UNIVERSITÉ DE LORRAINE | BIBLIOTHÈQUES UNIVERSITAIRES

AVERTISSEMENT
DROIT D'AUTEUR

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document. Toute contrefaçon, plagiat ou reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

UNIVERSITÉ DE LORRAINE
34 COURS LÉOPOLD — CS 25233
54052 NANCY CEDEX
WWW.UNIV-LORRAINE.FR

Ecole doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEURE DE L'UNIVERSITE DE LORRAINE

Mention : « Sciences de la Vie et de la Santé »

par **Magali TIRADO**

Contributions à une ingénierie efficace des polycétide synthases

Le 09 décembre 2025

Membres du jury :

Rapporteuses :	**M^me^ Cécile BON**	**Maître de conférence, Université de Toulouse, Toulouse (France)**
	M^me^ Sylvie LAUTRU	**Directrice de recherche, CNRS, Gif-sur-Yvette (France)**
Examinateur·rices :	**M. Stéphane DELAUNAY**	**Professeur, Université de Lorraine, Nancy (France), président du jury**
	M^me^ Joleen MASSCHELEIN	**Associate professor, KU Leuven, Leuven (Belgique)**
	M. Christophe JACOB	**Professeur, Université de Lorraine, Nancy, directeur de thèse (France)**
	M. Bertrand AIGLE	**Professeur, Université de Lorraine, Nancy, co-directeur de thèse (France)**
Membre invitée :	**M^me^ Kira WEISSMAN**	**Professeure, Université de Lorraine, Nancy (France)**

UMR 7365 CNRS-UL IMoPA, Biopôle de l'Université de Lorraine, Campus Brabois-Santé, 9 Avenue de la Forêt de Haye, BP 20199, 54505 Vandœuvre-lès-Nancy

UMR 1128 INRAE-UL DynAMic, Campus Faculté des Sciences et Technologies, Boulevard des Aiguillettes, BP 70239, 54506 Vandœuvre-lès-Nancy

Remerciements

Je souhaite dans un premier temps remercier Pr. Christophe Jacob et Pr. Bertrand Aigle pour m'avoir offert l'opportunité d'effectuer cette thèse sous leur direction, leur confiance, leur aide et leur soutien tout au long de ces trois années. Je remercie particulièrement le Pr. Bertrand Aigle de m'avoir donné l'opportunité en 2017 d'intégrer le cursus master en ingénierie de Nancy, à l'époque sous sa direction, et de vivre toutes les opportunités que cela a pu m'apporter, jusqu'à la fin de cette thèse.

Merci également à la Pr. Kira Weissman de m'avoir accueillie au sein de son équipe S2BEAM du laboratoire IMoPA et de m'avoir aidée, écoutée et soutenue durant ces trois années d'apprentissage de la recherche. Merci pour ton implication dans les différents projets de cette thèse et pour tes idées et analyses tout au long de ces trois années.

Je remercie également chaleureusement la Dr. Cécile Bon et la Dr. Sylvie Lautru d'avoir accepté de rapporter cette thèse. Vos regards pluridisciplinaires sont à n'en pas douter un atout pour ces travaux.

Mes remerciements sont également adressés à Pr. Stéphane Delaunay et Dr. Joleen Masschelein, pour avoir accepté d'examiner ces travaux de thèse. Vos regards croisés sur ces travaux seront un atout indéniable.

Je souhaite à nouveau remercier la Dr. Joleen Masschelein, ainsi que le Pr. Alain Hehn, d'avoir fait partie durant ces trois années de mon comité de suivi individuel. Nos échanges annuels ont pour moi été très enrichissants et motivants.

Ma gratitude va également envers au Pr. Jean-Yves Jouzeau puis au Pr. Bruno Charpentier, directeurs successifs du laboratoire IMoPA, de m'avoir accueillie au sein de cette structure. Je remercie également à nouveau le Pr. Bertrand Aigle, directeur de l'unité DynAMic, pour m'avoir permis d'intégrer ce laboratoire.

Gracias al Dr. Ángel Manteca por darme la bienvenida en su laboratorio. Las horas que pasé trabajando bajo el microscopío han sido mis favoritas de estos tres años. Gracias a la Dra. Marta Alonso Guervós por su amabilidad y su ayuda para processar las imágenes de microscopía. Disfruté mucho hablando de ciencia con ambos. Gracias también a la Dra. Gemma Fernández-García, Dr. Sergio Alonso Fernández y Paula Valdés Chiara por recibirme en su oficina durante estos dos meses.

Je remercie également tous les membres de l'équipe S2BEAM du laboratoire IMoPA. Merci à Dr. Arnaud Gruez et Dr. Benjamin Chagot pour toutes les discussions au fil de ces trois années qui m'ont autant nourrie scientifiquement que personnellement.

Je remercie également M^{me} Sabrina Collin, ingénieure d'étude au sein de l'équipe, pour son soutien et son écoute à toute épreuve. Merci d'avoir été là durant ces trois années et, au-delà d'avoir été une collègue remarquable, d'être devenue une amie.

Merci également à Hien Bonnin Tran, Christelle Kondratow et Ghislaine Charpentier pour leur aide dans toutes les démarches administratives compliquées liées à ces trois années. Merci aussi à Christophe Muller, sans qui le laboratoire IMoPA s'effondrerait.

Ma reconnaissance s'adresse également à tous les membres de l'équipe StrAda du laboratoire DynAMic. Merci aux Drs. Annabelle Thibessard, Claire Bertrand et Cyril Bontemps pour ces merveilleux moments partagés lors du 20ème symposium sur la biologie des Actinomycètes. Un merci particulier à Laurence Hôtel, assistante ingénieur, pour sa gentillesse et tout ce qu'elle a apporté à ce projet. Sans toi, ce projet de thèse aurait été bien plus lourd à porter.

Je remercie aussi tous mes collègues doctorant·es et technicien·nes, avec qui j'ai pu passer d'inoubliables moments et qui resteront des ami·es pour la vie. Merci à Jean-Malo de m'avoir accueillie dès mon stage de Master 2 est de m'avoir formé avec la rigueur et le sérieux qui te définissent si bien, et merci d'être devenu un ami précieux et un allié de taille durant ces trois années. Merci à Svenja, Rime, Arnaud, Romain, Serge et Andréa d'avoir si bien animé notre bureau. Merci à Elise de m'avoir aussi bien accueillie lorsque je descendais la colline. Merci à Jean-Malo et Delphine pour ces nombreuses soirées passées en votre compagnie. Merci aussi à Ghita, Roméo, Océane, Céline, Anne, Lauralie, Lucas, Amine Kévin, Quentin, Jade, Camille, Julie et Tiffany.

Merci également à Agathe Lemée, ma compagne dans cette aventure qu'a été la création et la gestion de notre club ORION *What health?*. Je remercie aussi toutes et tous les doctorant·es qui ont fait partie de cette aventure : Ghita, Arnaud, Martin, Antoine et Walid. Merci enfin à tous·tes les étudiant·es qui nous on rejoints et permis de monter tous ces beaux projets.

Je remercie aussi tous les membres du conseil de l'école doctorale BioSE et du Conseil Lorraine des Ecoles Doctorales (CLED) de m'avoir accueillie et permis de représenter au mieux les doctorant·es de l'Université de Lorraine.

Je remercie Lorraine Université d'Excellence (LUE) pour le financement de ces travaux de thèse, ainsi que pour la bourse *Doctor, Explore and Achieve More* (DrEAM) qui m'a permis d'effectuer une mobilité de deux mois à l'Université d'Oviedo, en Espagne.

Merci également à Lyne Hamed, étudiante que j'ai eu l'honneur d'encadrer lors de son stage de Licence 3. T'accompagner durant ces deux mois de stages m'ont permis de découvrir ma réelle passion : accompagner, écouter et aider les autres à s'épanouir et grandir. Merci pour ta gentillesse et ta confiance qui ont su me redonner de l'énergie.

J'exprime aussi ma gratitude envers mes plus proches ami·es, sans lesquel·les ces huit années à Nancy auraient été bien plus longues et pénibles. Merci à Margot, Mathilde, Lilian, Clarisse et Margaux.

Je ne serais pas en train d'écrire ces lignes si ma famille ne m'avait pas autant soutenue et épaulée. Merci à mes grands-parents maternels de m'avoir toujours accueillie à bras ouverts durant ces huit dernières années, pour les bons petits plats et le réconfort. Merci également à mes grands-parents paternels, qui ne pourront malheureusement jamais voir jusqu'où j'ai pu aller grâce à leur soutien et leur présence passée. Merci à mes parents de m'avoir aidée à prendre mon envol, et de m'avoir toujours soutenu malgré la distance. Merci à ma petite sœur, qui me complète à la perfection et sait me ramener à la réalité lorsque mon imagination m'emmène loin de la Terre ferme. Merci à vous d'avoir cru en moi, surtout lorsque je n'y croyais plus moi-même.

Je ressens une gratitude particulière pour Dr. Claire Rodrigues, docteure en psychologie et praticienne à Nancy. Merci de m'avoir écoutée et épaulée durant ces trois années. Votre cabinet a pour moi été un refuge dans les moments les plus compliqués.

Enfin, je remercie celui qui tient bon au milieu de tout ce tourbillon depuis plus d'un an maintenant, Martin. Merci d'écouter mes doutes et consoler mes peines, tu as su m'apaiser et m'aider à traverser cette longue épreuve avec plus de sérenité et de calme (ce qui n'était pas chose aisée). Je t'aime.

« *And if I fly, or if I fall*

Least I can say, I gave it all »

RuPaul

R	W	P	A	U	Q	J	E	R	B	A	M	P	A	E
S	T	R	E	P	T	O	M	Y	C	E	S	R	S	C
S	L	O	C	A	L	I	S	A	T	I	O	N	C	O
T	P	F	F	T	F	K	W	G	G	Y	E	F	I	E
A	O	F	B	N	C	R	U	G	K	I	N	L	E	L
M	L	H	R	D	I	A	Q	C	P	Z	A	K	N	I
B	Y	M	A	S	Q	P	M	O	G	Y	C	Y	C	M
O	C	O	P	K	N	D	C	P	X	X	D	E	E	Y
M	E	M	A	W	F	S	I	O	C	A	I	J	S	C
Y	T	J	M	C	O	J	K	H	U	F	Z	L	M	I
C	I	H	Y	R	S	F	V	E	L	S	R	Y	O	N
I	D	H	C	Z	F	U	S	C	G	P	C	Q	F	E
N	E	I	I	K	V	R	U	S	H	J	W	O	W	Z
E	M	O	N	F	D	Y	Q	I	Q	E	I	A	D	P
S	J	T	E	I	N	T	E	R	A	C	T	I	O	N

STREPTOMYCES
STAMBOMYCINES
LOCALISATION
MICROSCOPIE
COELIMYCINE
SCIENCES
RAPAMYCINE
INTERACTION
POLYCETIDE

Table des matières

Table des matières ..... **xi**

Table des figures ..... **xvii**

Table des tableaux ..... **xxv**

Liste des abréviations ..... **xxix**

Etat de l'art **1**

I. Les produits naturels 3

I.A. Les classes et fonctions de produits naturels 3

I.B. Les prémices de la découverte des produits naturels et leur utilisation 5

I.C. Le plus grand producteur de produits naturels d'intérêt : les bactéries du genre *Streptomyces* 7

I.D. A la recherche de nouvelles molécules thérapeutiques 12

II. Les polycétide synthases 19

II.A. Les polycétide synthases de type II 20

II.B. Les polycétide synthases de type III 21

II.C. Les polycétides synthases de type I 23

III. Structure, mécanisme d'action et évolution des domaines de polycétide synthases 31

III.A. Le domaine porteur d'acyle (ACP) 31

III.B. Le domaine acyltransférase (AT) 32

III.C. Le domaine cétosynthase (KS) 34

III.D. Le domaine cétoréductase (KR) 36

III.E. Le domaine déshydratase (DH) 39

III.F. Le domaine énoyl-réductase (ER) 40

III.G. Le domaine thioestérase (TE) 41

IV. Les interfaces d'interactions entre les sous-unités de polycétide synthases 43

IV.A. Les interfaces d'interaction des PKS de type *cis*-AT 44

IV.B. Les interfaces d'interactions des PKS de type *trans*-AT 46

V. L'organisation des méga-enzymes au sein des cellules bactériennes 48

V.A. L'organisation des PKS *trans*-AT 48

V.B. L'organisation des PKS *cis*-AT 53

V.C. La proximité entre les méga-complexes enzymatiques et la membrane bactérienne 54
VI. L'ingénierie des polycétide synthases pour la production de nouvelles molécules .. 55
VI.A. La modification des groupements portés par les polycétides connus 55
VI.B. La modification de la taille des polycétides 66
VI.C. L'ingénierie des polycétide synthases pour l'optimisation des rendements de production 73
Objectifs 85
Matériels et méthodes 91
I. Souches bactériennes, conditions de culture et plasmides 93
I.A. Souches bactériennes 93
I.B. Milieux de cultures 95
I.C. Vecteurs plasmidiques 96
I.D. Enzymes 97
II. Extraction de matériel génétique 98
II.A. Extraction d'ADN génomique bactérien 98
II.B. Mini-préparation de matériel plasmidique 99
II.C. Extraction de matériel génétique sur gel d'agarose 99
III. Construction des plasmides 99
III.A. Conception, amplification et assemblage des éléments nécessaires à l'édition des génomes de *Streptomyces* spp. 99
III.B. Clonage des matrices et ARN guides au sein des vecteurs d'intérêts pour l'ingénierie génétique de *Streptomyces* spp. 102
III.C. Conception et clonage des séquences codant les domaines de docking ... 104
IV. Préparation et transformation de cellules bactériennes 105
IV.A. Préparation et transformation des souches d'*Escherichia coli* chimio-compétentes 105
IV.B. Préparation et transformation des souches d'*Escherichia coli* électro-compétentes 106
IV.C. Conjugaison intergénérique des bactéries du genre *Streptomyces* 107

