Pour la formation du 2-aminocaprolactame (**6b**), après vérification, trois enzymes ont montré une activité, B3G250, K6P2L2 et Q97UW7, avec des rendements allant jusqu'à environ 19% avec Q97UW7 (Annexe 20). Ce résultat est très intéressant car pour cet exemple la quantité d'enzyme testée était faible (0,016 mg/mL).

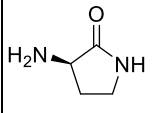
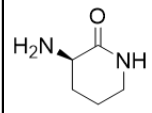
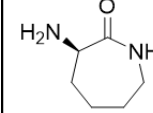
	 <b>4b</b>	 <b>5b</b>	 <b>6b</b>
A4IMJ1	n.d.		
Q3ABD3			
A4YDT1	n.d.		
B3G250	n.d.		10%
D3PTT7			
D6Y433			
D7BGN5			
D7BIL1			
F2NPU2			
K6P2L2			14%
Q97UW7			19%
E1QT70	n.d.	n.d.	
O29881			

Figure 28. Heat map des activités des CoA ligases sélectionnées pour le pool 2 suite à la détection des lactames attendus (**4b-6b**) par UHPLC-UV. Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 2 mM ATP, 2 mM  $MnCl_2$ , 2 mM  $MgCl_2$ , 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. Légende : n.d. pour non déterminé.

Pour les lactames formés avec le pool 3, constitué d'acides  $\omega$ -aminés substitués avec un groupement acide carboxylique, nous disposons du standard commercial du pyroglutamate (**8b**), le lactame résultant de la cyclisation du glutamate (**8a**). Nous avons pu déterminer des rendements pour ce lactame à partir d'une gamme de calibration (Annexe 22). Nous avons observé une forte cyclisation spontanée de l'acide glutamique, qui correspondait à environ 20% de rendement. Ce rendement de cyclisation spontanée a été soustrait des valeurs obtenues pour le rendement des CoA ligases actives. Comme pour le pool 2, nous n'avons pu que vérifier l'activité de manière qualitative pour les lactames **9b** et **10b** (Figure 29). Les aires obtenues pour ces deux lactames sont relativement faibles et l'activité est vérifiée pour 10 CoA ligases sur les 23 retenues (Annexe 21).

Pour la synthèse de pyroglutamate (**8b**) que nous avons pu quantifier, seules 3 enzymes ont vu leur activité confirmée, allant jusqu'à environ 33% de rendement avec la CoA ligase F2NPU2 de *Marinithermus hydrothermalis*. La CoA ligase K6P2L2, active pour la synthèse du lactame **6b**, a montré

aussi une bonne activité pour la synthèse de pyroglutamate avec environ 17% de rendement (Annexe 21).

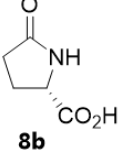
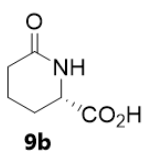
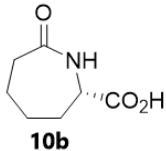
	 <b>8b</b>	 <b>9b</b>	 <b>10b</b>
A4IMJ1			
D3SQH8			
L0R9S0			
L0RAD3		n.d.	n.d.
Q3ABP3			
Q9HI39		n.d.	
D1A9C5			
D1ABC1			
D1C3U2		n.d.	
D1C448			
D6Y433		n.d.	n.d.
F2NLM4			
Q72L27		n.d.	
S9S7W1			
D1A3M5			
D1A5W6			
F2NPU2	33%	n.d.	
K6P2L2	17%		
K6Q029			
D9PUA1			
F2KPD0		n.d.	
O29881			n.d.
B0R4T6	9%		

Figure 29. Heat map des activités des CoA ligases sélectionnées pour le pool 3 suite à la détection des lactames attendus (**8b-10b**) par UHPLC-UV. Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl<sub>2</sub>, 1,5 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C sans agitation à l'étuve, 5h. Analyse par UHPLC-MS (ESI). Aires détectées à partir du m/z attendus pour les lactames correspondants. Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. Légende : n.d. pour non déterminé.

Enfin pour le pool 6, les lactames correspondants à la cyclisation des acides N-méthylbutyrique (**16a**) et N-éthylbutyrique (**17a**) ont pu être détectés en UHPLC-MS avec la même méthode que celle utilisée pour les lactames non fonctionnalisés du pool 1 (Annexe 23). Toutes les CoA ligases retenues sauf une, Q72L27, ont montré une activité, même parfois à l'état de traces, sur l'un des deux substrats, avec une nette préférence pour le substrat N-méthylé (**16a**) (Figure 30). Parmi toutes les CoA ligases actives, D1A9C5 a montré une activité bien plus forte que les autres enzymes sur les deux substrats testés.

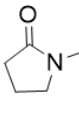
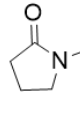
	 <b>16b</b>	 <b>17b</b>
A4IMJ1		
Q3ABP3		
Q3ABN2		
Q9HI39		
D1A9C5	++	++
F2NLM5		
Q72L27		
K6P2L2		
K6Q029		
D3RXJ3		
O28246		

Figure 30. Heat map des activités des CoA ligases sélectionnées pour le pool 6 suite à la détection des lactames attendus (**1b-3b**) par UHPLC-MS (APCI). Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl<sub>2</sub>, 1,5 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Détection des aires correspondants aux lactames attendus par UHPLC-MS (APCI). Pour chaque CoA ligases, les activités vérifiées sont indiquées en vert et les activités non vérifiées en rouge. Légende : ++ pour aire détectée forte.

Dans l'ensemble, les résultats UHPLC ont été cohérents avec les résultats obtenus pour le criblage avec la cascade PK-LDH, notamment pour les pool 1 et 6 où toutes les CoA ligases sélectionnées ont bien montré une activité avec les substrats du pool correspondant. Pour les pools 2 et 3, seulement 6 et 10 CoA ligases ont montré une activité sur les 13 et 20 enzymes retenues, respectivement. Le nombre plus faible est certainement dû à notre sélection très large de CoA ligases, en sélectionnant des activités très faibles avec le test d'activité PK-LDH.

En tout, 27 nouvelles enzymes ont montré une activité sur au moins un substrat fonctionnalisé (Annexe 16). Ce résultat est très positif et encourageant pour la suite. Sur la base de ces résultats, nous avons sélectionné les enzymes les plus prometteuses pour constituer un nouveau set de référence pour explorer encore plus profondément la biodiversité répertoriée, afin d'identifier de nouvelles enzymes ayant le potentiel de convertir une gamme de substrats plus diversifiés. Afin de choisir au mieux ces CoA ligases, nous avons analysé plus en détail ces enzymes en faisant un alignement de séquences ainsi qu'une superposition des sites actifs en utilisant la CoA ligase A4YDT1 comme référence.

### III.2. ANALYSE DES RESULTATS COA LIG1 ET NOUVEAU SET DE REFERENCE

Notre but étant de trouver des enzymes actives sur des acides  $\omega$ -aminés fonctionnalisés et/ou à longue chaîne, nous avons choisi des enzymes actives sur le pool 1 uniquement lorsqu'elles ont également montré une activité sur un autre pool. Ainsi nous avons retenu Q3ABN2 qui a montré une forte activité pour la synthèse de  $\gamma$ -butyrolactame (**1b**) et aussi sur le N-méthylbutyrolactame (**16b**), et K6Q029 que nous savions active sur l'acide 6-aminohexanoïque, précurseur du caprolactame, même si nous n'avons retrouvé cette activité dans nos tests, et aussi sur N-méthylbutyrolactame (**16b**). Pour le pool 2, nous avons sélectionné les CoA ligases Q97UW7 et K6P2L2, également retenue pour son activité

sur le pool 3 avec F2NPU2. Ensuite pour le pool 6, nous avons décidé de garder les CoA ligases D1A9C5 et O28246 qui avaient les aires les plus élevées pour la synthèse de lactames N-alkyles. A noter, parmi les enzymes sélectionnées, K6P2L2 et K6Q029 proviennent du même organisme *Thermaerobacter subterraneus*.

Tableau 13. Liste des CoA ligases constituant le nouveau set de référence pour la 2e collection de CoA ligases et le pool de substrats acides  $\omega$ -aminés sur lequel elles sont actives.

Code plaque	Code UniProt	Organisme	Annotation	Activité
CoA_1_H03	D1A9C5	<i>Thermomonospora curvata</i> (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	AMP-dependent synthetase and ligase	Pool 6
CoA_1_F05	Q3ABN2	<i>Carboxydotherrmus hydrogenoformans</i> (strain ATCC BAA-161 / DSM 6008 / Z-2901)	AMP-binding enzyme family protein	Pools 1 et 6
CoA_2_H02	K6P2L2	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II	Pools 2, 3 et 6
CoA_2_H05	K6Q029	<i>Thermaerobacter subterraneus</i> DSM 13965	Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II	Pools 1 et 6
CoA_3_D09	O28246	<i>Archaeoglobus fulgidus</i> (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	Medium-chain acyl-CoA ligase (AlkK-5)	Pool 6
CoA_2_G09	F2NPU2	<i>Marinithermus hydrothermalis</i> (strain DSM 14884 / JCM 11576 / T1)	O-succinylbenzoate--CoA ligase (EC 6.2.1.26)	Pool 3
CoA_3_A06	Q97UW7	<i>Sulfolobus solfataricus</i> (strain ATCC 35092 / DSM 1617 / JCM 11322 / P2)	Medium-chain-fatty-acid--CoA ligase (AlkK-4) (EC 6.2.1.-)	Pool 2

Une analyse détaillée de cette sélection a donc été réalisée pour valider le choix de ces CoA ligases pour former ce nouveau set de référence.

### III.2.1. ALIGNEMENT DE SEQUENCES DES COA LIGASES SELECTIONNEES

La matrice d'identité de séquences des CoA ligases sélectionnées pour constituer le nouveau set de référence fait apparaître une forte homologie de séquence (62%) entre K6P2L2 de *Thermaerobacter subterraneus* et D1A9C5 de *Thermomonospora curvata* (Figure 31). Les CoA ligases O28246 de *Archaeoglobus fulgidus* et Q97UW7 de *Sulfolobus solfataricus* partagent 50% d'homologie de séquence. En dehors de ces homologies fortes, les pourcentages d'identité de séquence entre les CoA ligases varient entre 25 à 39% environ, ce qui est positif pour continuer à explorer la biodiversité des CoA ligases. Pour rappel, nous souhaitons disposer de nouvelles CoA ligases actives sur les acides  $\omega$ -aminés fonctionnalisés. Une faible homologie entre des CoA ligases, qui ne sont pas actives sur les mêmes pools de substrats, pourrait nous ramener de nouvelles CoA ligases avec des spectres d'activité différentes également.



	K6Q029	F2NPU2	K6P2L2	D1A9C5	Q3ABN2	O28246	Q97UW7
K6Q029	100	35	38	36	34	26	25
F2NPU2	35	100	39	36	32	29	25
K6P2L2	38	39	100	62	31	25	26
D1A9C5	36	36	62	100	30	25	24
Q3ABN2	34	32	31	30	100	37	31
O28246	26	29	25	25	37	100	50
Q97UW7	25	25	26	24	31	50	100

Figure 31. Matrice d'identité avec les CoA ligases sélectionnées pour former le nouveau set de référence.

Nous avons réalisé un alignement de séquences protéiques de ces CoA ligases (Figure 32). On retrouve les motifs communs aux CoA ligases (voir chapitre précédent). Nous avons identifié les motifs A3 et A10 qui sont essentiels pour l'adénylation, ainsi que l'aspartate charnière présent dans le motif A8 qui se situe à l'interface des deux domaines. Dans le cas de la CoA ligase F2NPU2, nous avons noté la présence d'un résidu glutamate. Ce glutamate charnière est déjà observé chez d'autres O-succinylbenzoate CoA ligases (Chen, Sun et al. 2015, Chen, Jiang et al. 2016).

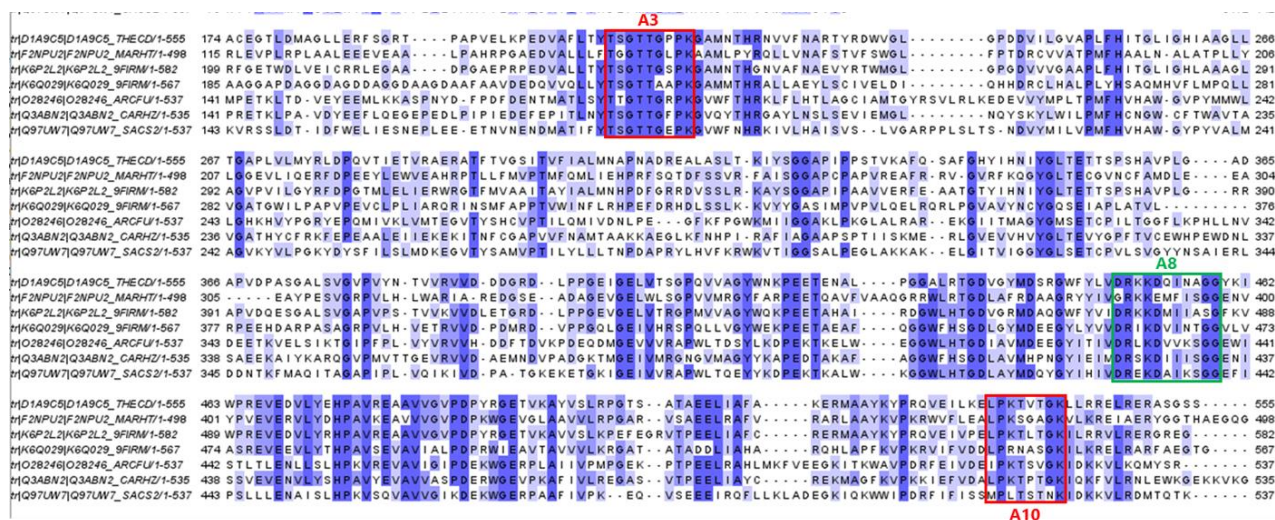


Figure 32. Alignement des séquences peptidiques des CoA ligases constituant le nouveau set de référence par ClustalOmega. Les motifs A3 et A10 essentiels pour l'adénylation sont indiqués en vert et la région charnière A8 est encadrée en vert.

Suite à cet alignement de séquences, une superposition des structures des CoA ligases a été réalisée pour identifier les résidus présents dans le site actif, notamment dans le site de liaison du substrat.

### III.2.2. SUPERPOSITION DES SITES ACTIFS

La superposition des structures des CoA ligases retenues a été réalisée à l'aide du logiciel PyMol et les structures sont des modèles générés par AlphaFold et récupérés depuis UniProt. Le but de cette superposition est de comparer les sites actifs et d'identifier leurs différences afin de déterminer quels pourraient être les éléments structuraux responsables de la spécificité de substrats. Nous avons utilisé comme base de comparaison la CoA ligase A4YDT1 (voir chapitre précédent, code PDB : 8BIT et 8BIQ), même si la comparaison avec des modèles AlphaFold n'est pas optimale car les modèles de CoA ligases sont générés en conformation d'adénylation, tandis que la structure d'A4YDT1 a été obtenue à partir de la protéine cristallisée en conformation de thioestérification (comme la plupart des structures de CoA



ligases). Cela a eu pour principale conséquence l'éloignement de résidus acides  $\omega$ -aminés impliqués dans la liaison de l'AMP.

Le tableau ci-dessous présente les résidus identifiés comme importants dans le site de liaison du substrat et de l'AMP pour chacune des CoA ligases, en prenant l'enzyme A4YDT1 comme référence (Tableau 14).

Tableau 14. Identification des résidus importants pour l'étape d'adénylation dans les sites actifs de chaque CoA ligase du nouveau set de référence. Les résidus ont été identifiés par superposition leur structure à celle de A4YDT1 comme enzyme référente (PDB : 8BIQ et 8BIT) dans le logiciel PyMOL. Les résidus sont colorés selon l'annotation ClustalOmega : en rouge pour les résidus petits et hydrophobes ; en bleu pour les résidus acides ; en magenta pour les résidus basiques ; en vert pour les résidus portant un groupement hydroxyl, thiol, amine et le résidu glycine. Les résidus conservés sont indiqués en gras.

	<b>A4YDT1</b> (référence)	<b>K6Q029</b>	<b>F2NPU2</b>	<b>K6P2L2</b>	<b>D1A9C5</b>	<b>Q3ABN2</b>	<b>O28246</b>	<b>Q97UW7</b>
<b>Site de liaison du substrat</b>	A255	S269	F193	I279	I254	C224	H230	H229
	K256	A270	H194	T280	T255	N225	V231	V230
	W259	H273	L197	I283	I258	W231	W234	W233
	A327	G339	G264	G349	G324	G294	G298	I299
	<b>G352</b>	<b>G365</b>	<b>G288</b>	<b>G374</b>	<b>G349</b>	<b>G317</b>	<b>G322</b>	<b>G324</b>
<b>Site de liaison de l'AMP</b>	G328	A340	G265	G350	G325	A295	G299	G300
	E329	S341	A266	A351	A326	A296	A300	G301
	P330	I342	P267	P352	P327	P297	K301	S302
	D349	N362	Q285	N371	N346	H314	G320	G321
	<b>Y351</b>	<b>Y364</b>	<b>Y287</b>	<b>Y373</b>	<b>Y348</b>	<b>Y316</b>	<b>Y321</b>	<b>Y323</b>
	Q353	Q366	L289	L375	L350	L318	M323	L325
	<b>D435</b>	<b>D446</b>	<b>D373</b>	<b>D461</b>	<b>D435</b>	<b>D410</b>	<b>D414</b>	<b>D415</b>
	K456	N467	I394	I482	N456	I431	K435	K436
R461	L472	N399	K487	K461	N436	W440	F441	

Tout d'abord, nous avons retrouvé des résidus conservés par toutes les CoA ligases, qui interviennent dans le bon alignement de l'AMP ou du substrat dans le site actif. Les CoA ligases qui partagent les pourcentages d'identité les plus élevés, K6P2L2 et D1A9C5 d'une part ; et O28246 et Q97UW7 d'autre part, sont composés des mêmes résidus dans leur site de liaison du substrat. En prenant en compte le pool de substrat pour lesquelles elles sont chacune actives, nous avons noté que K6P2L2 et D1A9C5 n'ont qu'un seul pool d'activité en commun tandis que Q97UW7 et O28246 ont été identifiées pour des pools différents (pool 2 et 6 respectivement).

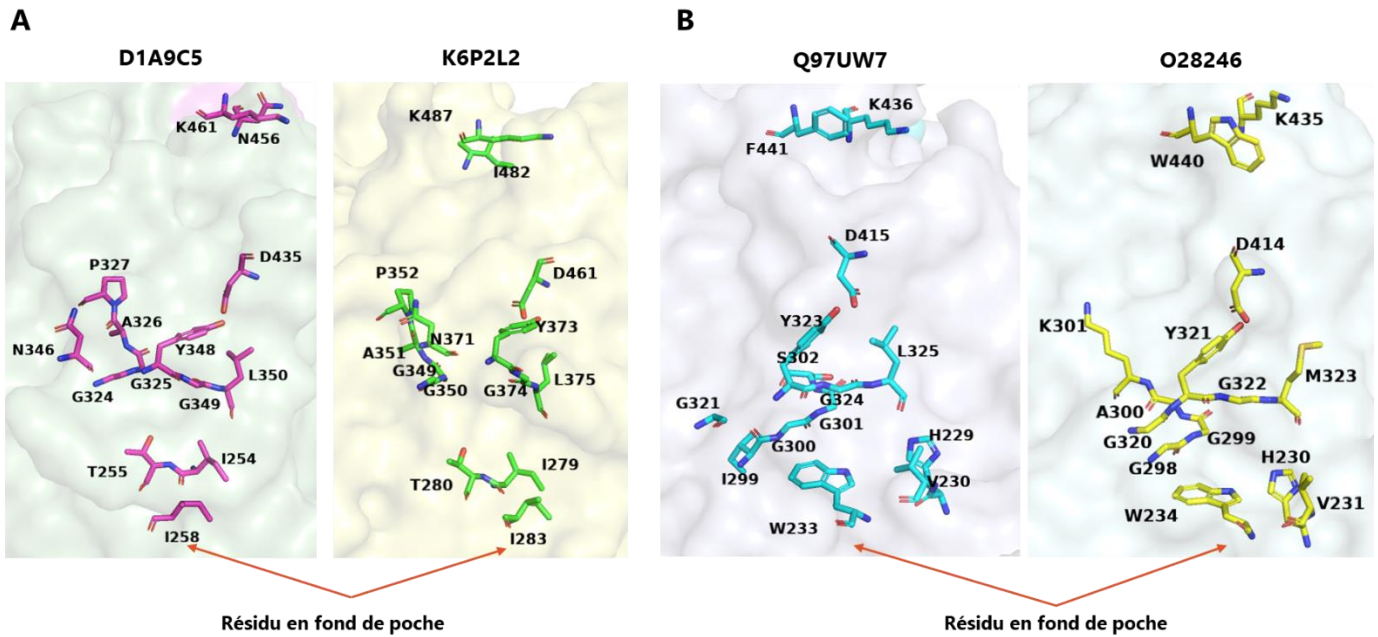


Figure 33. Comparaison des résidus constituant le site de liaison du substrat et de l'AMP pour les CoA ligases ayant les plus fortes homologues de séquence : (A) D1A9C5 et K6P2L2 avec 62% d'identité et (B) Q97UW7 et O28246 avec 50% d'identité. Identification des résidus effectuée par superposition des structures avec A4YDT1 comme CoA ligase de référence sur le logiciel PyMOL.

Une superposition des sites actifs entre les CoA ligases du nouveau set de référence a été réalisée en fonction du pool sur lesquels elles sont actives. Pour le pool 1, nous avons identifié K6Q029 et Q3ABN2. La superposition de leur sites actifs avec A4YDT1 ont montré que leur poche de liaison au substrat semble plus profonde que celle de A4YDT1, bien qu'elles contiennent aussi des résidus encombrants dans une position similaire à W259, le résidu plancher de A4YDT1 (Figure 34).

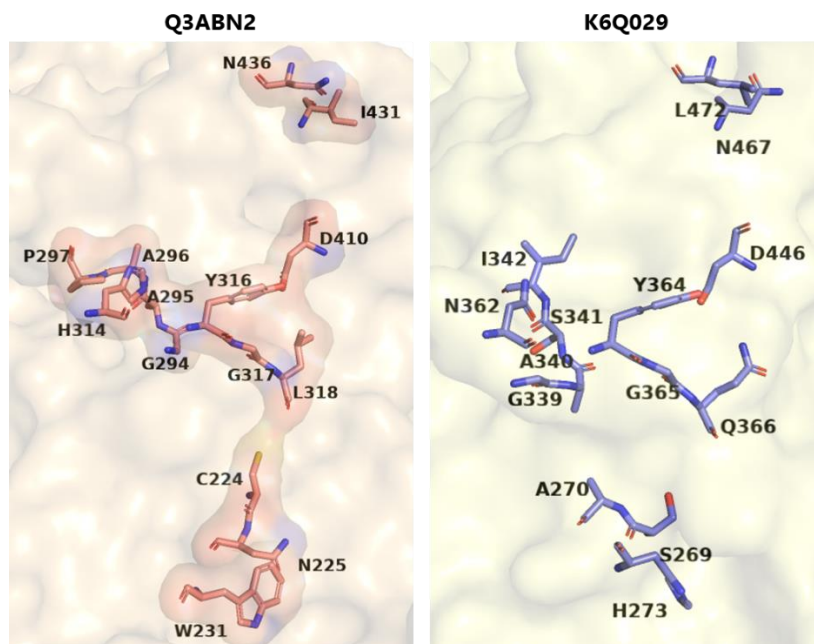


Figure 34. Résidus formant le site de liaison du substrat et de l'AMP dans le site actif des CoA ligases Q3ABN2 et K6Q029. Identification des résidus effectuée par superposition des structures avec A4YDT1 comme CoA ligase de référence sur le logiciel PyMOL.

Les enzymes actives sur le pool 2 ont des poches plus larges que celle de A4YDT1, d'où le fait qu'elles acceptent de plus larges substrats comme la lysine pour former du 2-aminocaprolactame. La CoA ligase Q97UW7 semble avoir un site actif légèrement moins profond et possède une poche plus hydrophobe comparé à K6P2L2 (Figure 33). La CoA ligase F2NPU2 identifiée aussi pour le pool 3, est annotée comme une O-succinylbenzoate CoA ligase. Son site actif contient deux résidus aromatiques, F193 et H194 qui peuvent établir des interactions  $\pi$ - $\pi$  avec son substrat natif. L'histidine pourrait aussi interagir avec le groupement acide carboxylique des substrats de ce pool, ce qui expliquerait également son activité sur le glutamate pour donner le pyroglutamate.

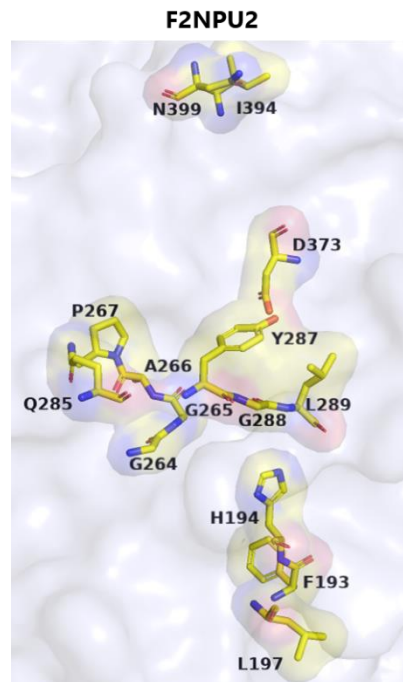


Figure 35. Résidus formant le site de liaison du substrat et de l'AMP dans le site actif de la CoA ligase F2NPU2. Identification des résidus effectuée par superposition des structures avec A4YDT1 comme CoA ligase de référence sur le logiciel PyMOL.

Enfin pour le pool 6, nous avons retenu les CoA ligases D1A9C5 et O28246. La superposition de leur site actif de ces deux CoA ligases a montré que D1A9C5 a une poche moins hydrophobe que O28246 (Figure 33).

Suite à cette analyse, nous avons décidé de continuer avec ces enzymes sélectionnées pour constituer le nouveau set de référence. A partir de ce nouveau set de référence, une nouvelle collection de CoA ligases a été construite, que nous avons nommé CoA Lig2 pour la suite du projet.

### III.3. CONSTRUCTION ET PRODUCTION DE COA LIG2

#### III.3.1. CONSTRUCTION DE LA DEUXIEME COLLECTION DE COA LIGASES

Une nouvelle collection de CoA ligases a été construite à partir du set de référence constitué des 7 CoA ligases retenues à l'issue du criblage et de l'analyse des résultats. Ce travail a été effectué à nouveau en collaboration avec Jean-Louis Petit de la plateforme FunDivEx de l'unité. Nous avons utilisé les paramètres standards (30% à 80% identité de séquences sur 80% de la longueur) pour continuer l'exploration de la biodiversité au sein de la famille des CoA ligases. La recherche dans les données

(méta)génomiques a permis de récupérer 211 796 enzymes candidates provenant de bactéries et d'archées. Cette liste a été réduite en prenant en compte le caractère thermophile des organismes d'où proviennent les CoA, défini à partir de 55°C. Au final, la nouvelle collection de CoA ligases CoA Lig2 est composée de 119 enzymes, 112 nouvelles CoA ligases identifiées et les 7 CoA ligases référentes.

### **III.3.2. PRODUCTION DE LA DEUXIEME COLLECTION DE COA LIGASES**

Nous avons rencontré de multiples difficultés pour la production de cette deuxième collection de CoA ligases. Nous avons déjà rencontré des difficultés d'expression lors de la production de la 1<sup>ère</sup> collection, seulement environ 30% des enzymes correctement exprimées, mais là nos problèmes ont été d'un tout autre ordre. Tout d'abord, il y a eu une contamination sur la plateforme FunDivEx. La source a pu être identifiée grâce à la ténacité de l'équipe, il s'est avéré que cette contamination était due à une lipase, utilisée dans un autre de nos projets. Le vecteur pET-22b(+) étant contaminé, toutes les enzymes produites correspondaient donc à cette lipase. De plus, nous avons rencontré d'autres problèmes : l'épuisement du stock de vecteur pET-22b(+) et la difficulté à en produire de nouveau en qualité et quantité notamment. Pour avancer au mieux, nous avons décidé de produire une collection réduite de 20 CoA ligases en passant par un prestataire extérieur, Twist Bioscience, pour la fourniture des plasmides incluant les gènes synthétiques codons optimisés. Nous avons ensuite travaillé avec l'équipe d'Egon Heuson de l'Unité de Catalyse et Chimie du Solide (UCCS) à Centrale Lille. Le but de cette collaboration était de produire ces 20 enzymes avec une mise en place d'automatisation de cette production, puis du criblage de la collection sur la plateforme REALCAT hébergée à Centrale Lille. Un tel procédé n'avait encore jamais été mis en place sur la plateforme et nos collaborateurs étaient très intéressés pour le faire. De notre côté, cela nous permettait d'avancer dans nos travaux, et à titre personnel je découvrais ainsi un nouvel univers avec la robotisation des criblages et des tests enzymatiques. Une automatisation de ces expériences nécessite au préalable de disposer d'un protocole fiable et robuste et donc une grande rigueur dans les manipulations expérimentales manuelles. C'est primordial afin d'adapter au mieux nos protocoles sur la plateforme, surtout pour les étapes clés ou critiques. Même si je n'ai pas eu le temps de pouvoir aller jusqu'à cette étape d'automatisation, j'ai été impressionnée de voir ces robots en action pour remplacer des tâches manuelles qui peuvent être laborieuses comme le pipetage en plaques de 96 puits. Cela permet de supprimer la plus grande source d'erreur qui est l'expérimentateur à mon sens, sans parler des possibles problèmes de calibration de la pipette multicanaux. La plateforme de criblage haut débit REALCAT a une partie dédiée à la biocatalyse équipée de robots pouvant sélectionner et repiquer des colonies de microorganismes en plaque 96 puits, mais aussi de robots de pipetage automatisé pour produire et cribler des biocatalyseurs.

Cette collection réduite doit refléter au mieux la diversité de la collection entière. Pour choisir les 20 CoA ligases nous nous sommes basés sur la composition du site actif en examinant la nature des résidus aux positions clés précédemment identifiées, comme nous l'avons déjà fait pour choisir les CoA ligases du set de référence.

Tout d'abord, nous avons classé les enzymes en fonction de la CoA ligase du set de référence pour laquelle elles ont été rapatriées (Annexe 25). Les enzymes K6Q029, Q3ABN2, O28246 et Q97UW7 sont les CoA ligases ramenant le moins d'enzymes dans leur groupe avec 8, 7, 10 et 6 enzymes respectivement. Les CoA ligases D1A9C5, F2NPU2 et K6P2L2 sont celles ayant le plus d'enzymes avec 29, 20 et 32 enzymes dans leur cluster respectif.

Afin de sélectionner au mieux nos CoA ligases, nous voulions une analyse fine des sites actif qu'une simple superposition PyMOL qui aurait été plus compliquée et plus longue à analyser. Nous nous

sommes tournés vers une ressource logiciel d'alignement de séquences appelée ASMC (pour *Active Sites Modeling and Clustering*). Mise au point dans l'unité puis développé par l'équipe LABGeM, l'ASMC classe les séquences d'enzymes en utilisant les informations structurales du site actif des protéines et réunit les séquences d'enzymes semblables dans des clusters (de Melo-Minardi, Bastard et al. 2010, Bastard, Smith et al. 2014, Bastard, Perret et al. 2017, Mayol, Bastard et al. 2019, Elisée, Ducrot et al. 2024). Même si pour avoir un résultat pleinement exploitable il est préférable de travailler sur un très grand nombre de séquences, avec l'aide du Dr Eddy Elisée, nous avons utilisé cette méthode en utilisant A4YDT1 comme enzyme référente puisque sa structure a été résolue.

Les résultats ASMC ont ressorti le même résidu en fond de poche que celui identifié par la superposition des structures sur PyMOL pour D1A9C5, F2NPU2, K6Q029 et K6P2L2 (Figure 36). Les CoA ligases O28246 et Q97UW7 ont un résidu valine en fond de poche tandis que Q3ABN2 a une cystéine, alors que par superposition nous avons un résidu tryptophane.

	Résidus du site actif de A4YDT1											
A4YDT1	A255	K256	W259	A327	P330	Y351	G352	T354	D435	F447	K456	R461
D1A9C5	I	T	I	G	P	Y	G	T	D	L	N	K
F2NPU2	F	H	L	G	P	Y	G	T	D	I	I	N
K6P2L2	-	L	I	G	P	Y	G	T	D	V	I	K
K6Q029	S	A	H	G	I	Y	G	S	D	V	N	L
O28246	P	M	V	G	K	G	M	E	D	I	K	W
Q3ABN2	-	M	C	A	A	Y	G	T	D	I	I	N
Q97UW7	P	M	V	I	S	Y	G	S	D	I	K	F

Figure 36. Identification des résidus constituant les sites actifs suite à l'analyse ASMC en utilisant la structure de la CoA ligase A4YDT1 comme enzyme référente. En rouge est indiqué les résidus situés en fond de poche pour chaque CoA ligase

Nous avons choisi de ne pas retenir les enzymes ayant un résidu encombrant (W, F, Y et P) à la position équivalente à W259 de A4YDT1, afin de nous éloigner de la structure de poche d'A4YDT1 et des enzymes connues (Shah, Ingram-Smith et al. 2009, Sofoe, Hart et al. 2019). D'autres critères de sélection ont été pris en compte, comme les résidus qui pourraient stabiliser le groupement amine terminal donc des résidus avec un groupement polaire sur leur chaîne latérale. Les enzymes ayant une glycine en fond de poche ont également été retenues car généralement cela correspond à des CoA ligases acceptant de larges substrats. De plus, nous avons sélectionné des enzymes ayant le même résidu en fond de poche que chaque CoA ligase de référence, puis nous avons regardé les autres résidus sur les positions adjacentes pour voir s'ils pourraient rentrer en interaction pour stabiliser d'autres substituants sur nos substrats. Suite à cette première sélection, lorsque deux protéines avaient la même séquence nous avons choisi celle dont le gène correspondant a le contenu en GC le plus faible, et enfin nous avons vérifié la diversité des séquences choisies en comparant les séquences avec une matrice d'identité.

La liste des 20 CoA ligases retenues ainsi la matrice de pourcentage d'identité de séquence associée sont présentées ci-dessous (Tableau 15, Figure 37). Globalement, l'homologie entre les 20 CoA ligases sélectionnées et celles du set de référence sont faibles. Cela reste un point positif pour continuer dans cette exploration de la diversité, car cela pourrait sûrement amener à avoir des CoA ligases avec différentes spécificités de substrat.

Tableau 15. Liste des 20 CoA ligases sélectionnées de la 2ème collection CoA Lig2.