V. Production et purification de protéines recombinantes 108
V.A. Production des protéines recombinantes 109
V.B. Purification des protéines recombinantes 109
VI. Caractérisation des interactions entre les domaines de docking 112
VI.A. Analyse de l'intégrité des domaines purifiés 112
VI.B. Etude des interactions entre les domaines de docking par titrage calorimétrique isotherme 113
VII. Localisation de la coelimycine synthase chez *S. coelicolor* par microscopie confocale 114
VII.A. Préparation des échantillons 114
VII.B. Acquisition des images par microscopie 114
VII.C. Traitement des clichés obtenus par microscopie confocale 115
Résultats 117
I. Effet de la scission de polycétide synthases : focus sur la rapamycine synthase et la stambomycine synthase 119
I.A. Etude de la rapamycine synthase : constructions plasmidiques et caractérisation biophysique des domaines de docking 119
I.B. Etude de la stambomycine synthase : construction des souches mutées 133
II. Localisation *in situ* de la PKS *cis*-AT coelimycine synthase 138
II.A. Contexte et résultats antérieurs à ce projet de thèse 138
II.B. Ingénierie génétique de *Streptomyces coelicolor* 140
II.C. Observation de la localisation de Cpk par microscopie confocale 145
Discussion 155
I. Investiguer l'effet de la scission sur les rendements de production 157
I.A. Etude de la rapamycine synthase 157
I.B. Etude de la stambomycine synthase 161
II. Etudier l'organisation d'une PKS *cis*-AT au sein de son hôte bactérien 163
Annexe xxxvii
Notes bibliographiques xxxiii
Résumés en français et anglais xlix

Table des figures

Figure 1. Répartition de la production de produits naturels connus par les différentes branches du vivant 3
Figure 2. Structures de membres représentatifs des familles de produits naturels 4
Figure 3. Affiche publicitaire pour l'utilisation de cachets de cocaïne contre les maux de dents aux Etats-Unis 4
Figure 4. Cycle de reproduction des bactéries du genre *Streptomyces* 8
Figure 5. Régulation de la production de métabolites spécialisés chez *Streptomyces* 9
Figure 6. Représentation schématique du chromosome de *Streptomyces coelicolor* et localisation de ses quatre principaux clusters de gènes biosynthétiques 10
Figure 7. Cluster de gènes lié à la production de la coelimycine chez *Streptomyces coelicolor* 11
Figure 8. Régulation de l'expression des quatre principaux antibiotiques produit par *Streptomyces coelicolor* 12
Figure 9. Nombre de médicaments approuvés aux Etats-Unis entre 1980 et 2014 13
Figure 10. Découverte du cluster de gènes biosynthétique lié à la production des stambomycines chez *Streptomyces ambofaciens* 15
Figure 11. Modification de l'érythromycine A par hémisynthèse 18
Figure 12. Structure de l'halichondrine B et de l'éribuline 18
Figure 13. Exemples de polycétides 20
Figure 14. Polycétide synthase de type II 21
Figure 15. Polycétide synthase de type III 22
Figure 16. Synthèse de polycétides par des polycétide synthases de type III 22
Figure 17. Exemples de polycétide synthases de type I itératives de champignons 24
Figure 18. Voie de biosynthèse proposée de la coelimycine chez *Streptomyces coelicolor* . 26
Figure 19. Voie de biosynthèse de la bacillaene par une PKS *trans*-AT retrouvée chez *Bacillus subtilis* 28
Figure 20. Cassette de β-méthylation de la bacillaene synthase 29
Figure 21. Représentation des deux premières sous-unités de l'azalomycine synthase 30
Figure 22. Alignement des structures de trois domaines porteurs d'acyl (ACP) et mécanisme de greffe du bras 4′-phosphopantéthéine (P-Pant) au domaine ACP 31
Figure 23. Structure d'un domaine acyltransférase (AT) de la splénocine synthase et mécanisme de transfert de l'unité acyl-coA du domaine AT au domaine porteur d'acyle 32
Figure 24. Arbre phylogénétique de domaines acyltransférases et représentation schématique des séquences pouvant permettre la recombinaison homologue de ces domaines pour la rapamycine synthase 34
Figure 25. Structure de domaines cétosynthase issus de polycétide synthases 35
Figure 26. Mécanisme catalytique des domaines cétosynthase 36

Figure 27. Structure d'un domaine cétoréductase de l'amphotéricine synthase 37
Figure 28. Mécanisme d'action des domaines cétoréductase 38
Figure 29. Structure et organisation du site actif du domaine déshydratase de la fluvirucin B1 synthase 39
Figure 30. Structure d'un domaine énoyl-réductase et représentation schématique de la spécificité catalytique de ces domaines 40
Figure 31. Structure de domaines thioestérase de polycétide synthases 42
Figure 32. Mécanisme catalytique proposé des domaines thioestérase 43
Figure 33. Structure de domaines de docking de type 1 44
Figure 34. Structure de domaines de docking de type 2 46
Figure 35. Structure des domaines de docking présents entre les sous-unités VirA et VirFG de la PKS *trans*-AT de la virginiamycine 47
Figure 36. Modèle d'un bidomaine cétosynthase – déhydratase, interagissant par le biais d'un domaine déshydratase docking 48
Figure 37. Localisation de la bacillaene synthase chez *Bacillus subtilis* par microscopie à fluorescence 50
Figure 38. Observation par microscopie électronique à transmission des souches sauvage et mutées de *Bacillus subtilis* 50
Figure 39. Analyse du domaine *Laterally-INteracting Ketosynthase Sequences* au sein d'un domaine cétosynthase de la bacillaene synthase 51
Figure 40. Macro-organisation d'un bi-module de polycétide synthase *trans*-acyltransférase 52
Figure 41. Alignements de séquences peptidiques et de structures secondaires de domaines cétosynthase et acyltransférase de polycétide synthase 53
Figure 42. Potentielles cibles d'ingénierie au sein des PKS pour modifier la structure des polycétides 55
Figure 43. Ingénierie de la 6-désoxyérythronolide B synthase par modification des domaines acyltransférase 56
Figure 44. Ingénierie de la polycétide synthase minimale DEBS1-TE 60
Figure 45. Représentation schématique de la répartition en modules et en unités échangeables de la 6-désoxyérythronolide B synthase (DEBS) 61
Figure 46. Ingénierie de la pikromycine synthase par échange de modules 62
Figure 47. Stratégies d'ingénierie d'une polycétide synthase pour modifier la stéréochimie de la molécule produite 65
Figure 48. Représentation schématique de la PKS des stambomycines 67
Figure 49. Structure des stambomycines et des mini-stambomycines obtenues par ingénierie génétique 68

Figure 50. Représentation des interfaces générées entre les sous-unités 4 et 9 de la PKS des stambomycines et rendements de production des différentes mini-stambomycines 69
Figure 51. Représentation schématique de la rapamycine synthase 71
Figure 52. Analyse des souches de *Streptomyces hygroscopicus* mutées pour le cluster de gènes codant la rapamycine synthase 72
Figure 53. Nombre de modules au sein de sous-unités de polycétide synthases répertoriées au sein de la base de données MIBiG 73
Figure 54. Ingénierie du domaine cétoréductase de la pikromycine synthase 75
Figure 55. Structure et diagramme en hélice de la paire de *Synthetic Zippers* 3:4 76
Figure 56. Structure du couple SpyCatcher:SpyTag 76
Figure 57. Représentation schématique des interfaces générées entre VemG et VemH 78
Figure 58. Scission de la sous-unité VemG de la vénémycine synthase 79
Figure 59. Scission de la butényl-spinosyne synthase (Bus) par ingénierie moléculaire et génétique 81
Figure 60. Carte du plasmide pCRISPomyces-2 qui permet l'expression du système CRISPR-Cas9 chez *Streptomyces* spp 96
Figure 61. Carte du plasmide pKCCpf1 qui permet l'expression du système CRISPR-Cpf1 chez *Streptomyces* spp 96
Figure 62. Carte du vecteur d'expression pBG102 97
Figure 63. Programme d'amplification par réaction en chaîne par polymérase et composition du mélange réactionnel des différentes polymérases utilisées lors de cette thèse 98
Figure 64. Principe de la méthode d'assemblage de type Golden Gate 103
Figure 65. Principe de la méthode d'assemblage de type Gibson 104
Figure 66. Spectres d'excitation et d'émission des protéines fluorescentes utilisées lors de cette étude 114
Figure 67. Alignement de séquences et structures secondaires des domaines de docking d'intérêt, de domaines porteurs d'acyl et de domaines cétosynthase 120
Figure 68. Analyse des profils d'interaction de charge des domaines de docking de type 1a chez DEBS et RAPS ainsi que du système RAPS scindé 121
Figure 69. Alignement de séquences et d'éléments de structure secondaire de domaines porteurs d'acyle et de domaine de docking de polycétide synthase 122
Figure 70. Représentation schématique du programme développé pour trouver des ARN guides potentiels pour l'utilisation de CRISPR-Cpf1 ou CRISPR-Cas9 dans un génome donné 123
Figure 71. Carte de la matrice d'édition pour l'insertion des DD de DEBS au sein de la PKS RAPS chez *S. hygroscopicus* pour la stratégie 1 par CRISPR-Cpf1 124

Figure 72. Profil de migration sur gel d'agarose (0,9 %) des étapes d'amplification et assemblage des éléments des matrices d'édition pour la scission de la rapamycine synthase 125
Figure 73. Chromatographes et gel SDS-PAGE des différentes étapes de purification du domaine de docking présent en position C-terminale de la sous-unité 2 de la 6-désoxyérythronolide B synthase 127
Figure 74. Analyse de l'intégrité du domaine de docking présent en position C-terminale de la sous-unité 2 de la 6-désoxyérythronolide B synthase 129
Figure 75. Analyse de l'intégrité du domaine de docking présent en position C-terminale de la sous-unité A de la rapamycine synthase 130
Figure 76. Spectres obtenus par titrage calorimétrique isotherme du couple de domaines de docking présents entre les sous-unités 2 et 3 de la 6-désoxyérythronolide B synthase 131
Figure 77. Spectres obtenus par dichroïsme circulaire pour le couple de domaines de docking présents entre les sous-unités 2 et 3 de la 6-désoxyérythronolide B synthase 132
Figure 78. Analyses du domaine porteur d'acyle du module 7 de la stambomycine synthase 134
Figure 79. Carte de la matrice d'édition pour l'insertion de la paire de SynZip 3:4 au sein de la stambomycine synthase chez *S. ambofaciens* par CRISPR-Cpf1 135
Figure 80. Criblage positif par PCR des exconjuguants de *Streptomyces ambofaciens* pour l'insertion des domaines SynZip au sein de la stambomycine synthase 136
Figure 81. Séquence nucléotidique et protéique résultant de la mutation apportée au génome de *S. ambofaciens* 137
Figure 82. Observation au microscope confocal de la souche FM145-JMM3 139
Figure 83. Carte de la matrice d'édition utilisée pour l'insertion de la protéine fluorescente mYPet fusionnée à la première sous-unité de la coelimycine synthase chez *Streptomyces coelicolor* 141
Figure 84. Criblage par PCR des exconjuguants de *Streptomyces coelicolor* pour l'insertion de la séquence codant la protéine mTurquoise2 en position 3′ de *cpkC* 142
Figure 85. Séquence nucléotidique et protéique résultant de l'insertion de la séquence codant mTurquoise2 en position 3′ de l'ORF *cpkC* chez *Streptomyces coelicolor* 143
Figure 86. Criblage par PCR des exconjuguants de *Streptomyces coelicolor* pour la délétion de *cpkC* 144
Figure 87. Séquence nucléotidique résultant de la délétion de l'ORF *cpkC* chez *Streptomyces coelicolor* 144
Figure 88. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145 pour l'optimisation des longueurs d'onde d'excitation 147

Figure 89. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L1 dans le but de valider les paramètres d'observation 148
Figure 90. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L1 148
Figure 91. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L2 149
Figure 92. Cliché obtenu par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L3 150
Figure 93. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L10 151
Figure 94. Cliché obtenu par microscopie confocale de la souche *Streptomyces coelicolor* FM145-JMM4 151

Table des tableaux

Tableau 1. Rendements de production de lactones des systèmes DEBS1-TE modifiées par introduction de domaines KR de type A2 et B2 66
Tableau 2. Liste des souches bactériennes utilisées au cours de cette thèse 93
Tableau 3. Liste des milieux de culture utilisés au cours de cette thèse 95
Tableau 4. Amorces oligonucléotidiques utilisées pour la construction des cassettes d'édition dans le cadre de la scission de polycétide synthases 100
Tableau 5. Liste des constructions plasmidiques nécessaires à l'étude de l'organisation de la coelimycine synthase chez *Streptomyces coelicolor* 101
Tableau 6. Amorces oligonucléotidiques utilisées pour la construction de la matrice d'édition *mYPet-cpkA* 102
Tableau 7. Amorces oligonucléotidiques, mélange réactionnel et programme utilisés pour le clonage de l'ARN guide au sein du vecteur pCRISPomyces-2 par Golden Gate 103
Tableau 8. Amorces oligonucléotidiques utilisées pour l'amplification des séquences codant les domaines de docking 105
Tableau 9. Solutions tampons utilisées pour la préparation des bactéries *Escherichia coli* chimio-compétentes 106
Tableau 10. Amorces oligonucléotidiques utilisées pour le criblage par PCR sur colonies et le génotypage des exconjuguants de *Streptomyces* spp 108
Tableau 11. Solutions tampons utilisées pour la purification des protéines d'intérêts 108
Tableau 12. Programme utilisé pour la chromatographie par chélation de métaux immobilisés 110
Tableau 13. Programme utilisé pour la chromatographie par chélation de métaux immobilisés inverse 111
Tableau 14. Séquences des protospacers potentiels identifiés dans le génome de *S. hygroscopicus* pour la scission de la rapamycine synthase 123
Tableau 15. Séquences protéiques des domaines de docking produits à partir des constructions plasmidiques dérivées du vecteur pBG102 128
Tableau 16. Séquences des protospacer potentiels identifiés dans le génome de *Streptomyces ambofaciens* pour la scission de la stambomycine synthase 135
Tableau 17. Liste des souches conçues dans le cadre de la scission de la stambomycine synthase durant ce travail de thèse 137
Tableau 18. Liste des souches conçues et étudiées dans le cadre de la localisation de la coelimycine synthase par le Dr. Jean-Malo Massicard 139
Tableau 19. Liste des plasmides utilisés pour l'édition génétique par CRISPR-Cas9 de *Streptomyces coelicolor* au cours de cette thèse 140
Tableau 20. Liste des souches construites au cours de ce travail pour la localisation de la coelimycine synthase 145

Liste des abréviations

A : domaine d'adénylation

ACN : acétonitrile

ACP : domaine porteur d'acyle

ACT : actinorhodine

ADN : acide désoxyribonucléique

ARN : acide ribonucléique

ARNcr : ARN CRISPR

ARNg : ARN guide

ARNt : ARN de transfert

ARNtracr : ARN transactivateur

ARO : aromatase

AT : domaine acyltransférase

ATd : domaine acyltransférase docking

BGC : cluster de gènes biosynthétiques

C : domaine de condensation

CDA : antibiotique calcium-dépendant

CFP : protéine fluorescente cyan

CLF : régulateur de taille de la chaîne polycétidique

CoA : coenzyme-A

CRISPR : *Clustered Regularly Interspaced Short Palindromic Repeats*

CSR : régulateur situé dans le cluster

CYC : cyclase

DC : dichroïsme circulaire

DD : domaine de docking

DEBS : 6-désoxyérythronolide B synthase

DH : domaine déshydratase

DHD : domaine déshydratase docking

DLS : diffusion dynamique de la lumière

DMSO : diméthylsulfoxyde

ECH : énoyl-CoA hydratase / isomerase

EMR : ellipticité moyenne par résidu

ER : domaine énoyl-réductase

FAS : acide gras synthase

FKBP12 : récepteur du tacrolimus

GBL : γ-butyrolactones

GS : gène synthétique

HCS : hydroxyméthylglutaryl-CoA synthase

HEPES : acide 4-(2-hydroxyéthyl)-1-pipérazineéthanesulfonique

HPLC : chromatographie en phase liquide à haute performance

HR : hautement réductrice

HRV 3C : protéase du rhinovirus humain 3C

HSII : HotStart II

Hyd : domaine hydrolase

iChip : isolation Chip

IMAC : chromatographie par chélation de métaux immobilisés

iPKS : polycétide synthase de type I itérative

IPTG : isopropyl β-D-1 thiogalactopyranoside

ITC : titrage calorimétrique isotherme

KASIII : 3-cétoacyl-ACP synthase

K_D : constante de dissociation

KR : domaine cétoréductase

KS : domaine cétosynthase

LINKS : Laterally-INteracting Ketosynthase Sequences

MALDI-TOF : désorption-ionisation laser assistée par matrice – temps de vol

mAT : AT spécifique du méthylmalonyl-CoA

MC : module de chargement

mFAS : acide gras synthase de mammifère

MIBiG : *Minimal Information about a Biosynthetic Gene cluster*

MOPS : acide 3-morpholino-1-propanesulfonique

MS : spectrométrie de masse

MT : domaine méthyltransférase

mTOR : complexe de la sérine/thréonine kinase cible de la rapamycine chez les mammifères

NADPH : nicotinamide adénosine dinucléotide phosphate

NR : non-réductrice

NRP : peptides non ribosomaux

NRPS : synthétase de peptides non ribosomaux

OMS : Organisation Mondiale pour la Santé

Ox : domaine oxydase

PAM : motif adjacent au protospacer

PCP : domaine porteur de peptide

Pdi : indice de polydispersité

PDIM : dimycocérosate de phthiocérol

PES : polyéthersulfone

PKS : polycétide synthase

P-Pant : 4′-phosphopantéthéine

PPTase : phosphopantéthéine transférase

PR : partiellement réductrice

RAPS : rapamycine synthase

RED : undecylprodigiosine

RH : région d'homologie

RiPP : peptides synthétisés par des ribosomes et modifiés de manière post-traductionnelle

SARP : protéines régulatrices des antibiotiques chez *Streptomyces*

SDS-PAGE : électrophorèse sur gel de polyacrylamide contenant du dodécylsulfate de sodium

SEC-GF : chromatographie d'exclusion de taille par gel filtration

SNAC : N-acétylcystéamine

SUMO : *Small Ubiquitin-like MOdifier*

SynZips : synthetic zippers

TE : domaine thioestérase

TEM : microscopie électronique à transmission

TFA : acide trifluoroacétique

TR : thioester-réductase

VEMS : vénémycine synthase

XU : unité échangeable

YFP : protéine fluorescente jaune

Etat de l’art

I. Les produits naturels

La survie des espèces des différents règnes du vivant repose sur leur capacité à se reproduire, croître, se défendre et communiquer. Elles produisent pour cela des composés biologiques appelés métabolites, qui selon leurs fonctions sont classés comme primaires ou spécialisés. Les métabolites primaires participent à la croissance, au développement et à la reproduction des organismes, tandis que les métabolites spécialisés sont principalement impliqués dans les interactions avec l'environnement et la régulation du métabolisme. Cependant, certains métabolites spécialisés peuvent également jouer un rôle dans la croissance et le développement d'un organisme. Les métabolites spécialisés sont également nommés produits naturels par la communauté et exhibent un potentiel prometteur en médecine et en agriculture notamment. Les produits naturels caractérisés à ce jour sont à 67 % produit par des plantes, puis par des animaux, des champignons et des bactéries, et à plus faible quantité par des protistes et des archées (Figure 1) (Chassagne et al., 2019).

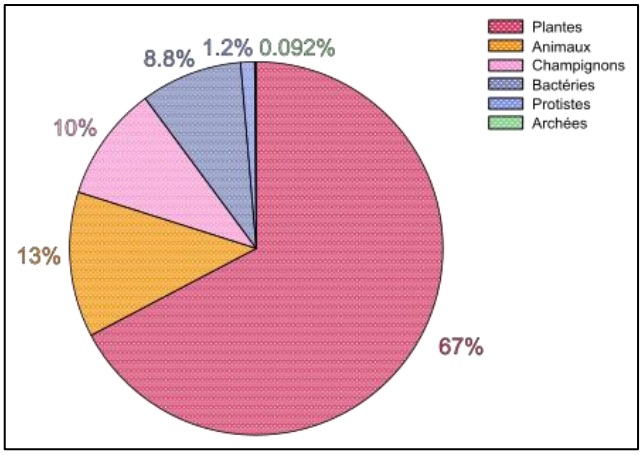


Figure 1. Répartition de la production de produits naturels connus par les différentes branches du vivant (adapté de Chassagne et al., 2019).

I.A. Les classes et fonctions de produits naturels

Ces produits regroupent plusieurs classes de molécules, déterminées par la manière dont ces composés sont produits. Parmi ces classes se retrouvent notamment les phénylpropanoïdes, les terpénoïdes, les alcaloïdes, les peptides, les glucides et les polycétides (Figure 2) (Chassagne et al., 2019; Vogt, 2010).

Les terpénoïdes sont des composés produits par des plantes, des animaux, des champignons, des bactéries, des protistes et des archées (Chassagne et al., 2019). Ils représentent la majorité des produits naturels connus et caractérisés à ce jour et dérivent d'unités isopréniques. Ces molécules représentent la majorité des composés volatiles secrétés par les plantes et leur permettent, par exemple, de se protéger des attaques des herbivores lors de leur alimentation ou leur ponte (Mumm et al., 2008).

Les molécules sécrétées ainsi par les plantes ne permettront pas activement de combattre l'invasion des herbivores, mais seront perçues par leurs prédateurs, qui sauront alors où trouver leur repas. Ce mécanisme indirect de protection des plantes est également accompagné d'un mécanisme direct de répulsion, également lié à la production de terpénoïdes (Mumm et al., 2008).

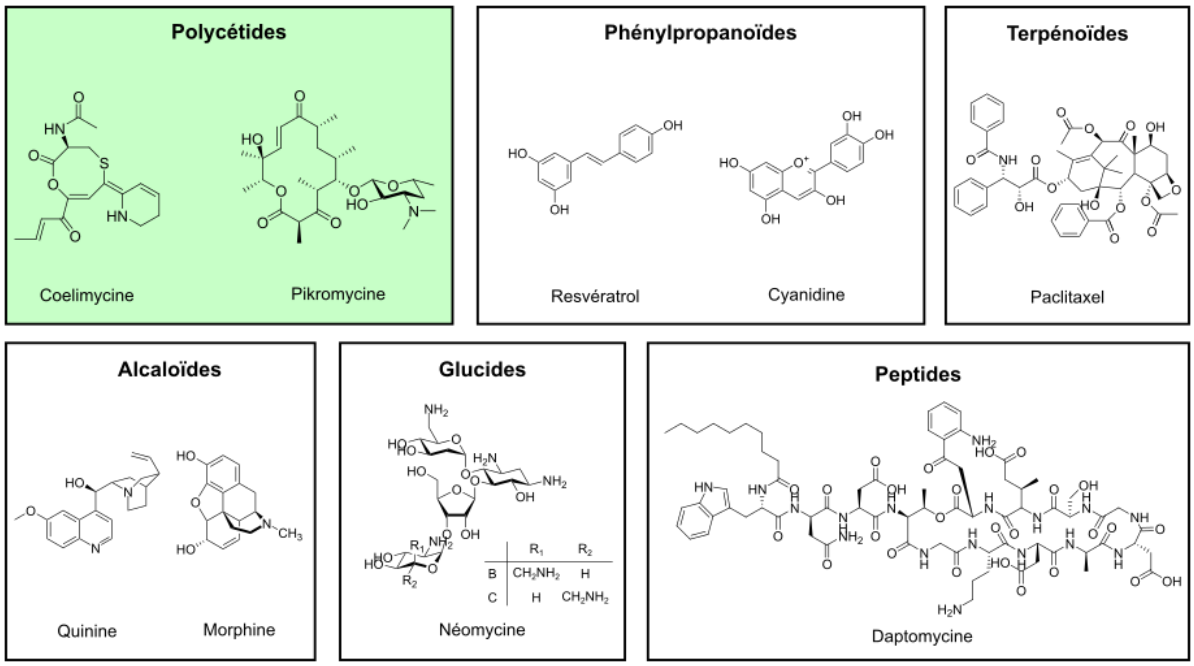


Figure 2. Structures de membres représentatifs des familles de produits naturels Les polycétides, représentés en vert, sont au cœur du sujet de cette thèse.

Les alcaloïdes représentent la deuxième classe des produits naturels la plus représentée. Ils sont également retrouvés au sein de tous les règnes du vivant (Chassagne et al., 2019). Parmi ces composés sont retrouvés, par exemple, la morphine et la cocaïne. Dans les années 1860, la cocaïne fait son entrée dans les pharmacies états-uniennes sous la forme de cachets, préconisés pour les enfants qui font leurs dents (Figure 3).



Figure 3. Affiche publicitaire pour l'utilisation de cachets de cocaïne contre les maux de dents aux Etats-Unis (http://parentpants.blogspot.com/2014/08/5-vintage-medicines-that-prove-doctors.html).

Les phénylpropanoïdes sont des métabolites retrouvés chez les plantes et ont différents rôles dans leur survie (Vogt, 2010). Ils sont produits à partir de résidus phénylalanines et regroupent plusieurs sous-groupes de métabolites qui sont les acides phénoliques, les flavonoïdes, les coumarines, les curcuminoïdes, les stilbènes, les quinones, les lignanes, les tannins, les terpénoïdes phénoliques et les alcaloïdes phénoliques (Neelam et al., 2020). Ces composés peuvent avoir des propriétés thérapeutiques tels que le resvératrol qui est un anti-cancéreux (Rauf et al., 2018), la cyanidine aux propriétés antioxydantes, anti-inflammatoires et anti-cancéreuses (Liang et al., 2021), et enfin la lutéoline qui est utilisée comme agent anti-cancéreux (Imran et al., 2019).

Les peptides du métabolisme spécialisé caractérisés à ce jour sont majoritairement synthétisés par des bactéries et des archées (Chassagne et al., 2019). Ils sont produits par l'assemblage d'acides aminés, et classés en deux catégories selon l'enzyme catalysant cet assemblage : les peptides synthétisés par des ribosomes et modifiés de manière post-traductionnelle (RiPP) d'une part et les peptides non-ribosomaux (NRP) produit par les NRP synthétases (NRPS) de l'autre. Ils ont des rôles variés au sein des cellules et peuvent avoir des propriétés thérapeutiques intéressantes comme la daptomycine qui est un antibactérien (Heidary et al., 2018).

Les glucides sont majoritairement des composés du métabolisme primaire des organismes, cependant certains d'entre eux peuvent être classés dans les métabolites spécialisés. La néomycine est un exemple d'antibactérien de la famille des glucides (Waksman & Lechevalier, 1949). De plus, certains sont également capables de se condenser avec d'autres métabolites spécialisés, formant ainsi des hétérosides. Un exemple d'hétéroside est la pikromycine, qui est un polycétide glycosylé ayant des propriétés antibiotiques (Suhren, 1951).

Enfin, les polycétides sont des polymères d'acyl-coenzyme A (acyl-CoA) produits par des bactéries, des champignons, des plantes et des animaux (Chassagne et al., 2019). La pikromycine évoquée précédemment est un exemple de polycétide, aux propriétés antibiotiques (Xue et al., 1998). Ces molécules sont produites par des enzymes analogues à celles impliquées dans la synthèse des acides gras (FAS) retrouvées au niveau du métabolisme primaire et sont appelées polycétide synthases (PKS).