<b>Code UniProt</b>	<b>Organisme</b>	<b>Annotation</b>
A0A0S3QRM5	<i>Thermosulfidibacter takaii</i>	Long-chain acyl-CoA synthetase
A0A1W1X0Y7	<i>Desulfacinum hydrothermale</i>	Fatty-acyl-CoA synthase
A0A1W1X6Z8	<i>Desulfacinum hydrothermale</i>	Long-chain acyl-CoA synthetase
A0A7J2THR7	<i>Archaeoglobus fulgidus</i>	Long-chain fatty acid--CoA ligase
A0A7J3M128	<i>Archaeoglobus fulgidus</i>	Fatty-acid--CoA ligase
A0A7R7YJS6	<i>Thermus thermophilus</i>	Acetyl-CoA synthetase
A9WAN2	<i>Chloroflexus aurantiacus</i>	AMP-dependent synthetase and ligase
C8WUV5	<i>Alicyclobacillus acidocaldarius subsp. acidocaldarius</i>	AMP-dependent synthetase and ligase
C8WWY7	<i>Alicyclobacillus acidocaldarius subsp. acidocaldarius</i>	AMP-dependent synthetase and ligase
C9R9F7	<i>Ammonifex degensii</i>	Acetyl-coenzyme A synthetase
D1C813	<i>Sphaerobacter thermophilus</i>	Acyl-CoA ligase (AMP-forming), exosortase system type 1 associated
F8IFB8	<i>Alicyclobacillus acidocaldarius</i>	acetate--CoA ligase
G0EDR0	<i>Pyrolobus fumarii</i>	AMP-dependent synthetase and ligase
G4RKH3	<i>Thermoproteus tenax</i>	Acetyl-coenzyme A synthetase
G7V9K7	<i>Thermovirga lienii</i>	AMP-dependent synthetase and ligase
O29585	<i>Archaeoglobus fulgidus</i>	Medium-chain acyl-CoA ligase (AlkK-3)
Q1ATL1	<i>Rubrobacter xylanophilus</i>	AMP-dependent synthetase and ligase
Q1AUW1	<i>Rubrobacter xylanophilus</i>	AMP-dependent synthetase and ligase
Q1AZ49	<i>Rubrobacter xylanophilus</i>	AMP-dependent synthetase and ligase
Q97UF6	<i>Saccharolobus solfataricus</i>	Long-chain-fatty-acid--CoA ligase (FadD-3)



	O29585	Q3ABN2	A0A7J3M128	O28246	Q97UW7	A9WAN2	F8IFB8	C9R9F7	D1C813	G4RKH3	G0EDR0	A0A7R7YJ56	A0A7J2THR7	Q97UF6	G7V9K7	F2NPU2	C8WUV5	Q1ATL1	D1A9C5	K6P2L2	A0A0S3QRM5	C8WWY7	A0A1W1X6Z8	Q1AUW1	K6Q029	Q1AZ49	A0A1W1X0Y7
O29585	100	29	35	34	37	21	22	22	23	25	24	25	28	25	25	25	25	23	23	26	24	23	21	25	26	30	28
Q3ABN2	29	100	32	37	31	21	23	24	26	25	26	26	32	31	29	32	29	29	31	31	30	31	31	28	34	37	33
A0A7J3M128	35	32	100	43	38	21	22	20	20	24	20	25	27	24	24	26	23	25	24	25	25	24	25	23	27	29	28
O28246	34	37	43	100	50	21	22	24	22	24	24	24	28	22	23	28	24	27	24	24	25	26	28	25	27	33	29
Q97UW7	37	31	38	50	100	23	21	22	23	25	22	22	28	24	23	25	27	25	23	26	26	26	27	26	25	29	26
A9WAN2	21	21	21	21	23	100	37	46	24	31	23	26	24	25	22	25	26	26	25	29	23	24	24	27	25	24	24
F8IFB8	22	23	22	22	21	37	100	47	22	30	24	32	24	24	22	28	24	26	28	28	23	26	23	26	27	25	24
C9R9F7	22	24	20	24	22	46	47	100	23	30	25	30	25	26	23	25	26	27	28	30	24	29	27	25	26	28	25
D1C813	23	26	20	22	23	24	22	23	100	24	27	27	26	28	25	29	29	29	31	29	26	31	29	27	31	30	29
G4RKH3	25	25	24	24	25	31	30	30	24	100	28	33	29	27	30	26	27	27	31	31	28	28	28	26	28	30	27
G0EDR0	24	26	20	24	22	23	24	25	27	28	100	27	31	28	24	30	31	32	29	33	29	33	29	28	25	30	27
A0A7R7YJ56	25	26	25	24	22	26	32	30	27	33	27	100	29	28	26	28	28	29	31	32	28	33	32	32	28	30	28
A0A7J2THR7	28	32	27	28	28	24	24	25	26	29	31	29	100	30	28	27	36	33	31	35	33	35	33	29	27	31	32
Q97UF6	25	31	24	22	24	25	24	26	28	27	28	28	30	100	32	31	34	31	38	40	34	35	35	31	29	31	29
G7V9K7	25	29	24	23	23	22	22	23	25	30	24	26	28	32	100	31	29	29	30	31	31	32	33	32	28	33	29
F2NPU2	25	32	26	28	25	25	28	25	29	26	30	28	27	31	31	100	34	34	36	39	33	37	37	32	35	35	34
C8WUV5	25	29	23	24	27	26	24	26	29	27	31	28	36	34	29	34	100	62	38	41	35	41	37	34	34	35	31
Q1ATL1	23	29	25	27	25	26	26	27	29	27	32	29	33	31	29	34	62	100	39	41	34	42	35	35	33	36	32
D1A9C5	23	31	24	24	23	25	28	28	31	31	29	31	31	38	30	36	38	39	100	62	35	42	39	34	36	36	33
K6P2L2	26	31	25	24	26	29	28	30	29	31	33	32	35	40	31	39	41	41	62	100	36	41	40	35	37	37	35
A0A0S3QRM5	24	30	25	25	26	23	23	24	26	28	29	28	33	34	31	33	35	34	35	36	100	42	43	31	29	33	32
C8WWY7	23	31	24	26	26	24	26	29	31	28	33	33	35	35	32	37	41	42	42	41	42	100	53	34	33	35	34
A0A1W1X6Z8	21	31	25	28	27	24	23	27	29	28	29	32	33	35	33	37	37	35	39	40	43	53	100	33	30	33	35
Q1AUW1	25	28	23	25	26	27	26	25	27	26	28	32	29	31	32	32	34	35	34	35	31	34	33	100	33	37	33
K6Q029	26	34	27	27	25	25	27	26	31	28	25	28	27	29	28	35	34	33	36	37	29	33	30	33	100	40	38
Q1AZ49	30	37	29	33	29	24	25	28	30	30	30	30	31	31	33	35	35	36	36	37	33	35	33	37	40	100	41
A0A1W1X0Y7	28	33	28	29	26	24	24	25	29	27	27	28	32	29	29	34	31	32	33	35	32	34	35	33	38	41	100

Figure 37. Matrice d'identité des 20 CoA ligases sélectionnées pour former la collection réduite de CoA Lig2 et les 7 CoA ligases de référence.

Le tableau de comparaison des résidus du site actif de chaque CoA ligases est mis en annexe et la figure ci-dessous présente les proportions des résidus présents en position équivalente à chaque résidu du site actif d'A4YDT1.

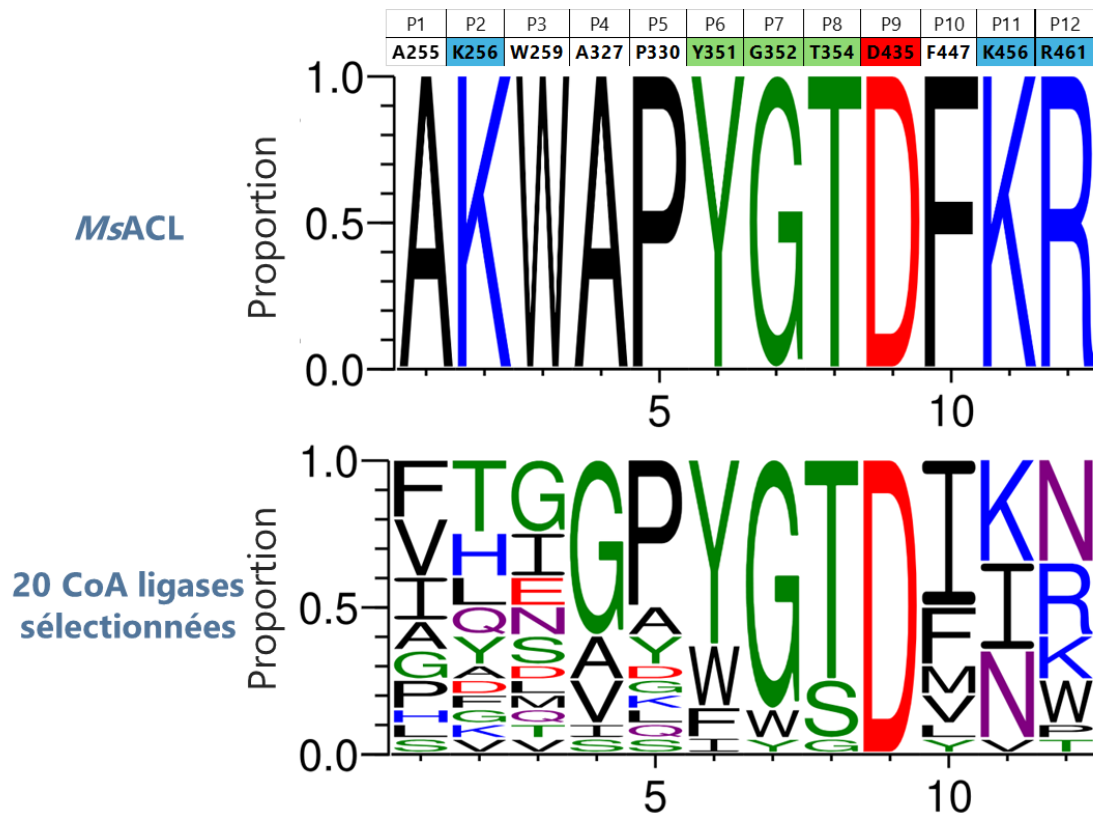


Figure 38. Proportions des résidus présents chez les 20 CoA ligases sélectionnées en position équivalente aux résidus présents dans le site actif d'A4YDT1 suite à l'alignement de séquences avec le logiciel ASMC. Figure faite sur Web Logo 3 par Dr Eddy Elisée (LABGeM).

Au final, cette collection réduite est composée de 5 CoA ligases ayant un résidu glycine en fond de poche, 5 d'entre elles ont un résidu leucine, isoleucine ou valine comme les enzymes de référence mais avec des résidus adjacents différents de ces derniers, ainsi qu'une CoA ligase avec un résidu méthionine (Figure 38). Il y a également 3 CoA ligases avec des résidus sérine et thréonine, 3 autres avec un résidu glutamate ou aspartate ainsi que 2 CoA ligases avec un résidu asparagine et un dernier avec un résidu glutamine. De manière plus globale, nous avons sélectionné des CoA ligases avec des sites de liaison du substrat divers et qui pourraient être actifs avec des substrats acides  $\omega$ -aminés de taille et de fonctionnalité diverses. Par ailleurs, cette figure met aussi en avant les résidus bien conservés par toutes comme le résidu aspartate impliqué dans l'ancrage de l'ATP dans le site actif. Un tableau détaillé de la composition des sites actifs des 20 CoA ligases sélectionnées a été mis en annexe (Annexe 26).

Le design des plasmides des 20 enzymes a été réalisé par Jean-Louis Petit et les plasmides ont été commandés chez Twist Bioscience, dont le vecteur d'expression est le vecteur pET-21(+).

### III.3.3. PRODUCTION DES 20 ENZYMES



### **III.3.3.1. PRODUCTION DE LA COLLECTION REDUITE DE CoA LIGASE A L'UCCS (CENTRALE LILLE)**

La surexpression individuelle des 20 enzymes et des 7 CoA ligases du set de référence (dont l'ADN a été fourni par la plateforme FunDivEx) a été réalisée en compagnie des doctorants Lucas Bisel et Oumaima Rabaoui de l'équipe d'Egon Heuson à Centrale Lille. Comme pour une production classique d'enzymes détaillée précédemment, les plasmides sont donc transformés dans des cellules compétentes d'expression BL21. Suite à la transformation, deux colonies par boîte ont été repiquées. Cela nous a permis de vérifier que le plasmide a bien été intégré dans nos cellules et que celles sélectionnées ne sont pas de fausses positives. Les colonies vérifiées ont été mise en culture et les cellules ont été lysées à la French Press. Le gel SDS-PAGE a montré qu'aucune CoA ligase n'a été surexprimée, contrairement aux enzymes de référence dont l'ADN provenait du Genoscope.

De retour au Genoscope, nous avons tenté de surexprimer une CoA ligase au hasard, Q1AZ49 en même temps que K6Q029 qui est non seulement une enzyme du set de référence mais qui sert aussi de témoin pour la production. Le même protocole a été utilisé et de nouveau il n'y a eu aucune surexpression de Q1AZ49. En parallèle, les constructions ont été vérifiées au Genoscope et nous avons contacté Twist Bioscience qui a affirmé de s'être assuré de la conformité des plasmides avant de les envoyer. Comme précisé précédemment le taux d'échec de production des CoA ligases de CoALig1 était très élevé avec seulement 31% de succès. Une autre possibilité serait que, par malchance, nous ayons choisi que des CoA ligases difficiles à exprimer avec le protocole générique.

Pour essayer de résoudre ce problème, nous avons tenté de faire un sous-clonage de nos enzymes : nous avons transféré les gènes de nos CoA ligases dans le vecteur d'expression pET-22b(+) utilisé en routine sur la plateforme.

### **III.3.3.2. SOUS-CLONAGE DES 20 CoA LIGASES DANS PET-22B(+)**

Pour réaliser ce sous-clonage en plaque de 96 puits, les amorces pour chaque CoA ligase ont été commandées préalablement. Tout d'abord une amplification PCR de nos fragments d'intérêt a été réalisée puis les amorces ont été utilisées en présence de la T4 polymérase pour générer des extrémités 5' et 3' cohésives pour faire le clonage par ligation indépendante (LIC) détaillée en annexe (Annexe 12). Les plasmides ont d'abord été transformés dans des cellules compétentes NovaBlue généralement employées pour augmenter le nombre de copies plasmidiques. Pour ces deux étapes, une vérification par électrophorèse sur gel d'agarose a été effectuée (Annexe 22). Après transformation dans les cellules compétentes d'expression BL21, les enzymes ont été mises en culture et surexprimées en utilisant le protocole habituel. Pour vérifier qu'il n'y a eu pas de problèmes au niveau de la production, 3 CoA ligases du set de référence (D1A9C5, K6P2L2 et K6Q029) ont été ajoutées sur la plaque comme contrôles positifs. Enfin, la purification a été réalisée en mini-colonne en plaque, un gel SDS-PAGE et la concentration en enzyme de chaque CoA ligase a été déterminée par la méthode de Bradford (Annexe 29). Au final, 8 CoA ligases ont été obtenues sous forme purifiée sur 20 au total, en plus des 3 CoA ligases de référence, soit 40% de succès de production, ce qui est légèrement supérieur aux taux de succès précédent pour la production de la collection CoA Lig1 (environ 30%). Par ailleurs, nous avons observé sur le gel SDS-PAGE des pertes de certaines enzymes entre la lyse et la purification. Cette collection réduite a été criblée comme détaillé précédemment.

Tableau 16. Liste des 8 CoA ligases obtenues sous forme purifiée.

Code UniProt	Organisme	Annotation
A0A1W1X0Y7	<i>Desulfacinum hydrothermale</i>	Fatty-acyl-CoA synthase
A0A7J2THR7	<i>Archaeoglobus fulgidus</i>	Long-chain fatty acid--CoA ligase
C8WUV5	<i>Alicyclobacillus acidocaldarius</i> <i>subsp. acidocaldarius</i>	AMP-dependent synthetase and ligase
C8WWY7	<i>Alicyclobacillus acidocaldarius</i> <i>subsp. acidocaldarius</i>	AMP-dependent synthetase and ligase
D1C813	<i>Sphaerobacter thermophilus</i>	Acyl-CoA ligase (AMP-forming), exosortase system type 1 associated
F8IFB8	<i>Alicyclobacillus acidocaldarius</i>	acetate--CoA ligase
G0EDR0	<i>Pyrolobus fumarii</i>	AMP-dependent synthetase and ligase
Q1AZ49	<i>Rubrobacter xylanophilus</i>	AMP-dependent synthetase and ligase

### III.4. CRIBLAGE DE LA COLLECTION REDUITE DE 20 COA LIGASES

Les 8 CoA ligases ont été criblées sur les pools de substrats 1 à 6 pendant une réaction de 5h à 60°C et à isovolume d'enzyme puis les résultats ont été analysés en utilisant la cascade PK-LDH. Comme pour le 1<sup>er</sup> criblage, cela amène donc à obtenir une réponse du type OUI/NON sur l'activité des CoA ligases. Plusieurs contrôles sans enzyme ont été ajoutés sur la plaque afin d'avoir une valeur moyenne seuil de blanc. Cette valeur a été soustraite des valeurs de pente des 8 CoA ligases pour chaque pool et nous avons considéré comme positives les activités de CoA dont les réponses issues de la cascade PK-LDH sont supérieures à cette valeur seuil. Les résultats bruts obtenus avec la cascade PK-LDH avec toutes les enzymes sont mis en annexe, seuls les résultats avec les 8 CoA ligases produites et les 3 CoA ligases de référence (D1A9C5, K6P2L2 et K6Q029) sont présentés sous forme de heat map.

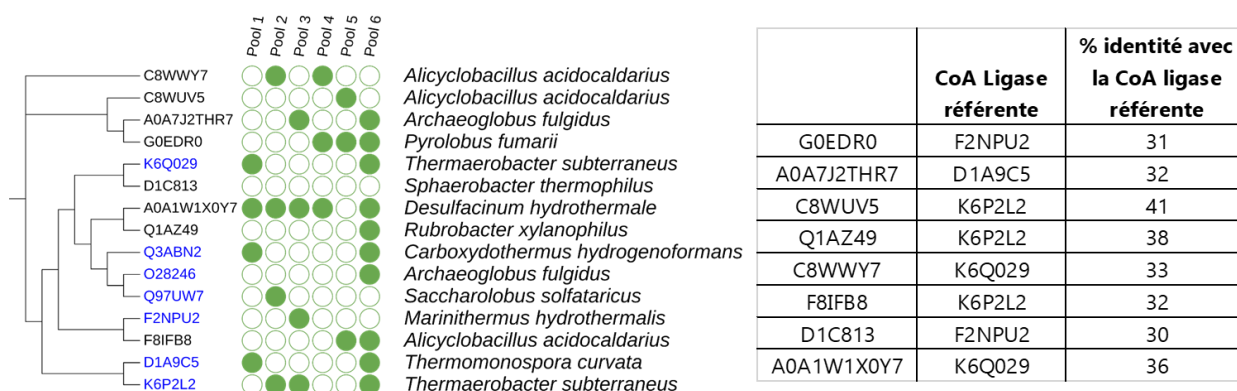


Figure 39. Heat map des activités des 8 CoA ligases de la collection réduite de CoA Lig2 suite à la cascade PK-LDH. Les enzymes du set de référence sont indiquées en bleu. Expérience réalisée en duplicat. Figure réalisée par Jean-Louis Petit de la plateforme FunDivEx.

Tout d'abord pour les enzymes de référence, nous avons retrouvé les activités attendues. Concernant les 8 enzymes produites, la CoA ligase A0A1W1X0Y7 est active pour la formation de lactames pour tous les pools sauf le pool 5. La CoA ligase G0EDR0 est active sur les substrats du pool 4, 5 et 6. C8WUV5 présente également une bonne activité sur le pool 5. Ces deux enzymes ont un résidu acide

en fond de poche (Glu et Asp respectivement). A0A7J2THR7 qui a aussi un résidu Glu en fond de poche, est active sur les substrats du pool 2 et 6. Enfin, la CoA ligase C8WWY7 a montré une activité sur les pools 2 et 4. Cette enzyme possède un résidu glycine en fond de poche ainsi que les même adjacents que Q1AZ49 qui n'a montré une activité que pour le pool 6.

Ces premiers résultats de criblage avec le test PK-LDH nécessitent d'être vérifiés et quantifiés le cas échéant.

### **III.5. CONCLUSION ET PERSPECTIVES SUR L'EXPLORATION DE LA BIODIVERSITE POUR LA RECHERCHE DE NOUVELLES COA LIGASES POUR LA SYNTHESE DE LACTAMES**

Le criblage de la première collection de CoA ligases avec divers substrats acides  $\omega$ -aminés nous a permis d'identifier de 55 nouvelles enzymes actives pour la synthèse biocatalytique de lactames fonctionnalisés. Suite à ce criblage, sept d'entre elles ont été retenues pour constituer un nouveau set de référence afin de construire une deuxième collection : K6Q029, D1A9C5, F2NPU2, K6P2L2, O28246, Q3ABN2 et Q97WU7. Ce processus itératif a permis d'explorer davantage la biodiversité répertoriée, afin d'identifier de nouvelles enzymes avec des activités prometteuses sur les substrats testés. Il faut noter que, confrontés à une série de problèmes expérimentaux, nous n'avons pas pu tester l'entièreté de la nouvelle collection CoA Lig2 (119 enzymes) et nous avons testé un nombre restreint d'enzymes en choisissant certaines d'entre elles sur des critères structuraux. Pour cela nous nous sommes appuyés sur les travaux menés sur l'enzyme A4YDT1 (voir chapitre II) qui nous ont permis de passer d'une enzyme inactive sur l'acide 6-aminohexanoïque, précurseur du caprolactame, à une enzyme permettant un rendement analytique de 30% grâce à une mutation ciblant le résidu du fond de poche. Ainsi, une collection de 8 CoA ligases a pu être criblée sur les pools de substrats fonctionnalisés et seule une n'a montré aucune activité dans les conditions du test. Ces résultats sont en cours de validation.

A noter que très récemment, nous avons réussi à produire la deuxième collection de CoA ligases dans son entier. Nous allons donc pouvoir cribler la collection sur les 6 pools de substrats.

## REFERENCES BIBLIOGRAPHIQUES

Bastard, K., A. Perret, A. Mariage, T. Bessonnet, A. Pinet-Turpault, J.-L. Petit, E. Darii, P. Bazire, C. Vergne-Vaxelaire, C. Brewee, A. Debard, V. Pellouin, M. Besnard-Gonnet, F. Artiguenave, C. Médigue, D. Vallenet, A. Danchin, A. Zaparucha, J. Weissenbach, M. Salanoubat and V. de Berardinis (2017). "Parallel evolution of non-homologous isofunctional enzymes in methionine biosynthesis." Nature Chemical Biology **13**(8): 858-866.

Bastard, K., A. A. T. Smith, C. Vergne-Vaxelaire, A. Perret, A. Zaparucha, R. De Melo-Minardi, A. Mariage, M. Boutard, A. Debard, C. Lechaplais, C. Pelle, V. Pellouin, N. Perchat, J.-L. Petit, A. Kreimeyer, C. Medigue, J. Weissenbach, F. Artiguenave, V. De Berardinis, D. Vallenet and M. Salanoubat (2014). "Revealing the hidden functional diversity of an enzyme family." Nature Chemical Biology **10**(1): 42-49.

Buller, R., S. Lutz, R. J. Kazlauskas, R. Snajdrova, J. C. Moore and U. T. Bornscheuer (2023). "From nature to industry: Harnessing enzymes for biocatalysis." Science **382**(6673): eadh8615.

Chen, Y., Y. Jiang and Z. Guo (2016). "Mechanistic Insights from the Crystal Structure of *Bacillus subtilis* o-Succinylbenzoyl-CoA Synthetase Complexed with the Adenylate Intermediate." Biochemistry **55**(48): 6685-6695.

Chen, Y., Y. Sun, H. Song and Z. Guo (2015). "Structural Basis for the ATP-dependent Configuration of Adenylation Active Site in *Bacillus subtilis* o-Succinylbenzoyl-CoA Synthetase\*." Journal of Biological Chemistry **290**(39): 23971-23983.

de Melo-Minardi, R. C., K. Bastard and F. Artiguenave (2010). "Identification of subfamily-specific sites based on active sites modeling and clustering." Bioinformatics **26**(24): 3075-3082.

Elisée, E., L. Ducrot, R. Méheust, K. Bastard, A. Fossey-Jouenne, G. Grogan, E. Pelletier, J.-L. Petit, M. Stam, V. de Berardinis, A. Zaparucha, D. Vallenet and C. Vergne-Vaxelaire (2024). "A refined picture of the native amine dehydrogenase family revealed by extensive biodiversity screening." Nature Communications **15**(1): 4933.

Kittilä, T., M. Schoppet and M. J. Cryle (2016). "Online Pyrophosphate Assay for Analyzing Adenylation Domains of Nonribosomal Peptide Synthetases." Chembiochem **17**(7): 576-584.

Mayol, O., K. Bastard, L. Beloti, A. Frese, J. P. Turkenburg, J.-L. Petit, A. Mariage, A. Debard, V. Pellouin, A. Perret, V. de Berardinis, A. Zaparucha, G. Grogan and C. Vergne-Vaxelaire (2019). "A family of native amine dehydrogenases for the asymmetric reductive amination of ketones." Nature Catalysis **2**(4): 324-333.

McQuade, T. J., A. D. Shallop, A. Sheoran, J. E. DelProposto, O. V. Tsodikov and S. Garneau-Tsodikova (2009). "A nonradioactive high-throughput assay for screening and characterization of adenylation domains for nonribosomal peptide combinatorial biosynthesis." Analytical Biochemistry **386**(2): 244-250.

Robinson, S. L., J. Piel and S. Sunagawa (2021). "A roadmap for metagenomic enzyme discovery." Natural Product Reports **38**(11): 1994-2023.

Scott, D. P., T. Yang, S. Valerie and D. M. Andrew (2010). "A Universal, Fully Automated High Throughput Screening Assay for Pyrophosphate and Phosphate Release from Enzymatic Reactions." Combinatorial Chemistry & High Throughput Screening **13**(1): 27-38.

Shah, M. B., C. Ingram-Smith, L. L. Cooper, J. Qu, Y. Meng, K. S. Smith and A. M. Gulick (2009). "The 2.1 Å crystal structure of an acyl-CoA synthetase from *Methanosarcina acetivorans* reveals an alternate acyl-binding pocket for small branched acyl substrates." Proteins: Structure, Function, and Bioinformatics **77**(3): 685-698.

Sofeo, N., J. H. Hart, B. Butler, D. J. Oliver, M. D. Yandeau-Nelson and B. J. Nikolau (2019). "Altering the Substrate Specificity of Acetyl-CoA Synthetase by Rational Mutagenesis of the Carboxylate Binding Pocket." ACS Synthetic Biology **8**(6): 1325-1336.

Vardakou, M., M. Salmon, J. A. Faraldos and P. E. O'Maille (2014). "Comparative analysis and validation of the malachite green assay for the high throughput biochemical characterization of terpene synthases." MethodsX **1**: 187-196.

Vergne-Vaxelaire, C., A. Mariage, J.-L. Petit, A. Fossey-Jouenne, C. Guérard-Hélaine, E. Darii, A. Debard, S. Nepert, V. Pellouin, M. Lemaire, A. Zaparucha, M. Salanoubat and V. de Berardinis (2018). "Characterization of a thermotolerant ROK-type mannofructokinase from *Streptococcus mitis*: application to the synthesis of phosphorylated sugars." Applied Microbiology and Biotechnology **102**(13): 5569-5583.

Zaparucha, A., V. de Berardinis and C. Vaxelaire-Vergne (2018). Genome Mining for Enzyme Discovery. Modern Biocatalysis: Advances Towards Synthetic Biological Systems. G. Williams and M. Hall, The Royal Society of Chemistry: 0.

## CONCLUSION ET PERSPECTIVES

Ces quatre années de thèse se sont déroulées à la suite des travaux initiés par Dr Chloé Lelièvre au cours de sa thèse de 2017-2020, portant sur la synthèse d'amides et de lactames par synthèse chemo-enzymatique en utilisant des CoA ligases dans une réaction détournée.

Dans un premier temps, l'approche guidée par la structure pour étendre le spectre de substrats d'une CoA ligase a été poursuivie. En collaboration avec Dr Nikolas Capra et les professeurs Dick Janssen et Andy-Mark Thunnissen de l'Université de Groningen (Pays-Bas), la structure de la CoA ligase thermophile A4YDT1 de *Metallosphaera sedulla* a été cristallisée. Au cours de la thèse de Chloé Lelièvre, il avait été montré qu'un résidu tryptophane situé en fond de poche était particulièrement important pour la spécificité de substrats. Suite à cette analyse, différents mutants avaient été conçus et produits. Ces travaux ont été complétés et les paramètres cinétiques des enzymes actives, sauvage et mutants, ont été déterminés. Grâce à ces travaux, nous avons pu passer d'une CoA ligase native dont l'activité était limitée à la formation de  $\gamma$ -butyrolactame, à une CoA ligase mutée catalysant la synthèse de lactames de 6 à 7 chaînons. Avec le mutant W259G, nous avons pu aller jusqu'à 30% de rendement analytique pour la synthèse de  $\epsilon$ -caprolactame, précurseur de nylon 6,6 très employé dans le monde du textile. Ces résultats restent tout de même modestes, notamment pour la synthèse de  $\delta$ -valérolactame avec seulement 18% de rendement avec le meilleur mutant, comparé à d'autres enzymes dépendantes de l'ATP comme les CARs où cette synthèse a atteint 85% de rendement (Qin, Zhang et al. 2022). Il est tout de même important de noter qu'il n'existe que peu d'exemples de CoA ligases étudiées pour la synthèse de lactames, seule la CoA ligase ORF27 ayant été décrite pour la synthèse de  $\epsilon$ -caprolactame (Zhang, Barajas et al. 2017). Ces travaux nous ont permis d'ajouter une nouvelle CoA ligase dans cette liste, encore restreinte pour le moment. Par ailleurs, nous avons accédé à la synthèse de plus larges amides grâce à l'élargissement du site de liaison du substrat, avec la prise en charge de plus larges acides, jusqu'à l'acide undécanoïque (C11). Ces résultats n'ont pu être que qualitatifs et nous n'avons pas été en mesure de quantifier les N-méthylamides obtenus afin de pouvoir comparer pleinement nos résultats avec d'autres CoA ligases. Ces travaux ont fait l'objet d'un article apparu dans *Frontiers in Catalysis* (Capra, Lelièvre et al. 2024).

La seconde partie de cette thèse s'est concentrée sur la recherche de nouvelles CoA ligases par exploration de la biodiversité. La première collection CoA Lig1 construite durant la thèse de C. Lelièvre, a été criblée avec 17 acides  $\omega$ -aminés fonctionnalisés, bio-sourcés ou bio-sourçables pour la plupart, pour la synthèse de lactames fonctionnalisés. Ces substrats ont été répartis en pools selon leurs caractéristiques structurales et/ou physicochimiques. Ce criblage nous a permis d'identifier 55 nouvelles CoA ligases pour la synthèse de lactames fonctionnalisés. Suite à ce criblage, 7 CoA ligases ont été retenues pour constituer un nouveau set de référence, afin d'explorer davantage la biodiversité répertoriée et de trouver des enzymes avec de meilleures activités sur nos substrats. Ainsi, par un processus itératif, nous avons constitué une collection CoA Lig2 constituée de 119 enzymes, 112 nouvelles CoA ligases homologues en plus de nos 7 CoA ligases de référence. Malheureusement, à cause de problèmes de production, nous avons dû réduire la collection à 20 enzymes. En nous basant sur nos travaux d'ingénierie sur la CoA ligase A4YDT1, nous avons sélectionné les 20 CoA ligases présentant les meilleurs paramètres structuraux pour avoir une activité sur des substrats  $\omega$ -amino acides fonctionnalisés. De façon très prometteuse, un premier criblage a permis d'identifier 7 nouvelles CoA ligases. Ces résultats obtenus avec la collection réduite restent à confirmer.

Au cours de cette thèse, nous avons recherché de nouvelles de CoA ligases pour la synthèse biocatalytique de lactames. D'une part, nous sommes partis de la structure native d'une enzyme connue afin de comprendre sa spécificité de substrat et orienter son activité par ingénierie protéique pour la synthèse de lactames cibles. D'autre part, nous avons criblé une collection de CoA ligases issues de la biodiversité sur divers acides  $\omega$ -aminés à la recherche d'enzymes actives. Ce travail a servi de point de départ pour étendre l'exploration de la biodiversité et récupérer de nouvelles CoA ligases. Les résultats obtenus dans le chapitre II nous ont aidés pour une analyse des CoA ligases identifiées en s'appuyant sur des données structurales. Il serait intéressant de combiner l'approche exploratoire avec l'approche guidée par la structure, qui sont complémentaires pour des travaux d'optimisation des CoA ligases actives nouvellement identifiées. Ainsi, nous pourrions bénéficier de nouvelles CoA ligases qui pourraient s'ajouter à la liste de CoA ligases connues pour la synthèse de lactames (Zhang, Kao et al. 2015, Zhang, Barajas et al. 2017, Shen, Jiang et al. 2024). Pour aller encore plus loin, nous pourrions également faire une montée en échelle de la production de lactames avec les nouvelles CoA ligases identifiées mais aussi envisager d'employer ces CoA ligases en cellule entière comme celles décrites dans la littérature.

Enfin, il est important de rappeler que ces travaux de thèse s'intègrent au projet ANR SMALA (« from Small to MacroLactam, an enzymatic approach », 2022-2026) coordonné par Anne ZAPARUCHA. Si dans l'unité nous nous concentrons sur la synthèse de petits lactames (cycle jusqu'à 7 chaînons), nos collaborateurs de l'UMR IMoPA à Nancy s'intéressent à la synthèse enzymatique de macrolactames. Ceci fait actuellement l'objet de la thèse d'Arnaud BUZON (2022-2025) sous la direction du Dr Arnaud GRUEZ. Ce projet couvre aussi un volet d'ingénierie protéique, géré par l'équipe du Pr Magali REMAUD-SIMÉON du Toulouse Biotechnology Institute (TBI). Dr Dipesh JADAV, post-doctorant, travaille sur l'ingénierie protéique de K6Q029, CoA ligase identifiée comme la plus prometteuse. Il est également prévu des travaux d'ingénierie sur les enzymes qui auront été identifiées durant la thèse d'Arnaud BOUZON. Enfin, une montée en échelle de la synthèse biocatalytique de lactames avec K6Q029 en bioréacteurs de 1 et 5 litres est en cours avec nos collaborateurs Dr Louis MOUTERDE et le post-doctorant Dr Vivien HERRSCHER de l'URD ABI (AgroParisTech, Pomacle). Un système de purification membranaire est également étudié pour l'intensification du procédé. Le but de ce projet est de disposer d'enzymes adaptées pour la synthèse biocatalytique de petits et grands lactames d'intérêt pouvant servir de constituants de base en chimie fine ou encore de précurseurs de polymères. De plus, la montée en échelle de cette synthèse en bioréacteurs pourrait constituer une preuve de faisabilité pouvant faciliter une potentielle implémentation de ce procédé à l'échelle industrielle.

## REFERENCES BIBLIOGRAPHIQUES

Capra, N., C. Lelièvre, O. Touré, A. Fossey-Jouenne, C. Vergne-Vaxelaire, D. B. Janssen, A.-M. W. H. Thunnissen and A. Zaparucha (2024). "Adapting an acyl CoA ligase from *Metallosphaera sedula* for lactam formation by structure-guided protein engineering." Frontiers in Catalysis **4**.

Qin, Z., X. Zhang, X. Sang, W. Zhang, G. Qu and Z. Sun (2022). "Carboxylic acid reductases enable intramolecular lactamization reactions." Green Synthesis and Catalysis.

Shen, X., X. Jiang, X. Sun, Q. Yuan and J. Wang (2024). "Identifying and charactering a 4-aminobutyryl-CoA ligase for the production of butyrolactam." Journal of Biotechnology **382**: 21-27.

Zhang, J., J. F. Barajas, M. Burdu, G. Wang, E. E. Baidoo and J. D. Keasling (2017). "Application of an Acyl-CoA Ligase from *Streptomyces aizunensis* for Lactam Biosynthesis." ACS Synth Biol **6**(5): 884-890.

Zhang, J., E. Kao, G. Wang, E. E. K. Baidoo, M. Chen and J. D. Keasling (2015). "Metabolic engineering of *Escherichia coli* for the biosynthesis of 2-pyrrolidone." Metabolic engineering communications **3**: 1-7.



## EXPERIMENTAL SECTION

### GENERALS

#### CHEMICALS

All reagents were purchased from commercial sources (Sigma Aldrich®, Enamine®, AmBeed®, Combi-Blocks®) and used without additional purification. Pyruvate Kinase, from rabbit muscle (PK, P9136), L-Lactic Dehydrogenase, from rabbit muscle (LDH, L1254), Myokinase, from rabbit muscle (MK, M3003), Phospho(enol)pyruvic acid monopotassium salt (PEP-K, 860077),  $\beta$ -Nicotinamide adenine dinucleotide, reduced dipotassium salt (NADH, N4505) were purchased from Sigma Aldrich.

#### EQUIPMENT

Biocatalytic reactions were carried out under stirring using Eppendorf Thermomixer® and screening assays in microplates were carried without stirring in an oven at fixed temperature.

UPLC-MS analyses were performed on UHPLC U3000 RS bar system (Thermo Fisher Scientific®) coupled to two different mass spectrophotometers:

- MSQ Plus™ Single Quadrupole with electrospray ESI in positive mode for amide detection, with the probe temperature at 450 °C and a cone voltage at 50 V. For lactam detection, the mass spectrophotometer was with APCI positive mode, the probe temperature was set up at 450 °C with a cone voltage at 60 V.
- ISQ EM Single Quadrupole with APCI (atmospheric-pressure chemical ionization) in positive mode with the probe current at 15  $\mu$ A, the source CID voltage at 60 V, the vaporization temperature was set up at 200°C and the Ion Transfert Tube (ITT) set up at 300 °C for simple lactams and the functionalized ones without a polar group. For lactams with a carboxylic group, the mass spectrophotometer was set up with ESI in positive mode.