I.B. Les prémices de la découverte des produits naturels et leur utilisation

En 1897, le médecin français Ernest Duchesne présente sa thèse intitulée « Contribution à l'étude de la concurrence vitale chez les micro-organismes : antagonisme entre les moisissures et les microbes » à la faculté de médecine et de pharmacie de Lyon (Duchesne, 1897).

Lors de ses travaux, il a pu mettre en évidence la compétition entre un champignon, *Penicillium glaucum* et des bactéries, et ainsi proposer l'utilisation de microorganismes en thérapie. Par la suite, le chercheur écossais Alexander Fleming extrait les métabolites produits par *Penicillium notatum*, dont il découvre des activités antibactériennes à la solution nommée pénicilline (Fleming, 1929). Cette découverte lui vaudra en 1945 le prix Nobel de physiologie ou médecine. Il partage ce prix avec le pharmacologue australien Howard Walter Florey et le biochimiste allemand Ernst Boris Chain, dont les efforts ont participé à l'isolation et la production du composé actif de la pénicilline (Chain et al., 1940).

A cette même période, la streptomycine est découverte par le microbiologiste états-unien Albert Schatz au sein du laboratoire du biochimiste et microbiologiste Selman Waksman (Schatz et al., 1944). Ce composé isolé de *Streptomyces griseus* a été identifié comme le premier antibiotique pouvant être utilisé contre la tuberculose. Cette découverte value à S. Waksman le prix Nobel de physiologie ou médecine de 1952, au détriment de son étudiant doctorant de l'époque, A. Schatz (Wainwright, 1989). Cependant cette découverte n'est pas la première concernant la mise en lumière des antibiotiques. En effet, en 1877 le pionner français de la microbiologie, Louis Pasteur, décrit le fait que certaines bactéries peuvent avoir un effet néfaste sur la croissance de *Bacillus anthracis*, et propose alors que cela puisse avoir un intérêt thérapeutique important. Par la suite, le dermatologue français Jean Gaté cosigne avec son collègue George Papacostas un livre faisant émerger la notion d'associations microbiennes, et par la même occasion les premières mentions de ce que nous connaissons aujourd'hui comme étant les antibiotiques (Papacostas & Gaté, 1928).

En 1950, le premier antibiotique de la classe des macrolides, qui font partie de la famille des polycétides, est découvert : la pikromycine (Aminov, 2017; Suhren, 1951). Même si ce composé n'a pas pu être commercialisé, un composé similaire est isolé en 1949 par un chercheur philippin Aberlardo Aguilar à partir de la bactérie *Saccharopolyspora erythraea*, pour la companie industrielle Eli Lily (Aminov, 2017). C'est donc en 1952 que ce composé, l'érythromycine, est commercialisé sous le nom d'Ilosone.

Même si les exemples cités montrent la découverte de composés d'utilité thérapeutique, certains autres ont pu être commercialisés à tort, comme la cocaïne précédemment citée. L'opium est également un alcaloïde, qui a été utilisé comme analgésique, antitussif et sédatif notamment en Asie Mineure. La morphine quant à elle a été utilisée par l'armée états-unienne lors de la guerre de Sécession (1861–1865) (López-Muñoz et al., 2011). Les soldats pouvaient alors s'auto-administrer des seringues de morphine pour atténuer les douleurs liées à leurs blessures.

A la fin de la guerre, un phénomène appelé le « mal militaire » a été observé, qui n'est autre que la dépendance développée par les soldats rentrés de la guerre. Il est estimé que ce syndrome aurait touché près de 45000 soldats (López-Muñoz et al., 2011). Afin de pallier ce souci d'addiction, le chercheur allemand Heinrich Dreser synthétise en 1874 un dérivé de la morphine qui sera commercialisé en 1899 : l'héroïne. Cela va sans dire que le problème n'était visiblement pas réglé, et en 1913 les effets addictifs de l'héroïne ont été caractérisés (López-Muñoz et al., 2011).

La grande famille des produits naturels est donc riche en composés à potentiel thérapeutique. Les quelques prémices précédemment décrits ont, entre autres, permis durant les XX^{ème} et XXI^{ème} siècles de poursuivre les efforts vers la découverte de nouvelles molécules thérapeutiques. Ces découvertes ont également été propulsées au début des années 2000 par l'avènement du séquençage de génomes complets, qui a permis de découvrir dans les génomes microbiens des clusters de gènes de biosynthèse de composés inconnus jusqu'à ce jour (Chassagne et al., 2019).

Afin d'identifier et optimiser de nouvelles molécules, il est nécessaire de s'intéresser aux producteurs de ces composés d'intérêts. Bien que produits par toutes les branches du vivant, les produits naturels exploités en médecine et agriculture à ce jour sont majoritairement produits par des bactéries du genre *Streptomyces*.

I.C. Le plus grand producteur de produits naturels d'intérêt : les bactéries du genre *Streptomyces*

Les bactéries du genre *Streptomyces* produisent des produits naturels dont 63,3 % ont une activité antibiotique et environ 24,5 % des produits naturels identifiés chez *Streptomyces* spp. sont des polycétides (Chassagne et al., 2019). Cela en fait des systèmes d'études prometteurs dans la recherche de nouvelles molécules et le développement de dérivés de ces composés. La production de ces composés est intimement liée au cycle de croissance et de vie de ces bactéries, qui leur est propre et particulier.

I.C.1. Le cycle de vie des Streptomyces

Les *Streptomyces* sont des bactéries ayant un cycle de reproduction semblable aux champignons (Figure 4) (Hopwood, 2019). La croissance de nouveaux individus démarre de spores, contenant chacune une copie unique du chromosome bactérien. Des tubes germinaux vont émerger des spores, pour former des hyphes végétatifs qui vont croître dans le milieu de culture afin d'accéder aux nutriments nécessaires au développement de la colonie. Ces hyphes contiennent de multiples copies du chromosome, au contraire des spores initiales, car le chromosome se réplique lors de la croissance, tandis que la présence de *septa* reste rare.

Une fois les hyphes végétatifs développés et les nutriments du milieu de culture épuisés, cette carence va provoquer le développement d'hyphes aériens par la lyse partielle des hyphes végétatifs, qui est la source de substances nutritives pour la croissance du mycélium aérien. Ce processus permet de recycler les nutriments et à terme de coloniser le milieu à la recherche de nouveaux nutriments. Ces hyphes contiennent également plusieurs copies du génome. C'est lors de cette étape de transition entre croissance végétative et croissance aérienne que la majorité des métabolites spécialisés connus sont produits par *Streptomyces* spp. Enfin, les hyphes aériens se divisent de manière synchrone par l'apparition de *septa* afin de former de nouvelles spores, à nouveau contenant une unique copie du chromosome.

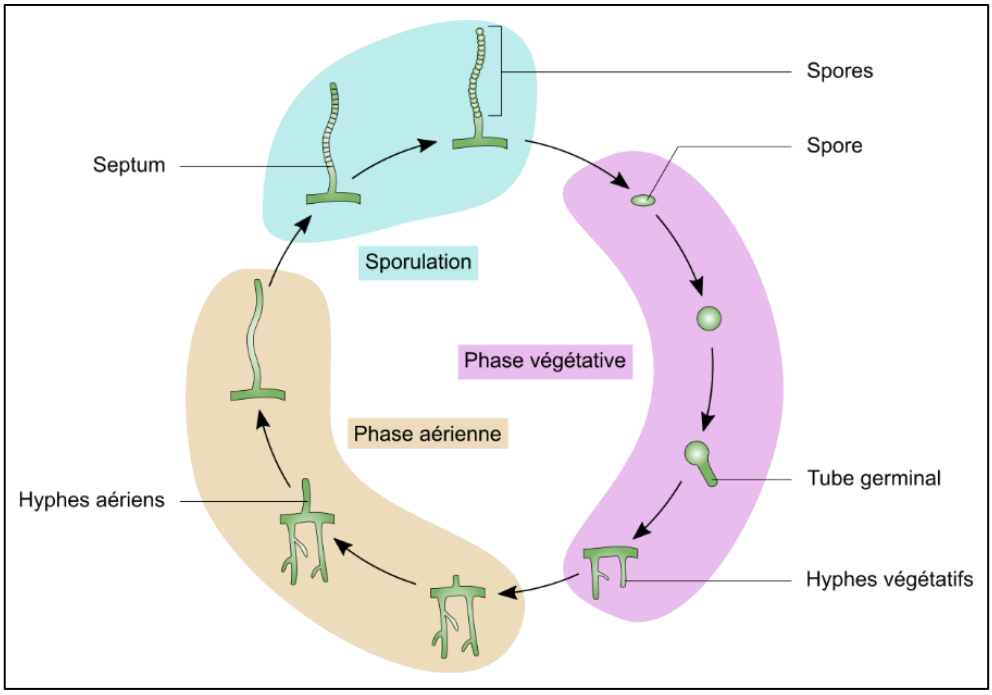


Figure 4. Cycle de reproduction des bactéries du genre *Streptomyces* (adapté de Kieser et al., 2000). Les *Streptomyces* ont un cycle de vie particulier pouvant ressembler à celui des champignons. A partir d'une spore haploïde, un tube germinatif croît afin d'explorer le milieu de culture. Celui-ci va donner naissance à des embranchements polyploïdes appelés hyphes végétatifs au moment de la phase végétative. Une fois les nutriments du milieu épuisés, la colonie entre dans la phase aérienne durant laquelle les hyphes végétatifs se dégradent et permettent la croissance d'hyphes aériens, également polyploïdes. Ces hyphes se cloisonnent ensuite avec l'apparition de septa et donnent naissance à de nouvelles spores haploïdes, qui permettront la croissance de nouvelles colonies lors de l'étape de sporulation.

I.C.2. Le génome des Streptomyces

Le génome des *Streptomyces* est linéaire et riche en paires de bases G+C (environ 72 %) (Kieser et al., 2000). La taille de leur chromosome est conséquente et peut aller de six à quinze millions de paires de bases. La réplication de ce chromosome se fait à partir de sa région centrale et de manière bidirectionnelle (Bury-Moné et al., 2023).

Le génome de *Streptomyces* spp. est compartimenté. La région centrale est composée des gènes communs à toutes les espèces et essentiels à leur survie. Cette région est hautement conservée entre les espèces, jusque dans l'ordre des gènes qui s'y trouvent. Cette caractéristique s'appelle la synténie. De part et d'autre de cette région sont retrouvés les bras chromosomiques, où se trouvent majoritairement des clusters de gènes biosynthétique (BGC) liés à la production de métabolites spécialisés. Les génomes de *Streptomyces* spp. contiennent en moyenne une quarantaine de ces BGC. Ces régions sont plus sensibles aux remaniements génétiques. Enfin, aux extrémités du chromosome se trouvent des télomères, associés à un complexe protéique, permettant de les protéger et ainsi limiter la perte d'information génétique (Bury-Moné et al., 2023). Plus les gènes sont éloignés de la région centrale du chromosome, et moins ils seront exprimés. Cela explique le caractère cryptique de certains BGC se trouvant éloignés de cette région et plus proche des extrémités du chromosome (Lioy et al., 2021).

La régulation de l'expression des BGC présents dans les génomes du genre *Streptomyces* est intimement liée à leur cycle de croissance (Figure 5). Comme précédemment présentée, la production de la majeure partie des métabolites spécialisés a lieu lors de la transition entre les phases de croissance végétative et aérienne (Hopwood, 2019).

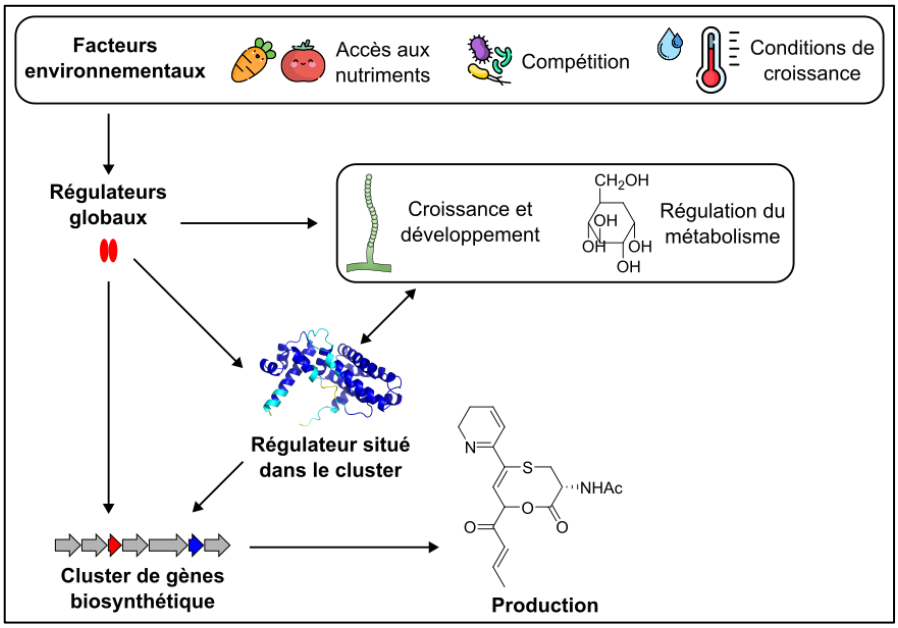


Figure 5. Régulation de la production de métabolites spécialisés chez *Streptomyces* (adapté de McLean et al., 2019). L'accès aux nutriments, la compétition avec les autres organismes et les conditions de croissances des *Streptomyces* influent sur l'expression des clusters de gènes biosynthétiques (BGC) contenus dans leur génome. L'activation de régulateurs globaux régule la croissance et le développement des colonies, ainsi que leur métabolisme. Ils influent aussi sur l'expression des BGC, au sein desquels se trouvent des régulateurs spécifiques pouvant également réguler la croissance et le métabolisme. Toutes ces régulations permettent d'activer ou inactiver la production des métabolites concernés.

Les gènes régulant la croissance bactérienne et l'expression de ces métabolites sont donc intimement liés et parfois même sont identiques. C'est le cas par exemple du gène *bldA* qui code un acide ribonucléique (ARN) de transfert (ARNt) reconnaissant un codon rare et dont l'expression permet à la fois la transition de phase de croissance de la colonie et la production de métabolites spécialisés chez différentes espèces de *Streptomyces*, comme l'actinorhodine et l'undecylprodigiosine chez *S. coelicolor* (Hackl & Bechthold, 2015). Le codon rare associé – UUA – est en effet retrouvé au sein des gènes *actII-orf4* qui code un activateur transcriptionnel du BGC de l'actinorhodine, ainsi que dans le gène *redZ* qui est un régulateur de la production de l'undecylprodigiosine (Li et al., 2007).

I.C.3. La régulation de la production de produits naturels chez Streptomyces coelicolor

S. coelicolor est un organisme modèle de l'étude de ce genre bactérien. Le séquençage de son génome a permis d'identifier qu'environ 12 % du génome code des protéines régulatrices, ce qui montre leur importance dans la croissance des colonies et leur survie (Bentley et al., 2002). Avant le séquençage complet du génome de *S. coelicolor*, seuls trois métabolites spécialisés étaient identifiés : l'actinorhodine (ACT), l'undecylprodigiosine (RED) et l'antibiotique calcium-dépendant (CDA). Une fois le séquençage effectué et analysé, dix-huit nouveaux BGC associés à la production de métabolites spécialisés ont pu être identifiés. C'est ainsi que le cluster lié à la production de la coelimycine A, un antibiotique produit par une PKS modulaire, a pu être identifié. Le génome de *S. coelicolor* et la localisation des quatre principaux BGC cités sont représentés dans la Figure 6.

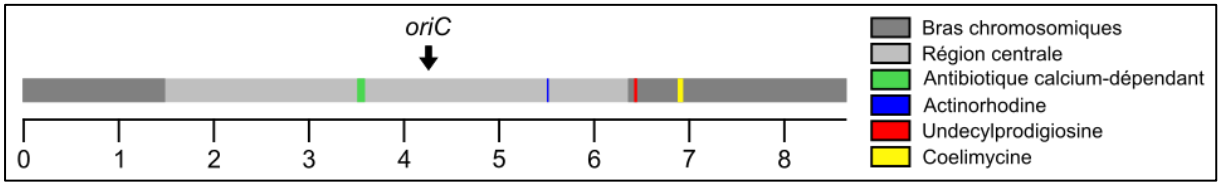


Figure 6. Représentation schématique du chromosome de *Streptomyces coelicolor* et localisation de ses quatre principaux clusters de gènes biosynthétiques. Le génome de *S. coelicolor* contient 8'667'507 pb. Les bras chromosomiques à gauche et à droite sont respectivement composés d'environ 1,5 Mb et 2,3 Mb. L'origine de réplication (*oriC*) se trouve au milieu du génome. Les clusters de gènes biosynthétiques responsables de la production de l'antibiotique calcium-dépendant (CDA) et de l'actinorhodine (ACT) sont retrouvés dans la région centrale du chromosome, tandis que ceux responsables de la production de l'undecylprodigiosine et de la coelimycine sont dans le bras chromosomique droit.

La production de ces quatre composés majeurs retrouvés chez *S. coelicolor* est hautement régulée (Liu et al., 2013). Les BGC contiennent en leur sein des gènes codant des régulateurs transcriptionnels appelés *Cluster-Situated Regulator* (CSR).

Certains régulateurs transcriptionnels sont spécifiques des bactéries du genre *Streptomyces* et sont alors regroupés sous le nom de protéines régulatrices d'antibiotique chez *Streptomyces* (SARP). Les gènes *actII-ORF4* retrouvé dans le BGC associé à la production d'actinorhodine et *redZ* au sein du BGC lié à la production de l'undecylprodigiosine sont des exemples de CSR spécifiques de *Streptomyces*. Le régulateur codé par le gène *actII-ORF4* par exemple régule la transcription de la totalité du BGC, en se liant à cinq régions promotrices *via* un domaine présent en position N-terminale de la protéine. Ce gène régulateur est lui-même activé par huit protéines régulatrices identifiées.

La régulation de la production des quatre principaux antibiotiques de *S. coelicolor* implique des molécules hormones-*like*. Parmi ces molécules sont retrouvés les γ-butyrolactones (GBL). Le cluster responsable de la production de la coelimycine par exemple contient notamment deux gènes codant des SARP : *cpkO* et *cpkN*. Il contient également un gène codant une GBL-synthase (*scbA*), une buténolide phosphate réductase (*scbB*), un récepteur aux GBL (*scbR*) et un pseudo-récepteur aux GBL (*scbR2*) (Figure 7) (Bednarz et al., 2021; Gomez-Escribano et al., 2012).

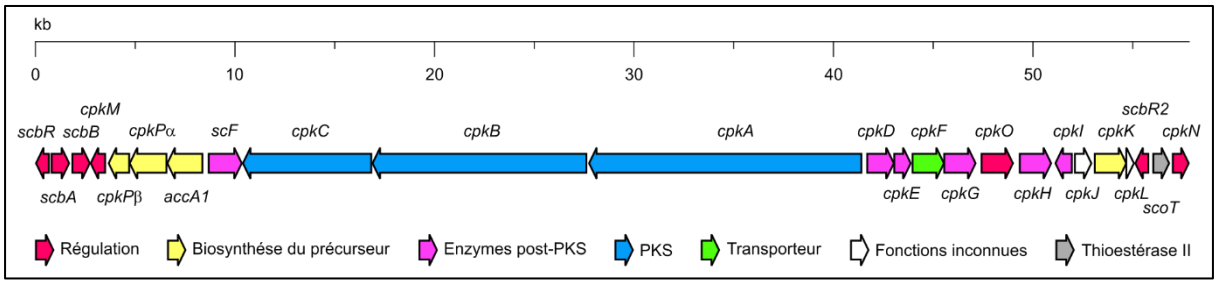


Figure 7. Cluster de gènes lié à la production de la coelimycine chez *Streptomyces coelicolor* (adapté de Gomez-Escribano et al., 2012). Ce cluster contient trois gènes codant des sous-unités de polycétide synthase (PKS) : *cpkA–C*. Il contient également un gène qui code le transporteur de la coelimycine (*cpkF*), ainsi que quatre gènes impliqués dans la biosynthèse des précurseurs du polycétide (*cpkP*α–β, *accA1* et *cpkK*), six gènes d'enzymes post-PKS (*scF*, *cpkD–E*, *cpkG*, *cpkH–I*), une séquence permettant la synthèse d'une thioestérase de type II (*scoT*) et deux gènes dont la fonction est inconnue (*cpkJ* et *cpkL*). Enfin, ce cluster contient sept gènes liés à sa régulation transcriptionnelle. Les séquences *cpkO* et *cpkN* codent des protéines régulatrices d'antibiotique chez *Streptomyces* (SARP). Leur transcription est régulée par un système de γ-butyrolactones (GBL) par le biais des gènes *scbA* (GBL-synthase), *scbB* (buténolide phosphate réductase), *scbR* (récepteurs de GBL) et *scbR2* (pseudo-récepteur de GBL). Le gène *cpkM* quant à lui est proposé comme codant une histidine kinase.

Lorsque ScbR est retrouvé au niveau du promoteur du gène cpkO, le cluster de gène est silencieux (Figure 8). L'activité de ScbA permet la synthèse du GBL spécifique de *S. coelicolor* (SCB), qui interagit par la suite avec ScbR permettant ainsi de lever l'inhibition (Takano et al., 2005). Enfin, ScbR2 est un pseudo-récepteur aux GBL. Il a été défini comme un répresseur tardif car il ne se lie pas directement aux GBL mais à l'actinorhodine (Gottelt et al., 2010; Liu et al., 2013).

L'actinorhodine est produite de façon plus tardive que la coelimycine par *S. coelicolor*, ainsi l'inhibition médiée par ScbR2 est levée plus tardivement que celle impliquant ScbR (Liu et al., 2013).

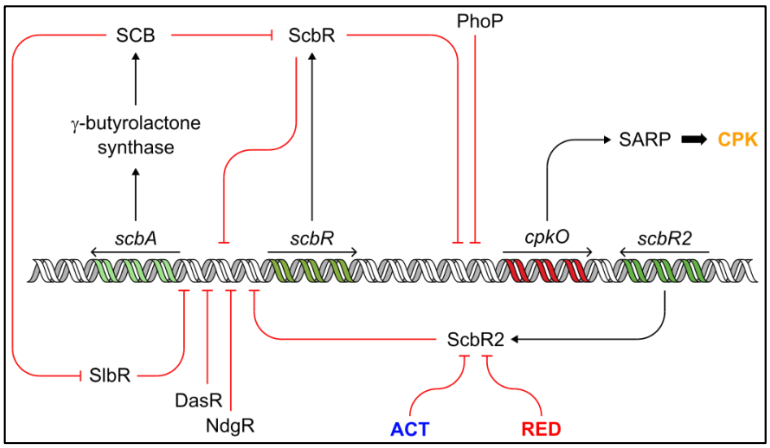


Figure 8. Régulation de l'expression des quatre principaux antibiotiques produit par *Streptomyces coelicolor* (adapté de G. Liu et al., 2013). Les flèches noires indiquent le rôle des gènes représentés et l'action des activateurs transcriptionnels, tandis que les lignes rouges qui se terminent par des barres indiquent des répressions. **Légende** – SARP : protéines régulatrices d'antibiotique chez *Streptomyces*, CPK : coelimycine, RED : undecylprodigiosine, ACT : actinorhodine.

I.D. A la recherche de nouvelles molécules thérapeutiques

L'utilisation des antibiotiques a été une révolution sanitaire majeure, qui a cependant un revers non négligeable : la dissémination à grande échelle de l'antibiorésistance. Cette problématique a vu le jour dès les prémices de la découverte des antibiotiques. En effet, dès la découverte de la pénicilline et sa première utilisation, Alexander Fleming prévenait en 1945 des effets néfastes d'une utilisation non contrôlée des antibiotiques (Institut Pasteur, 2018) :

« Cela aboutirait à ce que, au lieu d'éliminer l'infection, on apprenne aux microbes à résister à la pénicilline et à ce que ces microbes soient transmis d'un individu à l'autre, jusqu'à ce qu'ils en atteignent un chez qui ils provoqueraient une pneumonie ou une septicémie que la pénicilline ne pourrait guérir. »

Dans son rapport de la 77ème assemblée mondiale de la santé en 2024, l'organisation mondiale de la santé (OMS) rappelle qu'en 2019, le nombre de décès dans le monde liés à la résistance aux antibiotiques est estimé à 1,27 millions (Résistance aux antimicrobiens : accélérer les ripostes nationales et mondiales, 2024). Dans ce même rapport de 2024, l'OMS prédit que d'ici 2030, l'impact financier de cette résistance serait une perte annuelle d'entre 1 et 3,4 milliards de dollars américains sur le produit intérieur brut et 1 milliard de dollars de dépenses de santé d'ici à 2050.

Dans ce cadre, l'OMS publie également en 2024 quarante priorités de recherche autour de la résistance aux antibiotiques, liées à la prévention, le diagnostic, le traitement et les liens transversaux de cette résistance, ainsi que sur l'étude de la résistance plus particulière liée au traitement de la tuberculose (World Health Organization, 2024).

La crise liée à la résistance aux antibiotiques est d'autant plus préoccupante que les industries pharmaceutiques n'investissent plus dans la recherche pour de nouveaux antibiotiques, au profit de drogues anticancéreuses (Plackett, 2020). Cela s'observe notamment par la décroissance du nombre de nouveaux antibiotiques approuvés aux Etats-Unis depuis les années 1980, au profit de l'approbation de nouvelles molécules anticancéreuses (Figure 9).

Face au manque d'investissement des grandes entreprises et au manque à gagner que la recherche de nouveaux antibiotiques représente pour elles, un système de financement appelé « modèle Netflix » a émergé en Angleterre (National Institute for Health and Care Excellence, 2025). Cela repose sur un système de revenus fixes des entreprises pharmaceutiques concernant l'exploitation des antibiotiques, plutôt que sur un revenu « par-pilule » fluctuant, dans le but de favoriser la stabilité de ces entreprises et leur permettre d'investir à nouveaux et de manière plus importante dans la recherche de nouveaux antibiotiques.

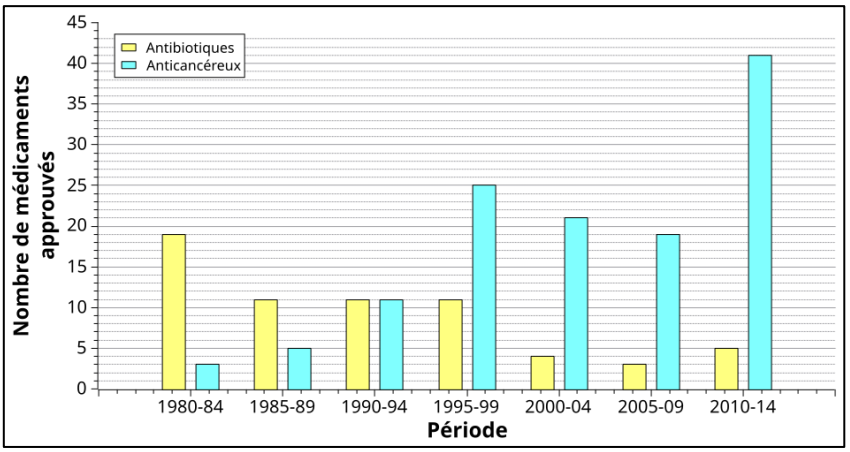


Figure 9. Nombre de médicaments approuvés aux Etats-Unis entre 1980 et 2014 (adapté de Plackett, 2020).