Enzymatic spectrophotometric screening assays in microplates were performed on a SpectraMax Plus384 (Molecular Devices ®). The kinetic data were recorded on a Safas UVMC2 (Safas®) thermostated with a refrigerated/heating circulator Polystat36 (Thermo Fisher Scientific ®) or CORIO CD-200F cryostat (JULABO GmbH®) and using high-precision quartz microcells with 10 mm light path (Hellma Analytics®).

Preparative flash column chromatography was performed on a CombiFlash Companion® using GraveResolv Silica cartridges.

NMR spectra were recorded on a Bruker 600 MHz spectrometer (Université Evry-Val d'Essonne, Université Paris-Saclay, France). Chemical shifts (expressed in ppm) of  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were referenced to the solvent peaks  $\delta(\text{H})=7.24$  ppm and  $\delta(\text{C})= 77.2$  ppm for  $\text{CDCl}_3$ .

### ENZYME PRODUCTION

#### GENERAL PROCEDURE FOR ENZYME OVEREXPRESSION

All the steps from primers purchase to purified enzymes in 96-microwell plates were conducted by the FunDivEx platform (Adrien Debard and Jean-Louis Petit, LGBM). The primers were designed with addition of specific extension for cloning into our in-house plasmid pET-22b(+) (Novagen®) modified

for ligation independent cloning (LIC). The oligonucleotides were purchased from Merck/Sigma-Aldrich. The forward primers contains a His-Tag after the initial methionine for enzyme purification. The verified plasmids were introduced in *E.coli* BL21 CodonPlus (DE3) competent cells for protein overexpression. The transformed cells were grown on Terrific Broth (TB) medium containing 0.5 M sorbitol, 5 mM betaine, and 100 g.ml<sup>-1</sup> carbenicillin at 37 °C until reaching OD<sub>600</sub> between 1.04 and 1.8. Isopropyl β-δ-thiogalactopyranoside (IPTG) was added at 0.5 mM final concentration to induce protein production and the cells further grown overnight at 20 °C. The culture medium was split into 50 ml aliquots for centrifugation. After centrifugation, the pellets were stored at -80 °C for 2 h to ease the cells lysis. The frozen pellets were resuspended in 4 ml of lysis buffer (50 mM MOPS buffer pH 8.0, 50 mM NaCl, 30 mM imidazole and 10 % glycerol) containing 1 mM Pefabloc®SC and 15 µl of Lyonase™ bioprocessing reagent (Novagen®) and agitated for 30 min at room temperature and sonicated using an Ultrasonic Processor. After centrifugation, the cell-free extract was recovered and stored at -80 °C until purification.

### **GENERAL METHOD FOR SMALL-SCALE PURIFICATION OF ENZYMES**

On a small-scale, the enzymes were purified by nickel affinity chromatography. After equilibration of the Ni-NTA column (QIAGEN), the cell-free extract was loaded onto the column according the manufacturer's instructions. The elution buffer contained 50 mM MOPS buffer pH 8.0, 50 mM NaCl, 250 mM imidazole and 10 % glycerol. The fractions were then desalted using 10 K Amicon® Ultra-4 column (Merck) with desalting buffer containing 50 mM MOPS buffer pH 8.0, 50 mM NaCl and 10 % glycerol. Protein concentration of the purified enzymes was determined by the Bradford method with bovine serum albumin standard (BSA) (Bio-RAD®). Samples of the purified enzymes were analyzed by SDS-PAGEs using Invitrogen NuPAGE system. The purified enzymes were stored at -80 °C until biocatalytic tests.

### **GENERAL METHOD FOR LARGE-SCALE PRODUCTION OF ENZYMES**

Large-scale purifications were conducted by Peggy Sirvain (LGBM), from 500 ml culture by nickel affinity chromatography in tandem with gel filtration (Hi load 16/600 Superdex 200 pg 17-1069-01). The storage buffer was 50 mM MOPS pH 8.0, 50 mM NaCl, 10 % glycerol and 1 mM DTT. Protein concentrations were determined by the Bradford method with bovine serum albumin (BSA) as the standard (Bio-Rad®) and samples were analyzed by SDS-PAGEs using the Invitrogen NuPage system. The purified enzymes were stored at -80 °C.

### **GENERAL METHOD FOR 96-WELL PLATE PURIFICATION**

Enzymes were cloned and produced following the described protocol (cf general method). After overexpression of the recombinant genes in *E.coli* BL21 strains, the cells were lysed in the lysis buffer (50 mM MOPS buffer pH 8.0, 50 mM NaCl, 10 % glycerol, 1 mM Pefabloc® and 15 µl Lyonase™) in a Branson 2510 Ultrasonic cleaner. After centrifugation, the cell extract concentrations were determined by Bradford method with bovine serum albumin (BSA) as standard (Bio-Rad®). The samples were analyzed by SDS-PAGEs using Invitrogen E-PAGE system.

### **GENERAL METHOD FOR SITE-DIRECTED MUTAGENESIS**

#### **PCR PRIMERS**

All the primers used for the mutants cloning are shown in the table below (Table 1):

Table 1. Primer sequences for variants cloning and PCR control. Mutated codons are highlighted in blue.

<b>A4YDT1 (UniProt ID) mutants</b>	<b>Primers</b>	<b>Sequences</b>
V238T	Forward	ACCACAGCCTCCATA <b>ACC</b> GGCGTCAGGGAGAGCGATCTCCACTTA
	Reverse	TAAGTGGAGATCGCTCTCCCTGACGCC <b>GGT</b> TATGGAGGCTGTGGT
W259G	Forward	TGGGCCAAGTTCGCG <b>GGC</b> AGCTCATTCTTCTCTCCCCTT
	Reverse	AAGGGGAGAGAAGAATGAGCT <b>GCC</b> CGGAACTTGGCCCA
F350A	Forward	CTTACCATAAGGGAC <b>GCCT</b> ACGGCCAGACTGAGACC
	Reverse	GGTCTCAGTCTGGCCGTA <b>GGC</b> GTCCCTTATGGTAAG
F350T	Forward	CTTACCATAAGGGAC <b>ACCT</b> ACGGCCAGACTGAGACC
	Reverse	GGTCTCAGTCTGGCCGTA <b>GGI</b> GTCCCTTATGGTAAG
F350V	Forward	CTTACCATAAGGGAC <b>GTG</b> TACGGCCAGACTGAGACC
	Reverse	GGTCTCAGTCTGGCCGTA <b>CAC</b> GTCCCTTATGGTAAG
F350Y	Forward	CTTACCATAAGGGAC <b>TAT</b> TACGGCCAGACTGAGACC
	Reverse	GGTCTCAGTCTGGCCGTA <b>ATA</b> GTCCCTTATGGTAAG

#### **SITE-DIRECTED MUTAGENESIS PROTOCOL**

Single and multiple site mutations were carried out using Quick-Change® II XL Site-directed Kit (Agilent®) according to manufacturer's instructions. Forward primers of the native enzyme A4YDT1 (UniProt ID) were used for each mutations in the case of the double mutants and forward primer of W259/V238T mutant was used in the case of the triple mutants. Successful cloning and creation of the target mutations were confirmed by sequencing. The enzyme mutants were purified as described previously.

#### **EXPRESSION AND PURIFICATION OF THE MUTANTS**

Mutants were expressed and purified using the same protocol previously described.

## **ANALYTICAL BIOCATALYTIC REACTIONS**

### **SAMPLE PREPARATION AND REACTION**

#### **ANALYTICAL BIOCATALYTIC REACTIONS IN CHAPTER II**

##### **Biocatalytic reactions for lactam synthesis with A4YDT1 and A4YDT1 mutants**

A 100- $\mu$ l reaction mixture containing 5 mM  $\omega$ -amino acid, 5 mM ATP, 5 mM MnCl<sub>2</sub>, 50 mM MOPS buffer pH 8.5 and 2 mg/ml purified enzyme was prepared. The reaction mixture was stirred at 400 rpm in Eppendorf Thermomixer®, for 24 h at 60 °C. In parallel, negative controls without enzyme and substrate were performed. Samples were treated as described in "Sample treatment" section below.

The calibration curve was set from 0.5 to 5 mM with commercial standards of the lactams, using 10 mM (for 0.5 - 2 mM calibration points) and 50 mM (5 mM calibration point) stock solutions of each standard, as described in the Table 2 below. The calibration was incubated under the same conditions as the biocatalytic reactions.

Table 2. Preparation of the calibration curve for lactam monitoring by UHPLC-MS (APCI mode). Legend: **[C]i** for initial concentration, **[C]f** for final concentration and **Vtot** for total volume.

	<b>[C]i (mM)</b>	<b>[C]f (mM)</b>			
		<b>0,5</b>	<b>1</b>	<b>2</b>	<b>5</b>
butyrolactam	10	2,5	5	10	5
valerolactam	10	2,5	5	10	5
caprolactam	10	2,5	5	10	5
MOPS buffer pH8,5	1000	2,5	2,5	2,5	2,5
H <sub>2</sub> O (μl)		40	32,5	17,5	32,5
<b>Vtot (μl)</b>		<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>

### Biocatalytic reactions for N-methylamide synthesis with A4YDT1 and A4YDT1 mutants

A 50-μl reaction mixture contained 5 mM carboxylic acid, 5 mM ATP, 5 mM MgCl<sub>2</sub>, 50 mM methylamine, phosphate buffer pH 8 and 0.2 mg/ml of purified enzyme was prepared. The reaction mixture was stirred at 400 rpm in Eppendorf Thermomixer®, for 24 h at 60 °C. In parallel, negative controls without enzyme and substrate were performed. Samples were treated as described in "Sample treatment" section below.

The calibration curve was set from 0.5 to 5 mM with synthesized standards of the N-methylamides (from C4 to C8), using 5 mM (for 0.5 - 2 mM calibration points) and 50 mM (2.5 and 5 mM calibration point) stock solutions of each standard, as described in Table 3 below. The calibration was incubated under the same conditions as the biocatalytic reactions.

Table 3. Preparation of the calibration curve for N-methylamide monitoring by UHPLC-MS (ESI positive mode). Legend: **[C]i** for initial concentration, **[C]f** for final concentration and **Vtot** for total volume.

	<b>[C]i (mM)</b>	<b>[C]f (mM)</b>				
		<b>0,5</b>	<b>1</b>	<b>2</b>	<b>2,5</b>	<b>5</b>
N-methylamide	5 ou 50	5	10	20	2,5	5
Phosphate buffer pH 8	100	25	25	25	25	25
H <sub>2</sub> O	1000	20	15	5	22,5	20
<b>Vtot (μl)</b>		<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>

### Analytical biocatalytic reactions for kinetic parameters and specific activity

Kinetic parameters of A4YDT1 and A4YDT1 variants W259G and W259G/V238T

Kinetic parameters of A4YDT1 and variants for the ω-amino acid substrates and ATP were determined using the same conditions, with varying concentrations of ω-amino carboxylic acid and without methylamine. Reaction mixtures (40 μl) were preincubated at 60 °C for 3 min prior to addition of enzymes, after which incubation was continued for 30 min at 60 °C. After cooling on ice, the reaction mixture was transferred to a 10 mm cuvette containing 110 μl of the NADH-pyrophosphate assay mixture, described below.

The saturating reaction conditions for the determination of the kinetic parameters for ATP were 50 mM Tris-HCl buffer pH 7.5, 5 mM MnCl<sub>2</sub>, 10 mM GABA for A4YDT1 WT and 6 M for A4YDT1 variants W259G and W259G/V238T.

The saturating reaction conditions for the determination of the kinetic parameters for  $\omega$ -amino acid substrate were 50 mM Tris-HCl buffer pH 7.5, 5 mM MnCl<sub>2</sub>, 10 mM ATP.

#### Specific activities of A4YDT1 variants for N-methylamide synthesis

The reactions were performed at 60 °C in Tris-HCl buffer, pH 7.5, in a final volume of 60  $\mu$ l containing 5 mM substrate, 2 mM ATP, 5 mM MnCl<sub>2</sub> and 50 mM methylamine. Tris-HCl buffer at pH 7.5 gave a lower absorbance background than other buffers or pH values (data not shown). Reactions were initiated by addition of the appropriate amount of enzyme. Aliquots (15  $\mu$ l) were withdrawn at different times and transferred to a 10 mm cuvette containing 85  $\mu$ l of the NADH-pyrophosphate assay mixture previously heated at 37 °C. Monitoring of NADH disappearance at 340 nm was used to determine the specific activity of the enzyme, according to the supplier's instructions described below.

### ANALYTICAL BIOCATALYTIC REACTIONS IN CHAPTER III

#### Screening of the collection CoA Lig1

A 50- $\mu$ l reaction mixture contained 0.5 mM of  $\omega$ -amino acid in each pool, an adjusted concentration between 1-2 mM ATP, MnCl<sub>2</sub> and MgCl<sub>2</sub> in order to get 1 molar equivalent per substrate for each pool, 50 mM MOPS buffer pH 8.5 were prepared in 96-well plates. The plates were incubated at 60 °C, for 5 h without stirring. Samples were treated as described in "Sample treatment" section and used for the screening analysis as described in Spectrophometric assays for enzymatic activities" section below.

For UHPLC verification, the same biocatalytic reaction conditions were used. The calibration curves were set from 0.01 to 0.5 mM with commercial standards of the lactams, using 10 mM (for 0.5-2 mM calibration points) and 50 mM (5 mM calibration point) stock solutions of each standard, as described in the Table 4, Table 5 and Table 6 below. The calibration was incubated under the same conditions as the biocatalytic reactions. Samples were treated as described in "Sample treatment" section below.

Table 4. Preparation of the calibration curve for simple lactams monitoring by UHPLC-MS (APCI mode). Legend: **[C]<sub>i</sub>** for initial concentration, **[C]<sub>f</sub>** for final concentration and **V<sub>tot</sub>** for total volume.

	[C] <sub>i</sub> (mM)	[C] <sub>f</sub> (mM)			
		0,01	0,1	0,2	0,5
butyrolactam	1	0,5	5	10	2,5
valerolactam	1	0,5	5	10	2,5
caprolactam	1	0,5	5	10	2,5
MOPS buffer pH 8,5		2,5	2,5	2,5	2,5
H <sub>2</sub> O ( $\mu$ L)		46	32,5	17,5	40
<b>V<sub>tot</sub> (<math>\mu</math>L)</b>		<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>

Table 5. Preparation of the calibration curve for 2-aminocaprolactam monitoring by UHPLC-UV. Legend: **[C]i** for initial concentration, **[C]f** for final concentration and **Vtot** for total volume.

	<b>[C]i (mM)</b>	<b>[C]f (mM)</b>			
		<b>0,01</b>	<b>0,1</b>	<b>0,2</b>	<b>0,5</b>
2-aminocaprolactam	1	0,5	5	10	25
MOPS buffer pH 8,5		2,5	2,5	2,5	2,5
H <sub>2</sub> O (μL)		47	42,5	37,5	22,5
<b>Vtot (μL)</b>		<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>

Table 6. Preparation of the calibration curve for pyroglutamate monitoring by UHPLC-MS (ESI positive mode). Legend: **[C]i** for initial concentration, **[C]f** for final concentration and **Vtot** for total volume.

	<b>[C]i (mM)</b>	<b>[C]f (mM)</b>			
		<b>0,01</b>	<b>0,1</b>	<b>0,2</b>	<b>0,5</b>
pyroglutamate	1	0,5	10	1	25
MOPS buffer pH 8,5		2,5	2,5	2,5	2,5
H <sub>2</sub> O (μL)		47	37,5	46,5	22,5
<b>Vtot (μL)</b>		<b>50</b>	<b>50</b>	<b>50</b>	<b>50</b>

### Screening of the collection CoA Lig2

A 50-μl reaction mixture contained 1 mM of ω-amino acid in each pool, an adjusted concentration between 2 - 4 mM ATP, MnCl<sub>2</sub> and MgCl<sub>2</sub> in order to get 1 molar equivalent per substrate for each pool, 50 mM MOPS buffer pH 8.5 were prepared in 96-well plates. The plates were incubated at 60 °C, for 5 h without stirring. Samples were treated as described in "Sample treatment" section and used for the screening analysis as described in "Spectrophometric assays for enzymatic activities" section below.

## SAMPLE TREATMENT

### FOR UHPLC-MS (ESI AND APCI) ANALYSES OF LACTAMS

A 30-μl aliquot was withdrawn from the reaction mixture, quenched with 2 μl of 6 M HCl solution and diluted with 110 μl MeOH LC-grade solution. After filtration on 0.22 μm filter, the samples were analyzed by UPLC-MS using conditions A or B for lactam monitoring.

### FOR UHPLC-MS (ESI) ANALYSES OF N-METHYLAMIDES

A 30-μl aliquot were withdrawn from the reaction mixture, quenched with 2 μl of 6 M HCl solution and diluted with 100 μl H<sub>2</sub>O. After filtration on 0.22 μm filter, the samples were analyzed by UPLC-MS using conditions for amide monitoring.

## UHPLC-UV/MS AND DERIVATIZATION PROCEDURE WITH BENZOYL CHLORIDE

### BENZOYL CHLORIDE DERIVATIZATION

Lactams containing primary amine function were derivatized according to the following protocol: to 20 μl of reaction mixture were added 10 μl of 1 M Na<sub>2</sub>CO<sub>3</sub>/NaHCO<sub>3</sub> aqueous solution pH 10, 40 μl H<sub>2</sub>O and 30 μl of a 50 mM BzCl solution in acetonitrile. The mixture was vortexed for 30 s and quenched with 20 μl of 1 M HCl aqueous solution and 30 μl of a 1/1 solution of H<sub>2</sub>O/CH<sub>3</sub>CN. After centrifugation

(10,000 rpm, 10 min, RT) and filtration on 0.22  $\mu\text{m}$  filter, the samples were analyzed by UPLC-UV/MS according to the conditions described below.

### **UPLC-UV/MS (ESI) ANALYSES**

The samples were analyzed using a Kinetex® EVO-C18 (Phenomenex) column (100 x 2.1 nm; 1.7  $\mu\text{m}$ ). The mobile phases were  $\text{H}_2\text{O}$  + 0.1 % formic acid as solvent A and  $\text{CH}_3\text{CN}$  as solvent B. A linear gradient 95/5 (A/B ratio) was applied for 2 min, then 95/5 to 0/100 in 7 min and hold for 1 min, then 0/100 to 95/5 in 1 min and hold for 1 min with re-equilibration time of 2 min; flow 0.3  $\text{ml}\cdot\text{min}^{-1}$ ; 25 °C; injection volume 3  $\mu\text{l}$ .

### **UHPLC-MS ANALYSES**

Lactam formation monitoring was performed using a Luna® Omega Polar-C18 (Phenomenex) column (100 x 2.1 nm; 1.6  $\mu\text{m}$ ).

### **NON-FUNCTIONALIZED LACTAMS AND FUNCTIONALIZED LACTAMS WITH NON-POLAR GROUPS (APCI)**

Condition A: the mobile phases were  $\text{H}_2\text{O}$  as solvent A and  $\text{CH}_3\text{OH}$  as solvent B. A linear gradient 95/5 to 50/50 (A/B ratio) was applied in 1 min and hold at 50/50 for 2 min, then 50/50 to 95/5 in 1 min and hold for 2 min (re-equilibration time).

### **FUNCTIONALIZED LACTAMS WITH CARBOXYLIC GROUP (ESI)**

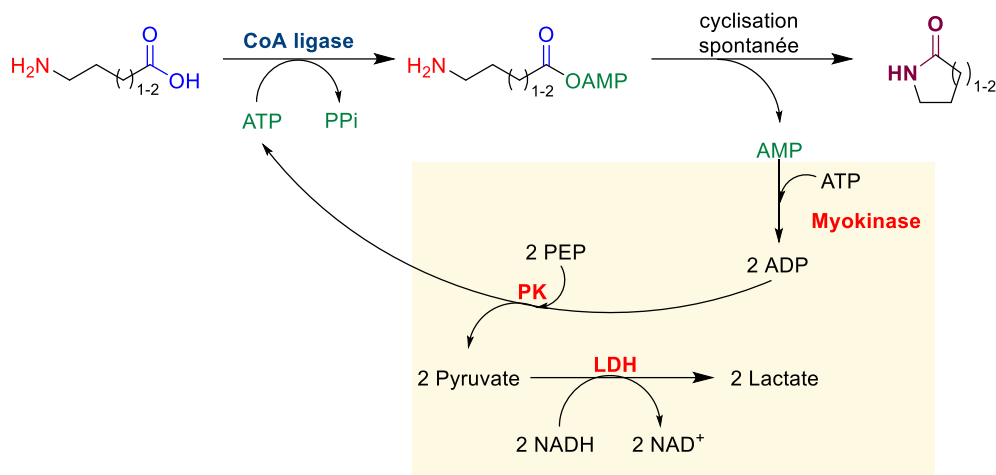
Condition B: the mobile phases were  $\text{H}_2\text{O}$  + 0.1 % formic acid as solvent A and  $\text{CH}_3\text{OH}$  as solvent B. A linear gradient 95/5 to 50/50 (A/B ratio) was applied in 1 min and hold at 50/50 for 2 min, then 50/50 to 95/5 in 1 min and hold for 2 min (re-equilibration time).

## **SPECTROPHOTOMETRIC ASSAYS FOR ENZYMATIC ACTIVITIES**

### **96-MICROWELL PLATE ENZYMATIC SCREENING ASSAYS**

#### **PYRUVATE KINASE – LACTATE DEHYDROGENASE (PK-LDH) CASCADE: LACTAM FORMATION MONITORING**

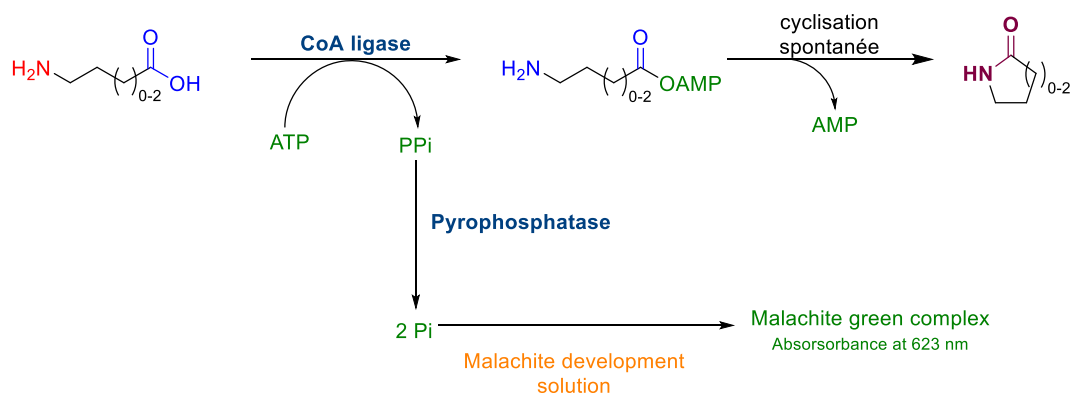
A 80- $\mu\text{l}$  reaction mixture containing 50 mM glycine-glycine buffer pH 7.5, 2.5 mM phosphoenol pyruvate (PEP) (pH adjusted to 7.5), 0.2 mM ATP, 1 mM  $\text{MgCl}_2$ , 1 mM  $\text{MnCl}_2$ , 0.4 mM NADH, 1 U pyruvate kinase (PK), 2.3 U lactate deshydrogenase (LDH), 1 U myokinase was prepared. To start the enzymatic cascade reaction, 20- $\mu\text{l}$  enzymatic reaction sample was added to the mix (100  $\mu\text{l}$  final volume). Absorbance at 340 nm (NADH consumption) was measured and monitored for 15 min. A background plate was set under the same conditions but with a mixture without substrate. An enzyme was considered biocatalytically active if it displayed a well higher slope compared to the background plate.



Scheme 1. PK-LDH cascade for activity screening monitoring.

### MALACHITE GREEN ASSAY

Malachite Green assay was performed following the protocol described in Vardakou's article (Vardakou, Salmon et al. 2014). A Malachite Green dye solution was prepared by adding 10 ml sulfuric acid and 0.074 g of malachite green oxalate salt in 50 ml of water. The solution was stirred and cooled down at room temperature and stored at 4°C. Prior to use, a fresh solution of Malachite Green development solution was prepared by adding 98.5  $\mu$ l of an aqueous solution of ammonium molybdate 7.5 % (v/v) and 1  $\mu$ l of TWEEN®20 11 % (v/v) into 395- $\mu$ l MG assay dye solution. 12  $\mu$ l of the MG development solution was added to 30- $\mu$ l enzymatic reaction sample and was let under incubation for 15 min at room temperature before monitoring the absorbance at 623 nm.



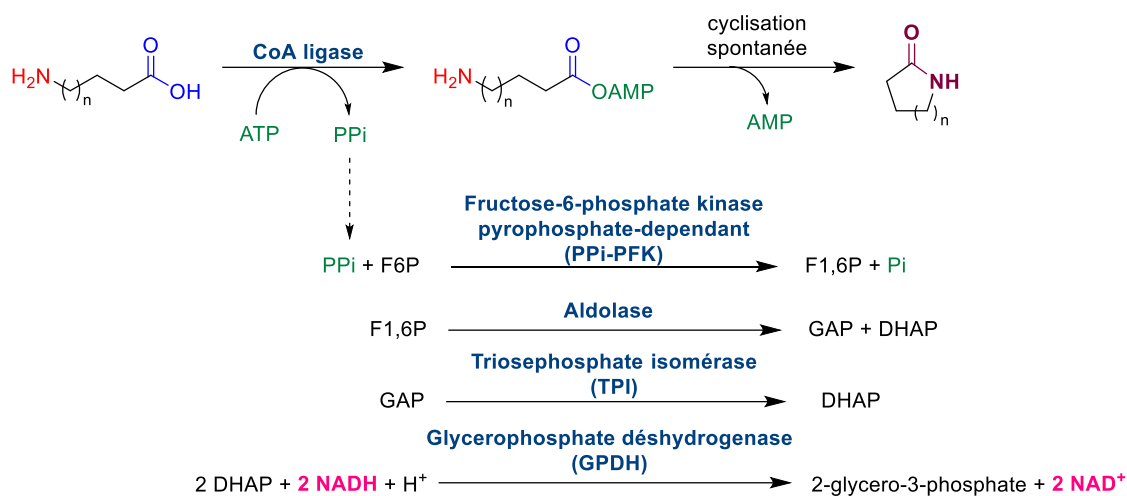
Scheme 2. Malachite Green Assay (MG assay) for activity screening monitoring.

### PYROPHOSPHATE REAGENT (SIGMA-ALDRICH®) FOR ENZYMATIC ACTIVITIES

Pyrophosphate reagent (Sigma-Aldrich®) was restituted and stored at -20 °C according to manufacturer's instructions. The kinetics were obtained using an endpoint method. In a high-precision quartz microcell, 33.3  $\mu$ l of Pyrophosphate reagent in 27.7  $\mu$ l H<sub>2</sub>O were incubated at 37 °C for 1 min before starting the reaction by the addition of 40  $\mu$ l biocatalytic reaction sample to the mixture (100  $\mu$ l final volume). The initial slope was measured at 340 nm (NADH consumption) and the specific activity



was determined according to the manufacturer's instructions (based on Beer-Lambert's law) after subtracting the slope obtained under the same conditions without substrate and/or enzyme.



Scheme 3. Pyrophosphate Reagent® for enzymatic activities.

## STANDARD SYNTHESIS

### SYNTHESIS OF N-METHYLHEXANAMIDE

To a solution of hexanoic acid (2 mmol, 250  $\mu\text{l}$ ) in dichloromethane (DCM, 10 ml), in round-bottom flask, under nitrogen at 0°C, were successively added N-methylamine hydrochloride (2.49 mmol, 1.25 eq., 168.7 mg), N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC.HCl, 2.24 mmol, 1.12 eq., 429.3 mg), 4-dimethylaminopyridine (DMAP, 0.43 mmol, 0.22 eq., 52.7 mg) and triethylamine (5 mmol, 2.5 eq., 693  $\mu\text{l}$ ). The reaction mixture was stirred at room temperature (RT) for 23 h then 20 ml of ammonium chloride saturated solution were added. The organic phase was extracted with 3 x 15 ml DCM, the combined organic phases were washed with 10 ml of brine and dried over magnesium sulfate. After filtration, the solvent were concentrated under reduced pressure to dryness. Purification of the product by flash chromatography on silica gel (gradient: ethyl acetate/ether petroleum 0/100 to 80/20) yielded the desired amide as a yellow oil (50 % yield; 128.4 mg).

**$^1\text{H NMR}$**  ( $\text{CDCl}_3$ , 298 K, 600 MHz):  $\delta$ [ppm]= 0.88 (t, 3H); 1.30 (t, 4H); 1.62 (m, 2H); 2.2 (t, 2H); 2.80 (d, 3H); 5.50 (s, 1H).

**$^{13}\text{C NMR}$**  ( $\text{CDCl}_3$ , 298 K 150 MHz):  $\delta$ [ppm]= 13.96 (s); 22.45 (s); 25.51 (s); 26.30 (s); 31.54 (s); 36.76 (s); 173.83 (s).

### SYNTHESIS OF N-METHYLHEPTANAMIDE

To a solution of heptanoic acid (2 mmol, 283  $\mu\text{l}$ ) in dichloromethane (DCM, 15 ml), in round-bottom flask, under nitrogen at 0°C, were successively added N-methylamine hydrochloride (2.53 mmol, 1.27 eq., 171.0 mg), N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC.HCl, 2.22 mmol, 1.11 eq., 424.9 mg), 4-dimethylaminopyridine (DMAP, 0.42 mmol, 0.21 eq., 51.3 mg) and

triethylamine (5 mmol, 2.5 eq., 693  $\mu$ l). The reaction mixture was stirred at room temperature for 23 h before being quenched with 20 ml of ammonium chloride saturated solution. Then, the organic phase was extracted with 3 x 15 ml DCM before washing the combined organic phases with 10 ml of brine. Finally, the product was dried over magnesium sulfate, filtrated and concentrated under reduced pressure to dryness. Purification of the product by flash chromatography on silica gel (gradient: ethyl acetate/ether petroleum 0/100 to 80/20) yielded the desired amide as a yellow oil (74.4 % yield; 213.1 mg).

**$^1\text{H}$  NMR** ( $\text{CDCl}_3$ , 298 K, 600 MHz):  $\delta$ [ppm]= 0.87 (t, 3H); 1.28 (m, 6H); 1.61 (m, 2H); 2.15 (t, 2H); 2.79 (d, 3H); 5.47 (s, 1H).

**$^{13}\text{C}$  NMR** ( $\text{CDCl}_3$ , 298 K, 150 MHz):  $\delta$ [ppm]= 14.07 (s); 22.55 (s); 25.80 (s); 26.30 (s); 29.06 (s); 31.60 (s); 36.80 (s); 173.87 (s).

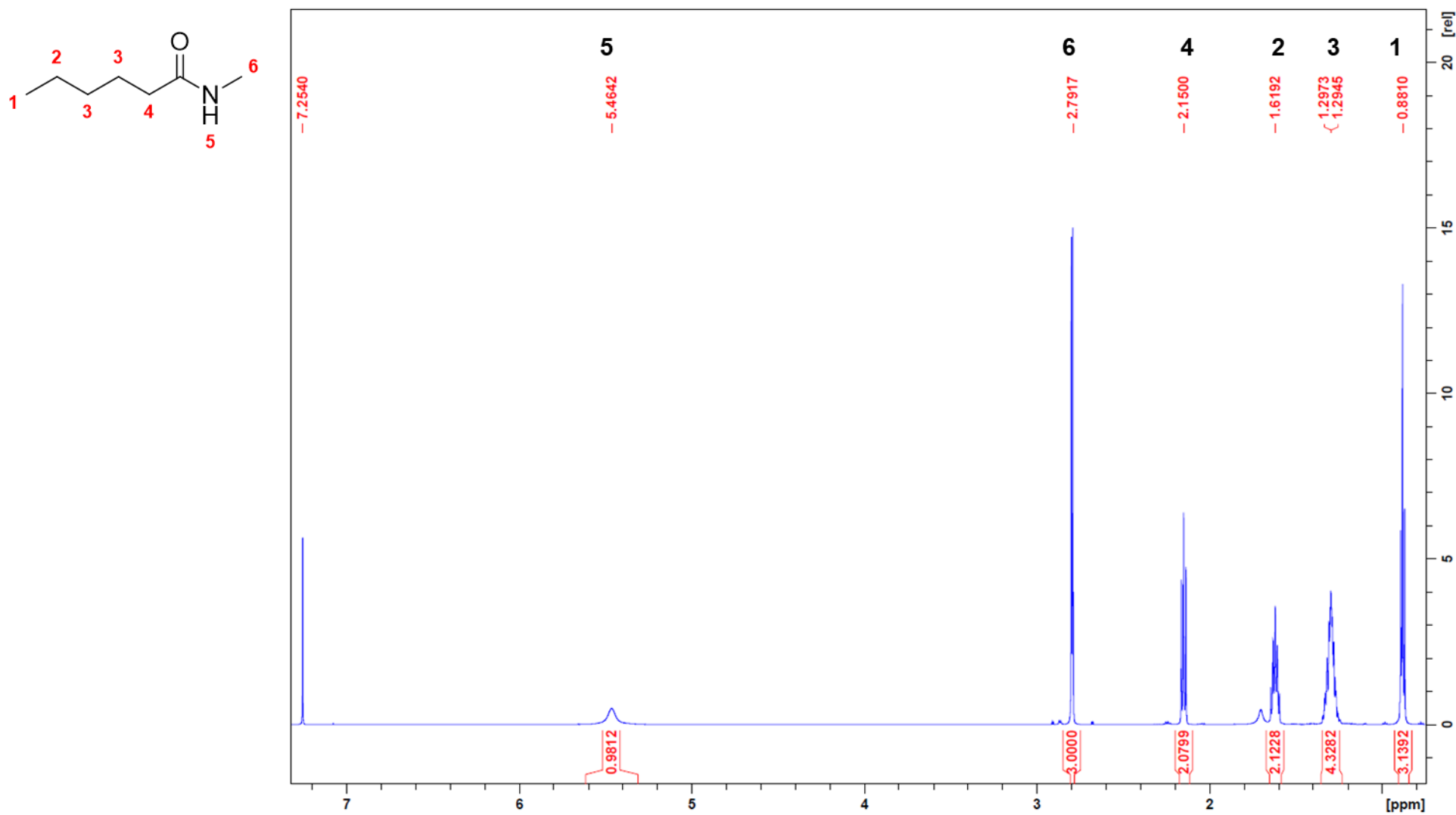
### SYNTHESIS OF N-METHYLOCTANAMIDE

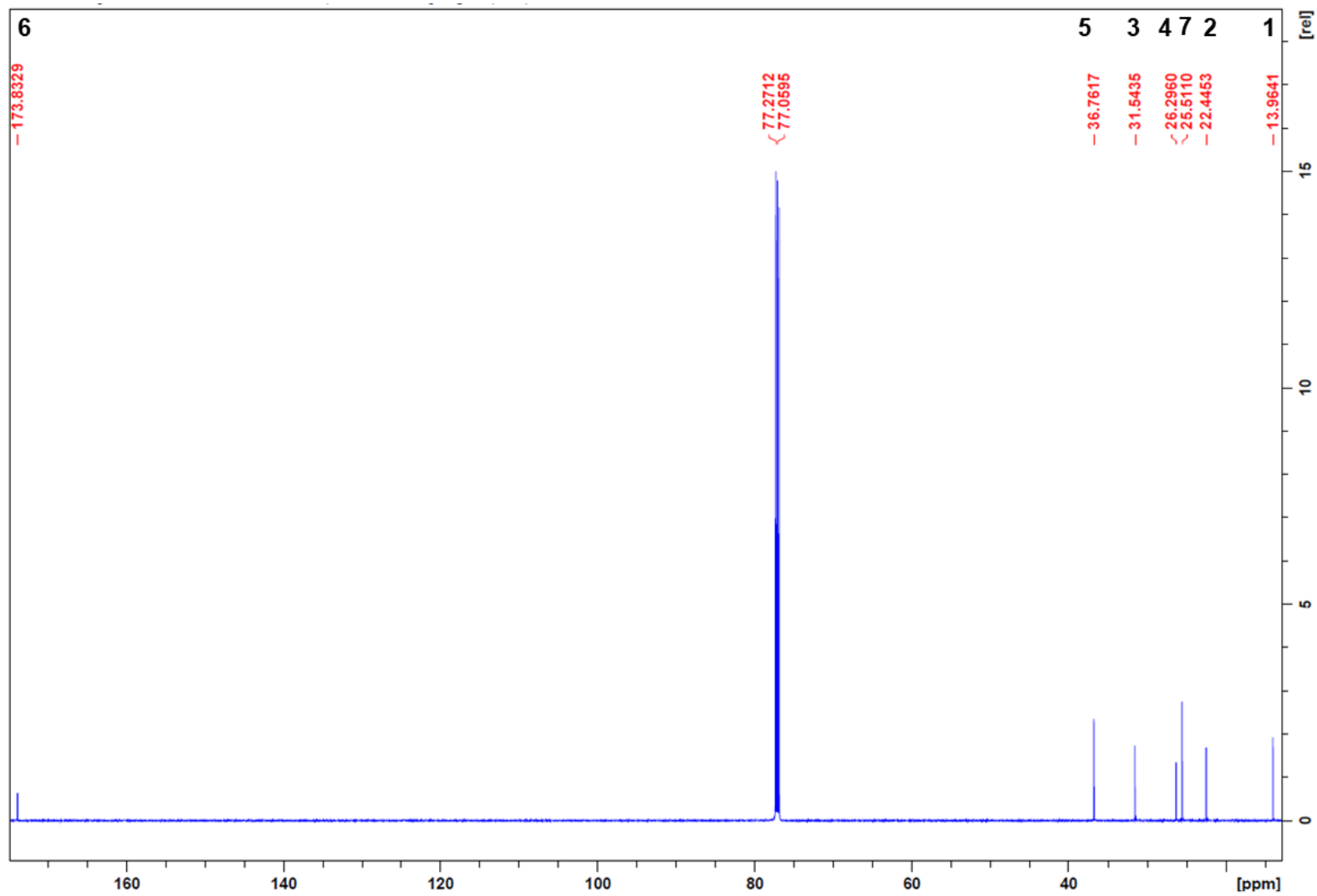
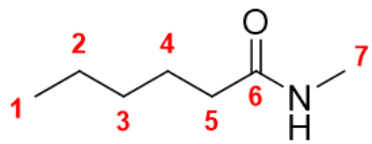
To a solution of heptanoic acid (2 mmol, 317  $\mu$ l) in dichloromethane (DCM, 15 ml), in round bottom flask, under nitrogen at 0  $^\circ\text{C}$ , were successively added N-methylamine hydrochloride (2.48 mmol, 1.24 eq., 167.6 mg), N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC.HCl, 2.23 mmol, 1.11 eq., 427.8 mg), 4-dimethylaminopyridine (DMAP, 0.44 mmol, 0.22 eq., 53.7 mg) and triethylamine (5 mmol, 2.5 eq., 693  $\mu$ l). The reaction mixture was stirred at room temperature for 23 h before being quenched with 20 ml of ammonium chloride saturated solution. Then, the organic phase was extracted with 3 x 15 ml DCM before washing the combined organic phases with 10 ml of brine. Finally, the product was dried over magnesium sulfate, filtrated and concentrated under reduced pressure to dryness. Purification of the product by flash chromatography on silica gel (gradient: ethyl acetate/ether petroleum 0/100 to 80/20) yielded the desired amide as a yellow oil (90 % yield; 283.6 mg).