Il existe différentes manières de lutter contre la dissémination des bactéries (multi)-résistantes. La vaccination est un levier important dans cette bataille, avec les gestes d'hygiène permettant de limiter la transmission de maladie infectieuse et la lutte contre l'automédication et l'utilisation démesurée et inappropriée des antibiotiques. En parallèle, la découverte de nouveaux composés thérapeutiques reste également un moyen de pallier l'absence de traitement possible contre certains pathogènes du fait de leur résistance.

Il existe différentes méthodes afin de dénicher ces nouveaux composés, notamment la recherche de nouveaux producteurs de molécules d'intérêts dans l'environnement. D'autres méthodes sont également employées permettant de produire des dérivés par hémisynthèse chimique ou par biologie synthétique.

I.D.1. La découverte de nouveaux producteurs et de nouvelles molécules

Afin qu'un composé produit par un organisme soit détectable et identifiable en laboratoire, plusieurs critères doivent être remplis (Dasí-Delgado et al., 2025). Dans un premier temps, le BGC lié à la production du composé doit être transcrit puis traduit de manière suffisante en conditions de laboratoire. Les substrats et les cofacteurs nécessaires au bon fonctionnement des enzymes doivent également être disponibles. Enfin, le produit ciblé doit être détectable – par des tests d'activité ou des méthodes analytiques – dans les conditions utilisées, et une méthode d'extraction appropriée doit être trouvée. Lorsqu'un BGC ne remplit pas ces critères, le composé qu'il produit n'est pas observable en conditions de laboratoire. Ce BGC est alors dit « silencieux », dans le sens où le composé qui lui est associé n'est pas observable soit car il n'est pas produit, soit car les méthodes utilisées ne permettent pas sa détection. Ces critères ne s'appliquent cependant qu'à la fraction de bactéries cultivables en laboratoire, ce qui représente environ 1 % des bactéries existantes. D'autres méthodes sont déployées afin d'accéder aux molécules produites par les 99 % de bactéries non cultivables, telles que la métagénomique.

Plusieurs méthodes peuvent être utilisées afin de découvrir les composés produits par les BGC silencieux. Il est, par exemple, possible de modifier l'expression des gènes d'intérêts au sein de l'hôte par ingénierie génétique. C'est cette technique qui a permis la découverte des stambomycines chez *Streptomyces ambofaciens* (Figure 10A) (Laureti et al., 2011). Le cluster de gènes liés à sa production – contenant notamment neuf gènes codant des sous-unités de PKS – a été découvert par séquençage du génome de la souche, cependant son expression était très faible voire nulle et ne permettait pas l'identification des métabolites associés.

Le gène *sam23877-7107* a été identifié comme pouvant coder une protéine régulatrice liant l'ATP, de la famille LuxR. Cette famille de protéines est également retrouvée au sein des systèmes PKS qui produisent la rapamycine et la pikromycine, respectivement RapH et PikD, dans lesquels elle agit comme activateur transcriptionnel (Kuscer et al., 2007; Wilson et al., 2001). La stratégie utilisée ici a été d'intégrer une copie de ce gène sous le contrôle du promoteur constitutif fort *ermE** au sein du vecteur intégratif pIB139 afin d'obtenir le dérivé pOE-0484 (Laureti et al., 2011; Wilkinson et al., 2002). La souche obtenue à la suite de la transformation de *S. ambofaciens* ATCC23877 a alors été nommée ATCC/OE484.

De plus, une souche contrôle délétée du gène *sam23877-7126* (*pks1*) codant la première sous-unité de la PKS des stambomycines a été conçue, ne permettant pas la production de ces composés. L'analyse par spectrométrie de masse des souches ainsi obtenues – surexprimant le gène codant l'activateur transcriptionnel, en présence ou non de la sous-unité initiatrice de la production des stambomycines – a permis de détecter la présence des composés d'intérêts (Figure 10B). Les stambomycines ont ainsi pu être caractérisées (Figure 10C) et leur potentiel anti-proliférateur découvert (Laureti et al., 2011).

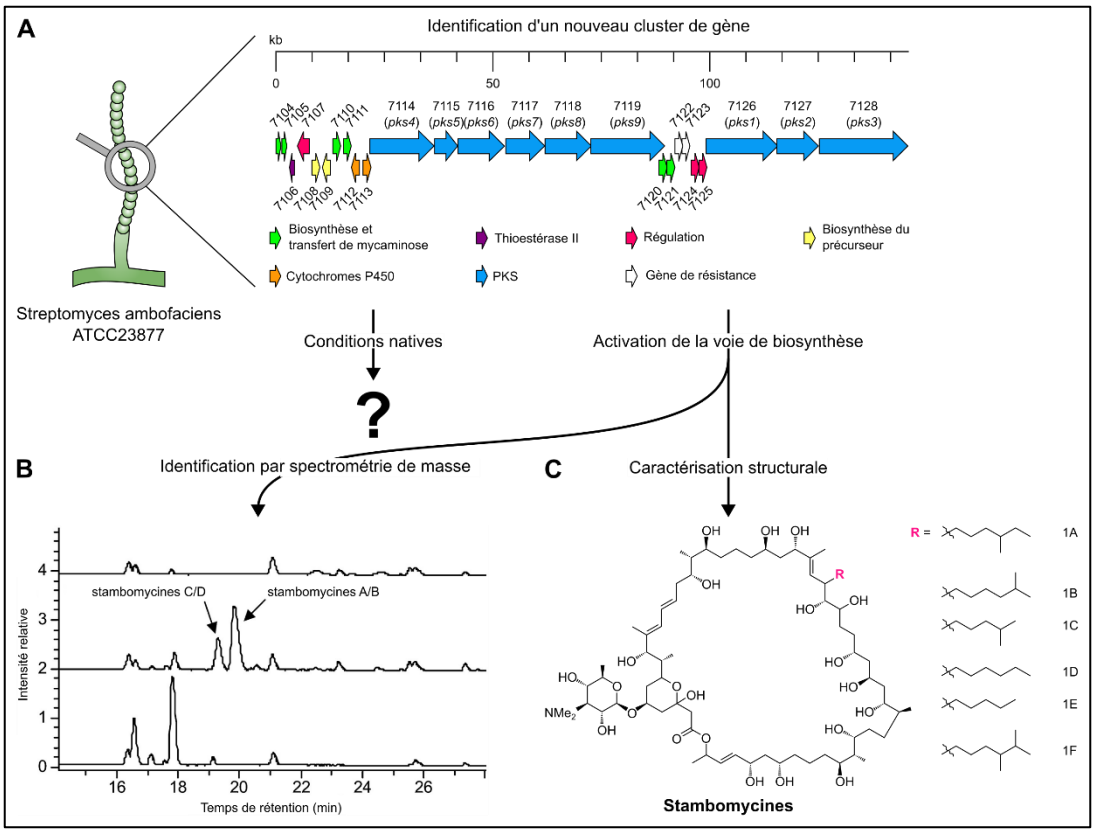


Figure 10. Découverte du cluster de gènes biosynthétique lié à la production des stambomycines chez *Streptomyces ambofaciens* (adapté de Laureti et al., 2011). **(A)** Cluster de gènes *sam23877-7104* – *sam23877-7128*, liés à la production des stambomycines. Le séquençage de la souche *S. ambofaciens* ATCC23877 a permis d'identifier un nouveau cluster de gènes lié à la production d'une polycétide synthase (PKS). Cependant ce cluster est silencieux en conditions de culture de laboratoire et aucun polycétide ne lui était donc attribué. L'activateur transcriptionnel identifié *sam23877-7107* a été cloné sous le contrôle du promoteur constitutif fort *ermE**, permettant l'activation du cluster et la production des polycétides associés. **(B)** Analyse par spectrométrie de masse des souches obtenues pour l'identification des stambomycines. Le chromatographe montre les analyses menées sur la souche ATCC23877 conjugué avec le vecteur pIB139 vide (*haut*), de la souche ATCC/OE484 (*milieu* et de la souche ATCC/OE484/Δ *sam23877-7126*(*pks1*). **(C)** Structure des stambomycines résolues à partir des données de spectrométrie de masse.

Il est également possible d'activer des BGC silencieux en réprimant leurs régulateurs négatifs. C'est ainsi par exemple que l'expression du cluster à l'origine de la production des totopotensamides a été activée chez *Streptomyces pactum* (Chen et al., 2017).

L'équipe a réprimé l'expression de plusieurs régulateurs négatifs et également exprimé un gène codant un activateur du cluster sous le contrôle du promoteur *ermE**, permettant ainsi la caractérisation des produits de ce cluster.

Une étude menée sur le cluster de gènes à l'origine de la production de l'atratumycine chez *Streptomyces atratus* SCSIO ZH16 a permis d'identifier ce métabolite (Sun et al., 2019). Pour cela, l'espèce bactérienne a subi une mutagenèse par utilisation de plasma. Cette méthode repose sur l'utilisation d'un jet d'hélium ionisé lancé sur les cellules afin d'y induire des mutations génétiques aléatoires. Les mutants ont ensuite été analysés par chromatographie en phase liquide à haute performance couplée à de la spectrométrie de masse (HP-LC-MS). Cette approche seule a permis d'obtenir une faible quantité de métabolite. Alors une optimisation du milieu de culture a été effectuée, permettant d'augmenter les rendements et caractériser le composé produit (Sun et al., 2019).

La co-culture de *Streptomyces* sp. PTY087I2 avec *Bacillus subtilis*, *Staphylococcus aureus* ou *Pseudomonas aeruginosa* a permis d'augmenter la production de granatocymine D, granaticine et dihydrogranaticine B (Sung et al., 2017). Dans une autre étude, la co-culture de *Rhodococcus* sp. et *Nocardiopsis* sp. a permis d'identifier des composants impliqués dans la réponse métabolique à cette mise en co-culture (Derewacz et al., 2015). A partir de cette observation, les ciromicines A et B ont pu être caractérisées par spectrométrie de masse. La co-culture constitue donc une méthode pouvant permettre l'activation de cluster de gènes silencieux en conditions de monocultures en laboratoire.

En 2002, l'équipe d'Epstein *et coll*. a développé une méthode permettant d'isoler des molécules produites par des micro-organismes « non cultivables » marins (Kaeberlein et al., 2002). Une partie du milieu naturel étudié est isolée, diluée puis capturée dans une gélose d'agar. Cette gélose est ensuite placée entre deux membranes de porosité 0,03 µm, qui permettent le passage des nutriments mais limite le mouvement de cellules. Ce dispositif a ensuite été placé dans un aquarium mimant le milieu naturel de ces cellules. Cela a permis d'obtenir des colonies pour en moyenne 22 ± 13 % des cellules inoculées.

Par la suite, ce dispositif a été amélioré afin de permettre l'isolation des bactéries présentes au sein de l'échantillon de manière à obtenir une unique cellule par puits sur une puce appelée *isolation Chip* (*iChip*) (Dasí-Delgado et al., 2025). L'utilisation de cette technologie permet d'augmenter d'un facteur 5 le taux de croissance de bactéries provenant d'un environnement donné comparé à la croissance en boîte de Pétri (Nichols et al., 2010). Afin d'améliorer encore le système, une étape de tri des cellules vivantes et mortes peut être effectué par cytométrie en flux afin d'optimiser la culture dans chaque puits de la puce (Liu et al., 2021).

Une autre méthode permettant de passer outre l'étape de culture des micro-organismes est la métagénomique. Dans ce cas, le matériel génétique total d'un échantillon donné est extrait puis séquencé dans le but de trouver de nouveaux gènes potentiellement impliqués dans la synthèse de molécules d'intérêts. En effet, certains micro-organismes ne peuvent croître en l'absence de leurs partenaires environnementaux. La métagénomique permet de pallier ce souci. C'est ainsi par exemple que la nospérine – un polycétide de la famille des pédérines – a été identifiée dans un échantillon de lichen *Peltigera membranacea* (Kampa et al., 2013). Ce lichen a été isolé en Islande et le matériel génétique total de l'échantillon a été extrait puis analysé par séquençage.

Ainsi, deux clusters de gènes ont été identifiés codant des systèmes PKS *trans*-AT, provenant probablement de cyanobactéries du genre *Nostoc*. Parmi ces deux BGC, le plus grand a été identifié comme analogue au BGC produisant des composés de la famille des pédérines. L'étude s'est donc intéressée à celui-ci et a ainsi élucidé la structure de la nospérine et identifié la PKS la produisant (Kampa et al., 2013).

I.D.2. L'hémisynthèse chimique de nouvelles molécules

L'érythromycine retrouvée nativement chez *Saccharopolyspora erythraea* est sensible à l'attaque nucléophile en condition acide des groupements hydroxyles portés par les atomes de carbones C6 et C12 sur le centre électrophile de la fonction cétone portée par le carbone C9. Cette attaque entraîne la formation de l'anhydroérythromycine A, qui est un composé inactif. Afin d'éviter cela, la solution explorée a été de protéger ce centre électrophile en modification la fonction cétone en fonction azane par l'addition d'un atome d'azote dans le cycle lactone de l'érythromycine A (Figure 11) (Mutak, 2007; Peltier, 1992). Cette modification donne naissance à l'azithromycine, qui est encore aujourd'hui l'un des antibiotiques les plus prescrits contre les infections bactériennes.

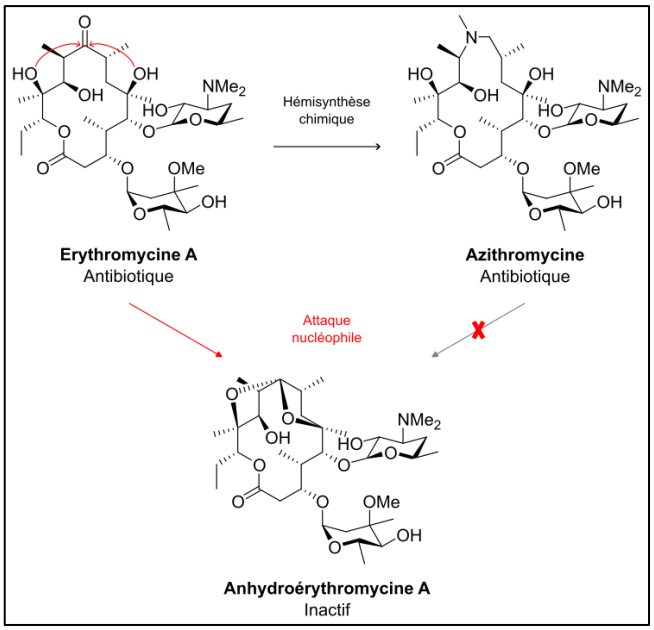


Figure 11. Modification de l'érythromycine A par hémisynthèse (adapté de Peltier, 1992). L'érythromycine A subit une attaque nucléophile en milieu acide, ce qui le transforme en anhydroérythromycine A inactive. Pour protéger la fonction cétone attaquée, celle-ci est modifiée en fonction azane, ce qui rend l'attaque nucléophile impossible.

I.D.3. La synthèse chimique de nouvelles molécules

Certains produits naturels montrent un potentiel thérapeutique, mais ne peuvent pas être extraits à partir de leur organisme productif. C'est le cas, par exemple, de l'halichondrine B, retrouvée chez l'éponge marine *Halichondria okadai* (Nicolaou et al., 2022).

Ce composé présente des propriétés antiprolifératives mais n'est cependant pas exploitable à partir de la souche productrice. L'objectif a donc été de synthétiser chimiquement un composé dérivé de l'halichondrine B dont la structure est plus petite mais conserve certains éléments de la molécule parentale, expliquant que l'activité du composé soit conservée (Figure 12). C'est ainsi que l'éribuline a pu être obtenue, puis commercialisée sous le nom de Halaven.

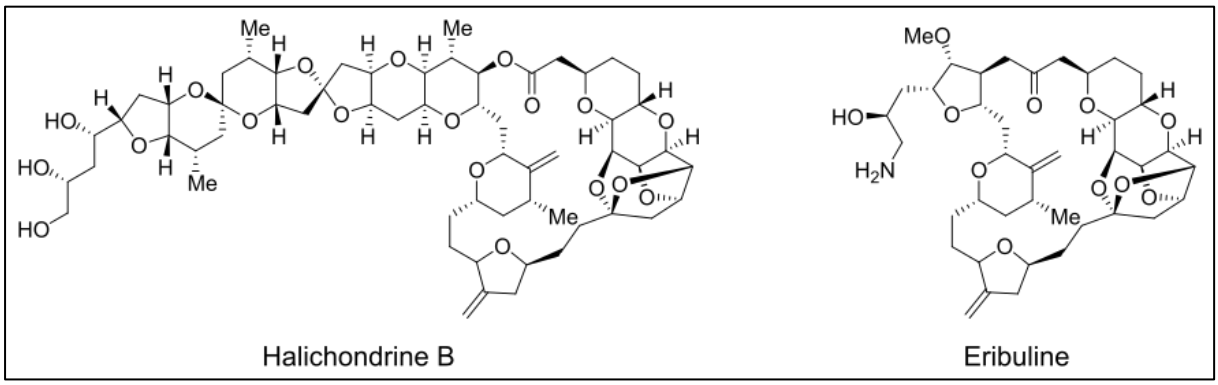


Figure 12. Structure de l'halichondrine B et de l'éribuline (adapté de Nicolaou et al., 2022). L'halichrondrine B est produite par une éponge marine mais peut difficilement en être extraite. Alors la synthèse chimique a permis de reconstruire le composé. L'éribuline est un dérivé plus petit de l'halichondrine B et donc plus simple à synthétisé.

Simplifier les structures des molécules d'intérêt thérapeutique permet alors de simplifier leur synthèse, et c'est dans cette optique que des études de simplification de polycétides ont été menées, qui seront présentés plus tard dans ce manuscrit.

I.D.4. L'ingénierie génétique des producteurs de produits naturels

Une troisième approche est également investiguée pour obtenir de nouvelles molécules : l'ingénierie génétique. En 1990, le cluster de gènes lié à la synthèse de l'érythromycine est caractérisé (Cortes et al., 1990). Ce cluster code une PKS modulaire, un complexe d'enzymes produisant les polycétides par addition d'unités acyl-CoA. Ces systèmes ont un mode de fonctionnement linéaire, dans le sens où le squelette du polycétide produit peut être déduit à partir du nombre de modules composant les PKS et la manière dont ils sont arrangés. Cette particularité permet d'envisager l'ingénierie de ces systèmes afin de produire des dérivés des produits existants (Weissman, 2016). Cette partie fera l'objet de plus amples explications par la suite.

II. Les polycétide synthases

En 1833, le journal Annales de chimie et de physique publie un article de Anselme Payen et Jean-François Persoz décrivant la présence de « diastase » dans des extraits de malt, transformant l'amidon en saccharides (Payen & Persoz, 1833). L'analyse de ces diastases dans les levains, et plus précisément les levures, a permis de les nommer enzymes, du grec *en* (dans) et *zumè* (levain).

Eduard Buchner découvre ensuite que ces enzymes peuvent fonctionner en dehors de leur organisme hôte, *in vitro*, ce qui lui valut le prix Nobel de Chimie en 1907 (Jaenicke, 2007). Il faut cependant attendre 1975 avant que la première enzyme cataloguée comme PKS soit découverte : la chalcone synthase. Celle-ci produit un flavonoïde et a une structure apparentée aux PKS de type III (Kreuzaler & Hahlbrock, 1975).

Les polycétides sont des métabolites spécialisés produits par des bactéries (Liu et al., 2013), des champignons (Herbst et al., 2018), des plantes (Katsuyama et al., 2009), des animaux (Torres & Schmidt, 2019) ou encore des parasites (D'Ambrosio et al., 2023). Au-delà de leurs rôles au sein des organismes hôtes, ils peuvent avoir des propriétés thérapeutiques diverses et intéressantes pour les industries agricole et pharmacologique : des antibiotiques (érythromycine A), des immunosuppresseurs (rapamycine), des antitumoraux (pédérine), des pesticides (spinosyne) ou encore des anticholestérolémiants (lovastatine).

Les polycétides sont produits par l'addition séquentielle d'unités acyl-CoA, catalysée par les PKS. La grande diversité structurale des polycétides est due à la nature des unités acyl-CoA sélectionnées, à la longueur des chaînes polycétidiques synthétisées, au degré de modification des fonctions cétones de chaque unité, à la diversité stéréochimique des composés et aux transformations chimiques effectuées par la suite sur le squelette polycétidique, par des enzymes dites post-PKS. Bien que les PKS soient retrouvées dans toutes les branches du règne du vivant, leurs structures et mécanismes réactionnels diffèrent selon le règne et les organismes étudiés (Figure 13).

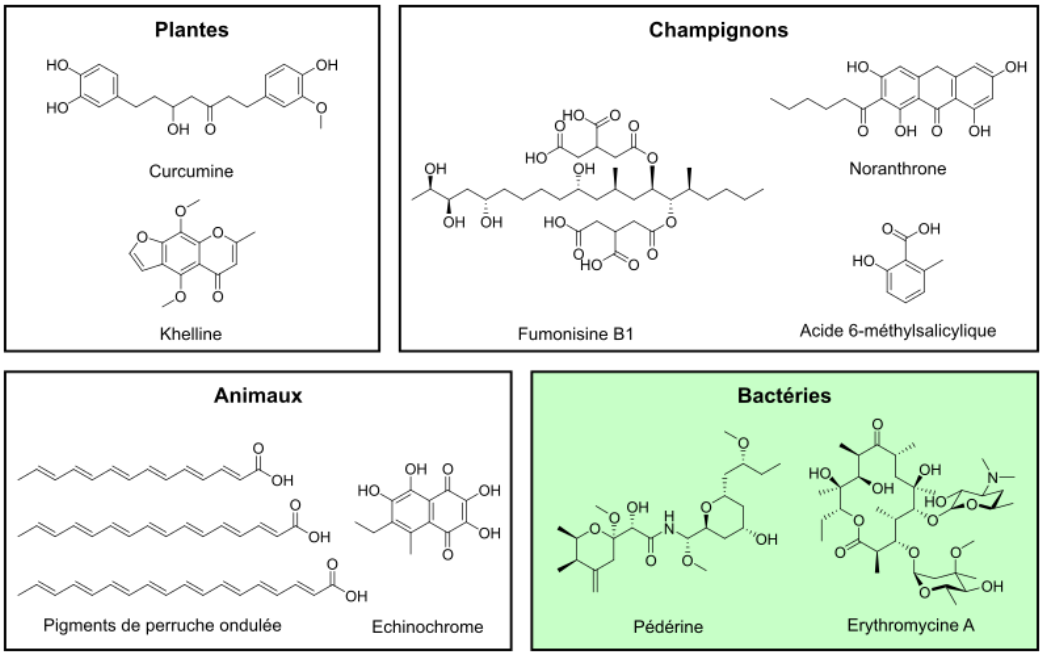


Figure 13. Exemples de polycétides. Les polycétides produits par des bactéries, indiqués en vert, sont au cœur du sujet de cette thèse.

La diversité structurelle des polycétides provient d'une diversité des enzymes PKS. Il existe plusieurs classes de PKS, selon leur mécanisme d'action et leur composition enzymatiques. Il existe des PKS itératives, composées d'un unique set d'enzymes catalysant l'ajout d'unités acylées, parmi lesquelles sont classées les PKS de type II, de type III et de type I itératives. Il existe également des PKS de type I modulaires, dont le mécanisme d'action n'est pas itératif et au sein desquelles chaque domaine agit une unique fois sur le polycétide en élongation.

II.A. Les polycétide synthases de type II

Les PKS de type II sont des enzymes constituées d'un hétérodimère de domaines cétosynthase (KS) α et β, ainsi que d'un domaine porteur d'acyl (ACP), agissant de manière itérative (Figure 14).

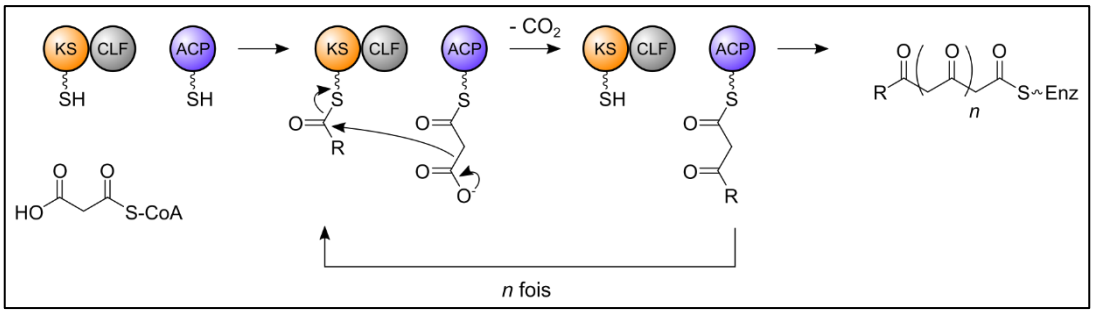


Figure 14. Polycétide synthase de type II (adapté de Chen et al., 2018 et Weissman, 2009). Les PKS de type II sont composées d'un duo d'enzymes cétosynthase (KSα / KSβ). L'unité KSα catalyse la condensation des unités acyl-CoA. L'unité KSβ quant à elle génère l'acétate qui initie la synthèse et régule la taille du polycétide produit. Il est appelé facteur d'élongation de la chaîne (CLF). Un domaine porteur d'acyle (ACP) discret porte le polycétide en élongation.

L'enzyme KSα a pour rôle de catalyser la condensation d'une nouvelle unité acyl-CoA au polycétide en cours d'élongation, porté par le domaine ACP par l'intermédiaire d'un bras 4'-phosphopantéthéine (P-Pant) (Weissman, 2009). La seconde enzyme KSβ quant à elle joue un rôle dans l'initiation de la synthèse en décarboxylant l'unité de départ et dans la régulation de la taille de la chaîne polycétidique (CLF).

L'unité de base de la synthèse des polycétides par les PKS de type II sont majoritairement des acétyl-CoA, mais elle peut également être un composé plus complexe comme des unités propyionyl-CoA ou butyryl-CoA, ou même des petites molécules d'acides gras. Le recrutement de telles unités de départ non-acylées est médié par des enzymes appelées 3-cétoacyl-ACP synthase (KASIII). Dans le cas de la synthèse de la spirocycline A chez *Streptomyces spectabilis*, une PKS de type II catalyse l'élongation d'une unité de départ provenant d'une PKS de type I itérative (Zhao et al., 2024).

A la fin de la voie de biosynthèse, le polycétide produit est libéré par l'action d'enzymes portant des activités cétoréductase (KR), cyclase (CYC) et aromatase (ARO). Les polycétides sont également modifiés par des oxygénases, des glycosyltransférases et des *N*–, *O*– et *C*–méthyltransférases.

II.B. Les polycétide synthases de type III

Les PKS de type III sont également des enzymes itératives, composées d'un unique domaine KS formant un homodimère, catalysant la condensation d'une à sept unités d'acyl-CoA (Figure 15) (Shimizu et al., 2017). Ces PKS sont majoritairement retrouvées au sein de bactéries et de plantes, et plus rarement au sein de champignons (Austin & Noel, 2003; Shimizu et al., 2017). La biosynthèse des polycétides par les PKS de type III est indépendante du domaine ACP. Ici, la KS possède une triade catalytique différente des KS des autres types de PKS. Chez les PKS de type III, la triade catalytique est composée des résidus Cys-His-Asn (CHN), caractéristique de la classe des thiolases d'initiation.