**$^1\text{H}$  NMR** ( $\text{CDCl}_3$ , 298 K, 600 MHz):  $\delta$ [ppm]= 0.87 (t, 3H); 1.29 (m, 8H); 1.62 (m, 2H); 2.15 (t, 2H); 2.80 (d, 3H); 5.45 (s, 1H).

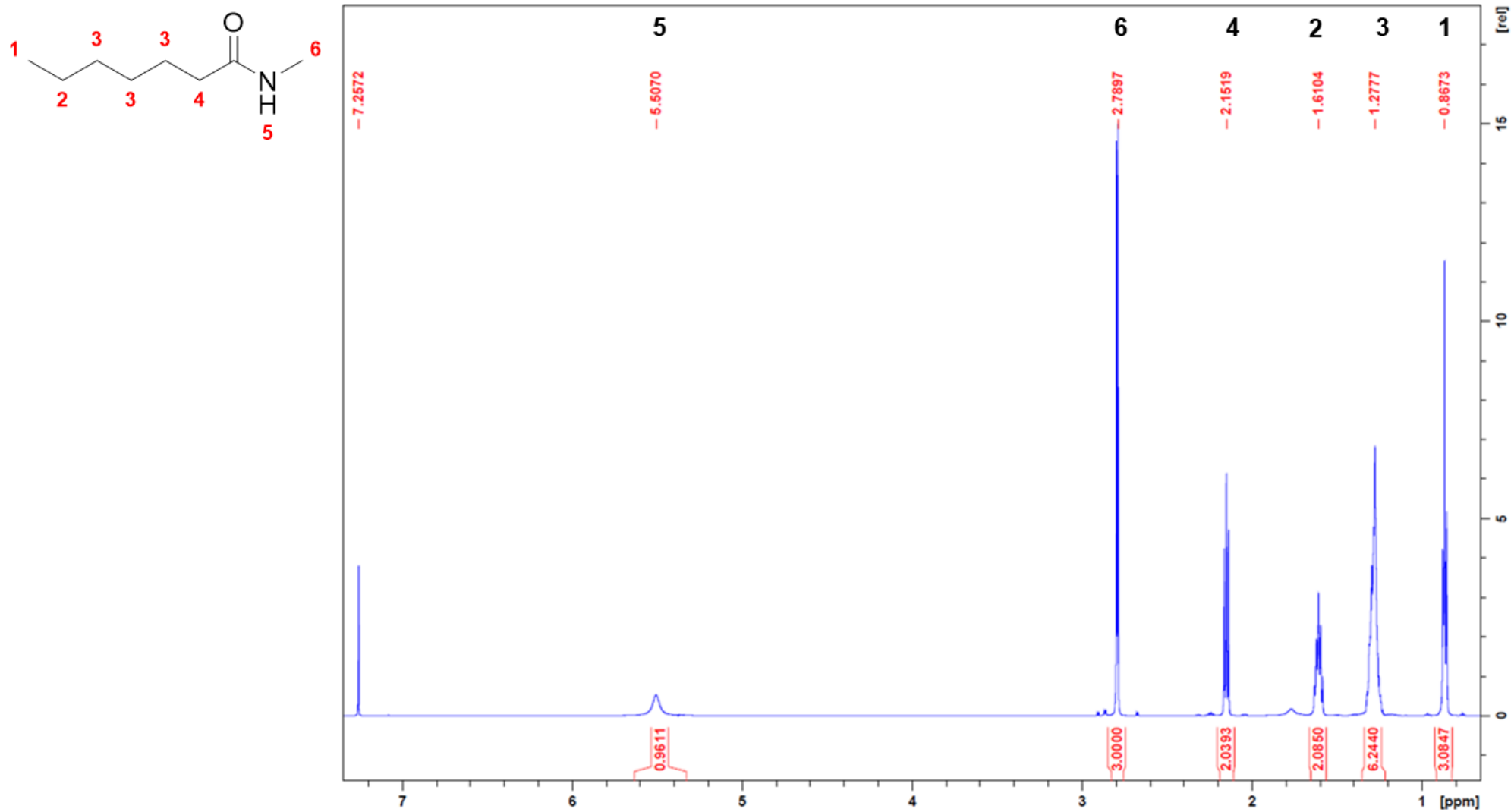
**$^{13}\text{C}$  NMR** ( $\text{CDCl}_3$ , 298 K, 150 MHz):  $\delta$ [ppm]= 14.10 (s); 22.65 (s); 25.85 (s); 26.30 (s); 29.06 (s); 29.35 (s); 31.74 (s); 36.80 (s); 173.86 (s).

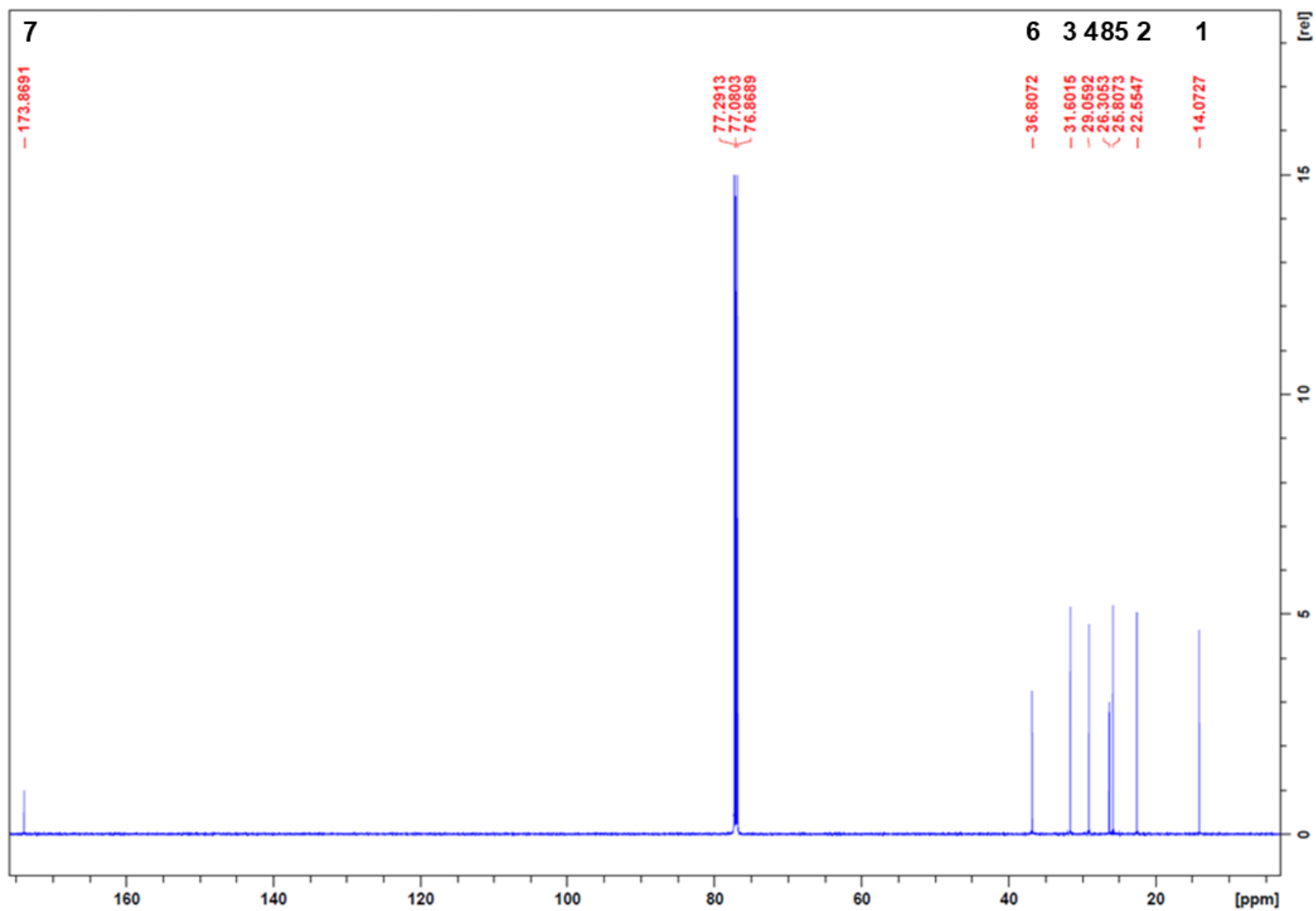
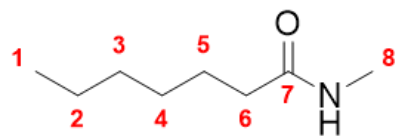
Appendix A. <sup>1</sup>H proton (CDCl<sub>3</sub>, 298 K, 600 MHz) and <sup>13</sup>C carbon (CDCl<sub>3</sub>, 298 K, 150 MHz) NMRs of N-methylhexanamide (peak assignments describe above).



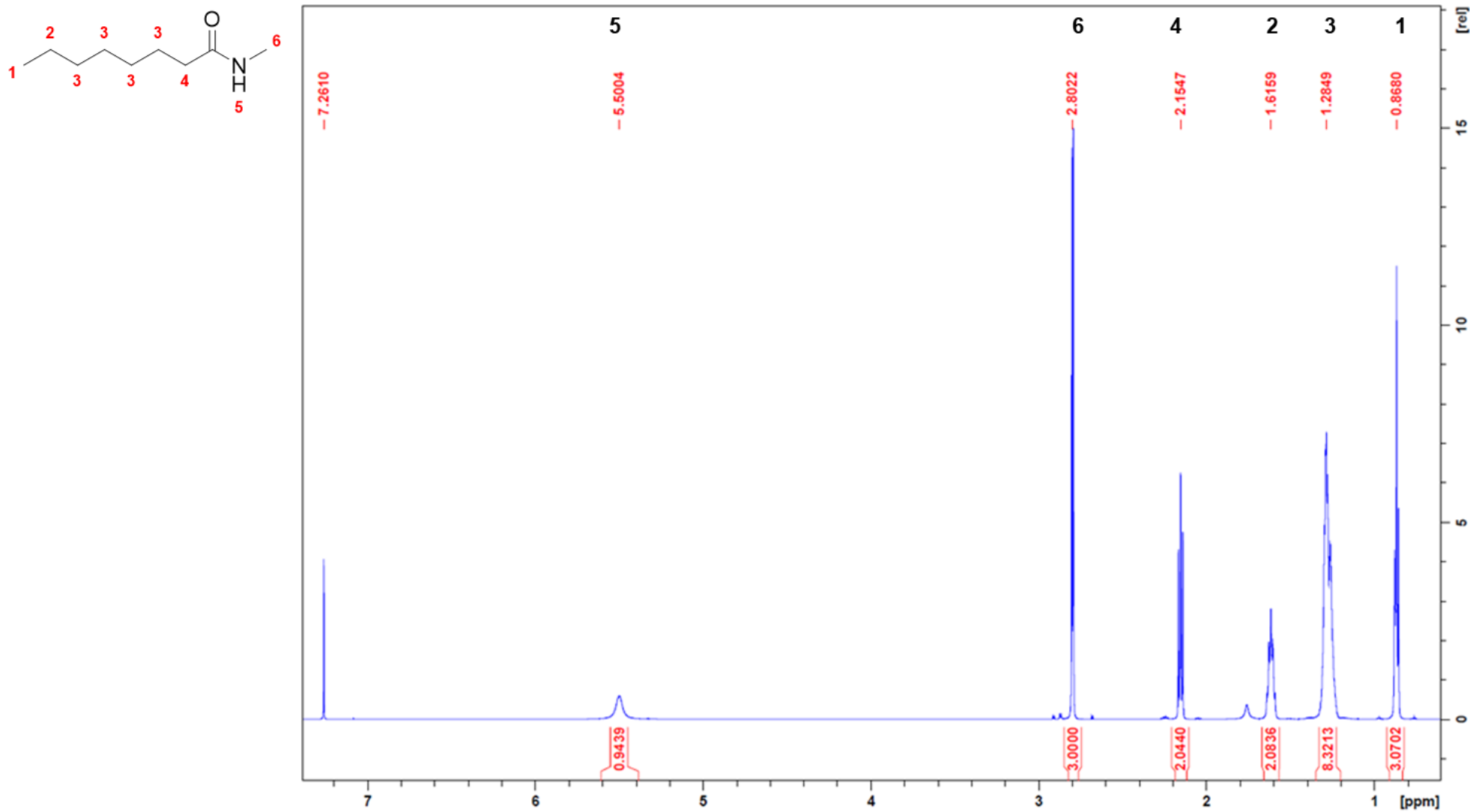


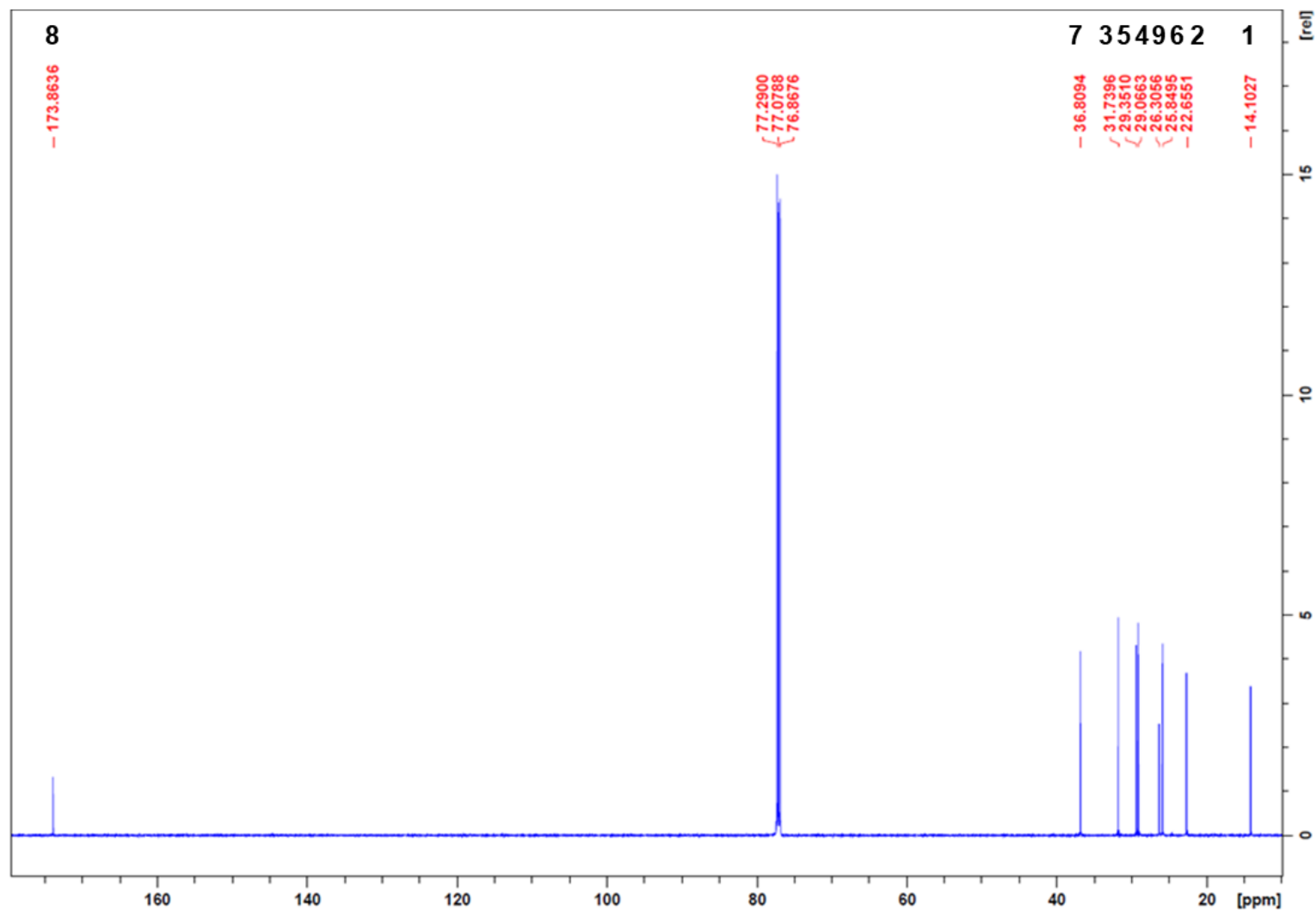
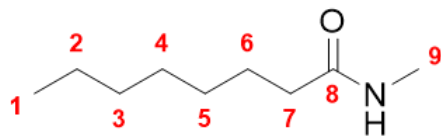
Appendix B.  $^1\text{H}$  proton ( $\text{CDCl}_3$ , 298 K, 600 MHz) and  $^{13}\text{C}$  carbon ( $\text{CDCl}_3$ , 298 K, 150 MHz) NMRs of N-methylheptanamide (peak assignments described above).





Appendix C. <sup>1</sup>H proton (CDCl<sub>3</sub>, 298 K, 600 MHz) and <sup>13</sup>C carbon (CDCl<sub>3</sub>, 298 K, 150 MHz) NMRs of N-methyloctanamide (peak assignments described above).







## REFERENCES BIBLIOGRAPHIQUES

Vardakou, M., M. Salmon, J. A. Faraldos and P. E. O'Maille (2014). "Comparative analysis and validation of the malachite green assay for the high throughput biochemical characterization of terpene synthases." MethodsX **1**: 187-196.

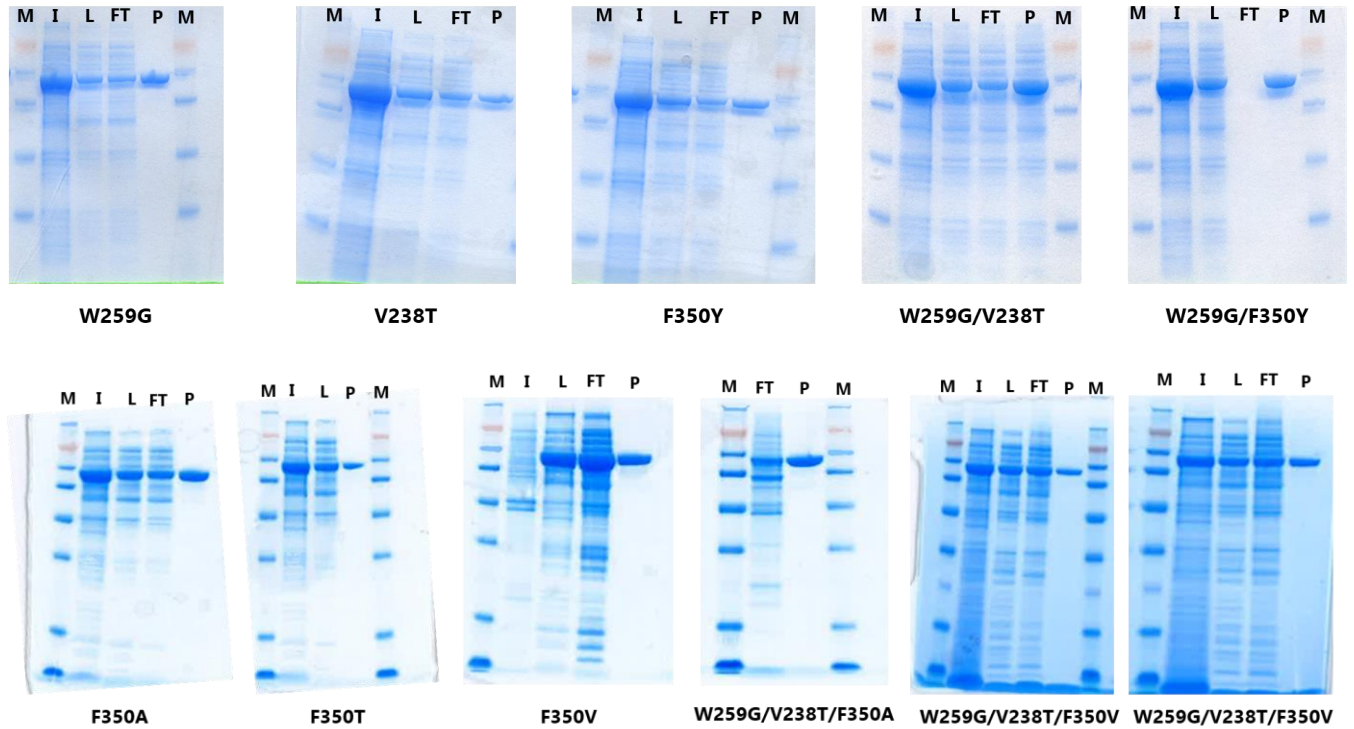
## ANNEXE

Annexe 1. Gel SDS-PAGE de la CoA ligase A4YDT1 WT et des mutants. M : Marqueur de poids moléculaire SeeBlue® Plus2 pre-strained standard (6 µl déposé), I : Induction (15 µg déposé), L : lysat (surpernageant, 15 µg déposé), FT : Flow Through NiNTA (15 µg déposé) and P : enzyme purifié (5 µg deposited). A. Purification à petite échelle. B. Purification à grande échelle. ....	173
Annexe 2. Chromatogrammes UHPLC-MS (APCI) des standards lactames et leur spectre de masse associé. ....	174
Annexe 3. Chromatogrammes UHPLC-MS (ESI) des N-méthylamides détectés de milieux réactionnels avec les enzymes A4YDT1 WT ou les mutants. ....	176
Annexe 4. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques de A4YDT1 WT. Les points correspondent à la moyenne des duplicats. ....	179
Annexe 5. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques du mutant A4YDT1 W259G. Les points correspondent à la moyenne des duplicats. ....	179
Annexe 6. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques du mutant A4YDT1 W259G/V238T. Les points correspondent à la moyenne des duplicats. ....	179
Annexe 7. Tableaux des paramètres cinétiques d'A4YDT1 et des mutants. ....	180
Annexe 8. Rendements analytiques de la formation de lactames à partir des acides ω-aminés avec A4YDT1 et ses mutants. ....	183
Annexe 9. Comparaison des aires relatives pour divers substrats par rapport au meilleur mutant pour chaque substrat, par détection du N-méthylamide correspondant par UHPLC-MS (ESI). ....	183
Annexe 10. Droite de calibration des standards lactames commerciaux utilisée pour calculer les rendements analytiques avec A4YDT1 et les mutants. ....	184
Annexe 11. Droites de calibration des standards amides synthétisés au laboratoire. ....	185
Annexe 12. Schéma du clonage par ligation indépendante (LIC). ....	186
Annexe 13. Gels SDS-PAGE de la production de la première collection de CoA ligase CoA Lig1. ....	187
Annexe 14. Concentration (en mg/ml) des CoA ligases purifiées de la collection CoA Lig1. Concentrations déterminées selon la méthode Bradford. ....	188
Annexe 15. Résultats du criblage de la collection CoA Lig1 suite à l'analyse avec la cascade PK-LDH. Légende : vert si la CoA ligase est active sur ce pool et rouge si la CoA ligase est inactive ce pool. Une CoA ligase est considérée comme active si son activité est supérieure ou égale à 2% d'activité par rapport au seuil de détection. ....	191
Annexe 16. Résultats de la vérification des CoA ligases sélectionnées par analyse UHPLC. Légende : vert si l'activité a été vérifiée et rouge si l'activité n'a pas été vérifiée. ....	192
Annexe 17. Chromatogrammes des lactames attendus par cyclisation des acides ω-aminés et leur spectre de masse associé. ....	193
Annexe 18. Rendements analytiques de la synthèse de lactames du pool 1 avec les CoA ligases sélectionnées suite à l'analyse avec la cascade PK-LDH. Conditions réactionnelles : 0,5 mM de chaque acide ω-aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl <sub>2</sub> , 1,5 mM MgCl <sub>2</sub> , 50 mM tampon MOPS pH 8,5, 10 µL d'enzyme purifiée. 60°C, sans agitation, 5h. Analyse par UHPLC-MS (APCI). Rendements déduits d'une gamme de calibration avec les standards commerciaux. ....	195
Annexe 19. Détection des aires correspondant aux lactames <b>4b</b> et <b>5b</b> attendus pour les substrats du pool 2 avec les CoA ligases sélectionnées suite à la cascade PK-LDH. Conditions réactionnelles : 0,5 mM de chaque acide ω-aminé dans le pool, 2 mM ATP, 2 mM MnCl <sub>2</sub> , 2 mM MgCl <sub>2</sub> , 50 mM tampon MOPS pH 8,5, 10 µL d'enzyme purifiée. 60°C, sans agitation, 5h. ....	195

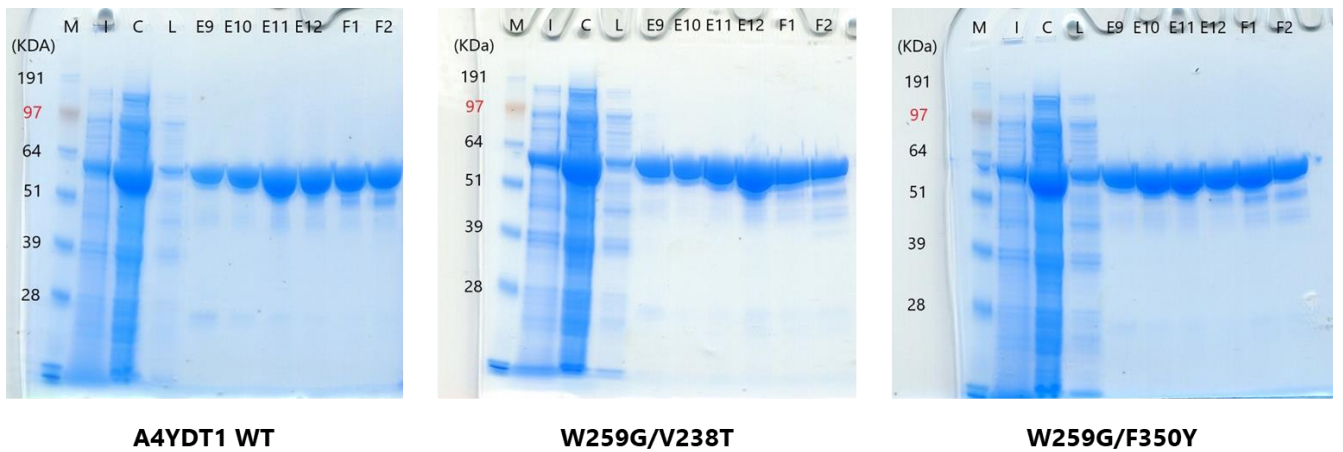
Annexe 20. Rendements analytiques pour la synthèse du 2-aminocaprolactame ( <b>6b</b> ) avec les CoA ligases sélectionnées pour le pool 2. Rendements déterminés à partir d'une gamme de calibration avec le standard commercial après dérivatisation au BzCl.....	196
Annexe 21. Détection des aires correspondant aux lactames <b>9b</b> et <b>10b</b> .....	196
Annexe 22. Rendements analytiques pour la synthèse de l'acide pyroglutamique ( <b>8b</b> ) avec les CoA ligases sélectionnées pour le pool 3. Rendements déterminés à partir d'une gamme de calibration avec le standard commercial.....	197
Annexe 23. ( <b>A</b> ) Détection des aires correspondant à la synthèse de N-méthylbutyrolactame ( <b>16b</b> ) et de N-éthylbutyrolactame ( <b>17b</b> ) avec les CoA ligases sélectionnées pour le pool 6. ( <b>B</b> ) Même histogramme en enlevant la forte réponse de la CoA ligase D1A9C5.....	198
Annexe 24. Droites de calibration des standards lactames pour le criblage de la collection CoA Lig1.....	199
Annexe 25. Liste des CoA ligases de la collection CoA Lig2 en fonction de la CoA ligase de référence pour laquelle elles ont été rapatriées.....	201
Annexe 26. Comparaison des résidus constituant le site actif des 20 CoA ligases sélectionnées pour former la collection réduite de CoA Lig2. Légende : résidus à chaîne latérale courte et hydrophobe en rouge (sauf Y), résidus acides en bleu, résidus basique en magenta et résidus à chaîne latérale polaire et résidu glycine en vert.....	205
Annexe 27. Plan de plaque des gènes des 20 CoA ligases de la collection réduite.....	206
Annexe 28. Gel d'agarose de la vérification de l'amplification PCR des gènes des 20 CoA ligases sélectionnées. Légende : de gauche à droite, les puits correspondent à A1 à H1 puis de A2 à H2 et A3 à D3.....	207
Annexe 29. Gels SDS-PAGE et concentration des CoA ligases de la production de la collection réduite de la deuxième collection CoA Lig2.....	208

Annexe 1. Gel SDS-PAGE de la CoA ligase A4YDT1 WT et des mutants. M : Marqueur de poids moléculaire SeeBlue® Plus2 pre-stained standard (6 µl déposé), I : Induction (15 µg déposé), L : lysat (surpernageant, 15 µg déposé), FT : Flow Through NiNTA (15 µg déposé) and P : enzyme purifié (5 µg deposited). A. Purification à petite échelle. B. Purification à grande échelle.

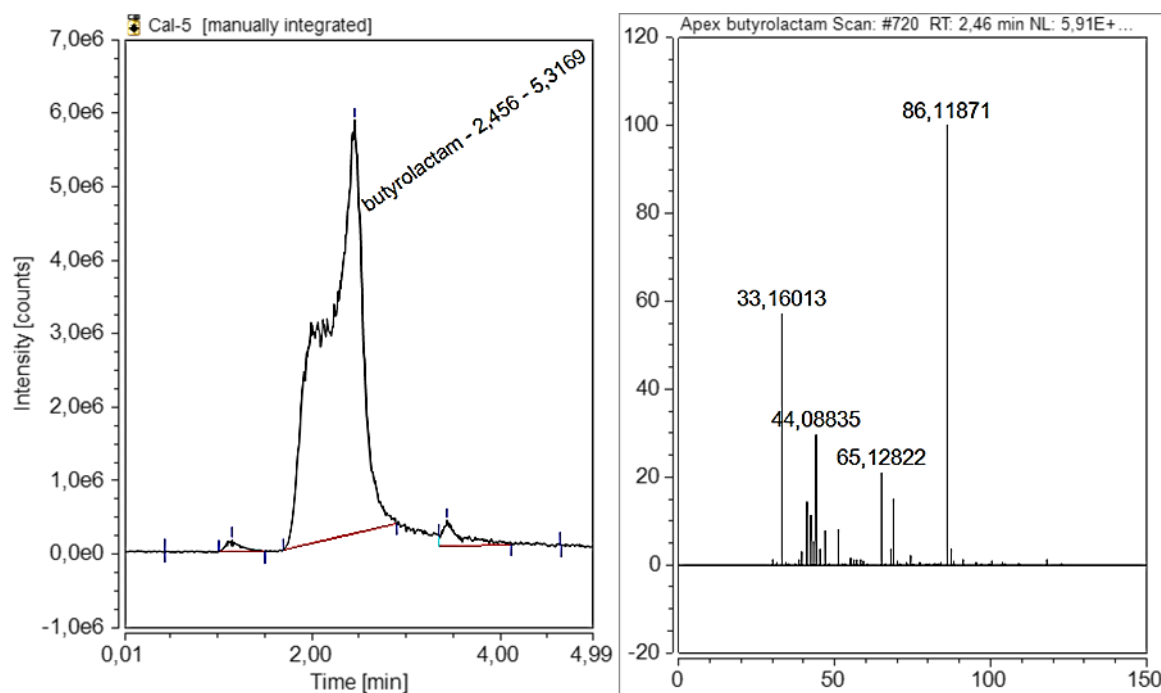
**A**



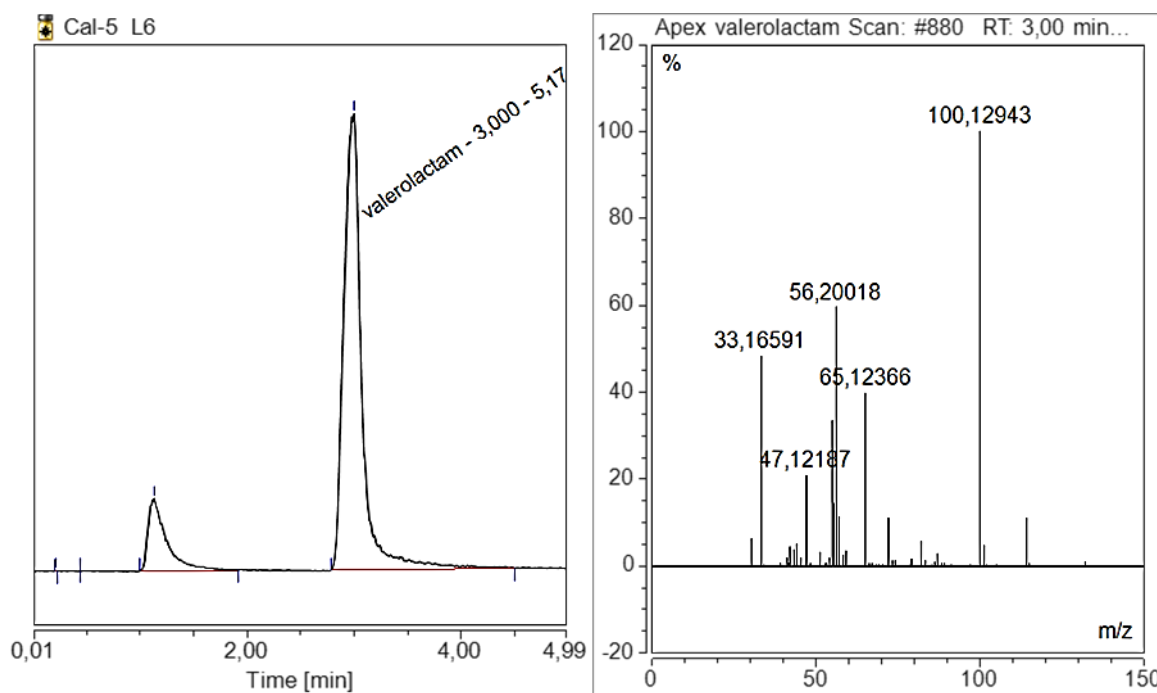
**B**



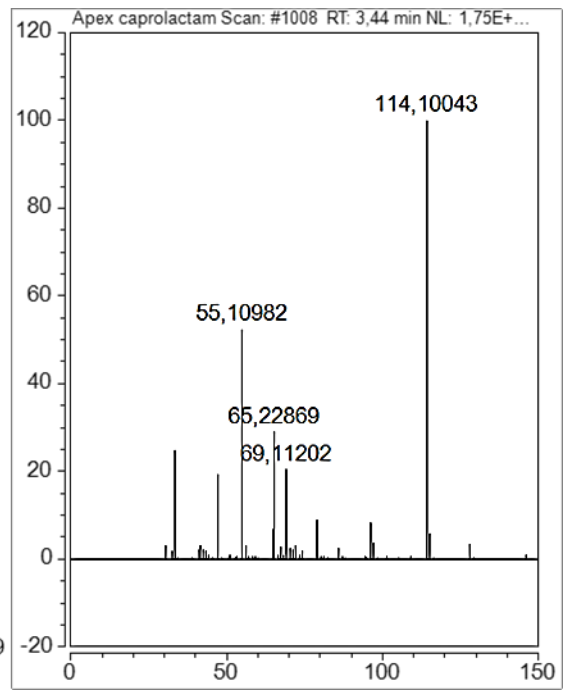
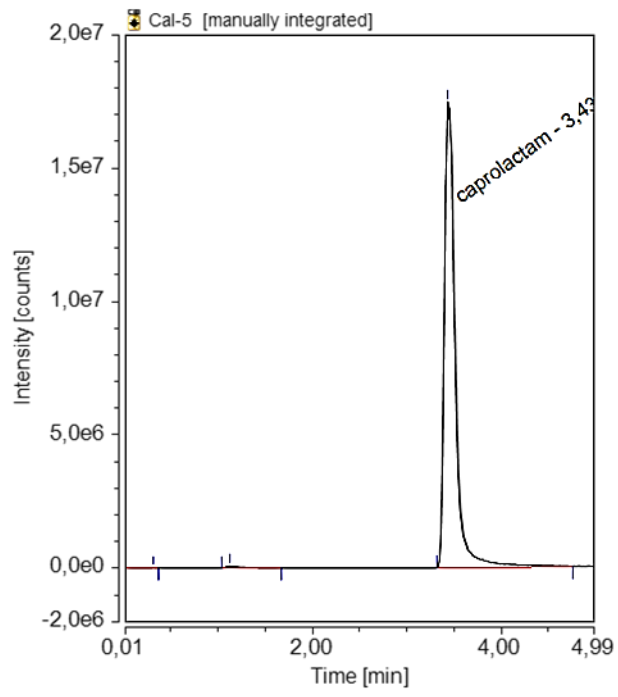
**(A)**  $\gamma$ -butyrolactame, extraction du  $m/z = 86$  du signal TIC,  $t_r=2,3$  min.



**(B)**  $\delta$ -valérolactame, extraction du  $m/z = 100$  du signal TIC,  $t_r= 3,0$  min.

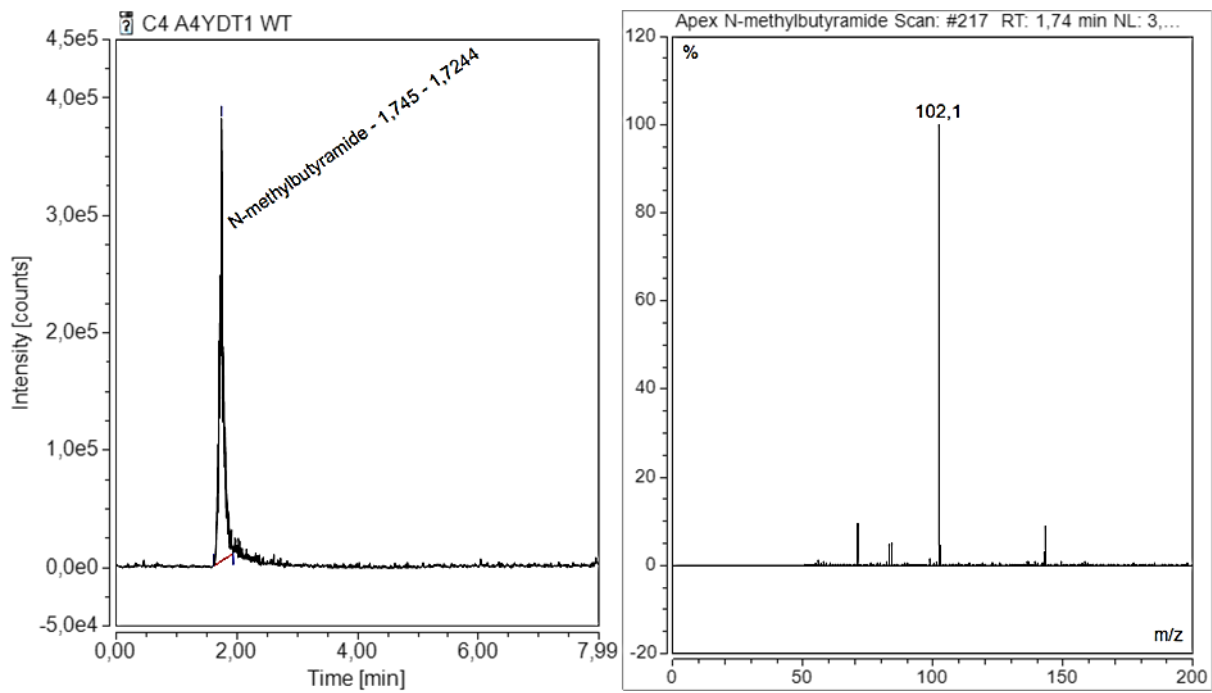


**(C)**  $\epsilon$ -caprolactame, extraction du  $m/z = 114$  du signal TIC,  $t_r=3,4$  min.

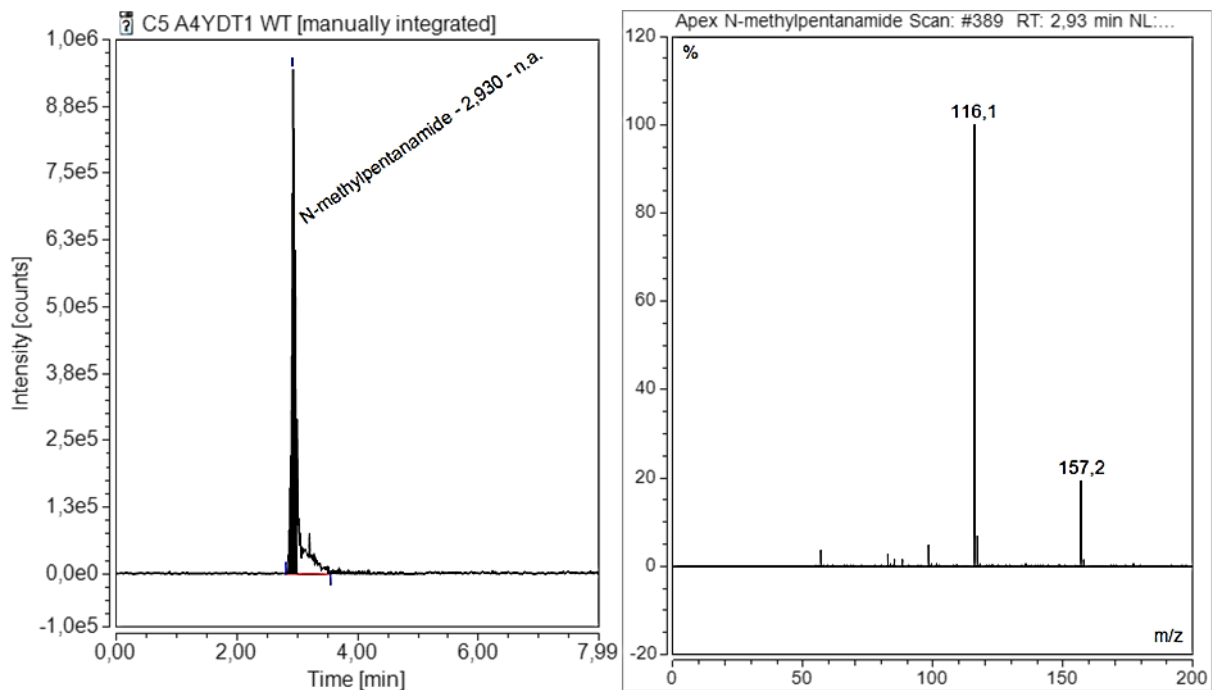


Annexe 3. Chromatogrammes UHPLC-MS (ESI) des N-méthylamides détectés de milieu réactionnels avec les enzymes A4YDT1 WT ou les mutants.

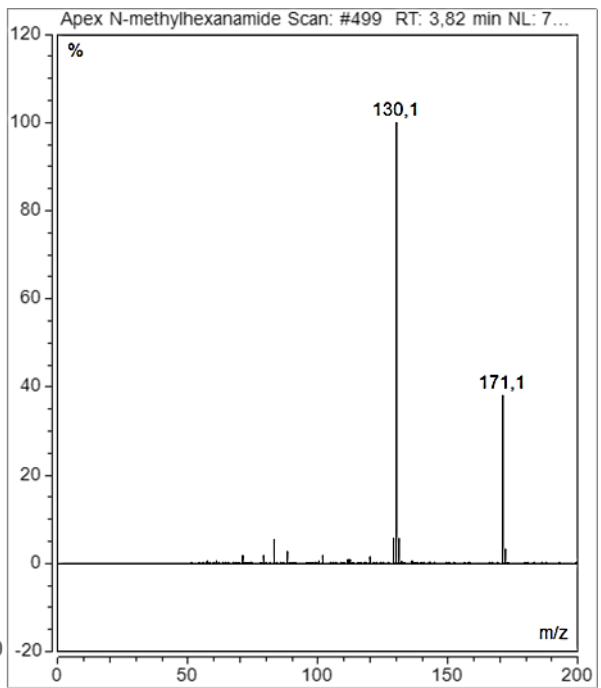
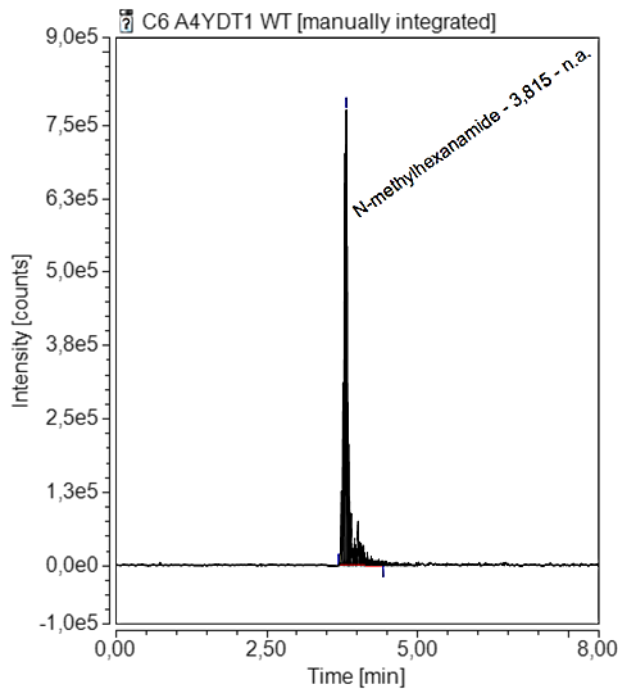
**(A)** N-méthylbutyramide ( $m/z = 102$ ,  $t_r = 1,74$  min)



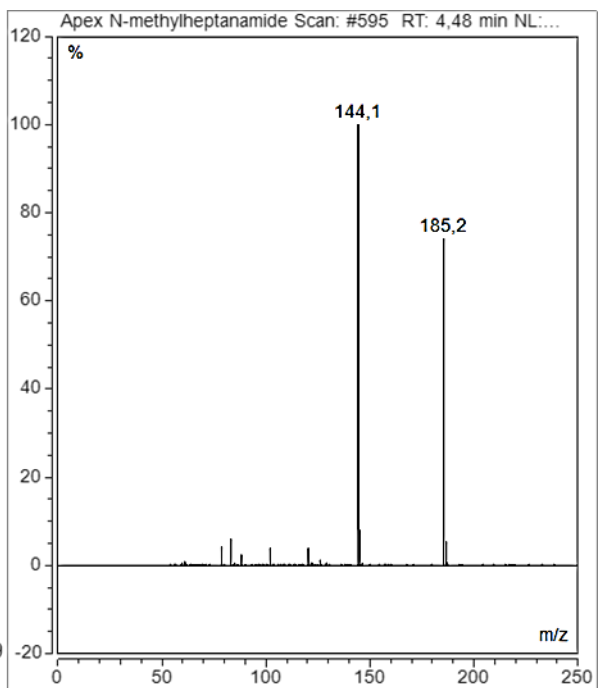
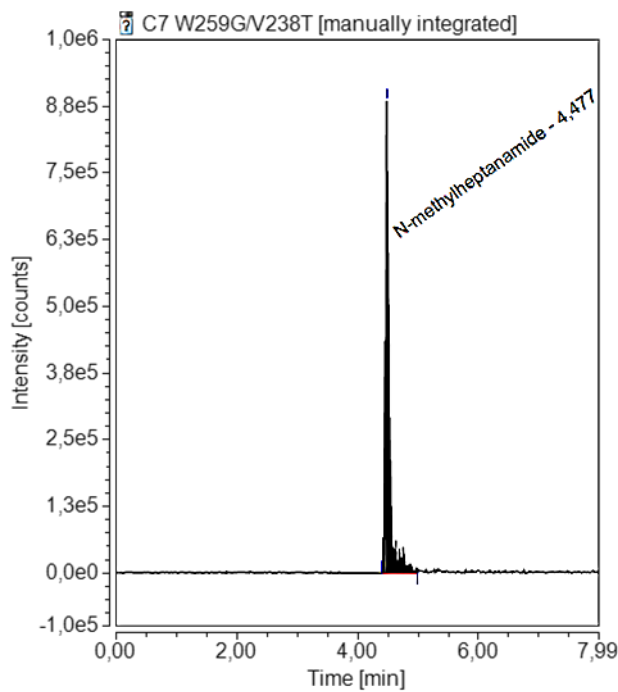
**(B)** N-méthylpentamide ( $m/z = 116$ ,  $t_r = 2,93$  min)



**(C)** N-méthylhexanamide ( $m/z = 130$ ,  $t_r = 3,82$ )

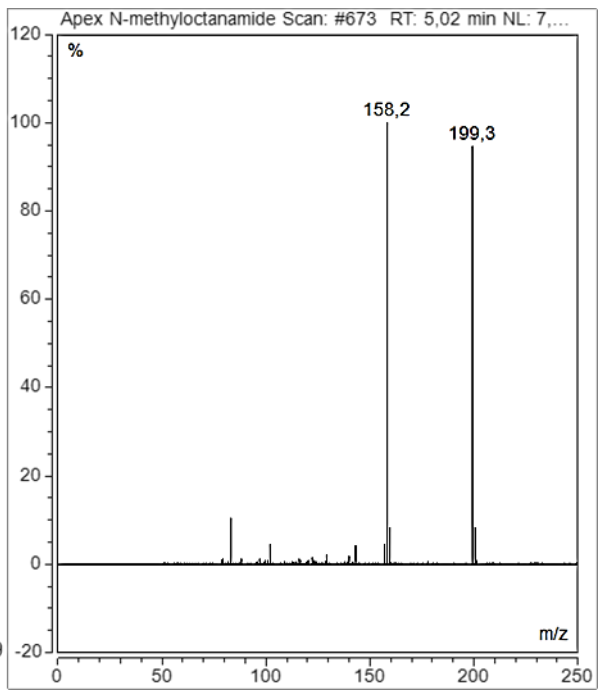
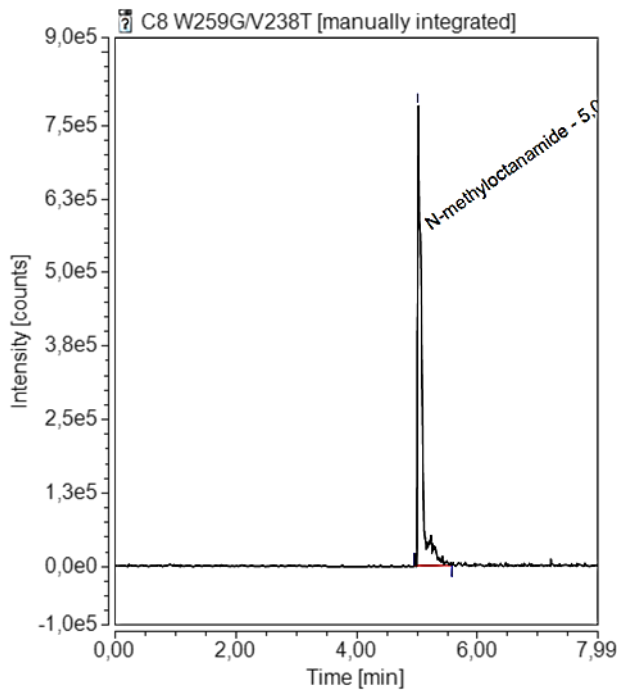


**(D)** N-méthylheptanamide ( $m/z = 144$ ,  $t_r = 4,48$  min)

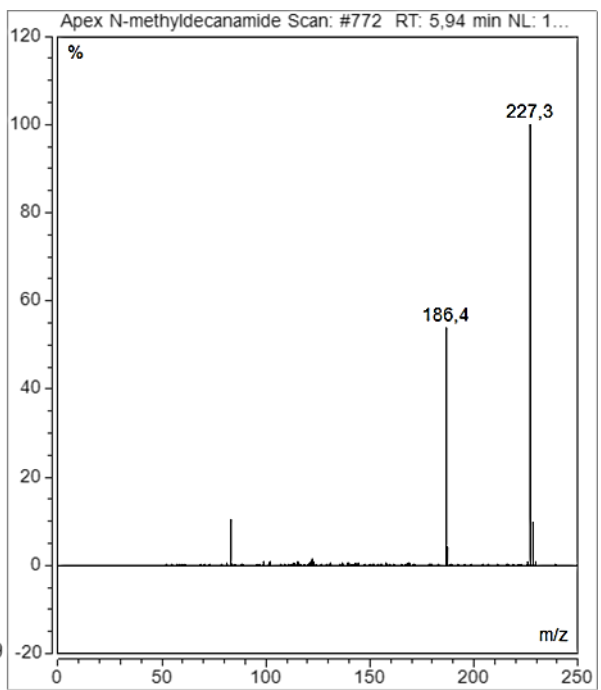
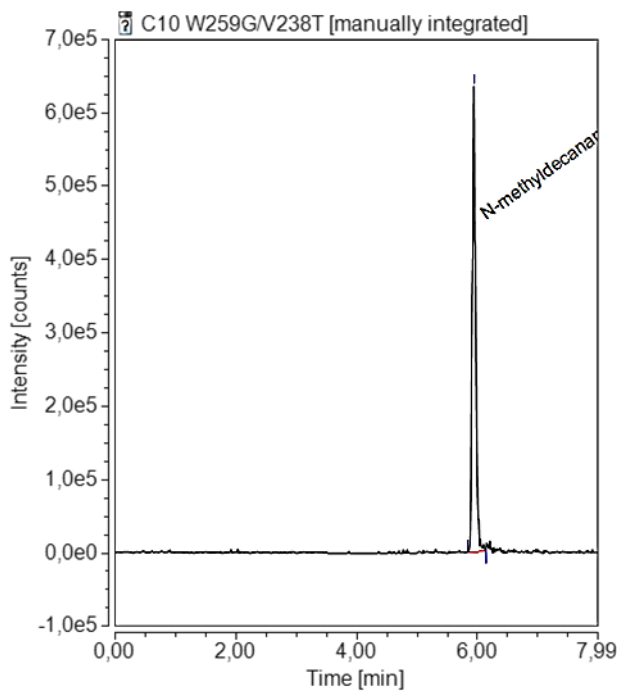


**(E)** N-méthyl-octanamide ( $m/z = 158$ ,  $t_r = 5,02$  min)

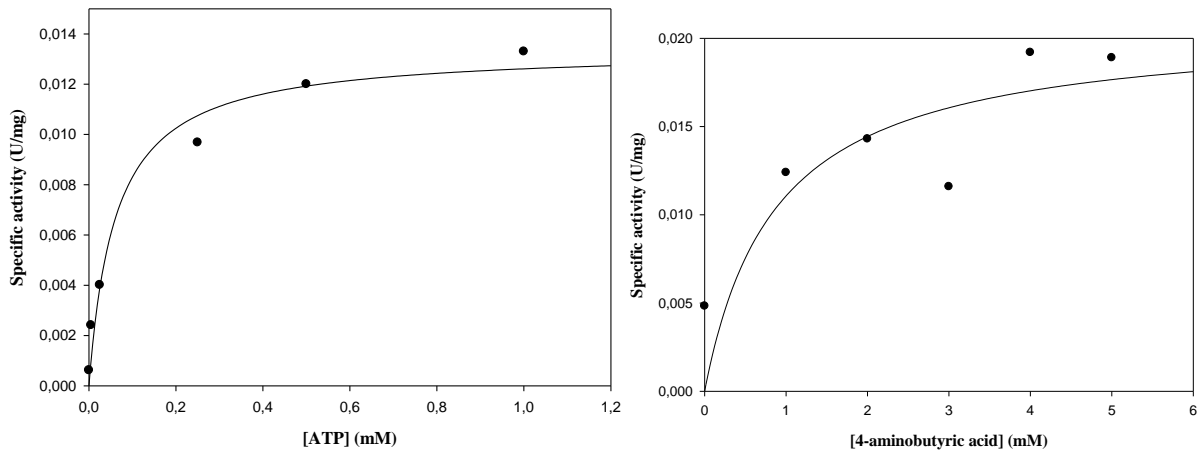




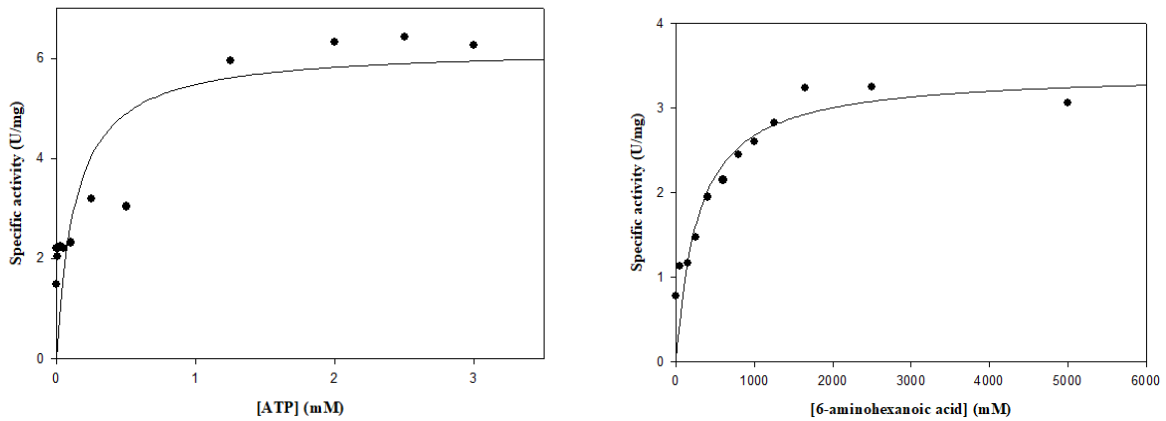
**(F)** N-méthyldecaneamide ( $m/z = 186$ ,  $t_r = 5,94$  min)



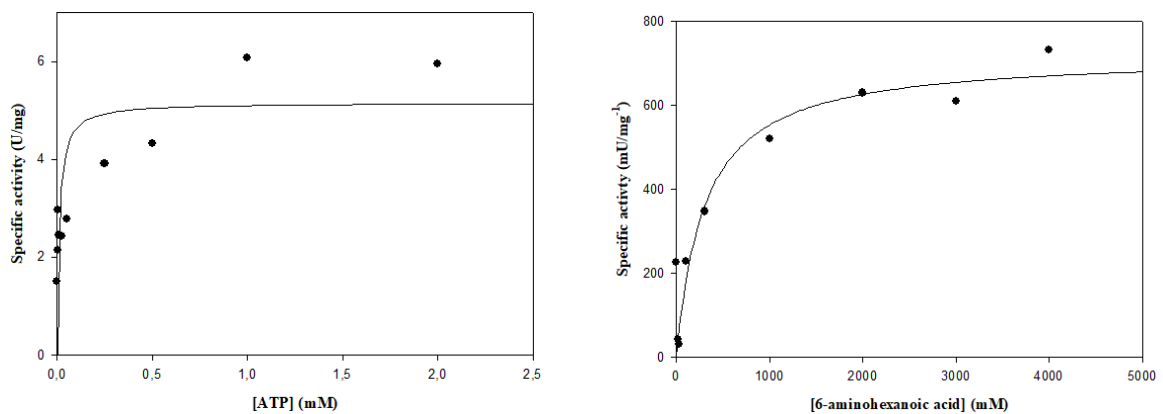
Annexe 4. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques de A4YDT1 WT. Les points correspondent à la moyenne des duplicats.



Annexe 5. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques du mutant A4YDT1 W259G. Les points correspondent à la moyenne des duplicats.



Annexe 6. Courbes de Michaelis-Menten pour la détermination des paramètres cinétiques du mutant A4YDT1 W259G/V238T. Les points correspondent à la moyenne des duplicats.



Annexe 7. Tableaux des paramètres cinétiques d'A4YDT1 et des mutants.

Expérience	Tampon	pH tampon	Substrat	[S] (mM)	[ATP] (mM)	Enzyme	MW (Da)	[E]initiale (mg/ml)	[E]finale (mg/ml)	V enzyme (μL)	Qté enzyme (mg)	[enzyme] (μmol/L)	V échantillon (μL)	V <sub>tot</sub> cuve (μL)	εNADH (L/mmol/cm)	Stoechio NADH	Longueur cuve (cm)	Pente (Abs/min)	VA pente (Abs/min)	[PPi] <sub>libéré</sub> /min (μmol/ml/min)	Activité spé (μmol/min/mg enz)	Activité spé (mU/mg enz)	Moy. Act spé (mU/mg enz)	SD Act spé
1	Tris	7,5	4-aminobutyric acid	10	0	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,002342	0,002342	0,000470659	0,000941318	0,941318328		
2	Tris	7,5	4-aminobutyric acid	10	0	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	0,000738	0,000738	0,000148312	0,000296624	0,296623794	0,618971061	0,45586788
3	Tris	7,5	4-aminobutyric acid	10	0,0025	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,007444	0,007444	0,001495981	0,002991961	2,991961415		
4	Tris	7,5	4-aminobutyric acid	10	0,0025	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,002191	0,002191	0,000440314	0,000880627	0,88062701	1,936294212	1,49293888
5	Tris	7,5	4-aminobutyric acid	10	0,005	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,007118	0,007118	0,001430466	0,002860932	2,860932476		
6	Tris	7,5	4-aminobutyric acid	10	0,005	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,004863	0,004863	0,000977291	0,001954582	1,954581994	2,407757235	0,64088657
7	Tris	7,5	4-aminobutyric acid	10	0,01	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,007778	0,007778	0,001563103	0,003126206	3,126205788		
8	Tris	7,5	4-aminobutyric acid	10	0,01	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,007579	0,007579	0,001523111	0,003046222	3,046221865	3,086213826	0,05655717
9	Tris	7,5	4-aminobutyric acid	10	0,025	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,009429	0,009429	0,001894895	0,003789791	3,789790997		
10	Tris	7,5	4-aminobutyric acid	10	0,025	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,010512	0,010512	0,00211254	0,00422508	4,225080386	4,007435691	0,30779608
11	Tris	7,5	4-aminobutyric acid	10	0,05	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,022223	0,022223	0,004466037	0,008932074	8,932073955		
12	Tris	7,5	4-aminobutyric acid	10	0,05	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,011951	0,011951	0,002401728	0,004803457	4,803456592	6,867765273	2,91937333
13	Tris	7,5	4-aminobutyric acid	10	0,1	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,019078	0,019078	0,003834003	0,007668006	7,668006431		
14	Tris	7,5	4-aminobutyric acid	10	0,1	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,018079	0,018079	0,00363324	0,007266479	7,2664791	7,467242765	0,2839227
15	Tris	7,5	4-aminobutyric acid	10	0,25	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,025441	0,025441	0,005112741	0,010225482	10,22548232		
16	Tris	7,5	4-aminobutyric acid	10	0,25	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,022709	0,022709	0,004563706	0,009127412	9,127411576	9,676446945	0,77645327
17	Tris	7,5	4-aminobutyric acid	10	0,5	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,034368	0,034368	0,006906752	0,013813505	13,81350482		
18	Tris	7,5	4-aminobutyric acid	10	0,5	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,025134	0,025134	0,005051045	0,01010209	10,10209003	11,95779743	2,62436657
19	Tris	7,5	4-aminobutyric acid	10	1	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,033188	0,033188	0,006669614	0,013339228	13,3392283		
20	Tris	7,5	4-aminobutyric acid	10	1	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,032901	0,032901	0,006611937	0,013223875	13,2238746	13,28155145	0,08156738
21	Tris	7,5	4-aminobutyric acid	0	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,012011	0,012011	0,002413786	0,004827572	4,827572347		
22	Tris	7,5	4-aminobutyric acid	0	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,012011	0,012011	0,002413786	0,004827572	4,827572347	4,827572347	0
23	Tris	7,5	4-aminobutyric acid	1	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,013765	0,013765	0,002766278	0,005532556	5,532556247		
24	Tris	7,5	4-aminobutyric acid	1	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,048111	0,048111	0,009668609	0,019337219	19,33721865	12,43488746	9,76137038
25	Tris	7,5	4-aminobutyric acid	2	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,019837	0,019837	0,003986535	0,007973071	7,97307074		
26	Tris	7,5	4-aminobutyric acid	2	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,051519	0,051519	0,010353497	0,020706994	20,70699357	14,34003215	9,00424318
27	Tris	7,5	4-aminobutyric acid	3	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,011615	0,011615	0,002334204	0,004668408	4,66840836		
28	Tris	7,5	4-aminobutyric acid	3	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,04586	0,04586	0,009216238	0,018432476	18,43247588	11,55044212	9,73266548
29	Tris	7,5	4-aminobutyric acid	4	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,022119	0,022119	0,00445137	0,008890273	8,890273312		
30	Tris	7,5	4-aminobutyric acid	4	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,073508	0,073508	0,014772508	0,029545016	29,54501608	19,21764469	14,6051087
31	Tris	7,5	4-aminobutyric acid	5	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,019304	0,019304	0,003879421	0,007758842	7,758842444		
32	Tris	7,5	4-aminobutyric acid	5	10	A4YDT1 WT	67307	5,68	0,5	3,52112676	0,00176056	7,42864784	40	100	6,22	2	1	-0,074621	0,074621	0,014996182	0,029992363	29,99236334	18,87560289	15,7214734

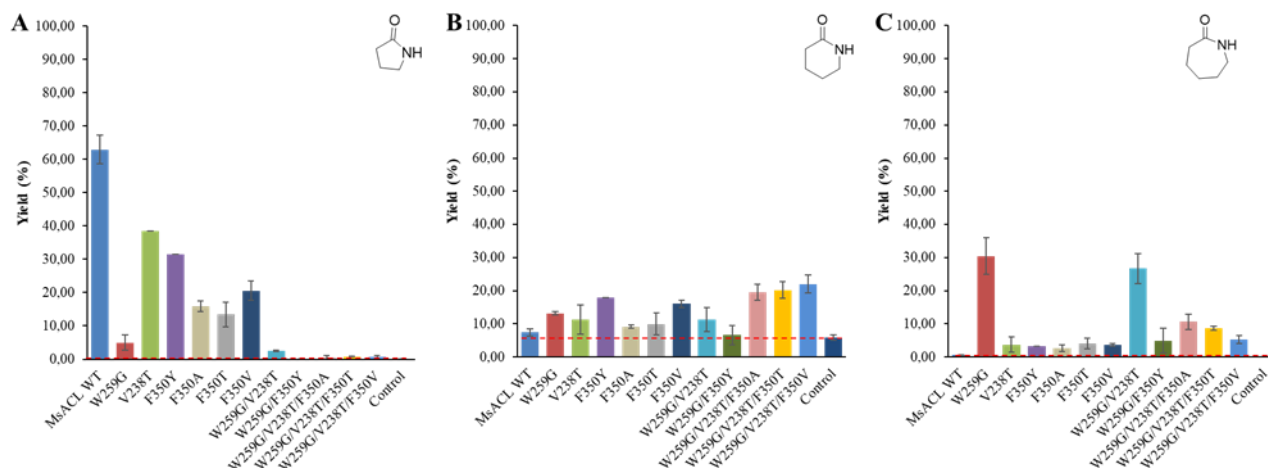
Expérience	Tampon	pH tampon	Substrat	[S] (mM)	[ATP] (mM)	Enzyme	MW (Da)	[E]initiale (mg/ml)	[E]finale (mg/ml)	V enzyme (μl)	Qté enzyme (mg)	[enzyme] (μmol/l)	V échantillon (μl)	V <sub>inc</sub> cue (μl)	eNADH (U/mmol/cm)	Stoechio NADH	Longueur cuve (cm)	Pente (Abs/min)	VA pente (Abs/min)	[PPI] <sub>inc</sub> /min (μmol/ml/min)	Activité spé (μmol/min/mg enz)	Activité spé (mU/mg enz)	Moy. Act spé (mU/mg enz)	SD Act spé
1	Tris		7,5 6-AHA	3000	0	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,035	0,035	0,007033762	0,937834941	937,8349411		
2	Tris		7,5 6-AHA	3000	0	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,035	0,035	0,007033762	0,937834941	937,8349411	937,8349411	0
3	Tris		7,5 6-AHA	3000	0,0025	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,049048	0,049048	0,009856913	1,314255091	1314,255091		
4	Tris		7,5 6-AHA	3000	0,0025	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,054826	0,054826	0,011018087	1,469078242	1469,078242	1391,666667	109,4765
5	Tris		7,5 6-AHA	3000	0,005	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,059665	0,059665	0,011990555	1,598740622	1598,740622		
6	Tris		7,5 6-AHA	3000	0,005	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,037117	0,037117	0,007459204	0,994560557	994,5605573	1296,650589	427,219821
7	Tris		7,5 6-AHA	3000	0,01	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,04334	0,04334	0,008709807	1,16130761	1161,30761		
8	Tris		7,5 6-AHA	3000	0,01	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,060763	0,060763	0,012211214	1,628161844	1628,161844	1394,734727	330,115794
9	Tris		7,5 6-AHA	3000	0,025	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,048421	0,048421	0,009730908	1,297454448	1297,454448		
10	Tris		7,5 6-AHA	3000	0,025	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,057673	0,057673	0,011590233	1,545364416	1545,364416	1421,409432	175,298819
11	Tris		7,5 6-AHA	3000	0,05	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,051663	0,051663	0,010382436	1,384324759	1384,324759		
12	Tris		7,5 6-AHA	3000	0,05	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,052098	0,052098	0,010469855	1,395980707	1395,980707	1390,152733	8,24200026
13	Tris		7,5 6-AHA	3000	0,1	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,055445	0,055445	0,011142484	1,485664523	1485,664523		
14	Tris		7,5 6-AHA	3000	0,1	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,053798	0,053798	0,010811495	1,44153269	1441,53269	1463,598607	31,2059182
15	Tris		7,5 6-AHA	3000	0,25	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,070549	0,070549	0,014717854	1,890380493	1890,380493		
16	Tris		7,5 6-AHA	3000	0,25	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,080066	0,080066	0,016090434	2,145391211	2145,391211	2017,885852	180,319808
17	Tris		7,5 6-AHA	3000	0,5	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,071562	0,071562	0,014381431	1,917524116	1917,524116		
18	Tris		7,5 6-AHA	3000	0,5	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,071562	0,071562	0,014381431	1,917524116	1917,524116	1917,524116	0
19	Tris		7,5 6-AHA	3000	1,25	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,119472	0,119472	0,024009646	3,201286174	3201,286174		
20	Tris		7,5 6-AHA	3000	1,25	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,16124	0,16124	0,032403537	4,320471597	4320,471597	3760,878885	791,383602
21	Tris		7,5 6-AHA	3000	2	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,159606	0,159606	0,032075161	4,275688103	4276,688103		
22	Tris		7,5 6-AHA	3000	2	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,13833	0,13833	0,027799437	3,70659164	3706,59164	3991,639871	403,119075
23	Tris		7,5 6-AHA	3000	2,5	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,143226	0,143226	0,02878336	3,83778135	3837,78135		
24	Tris		7,5 6-AHA	3000	2,5	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,159668	0,159668	0,032087621	4,278349411	4278,349411	4058,06538	311,528663
25	Tris		7,5 6-AHA	3000	3	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,153833	0,153833	0,030914992	4,121998928	4121,998928		
26	Tris		7,5 6-AHA	3000	3	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,141561	0,141561	0,028448754	3,793167203	3793,167203	3957,583065	232,519143
27	Tris		7,5 6-AHA	3000	4,5	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,10692	0,10692	0,021487138	2,864951768	2864,951768		
28	Tris		7,5 6-AHA	3000	4,5	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,170123	0,170123	0,034188706	4,558494105	4558,494105	7111,722937	1197,515127
29	Tris		7,5 6-AHA	3000	6	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,138037	0,138037	0,027740555	3,698740622	3698,740622		
30	Tris		7,5 6-AHA	3000	6	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,138037	0,138037	0,027740555	3,698740622	3698,740622	3698,740622	0
31	Tris		7,5 6-AHA	3000	8	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,141309	0,141309	0,028398111	3,786414791	3786,414791		
32	Tris		7,5 6-AHA	3000	8	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,172008	0,172008	0,034567524	4,609003215	4609,03215	4197,709003	581,657853
33	Tris		7,5 6-AHA	3000	10	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,144829	0,144829	0,029105506	3,880734191	3880,734191		
34	Tris		7,5 6-AHA	3000	10	W259G	67307	0,736	0,0075	0,4076087	3,0571E-06	0,11142972	40	100	6,22	2	1	-0,107467	0,107467	0,021597066	2,879608789	2879,608789	3380,17149	707,902561

35	Tris		7,5 6-AHA	0	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,031258	0,031258	0,006281752	0,837566988	837,5669882		
36	Tris		7,5 6-AHA	0	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,026867	0,026867	0,005399317	0,719908896	719,908896	778,7379421	83,1968348
37	Tris		7,5 6-AHA	50	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,043083	0,043083	0,008658159	1,154421222	1154,421222		
38	Tris		7,5 6-AHA	50	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,041602	0,041602	0,008360531	1,114737406	1114,737406	1134,579314	28,0606951
39	Tris		7,5 6-AHA	150	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,042182	0,042182	0,00847709	1,130278671	1130,278671		
40	Tris		7,5 6-AHA	150	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,045345	0,045345	0,009112741	1,215032154	1215,032154	1172,655413	95,9297628
41	Tris		7,5 6-AHA	250	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,051487	0,051487	0,010347066	1,379608789	1379,608789		
42	Tris		7,5 6-AHA	250	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,058445	0,058445	0,011745378	1,566050375	1566,050375	1472,829582	131,83411
43	Tris		7,5 6-AHA	400	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,067969	0,067969	0,013659365	1,82124866	1821,24866		
44	Tris		7,5 6-AHA	400	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,077947	0,077947	0,01566459	2,088612004	2088,612004	1954,930332	189,054434
45	Tris		7,5 6-AHA	600	10	W259G	67307	0,12266667	0,0075	2,44565217	1,8342E-05	0,11142972	40	100	6,22	2	1	-0,068023	0,068023	0,013670217	1,822695606	1822,695606		
46	Tris		7,5 6-AHA	600																				

Expérience	Tampon	pH tampon	Substrat	[S] (mM)	[ATP] (mM)	Enzyme	MW (Da)	[E]initiale (mg/ml)	[E]finale (mg/ml)	V enzyme (μL)	Qté enzyme (mg)	[enzyme] (μmol/L)	v échantillon (μL)	V <sub>ox</sub> cuve (μL)	εNADH (L/mmol/cm)	Stoechio NADH	Longueur cuve (cm)	Pente (Abs/min)	VA pente (Abs/min)	[PPi] <sub>libéré</sub> /min (μmol/ml/min)	Moy. [PPi] <sub>libéré</sub> /min (μmol/ml/min)	SD [PPi] <sub>libéré</sub> /min	Activité sp (μmol/min/mg enz)	Activité sp (μmol/mg enz)	Moy. Act sp (μmol/mg enz)	SD Act sp
1	Tris	7,5	6-aminohexanoic acid	3000	0	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,03942	0,03942	0,007922026			1,056270096	1056,270096		
2	Tris	7,5	6-aminohexanoic acid	3000	0	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,07303	0,07303	0,014676447	0,011299236	0,004776097	1,956859593	1956,859593	1506,564845	636,81294
3	Tris	7,5	6-aminohexanoic acid	3000	0,0025	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,094116	0,094116	0,018913987			2,521864952	2521,864952		
4	Tris	7,5	6-aminohexanoic acid	3000	0,0025	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,06205	0,06205	0,013304863	0,016109425	0,003966249	1,773981779	1773,981779	2147,923365	528,833263
5	Tris	7,5	6-aminohexanoic acid	3000	0,005	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,147248	0,147248	0,02959164			3,945551983	3945,551983		
6	Tris	7,5	6-aminohexanoic acid	3000	0,005	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,073681	0,073681	0,014807275	0,022199457	0,010454125	1,974303323	1974,303323	2959,927653	1393,8833
7	Tris	7,5	6-aminohexanoic acid	3000	0,01	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,081944	0,081944	0,016467846			2,195712755	2195,712755		
8	Tris	7,5	6-aminohexanoic acid	3000	0,01	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,101534	0,101534	0,020404743	0,018436294	0,002783807	2,720632369	2720,632369	2458,172562	371,174219
9	Tris	7,5	6-aminohexanoic acid	3000	0,025	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,098927	0,098927	0,019880828			2,650777063	2650,777063		
10	Tris	7,5	6-aminohexanoic acid	3000	0,025	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,082274	0,082274	0,016534164	0,018207496	0,002366449	2,204555198	2204,555198	2427,666131	315,526507
11	Tris	7,5	6-aminohexanoic acid	3000	0,05	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,155929	0,155929	0,031336214			4,178161844	4178,161844		
12	Tris	7,5	6-aminohexanoic acid	3000	0,05	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,052523	0,052523	0,010555265	0,02094574	0,01469435	1,407368703	1407,368703	2792,765273	1959,24662
13	Tris	7,5	6-aminohexanoic acid	3000	0,1	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,173512	0,173512	0,034869775			4,649303323	4649,303323		
14	Tris	7,5	6-aminohexanoic acid	3000	0,1	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,143575	0,143575	0,028853497	0,031861636	0,004254151	3,847132905	3847,132905	4248,218114	567,220142
15	Tris	7,5	6-aminohexanoic acid	3000	0,25	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,196064	0,196064	0,039401929			5,253590568	5253,90568		
16	Tris	7,5	6-aminohexanoic acid	3000	0,25	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,09665	0,09665	0,019423232	0,02941258	0,014127073	2,589764202	2589,764202	3921,677385	1883,60969
17	Tris	7,5	6-aminohexanoic acid	3000	0,5	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,195904	0,195904	0,040093248			5,345766345	5345,766345		
18	Tris	7,5	6-aminohexanoic acid	3000	0,5	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,12278	0,12278	0,024674437	0,023283842	0,010902745	3,289924973	3289,924973	4317,845659	1453,69938
19	Tris	7,5	6-aminohexanoic acid	3000	1	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,219191	0,219191	0,044049638			5,873285102	5873,285102		
20	Tris	7,5	6-aminohexanoic acid	3000	1	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,234523	0,234523	0,047130828	0,045590233	0,00217873	6,284110397	6284,110397	6078,697749	290,497352
21	Tris	7,5	6-aminohexanoic acid	3000	2	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,238798	0,238798	0,047989952			6,398662036	6398,662036		
22	Tris	7,5	6-aminohexanoic acid	3000	2	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,206507	0,206507	0,044110169	0,04470006	0,004652609	5,521355841	5521,355841	5960,008039	620,347886
23	Tris	7,5	6-aminohexanoic acid	3000	4	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,238839	0,238839	0,047998191			6,399758842	6399,758842		
24	Tris	7,5	6-aminohexanoic acid	3000	4	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,238839	0,238839	0,047998191	0,047998191	0	6,399758842	6399,758842	6399,758842	0
25	Tris	7,5	6-aminohexanoic acid	3000	6	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,224533	0,224533	0,045123191			6,016425509	6016,425509		
26	Tris	7,5	6-aminohexanoic acid	3000	6	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,248868	0,248868	0,050013666	0,047568428	0,003458088	6,668488746	6668,488746	6342,457128	461,078337
27	Tris	7,5	6-aminohexanoic acid	3000	8	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,258337	0,258337	0,05191166			6,92221329	6922,21329		
28	Tris	7,5	6-aminohexanoic acid	3000	8	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,258337	0,258337	0,05191166	0,05191166	0	6,92221329	6922,21329	6922,21329	0
29	Tris	7,5	6-aminohexanoic acid	3000	10	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,281095	0,281095	0,056490153			7,532020364	7532,020364		
30	Tris	7,5	6-aminohexanoic acid	3000	10	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,2392752	0,2392752	0,048085852	0,052288002	0,005942738	6,411446945	6411,446945	6971,733655	792,365063
31	Tris	7,5	6-aminohexanoic acid	3000	15	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,221172	0,221172	0,044447749			5,926366559	5926,366559		
32	Tris	7,5	6-aminohexanoic acid	3000	15	W259G/V238T	67307	0,0264	0,0075	11,3636364	8,5227E-05	0,11142972	40	100	6,22	2	1	-0,221172	0,221172	0,044447749	0,044447749	0	5,926366559	5926,366559	5926,366559	0
33	Tris	7,5	6-aminohexanoic acid	0	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,003942	0,003942	0,000792203			0,015844051	15,84405145		
34	Tris	7,5	6-aminohexanoic acid	0	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,007303	0,007303	0,001467645	0,001129924	0,00047761	0,029352894	29,35289389	22,59847267	9,5521941
35	Tris	7,5	6-aminohexanoic acid	15	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,011215	0,011215	0,002253818			0,045076367	45,07636656		
36	Tris	7,5	6-aminohexanoic acid	15	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,009715	0,009715	0,001952371	0,002103095	0,000213155	0,039047428	39,04742765	42,06189711	4,26310358
37	Tris	7,5	6-aminohexanoic acid	25	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,008711	0,008711	0,001750603			0,035012058	35,01205788		
38	Tris	7,5	6-aminohexanoic acid	25	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,006811	0,006811	0,00136877	0,001559686	0,000269997	0,027375402	27,37540193	31,1937299	5,39993121
39	Tris	7,5	6-aminohexanoic acid	100	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,05732	0,05732	0,011519293			0,230385852	230,3858521		
40	Tris	7,5	6-aminohexanoic acid	100	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,05611	0,05611	0,011276125	0,011397709	0,000171945	0,225522508	225,522508	227,9541801	3,43890356
41	Tris	7,5	6-aminohexanoic acid	300	10	W259G/V238T	67307	0,264	0,05	7,5757578	0,00037879	0,74286478	40	100	6,22	2	1	-0,088991	0,088991	0,017884						

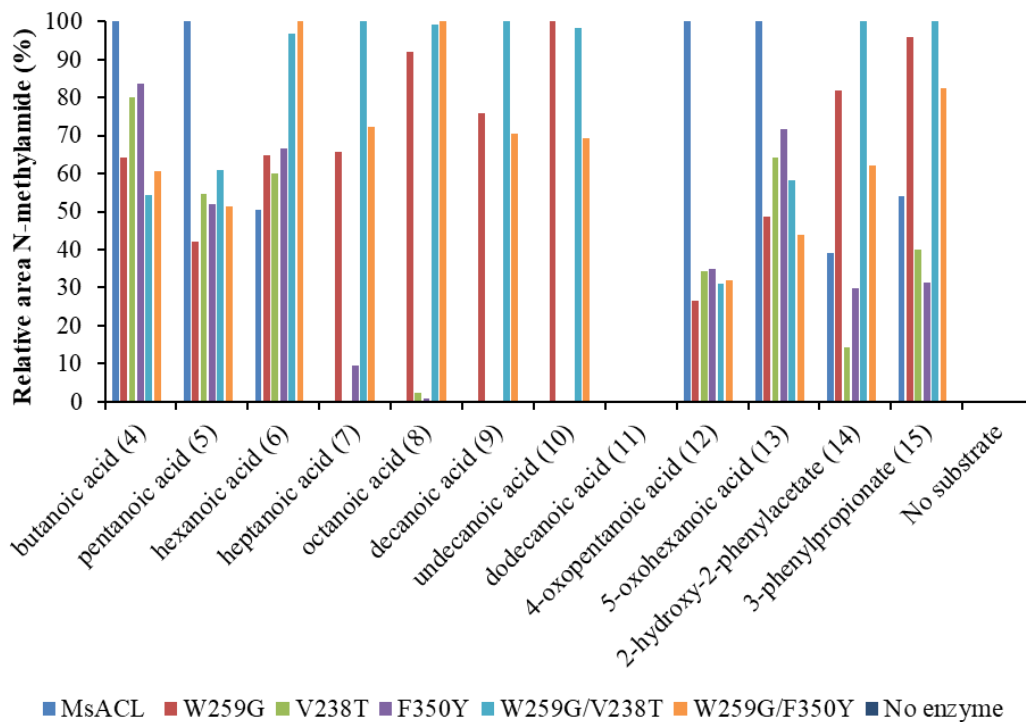
Annexe 8. Rendements analytiques de la formation de lactames à partir des acides  $\omega$ -aminés avec A4YDT1 et ses mutants.

Conditions réactionnelles : 5 mM acide  $\omega$ -aminé, 5 mM ATP, 5 mM  $MnCl_2$ , 50 mM tampon MOPS pH 8,5, 2 mg/mL enzyme purifiée. 400 rpm, 60°C, 24h. Les réactions contrôles correspondent à des réactions sans enzyme. Rendements analytiques déduits de courbes de calibrations par UHPLC-MS (APCI). *MsACL* correspond à A4YDT1.



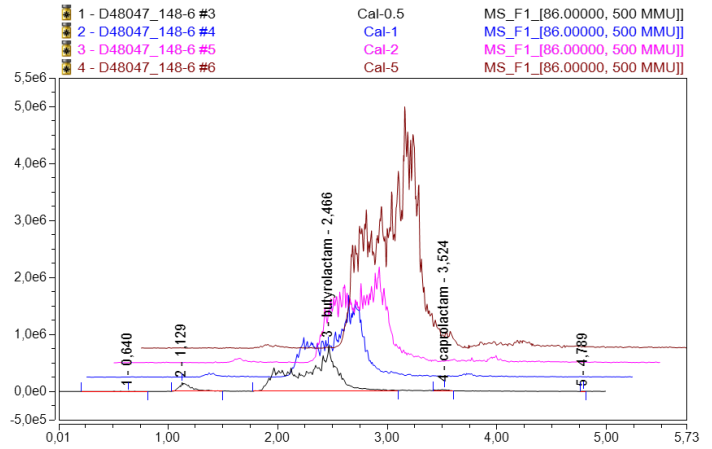
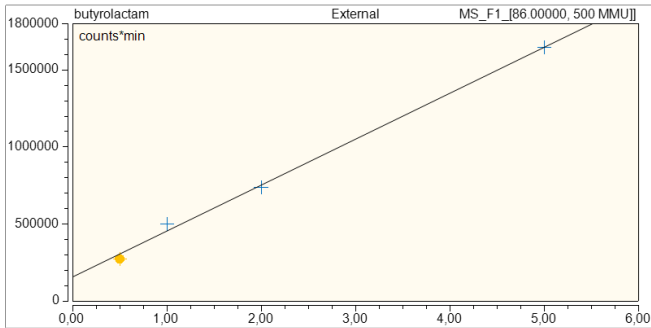
Annexe 9. Comparaison des aires relatives pour divers substrats par rapport au meilleur mutant pour chaque substrat, par détection du N-méthylamide correspondant par UHPLC-MS (ESI).

Conditions réactionnelles : 5 mM acide, 5 mM ATP, 5 mM  $MgCl_2$ , 50 mM  $MeNH_2$ , 50 mM tampon phosphate pH 8, 0.2 mg/mL enzyme purifié. 400 rpm, 60°C, 24h. *MsACL* correspond à A4YDT1.

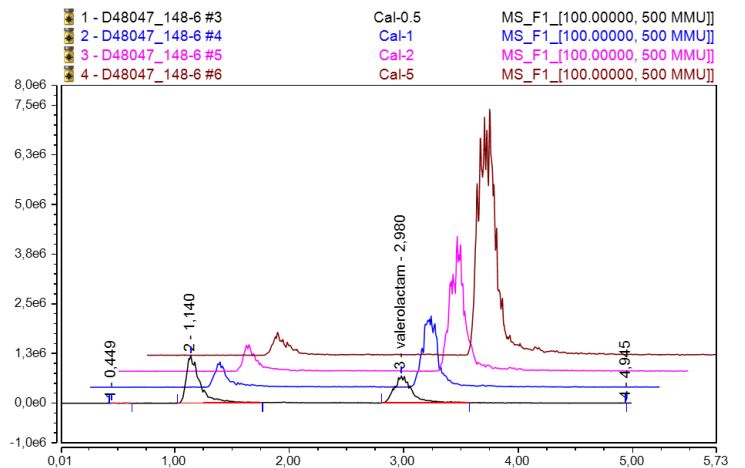
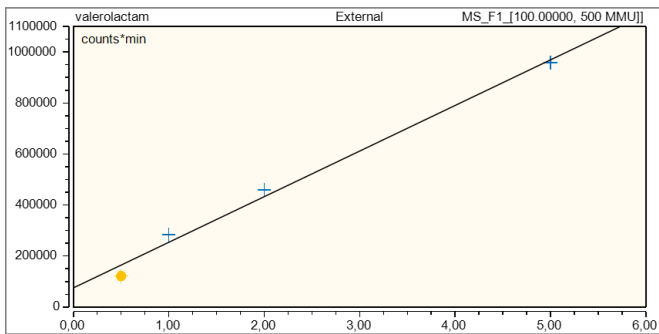


Annexe 10. Droite de calibration des standards lactames commerciaux utilisée pour calculer les rendements analytiques avec A4YDT1 et les mutants.

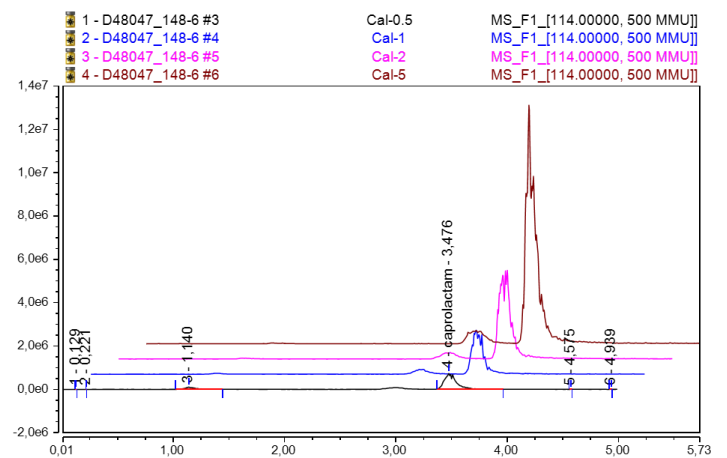
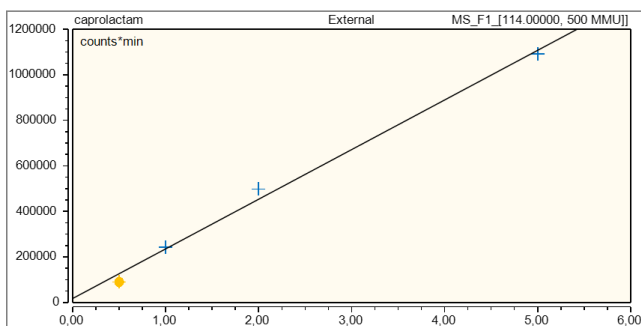
### A. Butyrolactame



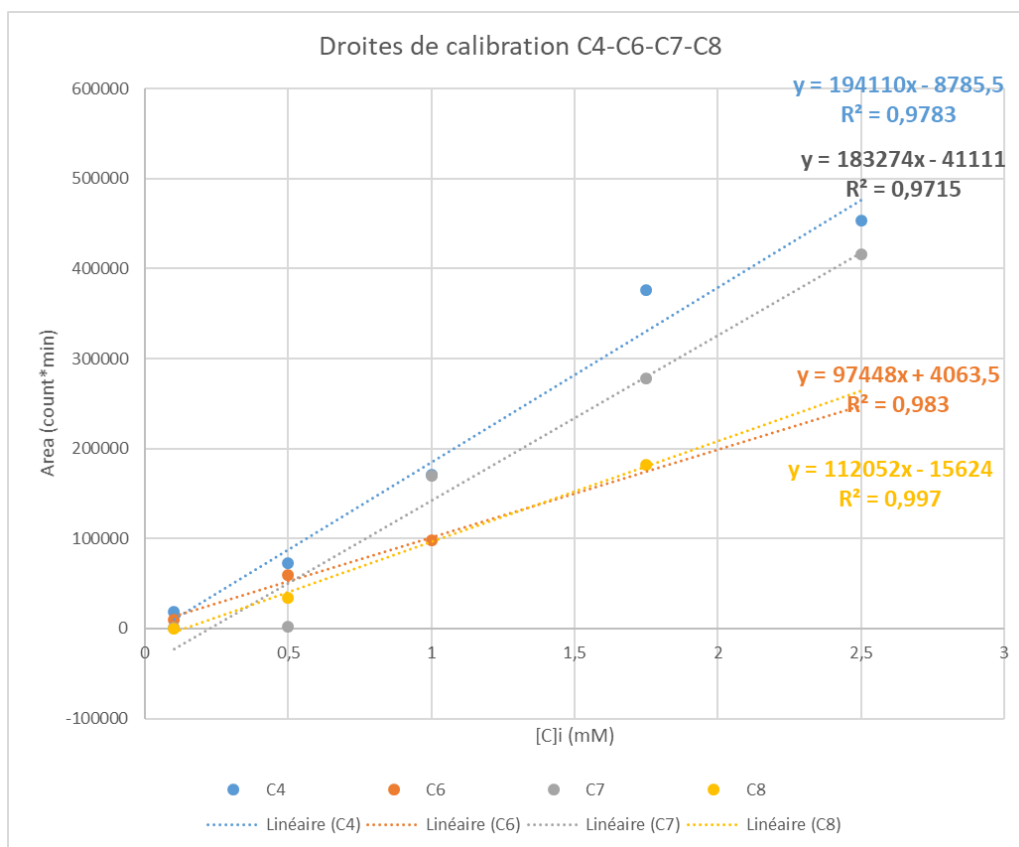
### B. Valérolactame



### C. Caprolactame

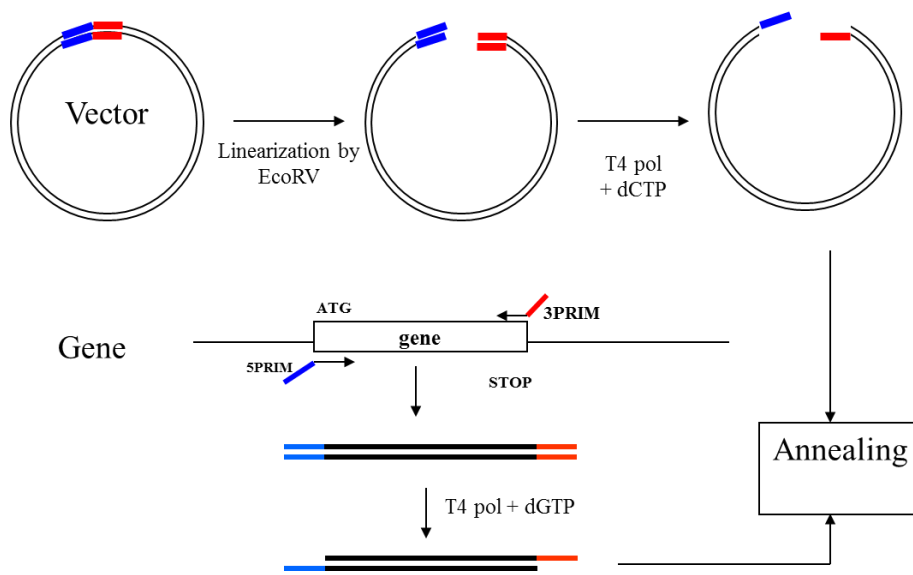


Annexe 11. Droites de calibration des standards amides synthétisés au laboratoire.





Annexe 12. Schéma du clonage par ligation indépendante (LIC).



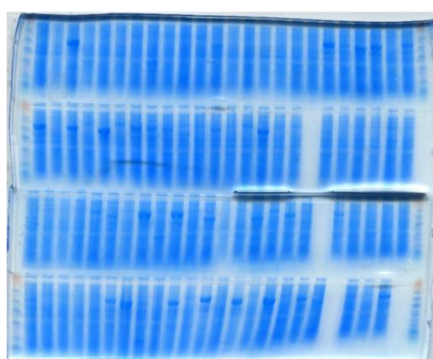
Annexe 13. Gels SDS-PAGE de la production de la première collection de CoA ligase CoA Lig1.

### Plaque CoA 1

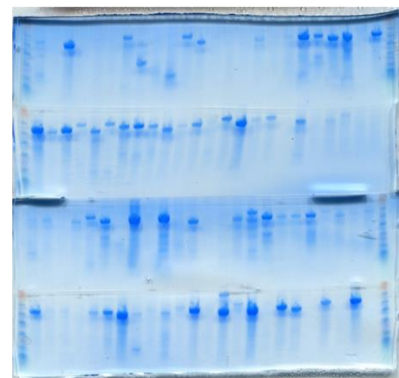
Induction



Lysat

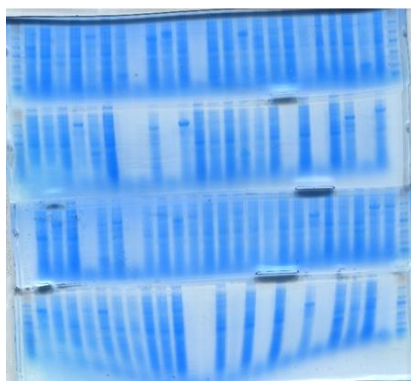


Purification (His-link)

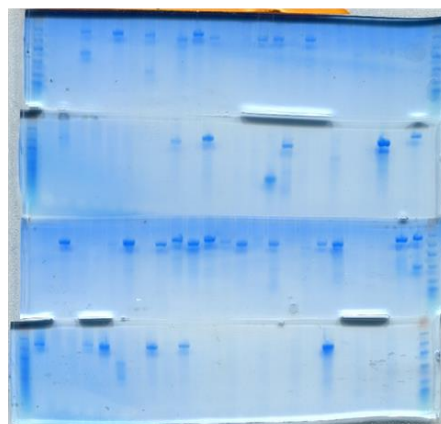


### Plaque CoA 2

Induction

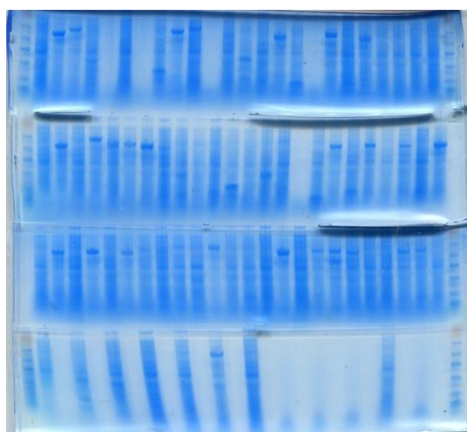


Purification (His-link)

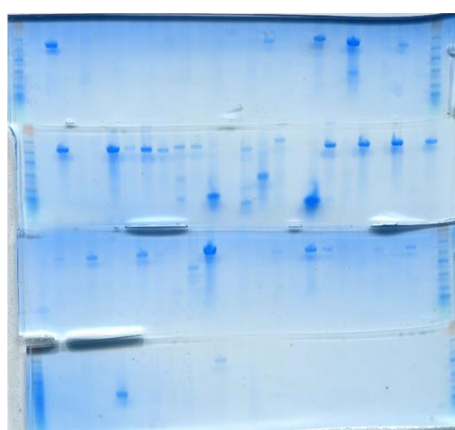


### Plaque CoA 3

Induction



Purification (His-link)



Annexe 14. Concentration (en mg/ml) des CoA ligases purifiées de la collection CoA Lig1. Concentrations déterminées selon la méthode Bradford.

Position plaque	Code UniPr	Organisme	Bradford
CoA_1_A01	A0A101FK9	Thermodesulfobacterium commune	0,177
CoA_1_A02	A0A231VLB0	Thermoanaerobacterium thermosaccharolyticum (Clostridium thermosaccharolyticum)	0,984
CoA_1_A04	B5YGV3	Thermodesulfovibrio yellowstonii (strain ATCC 51303 / DSM 11347 / YP87)	0,277
CoA_1_A06	F0S1S7	Desulfurobacterium thermolithotrophum (strain DSM 11699 / BSA)	0,278
CoA_1_A10	A0A0K9HT26	Geobacillus stearothermophilus (Bacillus stearothermophilus)	1,18
CoA_1_A11	A0A101E6E5	Caldanaerobacter subterraneus	0,665
CoA_1_B04	A0A150MS74	Geobacillus stearothermophilus (Bacillus stearothermophilus)	0,33
CoA_1_B06	A0A1D7X7U6	Moorella thermoacetica (Clostridium thermoacetum)	0,336
CoA_1_B08	A0A2X0QFM8	Brochothrix thermosphacta (Microbacterium thermosphactum)	0,131
CoA_1_B10	A3DK40	Clostridium thermocellum (strain ATCC 27405 / DSM 1237 / NBRC 103400 / NCIMB 10682 / NRRL B-4536 / VPI 7372) (Ruminiclostridium thermocellum)	0,58
CoA_1_B11	A4IKV3	Geobacillus thermodenitrificans (strain NG80-2)	Range?
CoA_1_B12	A4ILR6	Geobacillus thermodenitrificans (strain NG80-2)	0,733
CoA_1_C01	A4IMI8	Geobacillus thermodenitrificans (strain NG80-2)	Range?
CoA_1_C02	A4IMJ1	Geobacillus thermodenitrificans (strain NG80-2)	1,33
CoA_1_C03	A4IN03	Geobacillus thermodenitrificans (strain NG80-2)	0,683
CoA_1_C04	A4IN22	Geobacillus thermodenitrificans (strain NG80-2)	0,66
CoA_1_C05	A4INB3	Geobacillus thermodenitrificans (strain NG80-2)	0,359
CoA_1_C06	A4IPJ5	Geobacillus thermodenitrificans (strain NG80-2)	0,137
CoA_1_C08	A4IRS7	Geobacillus thermodenitrificans (strain NG80-2)	1,417
CoA_1_C09	A5D0C3	Pelotomaculum thermopropionicum (strain DSM 13744 / JCM 10971 / SI)	0,155
CoA_1_C10	A5D1Z4	Pelotomaculum thermopropionicum (strain DSM 13744 / JCM 10971 / SI)	0,292
CoA_1_C11	A5D1Z7	Pelotomaculum thermopropionicum (strain DSM 13744 / JCM 10971 / SI)	0,086
CoA_1_D01	B9L3G6	Thermomicrobium roseum (strain ATCC 27502 / DSM 5159 / P-2)	0,212
CoA_1_D02	C5A3C0	Thermococcus gammatolerans (strain DSM 15229 / JCM 11827 / EJ3)	0,185
CoA_1_D03	D3SQH8	Thermocrinis albus (strain DSM 14484 / JCM 11386 / HI 11/12)	0,298
CoA_1_D04	D7CKM4	Syntrophothermus lipocalidus (strain DSM 12680 / TGB-C1)	0,853
CoA_1_D05	D7CKY2	Syntrophothermus lipocalidus (strain DSM 12680 / TGB-C1)	0,445
CoA_1_D06	D7CPO9	Syntrophothermus lipocalidus (strain DSM 12680 / TGB-C1)	0,352
CoA_1_D07	F8A8M3	Thermodesulfatator indicus (strain DSM 15286 / JCM 11887 / CIR29812)	0,435
CoA_1_D08	F8A8Y9	Thermodesulfatator indicus (strain DSM 15286 / JCM 11887 / CIR29812)	0,129
CoA_1_D11	K8DYP0	Desulfotomaculum hydrothermale Lam5 = DSM 18033	0,107
CoA_1_E01	K8ELI8	Desulfotomaculum hydrothermale Lam5 = DSM 18033	0,135
CoA_1_E03	L0R9S0	Desulfovibrio hydrothermalis AM13 = DSM 14728	0,373
CoA_1_E08	M8D6Q9	Anoxybacillus flavithermus AK1	0,305
CoA_1_E09	O26753	Methanothermobacter thermautotrophicus (strain ATCC 29096 / DSM 1053 / JCM 10044 / NBRC 100330 / DeIta H) (Methanobacterium thermoautotrophicum)	0,936
CoA_1_E10	Q0PVA7	Methanothermobacter thermautotrophicus (Methanobacterium thermoformicum)	0,131
CoA_1_E11	Q2I0L6	Methanothermobacter thermautotrophicus (Methanobacterium thermoformicum)	0,075
CoA_1_F01	Q2RH11	Moorella thermoacetica (strain ATCC 39073 / JCM 9320)	0,143
CoA_1_F02	Q2RJ14	Moorella thermoacetica (strain ATCC 39073 / JCM 9320)	0,15
CoA_1_F03	Q3A9H6	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,701
CoA_1_F04	Q3ABD3	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Range?
CoA_1_F05	Q3ABN2	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	Range?
CoA_1_F06	Q3ABP3	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,388
CoA_1_F08	Q3AEC0	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,557
CoA_1_F09	Q3AEI5	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,155
CoA_1_F10	Q3AEK4	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,48
CoA_1_F11	Q3AEY6	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,121
CoA_1_F12	Q3AFG3	Carboxydotherrmus hydrogenoformans (strain ATCC BAA-161 / DSM 6008 / Z-2901)	0,188
CoA_1_G01	Q978T3	Thermoplasma volcanium (strain ATCC 51530 / DSM 4299 / JCM 9571 / NBRC 15438 / GSS1)	0,349
CoA_1_G02	Q97CH8	Thermoplasma volcanium (strain ATCC 51530 / DSM 4299 / JCM 9571 / NBRC 15438 / GSS1)	0,079
CoA_1_G03	Q9HI39	Thermoplasma acidophilum (strain ATCC 25905 / DSM 1728 / JCM 9062 / NBRC 15155 / AMRC-C165)	0,128
CoA_1_G04	A0A087LF85	Geobacillus stearothermophilus (Bacillus stearothermophilus)	1,113
CoA_1_G06	A0A1J1EL15	Thermus thermophilus	0,125
CoA_1_G08	B9KZ02	Thermomicrobium roseum (strain ATCC 27502 / DSM 5159 / P-2)	0,148
CoA_1_G10	B9L3E4	Thermomicrobium roseum (strain ATCC 27502 / DSM 5159 / P-2)	0,374
CoA_1_G11	B9L4S4	Thermomicrobium roseum (strain ATCC 27502 / DSM 5159 / P-2)	0,243
CoA_1_G12	B9L4W4	Thermomicrobium roseum (strain ATCC 27502 / DSM 5159 / P-2)	1,205
CoA_1_H01	B9L5H0	Thermomicrobium roseum (strain ATCC 27502 / DSM 5159 / P-2)	0,079
CoA_1_H03	D1A9C5	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	0,439
CoA_1_H04	D1ABC1	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	0,19
CoA_1_H05	D1ADK0	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	0,301
CoA_1_H06	A4YDT1	temoin	1,168
CoA_1_H07	B3G250	temoin	1,782
CoA_1_H08	A0A0A8PCP0	temoin	Range?
CoA_1_H09	A4YDR9	temoin	0,652

Position plaque	Code UniProt	Genome	Bradford
CoA_2_A03	D1C448	Sphaerobacter thermophilus (strain DSM 20745 / S 6022)	0,251
CoA_2_A04	D1C5Q3	Sphaerobacter thermophilus (strain DSM 20745 / S 6022)	0,205
CoA_2_A05	D1C738	Sphaerobacter thermophilus (strain DSM 20745 / S 6022)	0,283
CoA_2_A08	D1C9B6	Sphaerobacter thermophilus (strain DSM 20745 / S 6022)	0,09
CoA_2_A09	D1CAS9	Sphaerobacter thermophilus (strain DSM 20745 / S 6022)	0,174
CoA_2_B05	D7BFC2	Meiothermus silvanus (strain ATCC 700542 / DSM 9946 / VI-R2) (Thermus silvanus)	0,141
CoA_2_B09	E4U4H7	Oceanithermus profundus (strain DSM 14977 / NBRC 100410 / VKM B-2274 / 506)	0,051
CoA_2_B10	E4U5A5	Oceanithermus profundus (strain DSM 14977 / NBRC 100410 / VKM B-2274 / 506)	0,16
CoA_2_C05	G2SH91	Rhodothermus marinus SG0.5JP17-172	0,125
CoA_2_C06	G2SHU3	Rhodothermus marinus SG0.5JP17-172	0,349
CoA_2_D08	Q47L31	Thermobifida fusca (strain YX)	0,573
CoA_2_D11	Q5SKN9	Thermus thermophilus (strain HB8 / ATCC 27634 / DSM 579)	1,419
CoA_2_D12	Q72J95	Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039)	0,312
CoA_2_E05	D1A5F6	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,622
CoA_2_E06	D1A5K7	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,667
CoA_2_E07	D1A5P2	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,293
CoA_2_E08	D1A5V8	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,379
CoA_2_E10	D1A5W6	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,761
CoA_2_E12	D1A6J8	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,381
CoA_2_F01	D1A763	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,394
CoA_2_F03	D1A926	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,455
CoA_2_F04	D1AAQ7	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,275
CoA_2_F05	D1AB70	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,678
CoA_2_F06	D1AB87	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Her	0,161
CoA_2_F07	D1ABL9	Thermomonospora curvata (strain ATCC 19995 / DSM 43183 / JCM 3096 / NBRC 15933 / NCIMB 10081 / Henssen B9)	0,168
CoA_2_F09	D1C9C5	Sphaerobacter thermophilus (strain DSM 20745 / S 6022)	0,239
CoA_2_F12	D6Y4P1	Thermobispora bispora (strain ATCC 19993 / DSM 43833 / CBS 139.67 / JCM 10125 / NBRC 14880 / R51)	0,699
CoA_2_G01	D6Y6S5	Thermobispora bispora (strain ATCC 19993 / DSM 43833 / CBS 139.67 / JCM 10125 / NBRC 14880 / R51)	0,46
CoA_2_G03	D6YAV1	Thermobispora bispora (strain ATCC 19993 / DSM 43833 / CBS 139.67 / JCM 10125 / NBRC 14880 / R51)	0,283
CoA_2_G10	F2NQX1	Marinithermus hydrothermalis (strain DSM 14884 / JCM 11576 / T1)	0,66
CoA_2_H02	K6P2L2	Thermaerobacter subterraneus DSM 13965	0,127
CoA_2_H04	K6PNF8	Thermaerobacter subterraneus DSM 13965	0,392
CoA_2_H05	K6Q029	Thermaerobacter subterraneus DSM 13965	0,273
CoA_2_H06	A4YDT1	temoin	0,005
CoA_2_H07	B3G250	temoin	0,031
CoA_2_H08	A0A0A8PCP0	temoin	-0,011
CoA_2_H09	A4YDR9	temoin	-0,001

Position plaque	Code UniProt	Organisme	Bradford
CoA_3_A08	Q97VT6	Sulfolobus solfataricus (strain ATCC 35092 / DSM 1617 / JCM 11322 / P2)	0,087
CoA_3_A12	A0A088E3M2	Metallosphaera sedula	0,091
CoA_3_B01	A0A088E5G6	Metallosphaera sedula	0,706
CoA_3_B09	C5A3C0	Thermococcus gammatolerans (strain DSM 15229 / JCM 11827 / EJ3)	0,427
CoA_3_B10	D2RDA5	Archaeoglobus profundus (strain DSM 5631 / JCM 9629 / NBRC 100127 / Av18)	0,863
CoA_3_C03	D3RZ03	Ferroglobus placidus (strain DSM 10642 / AEDII12DO)	0,694
CoA_3_C04	D3RZA8	Ferroglobus placidus (strain DSM 10642 / AEDII12DO)	0,402
CoA_3_C05	D3SOR8	Ferroglobus placidus (strain DSM 10642 / AEDII12DO)	0,136
CoA_3_C06	D3SOX4	Ferroglobus placidus (strain DSM 10642 / AEDII12DO)	0,829
CoA_3_C07	D3S144	Ferroglobus placidus (strain DSM 10642 / AEDII12DO)	0,222
CoA_3_C08	D7E893	Methanohalobium evestigatum (strain ATCC BAA-1072 / DSM 3721 / NBRC 107634 / OCM 161 / Z-7303)	0,202
CoA_3_D01	D9Q0T0	Acidilobus saccharovorans (strain DSM 16705 / JCM 18335 / VKM B-2471 / 345-15)	0,652
CoA_3_D03	E1QNN2	Vulcanisaeta distributa (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)	0,083
CoA_3_D04	E1QQN7	Vulcanisaeta distributa (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)	0,155
CoA_3_D05	E1QRU7	Vulcanisaeta distributa (strain DSM 14429 / JCM 11212 / NBRC 100878 / IC-017)	0,052
CoA_3_D09	O28246	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,496
CoA_3_D10	O28347	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,639
CoA_3_D11	O28502	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,935
CoA_3_D12	O28762	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,221
CoA_3_E04	O29570	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,219
CoA_3_E06	O29881	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	1,439
CoA_3_E08	O30043	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,116
CoA_3_E09	O30302	Archaeoglobus fulgidus (strain ATCC 49558 / VC-16 / DSM 4304 / JCM 9628 / NBRC 100126)	0,281
CoA_3_E11	Q4J6T2	Sulfolobus acidocaldarius (strain ATCC 33909 / DSM 639 / JCM 8929 / NBRC 15157 / NCIMB 11770)	0,119
CoA_3_E12	Q4J776	Sulfolobus acidocaldarius (strain ATCC 33909 / DSM 639 / JCM 8929 / NBRC 15157 / NCIMB 11770)	0,086
CoA_3_F01	Q4J801	Sulfolobus acidocaldarius (strain ATCC 33909 / DSM 639 / JCM 8929 / NBRC 15157 / NCIMB 11770)	0,115
CoA_3_F02	Q4J9Q2	Sulfolobus acidocaldarius (strain ATCC 33909 / DSM 639 / JCM 8929 / NBRC 15157 / NCIMB 11770)	0,171
CoA_3_F04	Q4J9W4	Sulfolobus acidocaldarius (strain ATCC 33909 / DSM 639 / JCM 8929 / NBRC 15157 / NCIMB 11770)	0,003
CoA_3_F05	Q53623	Sulfolobus acidocaldarius	0,055
CoA_3_F09	Q97CH8	Thermoplasma volcanium (strain ATCC 51530 / DSM 4299 / JCM 9571 / NBRC 15438 / GSS1)	0,072
CoA_3_G03	A8A939	Ignicoccus hospitalis (strain KIN4/I / DSM 18386 / JCM 14125)	0,429
CoA_3_G06	D9Q0V8	Acidilobus saccharovorans (strain DSM 16705 / JCM 18335 / VKM B-2471 / 345-15)	0,047

Annexe 15. Résultats du criblage de la collection CoA Lig1 suite à l'analyse avec la cascade PK-LDH. Légende : vert si la CoA ligase est active sur ce pool et rouge si la CoA ligase est inactive ce pool. Une CoA ligase est considérée comme active si son activité est supérieure ou égale à 2% d'activité par rapport au seuil de détection.

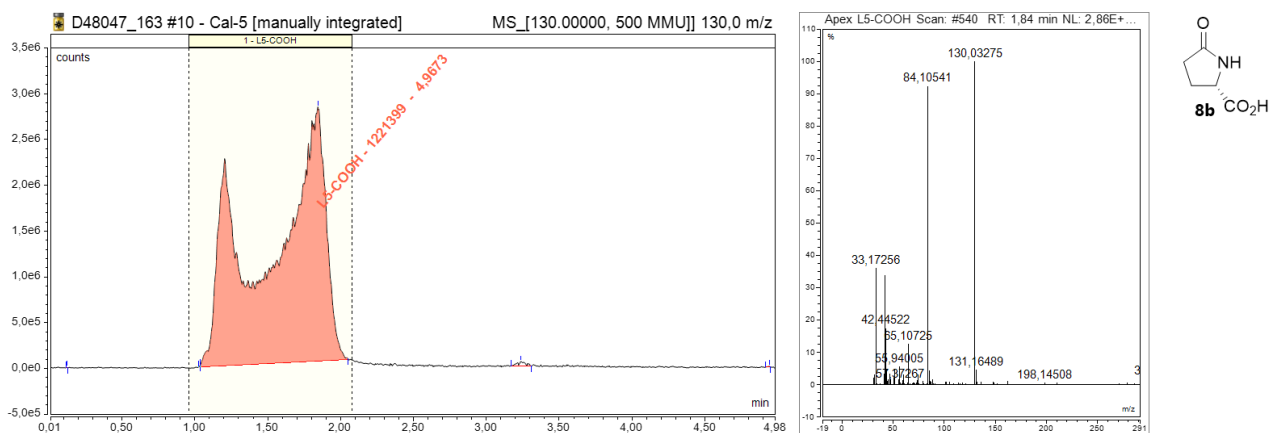
Enzyme	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5	Pool 6
AOA150LLT2						
A4IMI8						
A4IMJ1						
A5D1Z4						
D3SQH8						
F8A8M3						
K8DYP0						
LOR9S0						
LORAD3						
Q3ABD3						
Q3ABN2						
Q3ABP3						
Q3AEI5						
Q9HI39						
B9L4W4						
D1A9C5						
D1ABC1						
A4YDT1						
B3G250						
D1C3U2						
D1C448						
D3PTT7						
D6Y433						
D7BGN5						
D7BIL1						
F2NLM4						
F2NLM5						
G2SH91						
G2SHU3						
Q72L27						
S9S7W1						
D1A3M5						
D1A5W6						
D1AB70						
D1AB87						
F2NPU2						
K6P2L2						
K6Q029						
Q97UW7						
Q97VT6						
C5A3C0						
D3RXJ3						
D3RYS2						
D3RZ03						
D3RZA8						
D9PUA1						
E1QQ97						
E1QT70						
F2KPD0						
O28246						
O29881						
Q9HLD6						
V9S912						
A8A939						
B0R4T6						

Annexe 16. Résultats de la vérification des CoA ligases sélectionnées par analyse UHPLC. Légende : vert si l'activité a été vérifiée et rouge si l'activité n'a pas été vérifiée.

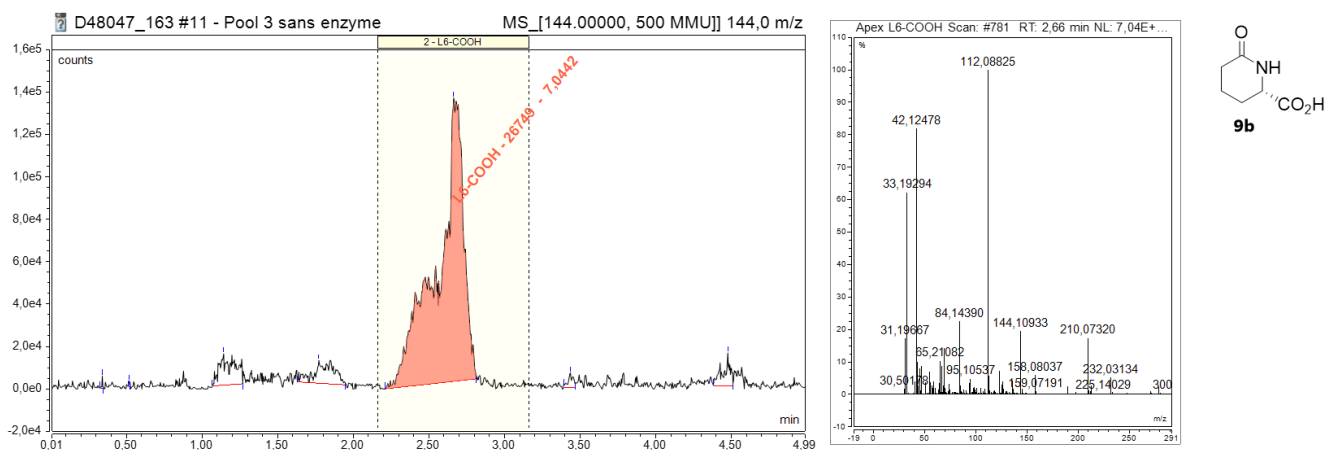
Enzyme	Pool 1	Pool 2	Pool 3	Pool 6
AQA150LLT2				
A4IM8				
A4IMU1				
A5D1Z4				
D3SQH8				
F8A8M3				
K8DYPO				
L0R9S0				
L0RAD3				
Q3ABD3				
Q3ABN2				
Q3ABP3				
Q3AEI5				
Q9HI39				
B9L4W4				
D1A9C5				
D1ABC1				
A4YDT1				
B3G250				
D1C3U2				
D1C448				
D3PTT7				
D6Y433				
D7BGN5				
D7BIL1				
F2NLM4				
F2NLM5				
G2SH91				
G2SHU3				
Q72L27				
S9S7W1				
D1A3M5				
D1A5W6				
D1AB70				
D1AB87				
F2NPU2				
K6P2L2				
K6Q029				
Q97UW7				
Q97VT6				
C5A3C0				
D3RXJ3				
D3RYS2				
D3RZ03				
D3RZA8				
D9PUA1				
E1QQ97				
E1QT70				
F2KPDO				
O28246				
O29881				
Q9HLD6				
V9S912				
A8A939				
B0R4T6				

Annexe 17. Chromatogrammes des lactames attendus par cyclisation des acides  $\omega$ -aminés et leur spectre de masse associé.

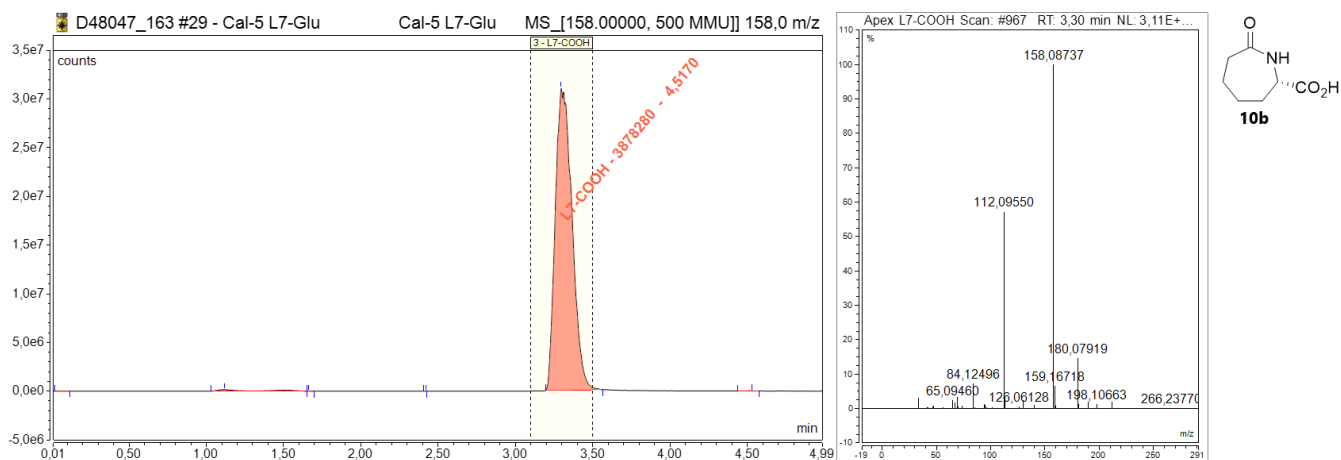
Pyroglutamate (**8b**) ( $m/z = 130$ ,  $t_r = 1.84$  min)



Acide (*S*)-6-oxopiperidine-2-carboxylique (**9b**) ( $m/z = 144$ ,  $t_r = 2.66$  min)

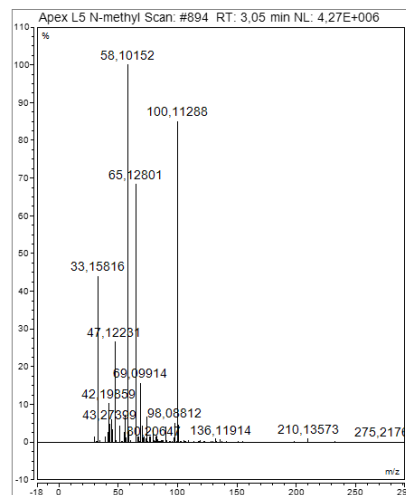
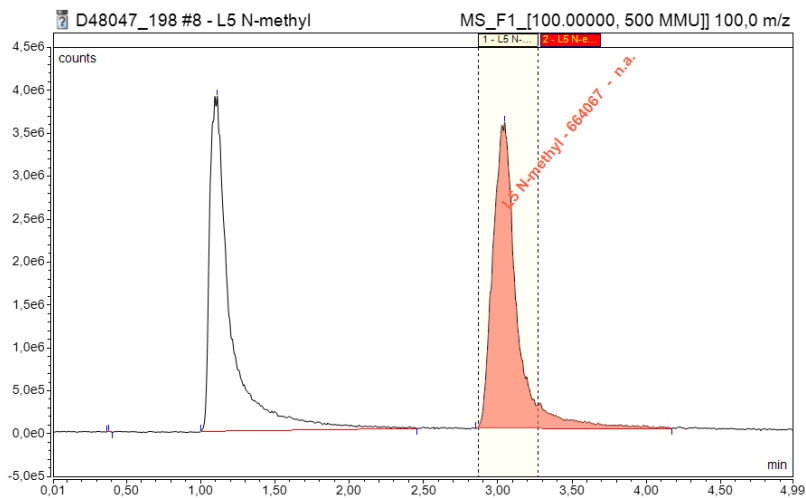


Acide (*S*)-7-oxoazépane-2-carboxylique (**10b**) ( $m/z = 158$ ,  $t_r = 3.30$  min)



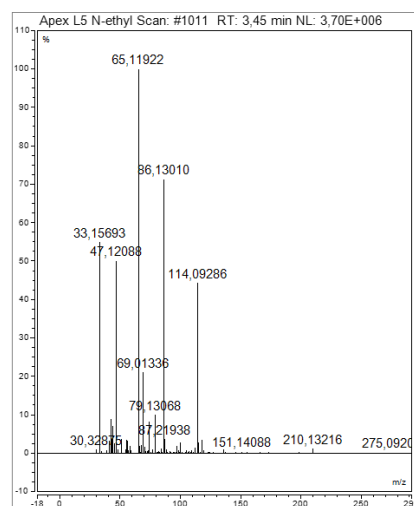
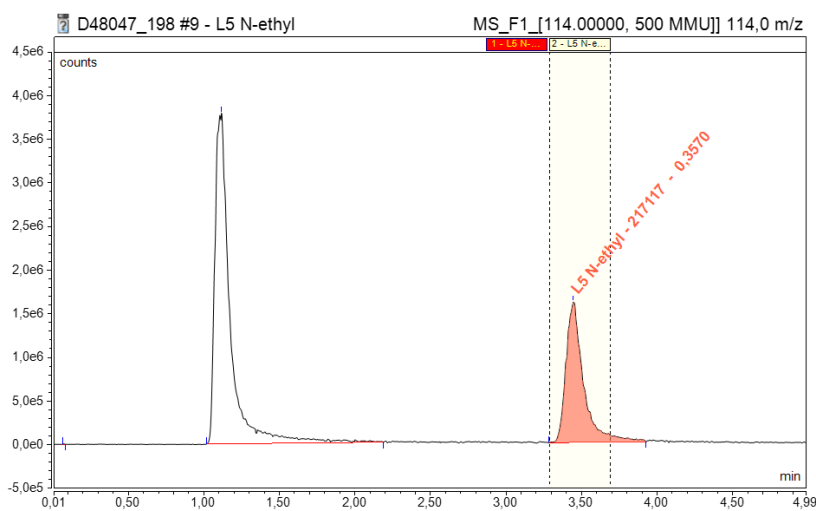


N-méthylbutyrolactame (16b) ( $m/z = 100$ ,  $t_r = 3.05$  min)



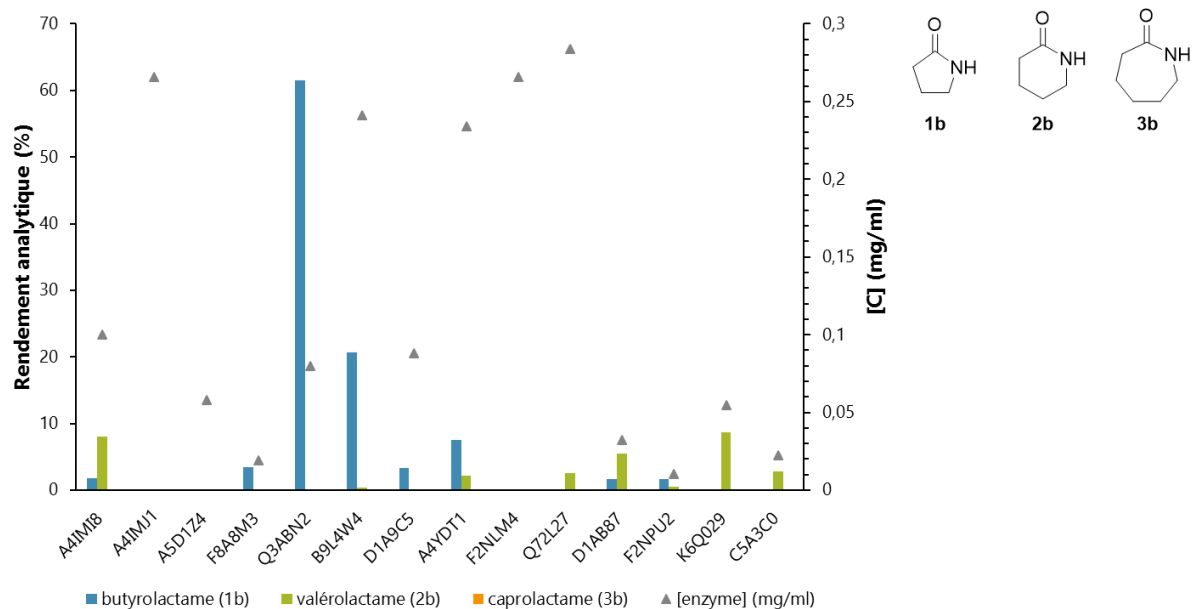
**16b**

N-éthylbutyrolactame (17b) ( $m/z = 114$ ,  $t_r = 3.45$  min)

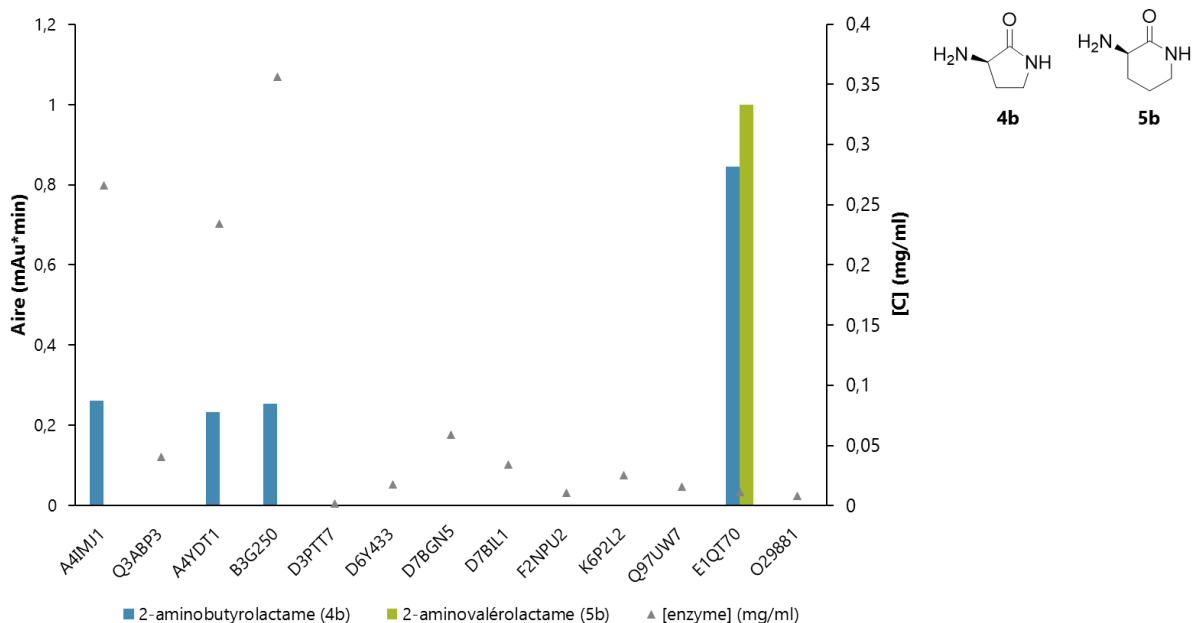


**17b**

Annexe 18. Rendements analytiques de la synthèse de lactames du pool 1 avec les CoA ligases sélectionnées suite à l'analyse avec la cascade PK-LDH. Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl<sub>2</sub>, 1,5 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h. Analyse par UHPLC-MS (APCI). Rendements déduits d'une gamme de calibration avec les standards commerciaux.

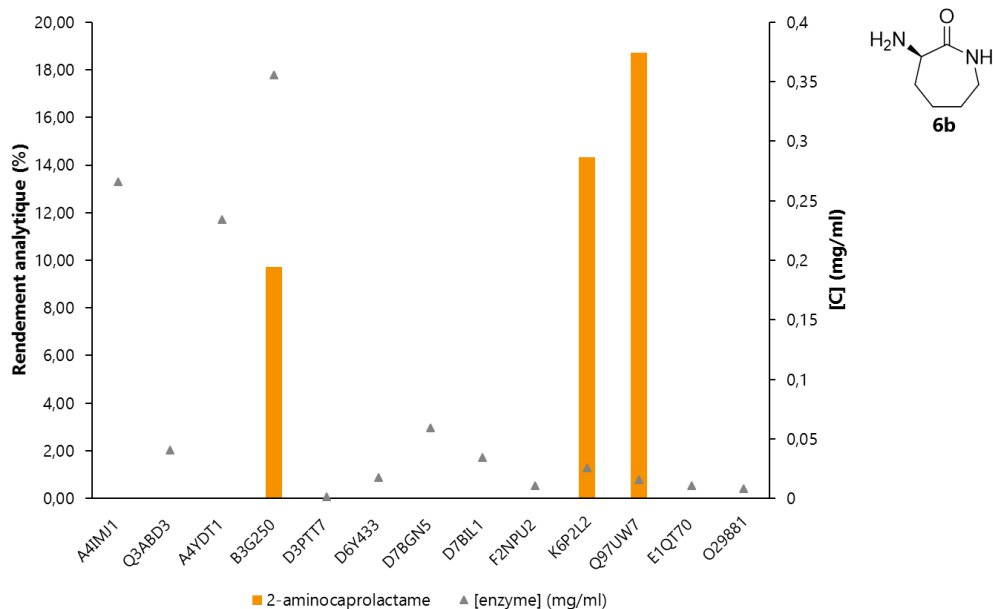


Annexe 19. Détection des aires correspondant aux lactames **4b** et **5b** attendus pour les substrats du pool 2 avec les CoA ligases sélectionnées suite à la cascade PK-LDH. Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 2 mM ATP, 2 mM MnCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h.



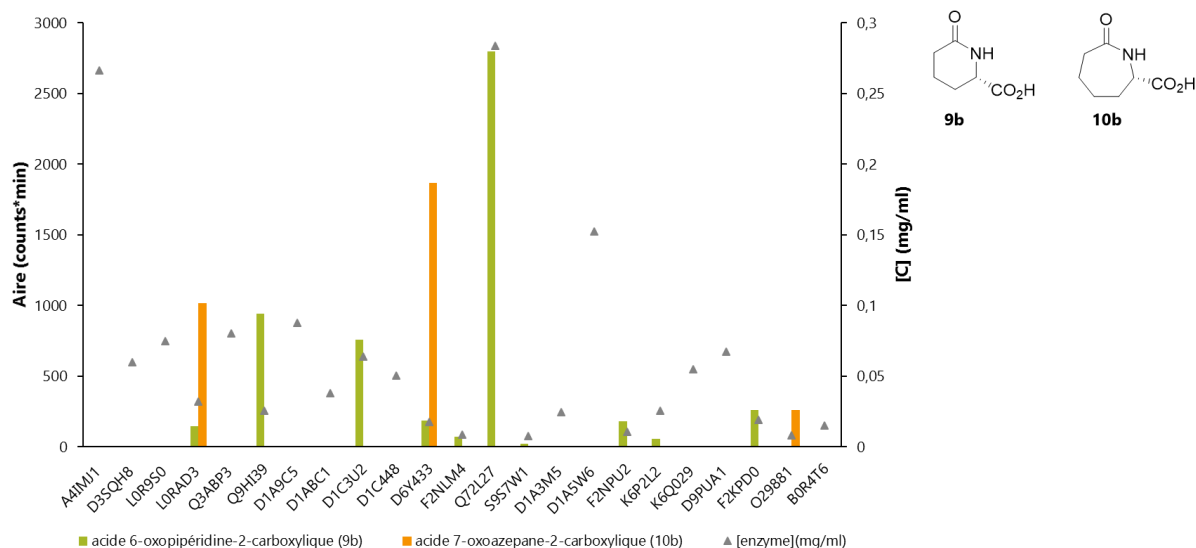
Annexe 20. Rendements analytiques pour la synthèse du 2-aminocaprolactame (**6b**) avec les CoA ligases sélectionnées pour le pool 2. Rendements déterminés à partir d'une gamme de calibration avec le standard commercial après dérivatisation au BzCl.

Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 2 mM ATP, 2 mM MnCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C, sans agitation, 5h.



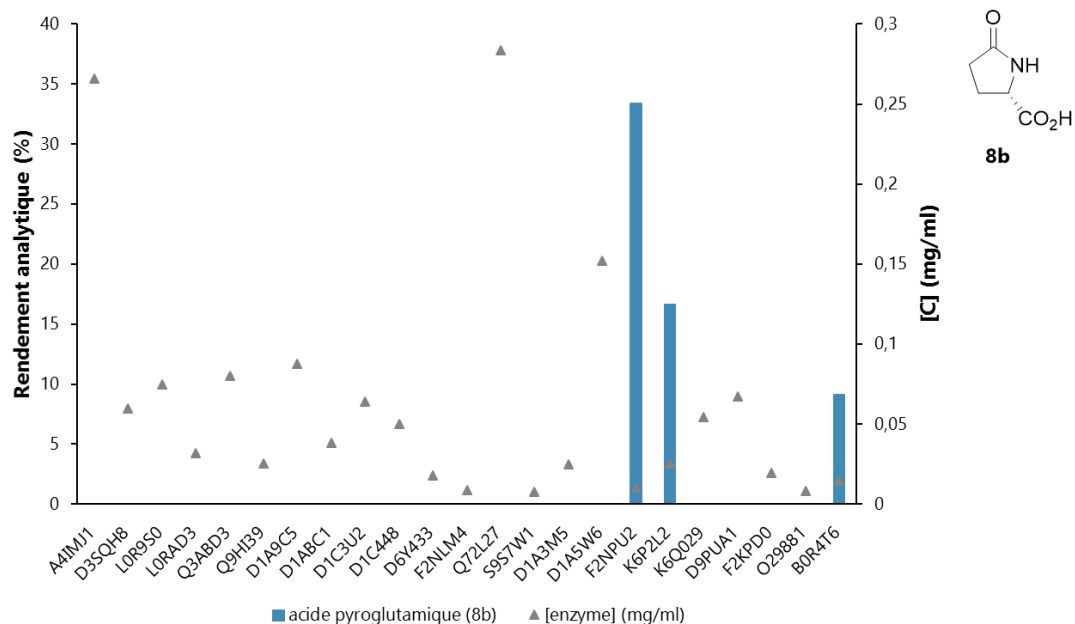
Annexe 21. Détection des aires correspondant aux lactames **9b** et **10b**.

Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM MnCl<sub>2</sub>, 1,5 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C sans agitation à l'étuve, 5h. Analyse par UHPLC-MS (ESI). Aires détectées à partir du m/z attendu pour les lactames correspondants.



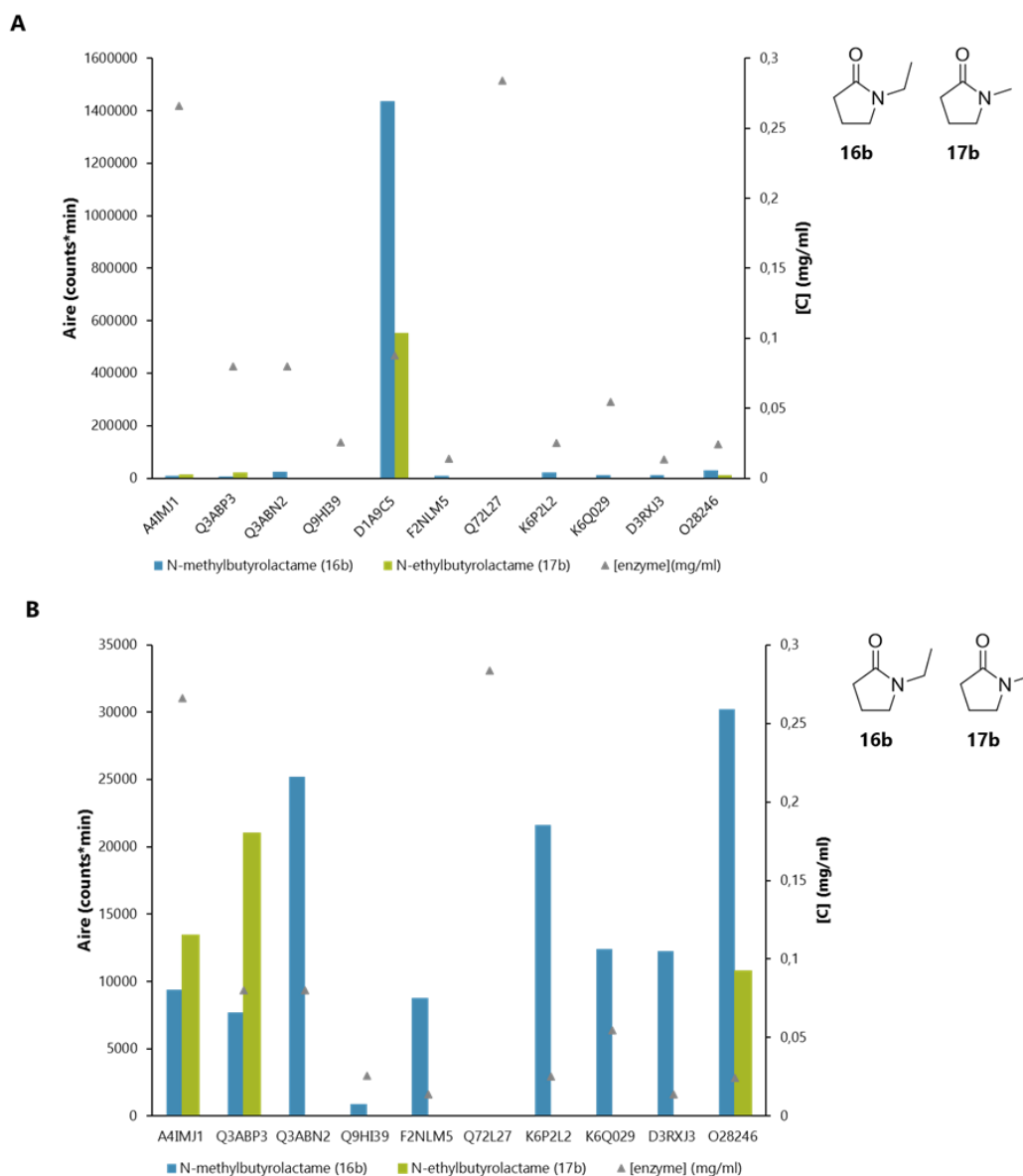
Annexe 22. Rendements analytiques pour la synthèse de l'acide pyroglutamique (**8b**) avec les CoA ligases sélectionnées pour le pool 3. Rendements déterminés à partir d'une gamme de calibration avec le standard commercial.

Conditions réactionnelles : 0,5 mM de chaque acide  $\omega$ -aminé dans le pool, 1,5 mM ATP, 1,5 mM  $MnCl_2$ , 1,5 mM  $MgCl_2$ , 50 mM tampon MOPS pH 8,5, 10  $\mu$ L d'enzyme purifiée. 60°C sans agitation à l'étuve, 5h. Analyse par UHPLC-MS (ESI). Aires détectées à partir du m/z attendus pour les lactames correspondants.



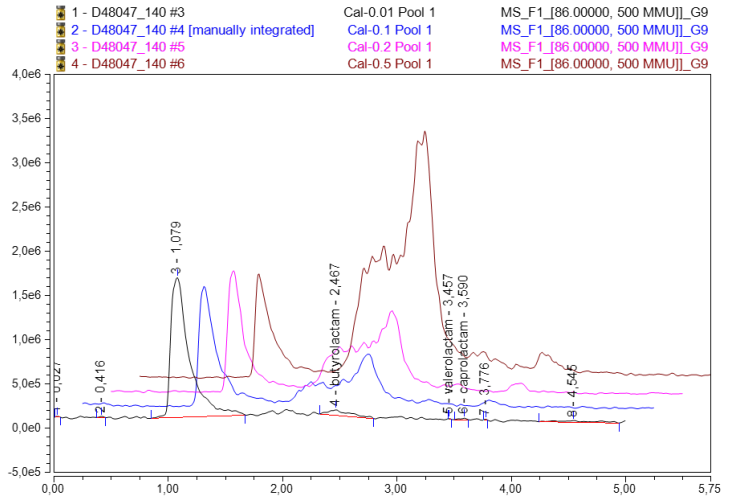
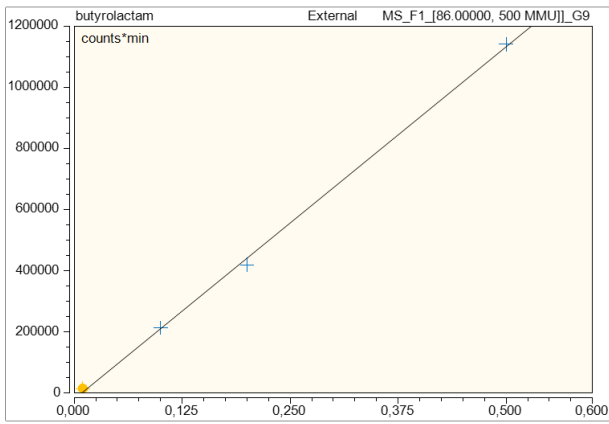
Annexe 23. (A) Détection des aires correspondant à la synthèse de N-méthylbutyrolactame (**16b**) et de N-éthylbutyrolactame (**17b**) avec les CoA ligases sélectionnées pour le pool 6. (B) Même histogramme en enlevant la forte réponse de la CoA ligase D1A9C5.

Conditions réactionnelles : 0,5 mM acide  $\omega$ -aminés, 1 mM ATP, 1 mM MnCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 50 mM tampon MOPS pH8,5, 10  $\mu$ L d'enzyme. 60°C à l'étuve sans agitation, 5h. Analyse par UHPLC-MS en mode APCI, aires détectées par extraction du m/z des lactames attendus.

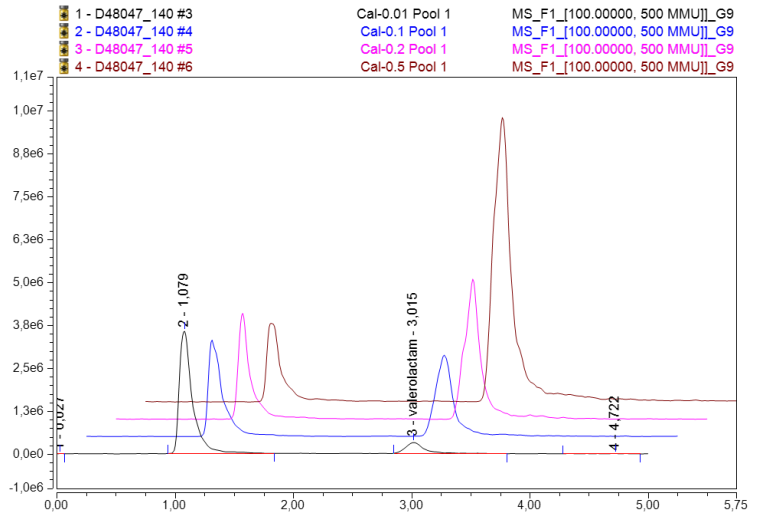
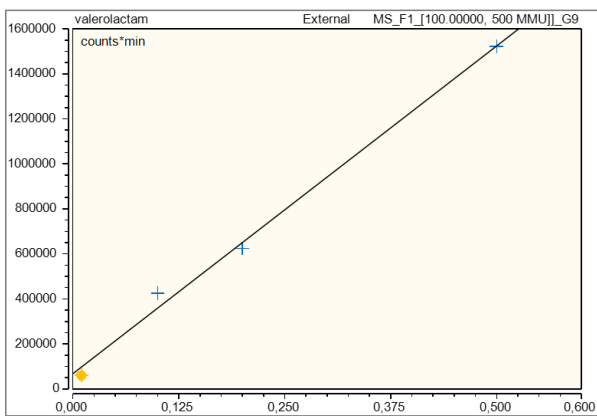


Annexe 24. Droites de calibration des standards lactames pour le criblage de la collection CoA Lig1.

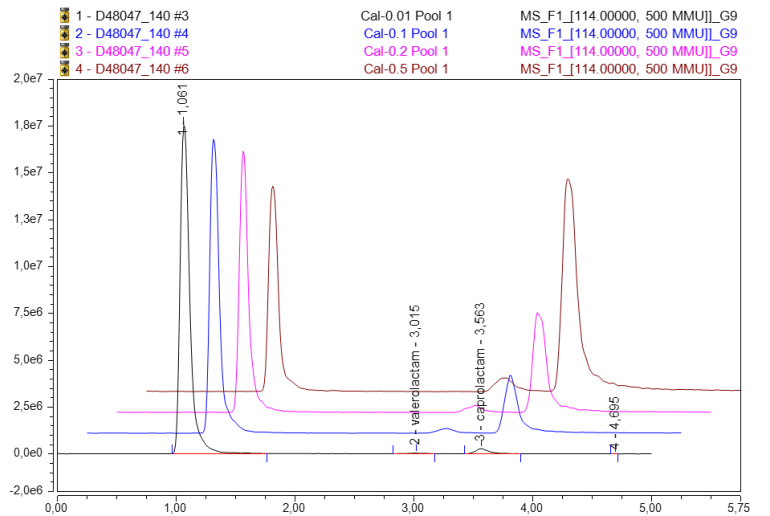
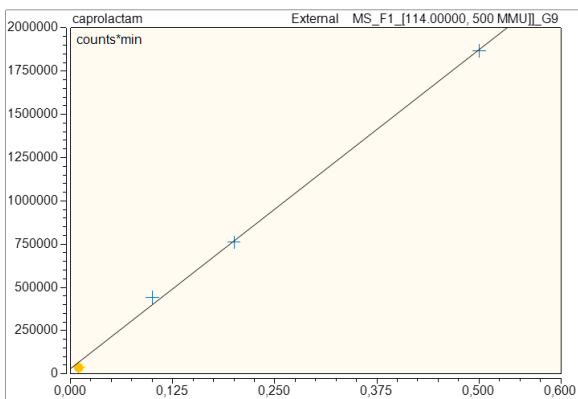
A. Butyrolactame



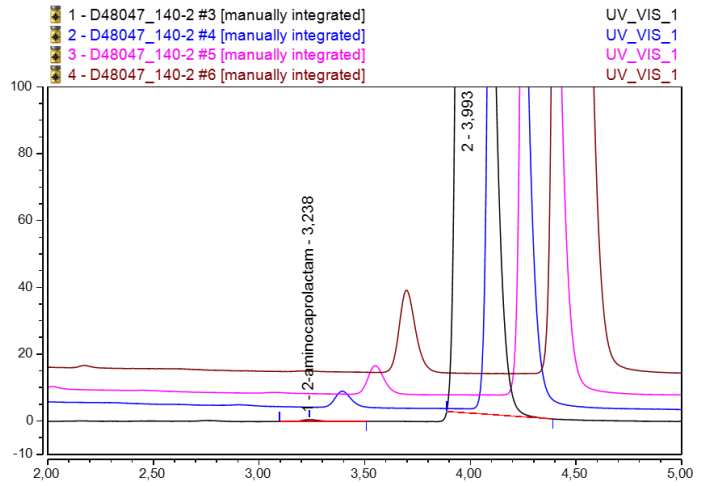
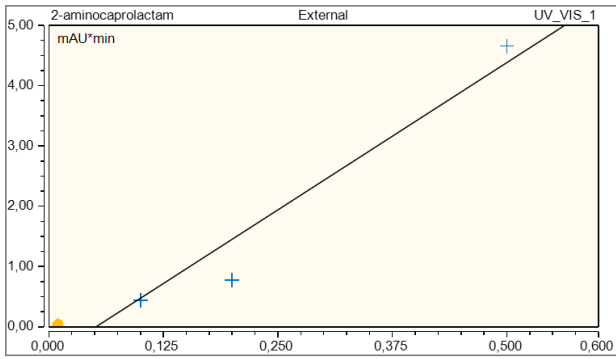
B. Valérolactame



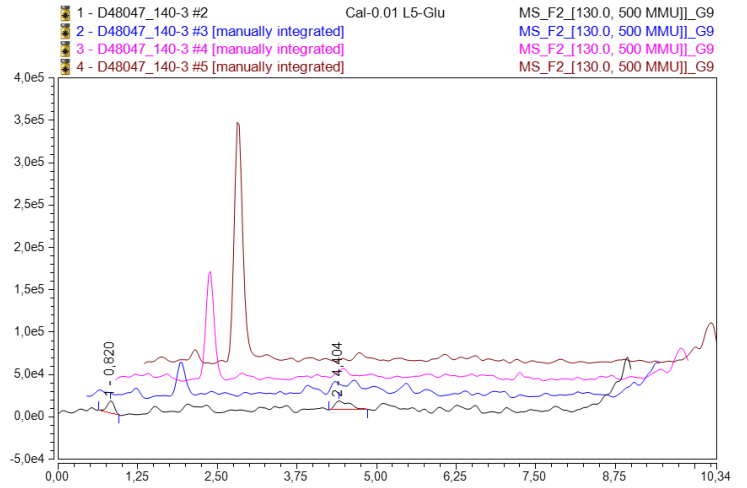
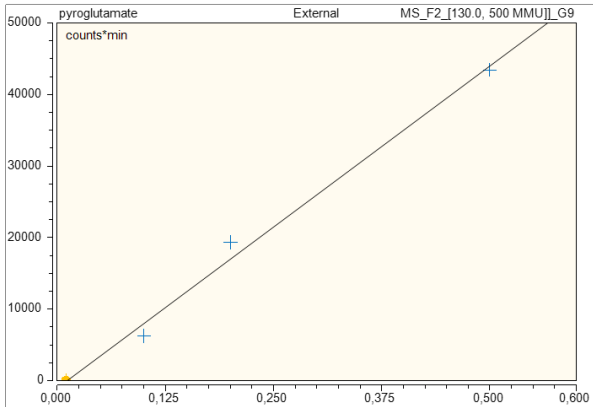
C. Caprolactame



### D. 2-aminocaprolactame



### E. Pyroglutamate



Annexe 25. Liste des CoA ligases de la collection CoA Lig2 en fonction de la CoA ligase de référence pour laquelle elles ont été rapatriées.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12		
<b>A4YDT1</b>	<b>A255</b>	<b>K256</b>	<b>W259</b>	<b>A327</b>	<b>P330</b>	<b>Y351</b>	<b>G352</b>	<b>T354</b>	<b>D435</b>	<b>F447</b>	<b>K456</b>	<b>R461</b>	<b>CoA ligase de référence</b>	<b>%ID avec l'enzyme de référence</b>
A0A1M4S4T0	C	V	S	G	P	Y	G	T	D	I	I	N	D1A9C5	32.5
A0A1M4SBK9	P	M	I	G	P	Y	G	T	D	L	I	N	D1A9C5	34.6
A0A1M4W1E8	H	A	W	G	A	W	G	T	D	I	K	W	D1A9C5	30.3
A0A1W1WZD6	M	A	G	G	P	W	G	N	D	F	I	N	D1A9C5	30.0
A0A1W1XRM1	F	A	F	G	A	L	G	T	D	F	K	R	D1A9C5	32.1
A0A510HJ46	I	G	V	G	A	F	G	T	D	I	I	N	D1A9C5	35.1
A0A7C0Y9D9	A	A	G	G	A	L	G	T	D	F	K	R	D1A9C5	31.5
A0A7J2THR7	F	H	E	G	P	W	G	T	D	F	K	T	D1A9C5	31.7
A0A7J2TIH9	I	T	L	G	A	F	G	T	D	F	K	R	D1A9C5	31.1
A5D415	C	F	S	A	P	Y	G	T	D	I	I	N	D1A9C5	32.0
B7GKW0	V	F	T	G	S	Y	G	S	D	I	I	N	D1A9C5	35.8
B7GM40	H	V	W	A	A	Y	G	T	D	I	I	N	D1A9C5	31.2
B9MQB2	C	F	V	A	P	Y	G	T	D	I	I	N	D1A9C5	31.5
C8WQA2	I	A	L	F	P	Y	G	T	D	F	K	S	D1A9C5	35.1
C8WS70	I	G	Y	G	P	Y	G	T	D	F	T	R	D1A9C5	33.8
C8WW11	I	T	T	V	P	W	W	T	D	I	N	R	D1A9C5	30.4
C8WXX0	H	V	M	G	G	Y	G	T	D	V	V	N	D1A9C5	33.1
C9R9F7	I	T	S	V	P	W	W	T	D	I	N	R	D1A9C5	30.0
C9R9F9	V	T	A	V	A	W	F	T	D	Y	K	R	D1A9C5	33.3
D3PRI0	Y	Q	R	A	T	L	A	T	D	L	K	R	D1A9C5	30.8
D3PRI1	I	T	S	V	P	W	W	T	D	I	N	R	D1A9C5	30.3
E0RP80	T	A	L	G	A	Y	G	T	D	I	V	N	D1A9C5	30.4
F8AK06	I	T	S	V	P	Y	W	T	D	I	N	R	D1A9C5	30.3
G4RKH3	I	G	G	A	S	Y	G	T	D	F	K	R	D1A9C5	32.7
H0UNC9	A	L	T	G	R	Y	G	T	D	L	I	N	D1A9C5	33.0
Q1ATG8	A	P	H	G	P	Y	G	T	D	I	N	K	D1A9C5	56.7
Q1AV80	A	P	H	G	S	Y	G	S	D	I	N	K	D1A9C5	39.7
Q47QD1	G	N	F	A	P	F	G	T	D	F	R	N	D1A9C5	30.2
S9QMV7	T	H	F	G	P	Y	G	T	D	I	I	N	D1A9C5	35.3



	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12		
<b>A4YDT1</b>	<b>A255</b>	<b>K256</b>	<b>W259</b>	<b>A327</b>	<b>P330</b>	<b>Y351</b>	<b>G352</b>	<b>T354</b>	<b>D435</b>	<b>F447</b>	<b>K456</b>	<b>R461</b>	<b>CoA ligase de référence</b>	<b>%ID avec l'enzyme de référence</b>
A0A0S3QRN4	F	H	G	G	P	Y	G	S	D	I	I	N	F2NPU2	36.3
A0A1W6VME1	N	F	S	G	K	Y	G	A	D	V	N	K	F2NPU2	30.7
A0A7C1JX37	H	V	L	G	P	Y	G	T	D	V	V	N	F2NPU2	37.1
A0A7R7YJS6	A	Y	G	A	P	I	G	T	D	Y	K	W	F2NPU2	30.1
A5D0Z4	R	G	H	M	P	Y	G	T	D	V	I	N	F2NPU2	30.2
A9WEJ4	I	H	V	S	P	Y	G	T	D	L	N	K	F2NPU2	31.8
B7GK12	H	V	L	G	P	Y	G	T	D	V	I	N	F2NPU2	33.7
C0QSN9	T	F	T	G	P	Y	G	S	D	I	I	N	F2NPU2	31.2
C8WQM8	S	A	F	G	P	Y	G	S	D	I	I	N	F2NPU2	37.0
D1A5W3	A	P	H	G	P	Y	G	T	D	I	L	N	F2NPU2	30.6
D1AE87	A	H	T	A	P	Y	A	T	D	L	I	N	F2NPU2	30.3
D1B797	A	L	T	G	R	Y	G	T	D	L	I	N	F2NPU2	32.8
D1C813	F	D	M	S	Q	Y	G	T	D	F	K	R	F2NPU2	30.2
G0EDR0	V	L	D	G	Y	Y	G	T	D	F	K	P	F2NPU2	30.5
G4RQ24	G	V	W	G	P	Y	G	T	D	I	K	S	F2NPU2	31.8
G7V9K7	A	L	T	G	K	Y	G	T	D	I	I	N	F2NPU2	32.4
Q1AT30	A	K	W	A	P	Y	G	T	D	F	L	R	F2NPU2	30.8
Q1ATL1	V	T	Q	G	P	Y	G	S	D	M	N	K	F2NPU2	33.9
Q1AUW1	H	Q	L	G	A	Y	G	T	D	I	I	N	F2NPU2	33.3
Q1AV60	A	Y	G	A	A	I	G	T	D	Y	K	W	F2NPU2	30.9

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12		
<b>A4YDT1</b>	<b>A255</b>	<b>K256</b>	<b>W259</b>	<b>A327</b>	<b>P330</b>	<b>Y351</b>	<b>G352</b>	<b>T354</b>	<b>D435</b>	<b>F447</b>	<b>K456</b>	<b>R461</b>	<b>CoA ligase de référence</b>	<b>%ID avec l'enzyme de référence</b>
A0A0S3QRM5	G	Y	N	G	P	Y	G	T	D	I	I	N	K6P2L2	35.2
A0A0S3QUA3	F	A	Y	G	A	L	G	T	D	F	K	R	K6P2L2	31.9
A0A1W1WZ81	S	S	N	G	P	Y	G	T	D	I	V	N	K6P2L2	30.9
A0A1W1X6Z8	G	K	V	G	P	Y	G	S	D	I	I	N	K6P2L2	40.0
A0A401H948	M	M	W	T	P	I	N	S	D	I	K	R	K6P2L2	30.1
A0A7C5RTB1	M	M	W	T	P	S	G	T	D	I	K	R	K6P2L2	30.6
A0B5I1	A	K	W	A	P	Y	G	T	D	F	K	R	K6P2L2	30.6
A3DBZ4	A	T	F	L	K	Y	N	I	D	I	I	N	K6P2L2	33.0
A9WAN2	V	V	S	A	P	Y	W	T	D	M	N	R	K6P2L2	30.4
A9WBE3	-	L	V	G	P	Y	G	S	D	V	I	N	K6P2L2	34.9
C8WRE0	C	F	C	G	P	Y	G	T	D	A	I	N	K6P2L2	37.4
C8WSB8	Q	T	L	A	P	F	G	S	D	F	N	K	K6P2L2	32.3
C8WSC2	Q	T	L	T	A	W	G	T	D	I	N	K	K6P2L2	31.4
C8WTN0	I	N	V	A	P	Y	G	T	D	I	N	K	K6P2L2	30.8
C8WUV5	V	T	E	G	A	Y	G	S	D	I	N	K	K6P2L2	40.9
C8WWL0	A	K	W	A	P	Y	G	T	D	F	I	T	K6P2L2	32.2
D1AB91	H	G	S	G	I	F	G	S	D	V	N	K	K6P2L2	30.6
D1AD60	P	L	S	A	P	Y	G	T	D	L	I	N	K6P2L2	30.4
D1AE83	P	L	G	T	I	F	G	T	D	V	N	K	K6P2L2	32.0
D1C7F2	M	M	W	T	P	S	G	T	D	I	K	R	K6P2L2	31.8
F8IFB8	L	T	I	V	P	W	G	T	D	V	N	R	K6P2L2	31.6
F9VNV4	I	T	I	G	P	W	G	T	D	I	N	K	K6P2L2	37.9
G4RQ25	L	V	F	V	A	Y	G	T	D	I	K	W	K6P2L2	30.8
K6PQY6	L	T	I	V	P	Y	G	T	D	V	N	R	K6P2L2	30.4
O30147	A	E	L	G	P	W	G	T	D	F	K	T	K6P2L2	32.9
Q1ATJ2	H	V	W	G	A	W	G	T	D	I	K	W	K6P2L2	30.9
Q1ATR9	F	T	L	A	P	I	G	T	D	F	I	N	K6P2L2	35.0
Q1AVA7	L	T	I	V	P	Y	G	T	D	V	N	R	K6P2L2	30.5
Q1AVB4	I	G	F	G	P	Y	G	T	D	F	T	R	K6P2L2	36.1
Q1AZ31	A	V	F	G	S	Y	G	S	D	I	N	K	K6P2L2	39.8
Q1AZ49	F	H	G	G	L	F	G	T	D	V	I	N	K6P2L2	37.8
Q97YK9	A	Y	G	A	P	I	G	T	D	Y	K	W	K6P2L2	30.5

<b>A4YDT1</b>	<b>A255</b>	<b>K256</b>	<b>W259</b>	<b>A327</b>	<b>P330</b>	<b>Y351</b>	<b>G352</b>	<b>T354</b>	<b>D435</b>	<b>F447</b>	<b>K456</b>	<b>R461</b>	<b>CoA ligase de référence</b>	<b>%ID avec l'enzyme de référence</b>
A0A1M4XB60	H	T	L	A	P	F	G	T	D	F	K	N	K6Q029	30.8
A0A1W1WZK3	F	H	Q	G	I	Y	G	T	D	V	K	N	K6Q029	41.8
A0A1W1X0Y7	S	A	N	G	Y	Y	G	T	D	L	V	N	K6Q029	36.2
A0A510HH66	L	L	P	G	S	Y	G	T	D	V	I	N	K6Q029	30.1
A9WAM1	H	T	L	G	P	Y	G	T	D	I	I	N	K6Q029	30.8
C8WWY7	F	H	G	G	P	Y	G	T	D	I	I	N	K6Q029	33.2
Q72IN3	P	F	V	G	P	Y	G	S	D	I	I	N	K6Q029	31.6
Q8ZXA2	G	L	W	G	P	Y	G	T	D	V	K	S	K6Q029	30.3
A0A0S3QRM1	P	M	V	G	A	Y	G	S	D	I	K	W	O28246	55.0
A0A1M4WWB6	-	M	C	A	P	Y	G	T	D	I	I	N	O28246	32.2
A0A1W1X5E3	P	M	V	G	A	Y	G	S	D	I	K	W	O28246	55.8
A0A7J3M0M4	A	Q	W	S	T	Y	G	T	D	I	K	W	O28246	39.2
A0A7J3M128	P	F	G	V	G	F	G	S	D	I	K	W	O28246	43.6
C8WQP5	H	V	W	G	A	Y	G	T	D	I	K	W	O28246	47.9
F8ICT4	H	I	L	A	P	Y	G	T	D	I	N	K	O28246	31.2
Q8ZVD8	P	M	V	V	A	Y	G	T	D	I	K	W	O28246	47.8
Q97XW4	G	L	Y	G	P	Y	G	T	D	I	K	W	O28246	33.5
Q9HKD4	S	L	F	G	A	Y	G	T	D	I	K	F	O28246	31.5
A0A1W6VN32	Q	T	L	I	P	W	G	T	D	I	I	N	Q3ABN2	32.5
A0A401HAF0	I	V	F	G	A	Y	G	T	D	I	K	W	Q3ABN2	35.4
A4YFI1	V	T	V	G	P	W	G	T	D	V	V	K	Q3ABN2	31.9
D1A2E7	F	A	C	G	K	F	G	A	D	V	N	K	Q3ABN2	32.1
Q1AUD1	-	M	C	A	P	Y	G	T	D	I	I	N	Q3ABN2	53.6
Q976X7	H	I	W	G	A	W	G	T	D	I	K	W	Q3ABN2	31.9
Q97UF6	I	T	I	G	P	W	G	T	D	I	N	K	Q3ABN2	31.1
A0A1M4VJX1	H	A	W	G	Q	W	G	T	D	I	K	W	Q97UW7	34.1
A0A1W1WZV0	A	Q	W	G	E	Y	G	T	D	I	K	W	Q97UW7	32.7
F8IDN4	P	F	I	G	Q	Y	G	S	D	V	K	W	Q97UW7	34.3
F8IJJ8	H	V	W	A	P	Y	G	T	D	I	I	N	Q97UW7	33.1
O29585	P	Q	I	I	D	-	Y	G	D	I	K	W	Q97UW7	36.6
Q97AY9	P	L	F	G	I	T	D	L	D	I	K	F	Q97UW7	37.4

Annexe 26. Comparaison des résidus constituant le site actif des 20 CoA ligases sélectionnées pour former la collection réduite de CoA Lig2. Légende : résidus à chaîne latérale courte et hydrophobe en rouge (sauf Y), résidus acides en bleu, résidus basique en magenta et résidus à chaîne latérale polaire et résidu glycine en vert.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
<b>A4YDT1</b>	<b>A255</b>	<b>K256</b>	<b>W259</b>	<b>A327</b>	<b>P330</b>	<b>Y351</b>	<b>G352</b>	<b>T354</b>	<b>D435</b>	<b>F447</b>	<b>K456</b>	<b>R461</b>
G0EDR0	V	L	D	G	Y	Y	G	T	D	F	K	P
A0A7J2THR7	F	H	E	G	P	W	G	T	D	F	K	T
C8WUV5	V	T	E	G	A	Y	G	S	D	I	N	K
G4RKH3	I	G	G	A	S	Y	G	T	D	F	K	R
A0A7R7YJ56	A	Y	G	A	P	I	G	T	D	Y	K	W
Q1AZ49	F	H	G	G	L	F	G	T	D	V	I	N
C8WWY7	F	H	G	G	P	Y	G	T	D	I	I	N
A0A7J3M128	P	F	G	V	G	F	G	S	D	I	K	W
F8IFB8	L	T	I	V	P	W	G	T	D	V	N	R
Q97UF6	I	T	I	G	P	W	G	T	D	I	N	K
O29585	P	Q	I	I	D	-	Y	G	D	I	K	W
Q1AUW1	H	Q	L	G	A	Y	G	T	D	I	I	N
D1C813	F	D	M	S	Q	Y	G	T	D	F	K	R
A0A0S3QRM5	G	Y	N	G	P	Y	G	T	D	I	I	N
A0A1W1X0Y7	S	A	N	G	Y	Y	G	T	D	L	V	N
Q1ATL1	V	T	Q	G	P	Y	G	S	D	M	N	K
C9R9F7	I	T	S	V	P	W	W	T	D	I	N	R
A9WAN2	V	V	S	A	P	Y	W	T	D	M	N	R
G7V9K7	A	L	T	G	K	Y	G	T	D	I	I	N
A0A1W1X6Z8	G	K	V	G	P	Y	G	S	D	I	I	N
<b>D1A9C5</b>	I	T	I	G	P	Y	G	T	D	L	N	K
<b>F2NPU2</b>	F	H	L	G	P	Y	G	T	D	I	I	N
<b>K6P2L2</b>	-	L	I	G	P	Y	G	T	D	V	I	K
<b>K6Q029</b>	S	A	H	G	I	Y	G	S	D	V	N	L
<b>O28246</b>	P	M	V	G	K	G	M	E	D	I	K	W
<b>Q3ABN2</b>	-	M	C	A	A	Y	G	T	D	I	I	N
<b>Q97UW7</b>	P	M	V	I	S	Y	G	S	D	I	K	F

Annexe 27. Plan de plaque des gènes des 20 CoA ligases de la collection réduite.

Well Location	Name	Insert Length	Yield (ng)
A1	A0A0S3QRM5-opt	1722	815
A2	C8WWY7.opt	1686	836
A3	Q1ATL1.opt	1698	813
B1	A0A1W1X0Y7.opt	1575	851
B2	C9R9F7.opt	2007	1054
B3	Q1AUW1.opt	1560	796
C1	A0A1W1X6Z8.opt	1728	851
C2	D1C813.opt	1611	1442
C3	Q1AZ49.opt	1599	1237
D1	A0A7J2THR7.opt	1689	802
D2	F8IFB8.opt	2010	1098
D3	Q97UF6.opt	1578	1448
E1	A0A7J3M128.opt	1671	1217
E2	G0EDR0.opt	1524	1370
F1	A0A7R7YJS6.opt	1587	1062
F2	G4RKH3.opt	1734	698
G1	A9WAN2.opt	1938	711
G2	G7V9K7.opt	1512	487
H1	C8WUV5.opt	1716	922
H2	O29585.opt	1551	835

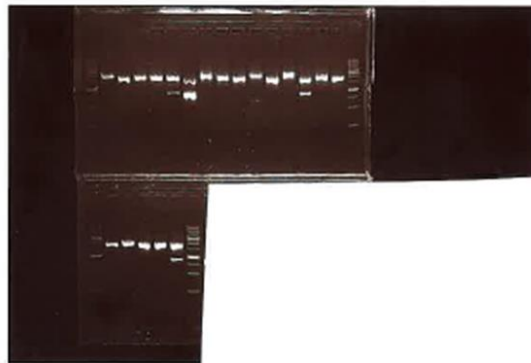
Annexe 28. Gel d'agarose de la vérification de l'amplification PCR des gènes des 20 CoA ligases sélectionnées.  
Légende : de gauche à droite, les puits correspondent à A1 à H1 puis de A2 à H2 et A3 à D3.

**Après amplification  
par PCR**



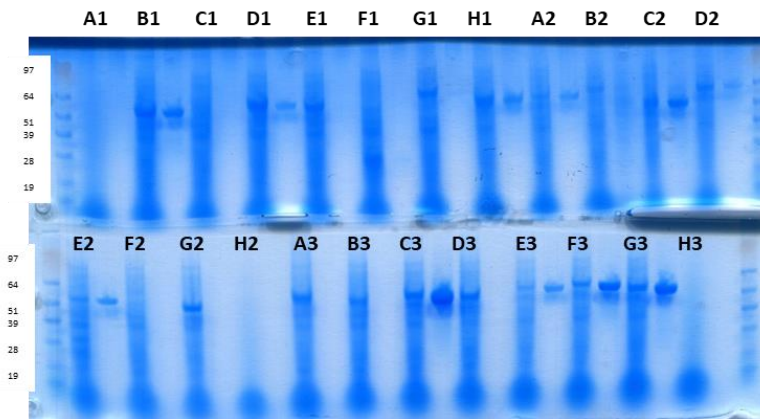
Colonnes 2 et 3

**Après extraction du  
plasmide des cellules  
NovaBlue**

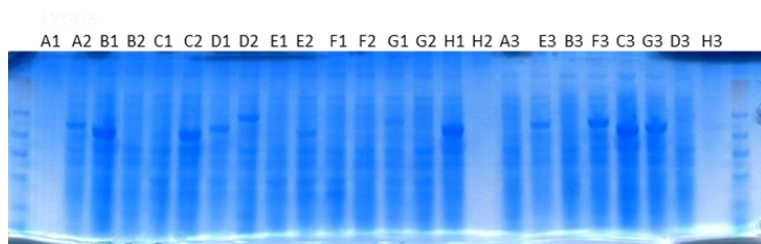


Annexe 29. Gels SDS-PAGE et concentration des CoA ligases de la production de la collection réduite de la deuxième collection CoA Lig2.

### Induction (1<sup>ère</sup> colonne) et enzymes purifiées (2<sup>ème</sup> colonne)



### Lysats



### Concentration des CoA ligases purifiées

Wells	Uniprot	mg/ml	kDa
A01	A0A0S3QRM5-opt	-0,011	63,181
B01	A0A1W1X0Y7.opt	0,721	56,552
C01	A0A1W1X6Z8.opt	0,003	62,936
D01	A0A7J2THR7.opt	0,223	62,42
E01	A0A7J3M128.opt	0,023	62,616
F01	A0A7R7YJS6.opt	0,076	58,475
G01	A9WAN2.opt	0,033	70,341
H01	C8WUV5.opt	0,472	61,864
A02	C8WWY7.opt	0,204	61,584
B02	C9R9F7.opt	0,027	74,09
C02	D1C813.opt	0,66	57,188
D02	F8IFB8.opt	0,173	73,132
E02	G0EDR0.opt	0,347	54,648
F02	G4RKH3.opt	0,018	62,745
G02	G7V9K7.opt	0,031	53,955
H02	O29585.opt	-0,011	56,154
A03	Q1ATL1.opt	-0,023	60,968
B03	Q1AUW1.opt	0,011	54,17
C03	Q1AZ49.opt	2,149	56,817
D03	Q97UF6.opt	0,017	58,913
E03	K6P2L2	0,351	62,716
F03	K6Q029	1,06	60,823
G03	D1A9C5	1,287	59,909
H03	vide	Blanc	

\*: obtenu en PCR mais problème lors de la production





**Titre :** Approche biocatalytique pour la synthèse de lactames

**Mots clés :** biocatalyse, amide, lactames, CoA ligase, structure 3D protéine

**Résumé :** La biotechnologie est maintenant bien établie dans le paysage de l'industrie chimique. Dans le domaine de la biocatalyse, il existe de nombreux exemples de mise en œuvre réussie d'étapes enzymatiques dans des procédés de synthèse ou de développement de voies enzymatiques complètes, en particulier dans l'industrie pharmaceutique.

Les lactames sont des composés qui se retrouvent dans de nombreux produits naturels et synthétiques, tels que les ingrédients pharmaceutiques actifs (API) et sont des précurseurs dans la chimie des polymères. Pour accéder aux lactames à partir de squelettes simples d'acides  $\omega$ -aminés, la synthèse nécessite un processus en deux étapes : tout d'abord, l'activation de la fonction acide, puis, l'addition intramolécule d'amine nucléophile. Pour les petits lactames (5-7 chaînons), la cyclisation intramolécule se produit spontanément une fois l'acide activé.

Dans le cadre de nos travaux sur des voies d'accès enzymatiques aux amides, nous avons récemment mis au point une synthèse impliquant une activation par CoA ligases, en absence de CoASH, et utilisant uniquement leur aptitude d'activation de l'acide par adénylation par l'ATP, introduit à seulement 5% molaire grâce à la mise en place d'un système de régénération.

Jusqu'à présent, le type de lactames obtenus par biocatalyse est limité aux cycles nus, dépourvus de fonctions. L'objectif du projet est de disposer d'enzymes permettant la synthèse par approche biocatalytique de lactames fonctionnalisés de 5 à 7 chaînons. Pour cela deux types d'approche seront menés :

1) Exploration de la biodiversité. Une collection d'enzymes de type CoA ligases, construite par une approche génomique basée sur l'identité de séquences, est disponible au laboratoire. Cette collection, constituée d'environ 250 enzymes, sera criblée sur des substrats précurseurs de lactames d'intérêt.

2) Conception rationnelle. En collaboration avec l'équipe des Pr. Dick Janssen et Dr. Andy-Mark Thunnissen (Université de Groningen, Pays-Bas), les données structurales seront utilisées pour générer des bibliothèques ciblées. Des expériences de docking avec les substrats seront réalisées pour cibler les mutations.

**Title :** Biocatalytic approach for lactam synthesis

**Keywords :** biocatalysis, amide, lactams, CoA ligase, 3D structure protein

**Abstract:** Biotechnology is now a well-established part of the chemical industry landscape. In the field of biocatalysis, there are numerous examples of the successful implementation of enzymatic steps in synthesis processes or the development of complete enzymatic pathways, particularly in the pharmaceutical industry.

Lactams are compounds found in many natural and synthetic products, such as active pharmaceutical ingredients (APIs), and are precursors in polymer chemistry. To access lactams from simple  $\omega$ -amino acid backbones, synthesis requires a two-step process: first, activation of the acid function, followed by intramolecular nucleophilic amine addition. For small lactams (5-7 members), intramolecular cyclization occurs spontaneously once the acid has been activated.

As part of our work on enzymatic access routes to amides, we recently developed a synthesis involving activation by CoA ligases, in the absence of CoASH, and using only their ability to activate the acid by adenylation with ATP, introduced at only 5 mol% thanks to the implementation of a regeneration system.

Until now, the type of lactams obtained by biocatalysis has been limited to bare, functionless rings. The aim of the project is to provide enzymes for the biocatalytic synthesis of 5- to 7-membered functionalized lactams. To achieve this, two approaches will be pursued:

1) Exploration of biodiversity. A collection of CoA ligase enzymes, built up using a genomic approach based on sequence identity, is available in the laboratory. This collection, comprising around 250 enzymes, will be screened on lactam precursor substrates of interest. Adenylation enzymes other than CoA ligases have been identified and will also be tested.

2) Rational design. In collaboration with the team of Prof. Dick Janssen and Dr. Andy-Mark Thunnissen (University of Groningen, Netherlands), structural data will be used to generate targeted libraries. Substrate docking experiments will be carried out to target mutations.