Au sein des autres types de PKS, cette triade est composée des résidus Cys-His-His (CHH), et dans ce cas ces KS appartiennent à la famille des thiolases d'élongation.

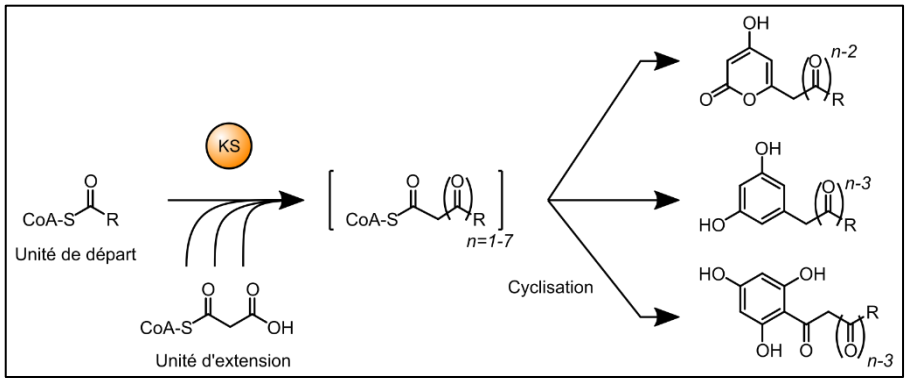


Figure 15. Polycétide synthase de type III (adapté de Weissman, 2009). Les PKS de type III sont composées d'un unique domaine cétosynthase (KS), qui catalyse l'addition d'une à sept unités acyl-CoA à partir d'une unité de départ.

Malgré l'apparence simple de ces enzymes, plusieurs mécanismes permettent la diversité des produits obtenus. Ils peuvent varier par la nature de l'unité initiale recrutée, le nombre de cycles d'élongation et enfin la variabilité de la cyclisation intramoléculaire (Figure 16) (Austin & Noel, 2003).

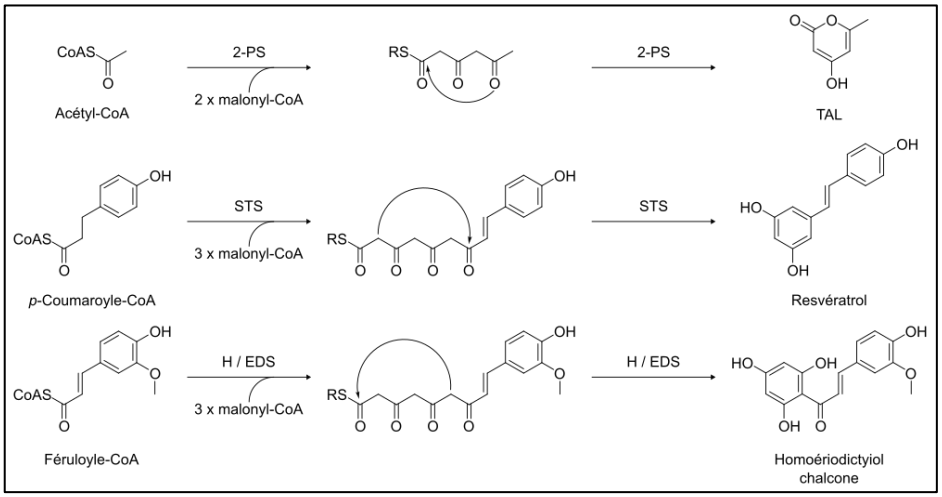


Figure 16. Synthèse de polycétides par des polycétide synthases de type III (adapté de Austin & Noel, 2003). La variété au sein de la structure des polycétides produits par les PKS de type III provient de l'unité initiale sélectionnée, le nombre de cycles d'élongation et le mécanisme de cyclisation final. **Légende** – 2-PS : 2-pyrone synthase, STS : stilbene synthase, H/EDS : homoériodictyol / ériodictyol synthase, TAL : acide lactone triacétique.

II.C. Les polycétides synthases de type I

Les PKS de type I sont des enzymes plus complexes, composés de plusieurs domaines. Il existe deux familles selon que la synthèse des polycétides se fasse de manière itérative ou non.

II.C.1. Les PKS de type I itératives

Les PKS itératives de type I (iPKS) sont semblables aux FAS retrouvées chez les mammifères (mFAS), tant dans leur structure que leur mécanisme enzymatique (Herbst et al., 2018). La différence majeure entre ces deux systèmes enzymatiques est la nature de la molécule produite. Les mFAS produisent des acides gras, qui sont des composés complètement saturés, tandis que les iPKS sont capables de produire des composés plus ou moins saturés et d'apporter des modifications aux carbones α et β des fonctions cétones. De plus, les acides gras sont des composants du métabolisme primaire des organismes, tandis que les polycétides sont classés au sein du métabolisme spécialisé.

Les iPKS sont des complexes dimériques composés d'un domaine KS et ACP, mais également d'autres domaines spécifiques aux PKS de type I. Ici, le recrutement de la nouvelle unité acylée se fait pas le biais d'un domaine acyltransférase (AT). L'ensemble des domaines KS, AT et ACP constitue ce qui est appelé un module. Un module est composé *a minima* de ces trois domaines, mais peut également contenir des domaine supplémentaires et optionnels, qui permettent notamment de modifier la fonction cétone de l'unité acylée recrutée.

Les iPKS sont des enzymes ubiquitaires et sont divisées en trois catégories : les iPKS bactériennes, les iPKS de type acide mycocérosique synthase retrouvées chez les Mycobactéries et les iPKS de champignons. Parmi les iPKS de champignons, trois sous-classes ont pu être caractérisées selon leur activité réductrice : hautement réductrices (HR), partiellement réductrices (PR) ou non-réductrices (NR) (Herbst et al., 2018). Les iPKS HR sont composées des trois domaines essentiels KS, AT et ACP, mais également de trois domaines supplémentaires assurant la réduction des fonctions cétones (Figure 17A). Ces activités sont portées par des domaines cétoréductase (KR), énoyl-réductase (ER) et déshydratase (DH). Les polycétides obtenus sont donc complètement saturés et peuvent aussi servir de précurseurs à la synthèse d'acide gras rares. Les iPKS PR quant à elles ne contiennent pas le domaine ER permettant la réduction totale de la fonction cétone, et sont ainsi partiellement réductrices (Figure 17B). Enfin, les iPKS NR ne possèdent pas de domaines permettant la modification de la fonction cétone (Figure 17C). La plupart de ces PKS contiennent des domaines – actifs ou non – qui catalysent la méthylation du Cα (CMeT). A la fin de la voie de biosynthèse, la libération du polycétide peut être catalysée par l'action d'un domaine thioestérase (TE) (Figure 17C).

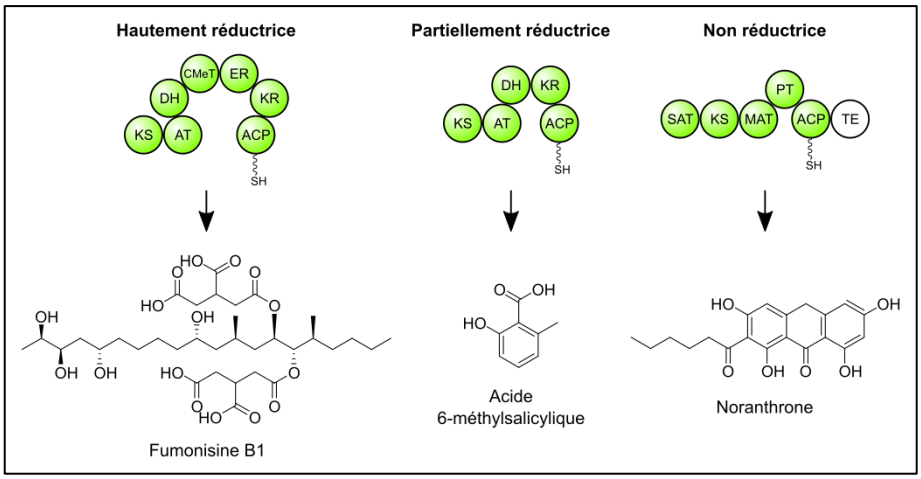


Figure 17. Exemples de polycétide synthases de type I itératives de champignons (adapté de Herbst et al., 2018). Les PKS de type I itératives (iPKS) sont composées d'un unique module, qui catalyse l'addition successive d'unités acyl-CoA. Un module minimal est composé d'un domaine cétosynthase (KS) qui catalyse la formation de la liaison entre le polycétide naissant et l'unité d'acyl-CoA incorporée, un domaine acyltransférase (AT) qui sélectionne cette unité et un domaine porteur d'acyle (ACP) auquel le polycétide en élongation est rattaché. Dans le cas des iPKS non réductrices, le domaine AT qui sélectionne la première unité acyl-CoA est nommé *starter* AT (SAT) et lorsque le domaine AT recrute un malonyl-CoA, il est nommé MAT. Des domaines supplémentaires peuvent être retrouvés, permettant, par exemple la modification de la fonction cétone – cétoréductase (KR), déshydratase (DH) et énoyl-réductase (ER) – mais aussi la C-méthylation des unités nouvellement incorporées (CMeT) ou la cyclisation du produit final (PT). Enfin, au sein des PKS NR, un domaine thioestérase (TE) libère le polycétide.

II.C.2. Les PKS de type I modulaires

Parmi les PKS de type I, sont également retrouvées des enzymes modulaires (Weissman, 2009). Celles-ci sont dimériques et composées de plusieurs modules, chacun catalysant l'ajout d'une unité acyl-CoA. Le mécanisme n'est donc plus itératif, comme c'est le cas pour les iPKS, les PKS de type II et de type III. Il existe au sein des PKS modulaires de type I deux types, les *cis*-AT et les *trans*-AT, selon que l'activité AT est intrinsèque à la chaîne polypeptidique (*cis*-AT) (Aparicio et al., 1996; Gomez-Escribano et al., 2012; Laureti et al., 2011) ou portée par une enzyme discrète (*trans*-AT) (Butcher et al., 2007; Pulsawat et al., 2007). Afin d'assurer la voie de biosynthèse des polycétides, les sous-unités des PKS modulaires doivent interagir entre elles dans un ordre prédéfini et unique. Pour cela, des domaines peuvent être retrouvés à la jonction des sous-unités qui médient ces interactions et sont appelés des domaines de docking (DD). D'autres interactions peuvent permettre l'assemblage des sous-unités chez les PKS *trans*-AT, celles-ci seront discutées par la suite. Les polycétides peuvent être libérés sous forme linéaire ou cyclique, et dans ce dernier cas ils sont alors nommés macrolides.

Certaines PKS de type I peuvent également contenir des modules de type NRPS et sont alors dites hybrides (Chen et al., 2017; Pulsawat et al., 2007). Ces modules permettent l'ajout d'acides aminés au sein de la structure du polycétide.

Ils sont composés *a minima* d'un domaine d'adénylation (A) qui recrute l'acide aminé à ajouter, d'un domaine de condensation (C) qui catalyse la formation de la liaison entre le composé en production et le résidu sélectionné, et enfin d'un domaine porteur de peptide (PCP), analogue au domaine ACP des PKS.

II.C.2.a. Les PKS de type *cis*-AT

Les PKS modulaires *cis*-AT sont des enzymes qui possèdent des domaines AT intrinsèques aux chaînes polypeptidiques des modules. Un exemple de polycétide produit par une PKS *cis*-AT est la coelimycine. La coelimycine est produite par une PKS nommée Cpk, retrouvée chez *S. coelicolor*. Elle est composée de cinq modules d'élongation répartis en trois sous-unités (Figure 18) (Gomez-Escribano et al., 2012).

Le module de chargement (MC) recrute la première unité – dans ce cas un acétyl-CoA – qui sera transférée au domaine ACP du module 1. Le domaine KS du module MC est ici inactif car à la place de la cystéine catalytique se trouve un acide glutamique. Au sein du module 1, une unité d'élongation est sélectionnée par le domaine AT. Le domaine KS va ensuite catalyser la formation d'une liaison carbone-carbone entre l'unité portée par le domaine ACP et celle recrutée par le domaine AT, par une réaction de condensation dite de Claisen.

Cette réaction consiste en l'assemblage de deux β-céto-esters, catalysée par des résidus histidine et cystéine présent dans la poche catalytique des domaines KS. Le domaine KR présent au sein du module 1 réduit la fonction cétone portée par l'unité recrutée par le précédent module en fonction hydroxyle.

Enfin, le domaine DH entraîne la déshydratation de la fonction hydroxyle et la formation d'une double liaison carbone-carbone. Un domaine ER peut être présent en complément des domaines KR et DH, qui permet la réduction de cette double liaison (absent au sein de la coelimycine synthase). Le polycétide en cours de formation, alors constitué de deux unités, est ensuite transféré au prochain module d'élongation pour poursuivre sa synthèse. A la fin de la voie de biosynthèse, un domaine thioester-réductase (TR) catalyse la libération du précurseur produit par la PKS. La libération d'un polycétide par un domaine TR permet d'obtenir un composé aldéhyde ou alcool, tandis qu'un domaine TE catalyse la formation d'un acide carboxylique. La coelimycine A est un exemple de polycétide libéré par un domaine TR, amenant la formation d'un groupement aldéhyde.

Ce composé est ensuite pris en charge par des enzymes post-PKS afin d'obtenir la coelimycine A (Gomez-Escribano et al., 2012). CpkE est une α,β hydrolase considérée comme catalyseur potentiel pour la première modification post-PKS : l'isomérisation du composé. CpkG est une amino-transférase qui substitue la fonction cétone par un groupement amine.

CpkD, Scf et CpkH sont des époxydases permettant la formation de structure époxydes au sein du polycétide et la cyclisation spontanée du composé. Toutes ces modifications permettent d'obtenir la coelimycine A, qui est un antibiotique incolore.

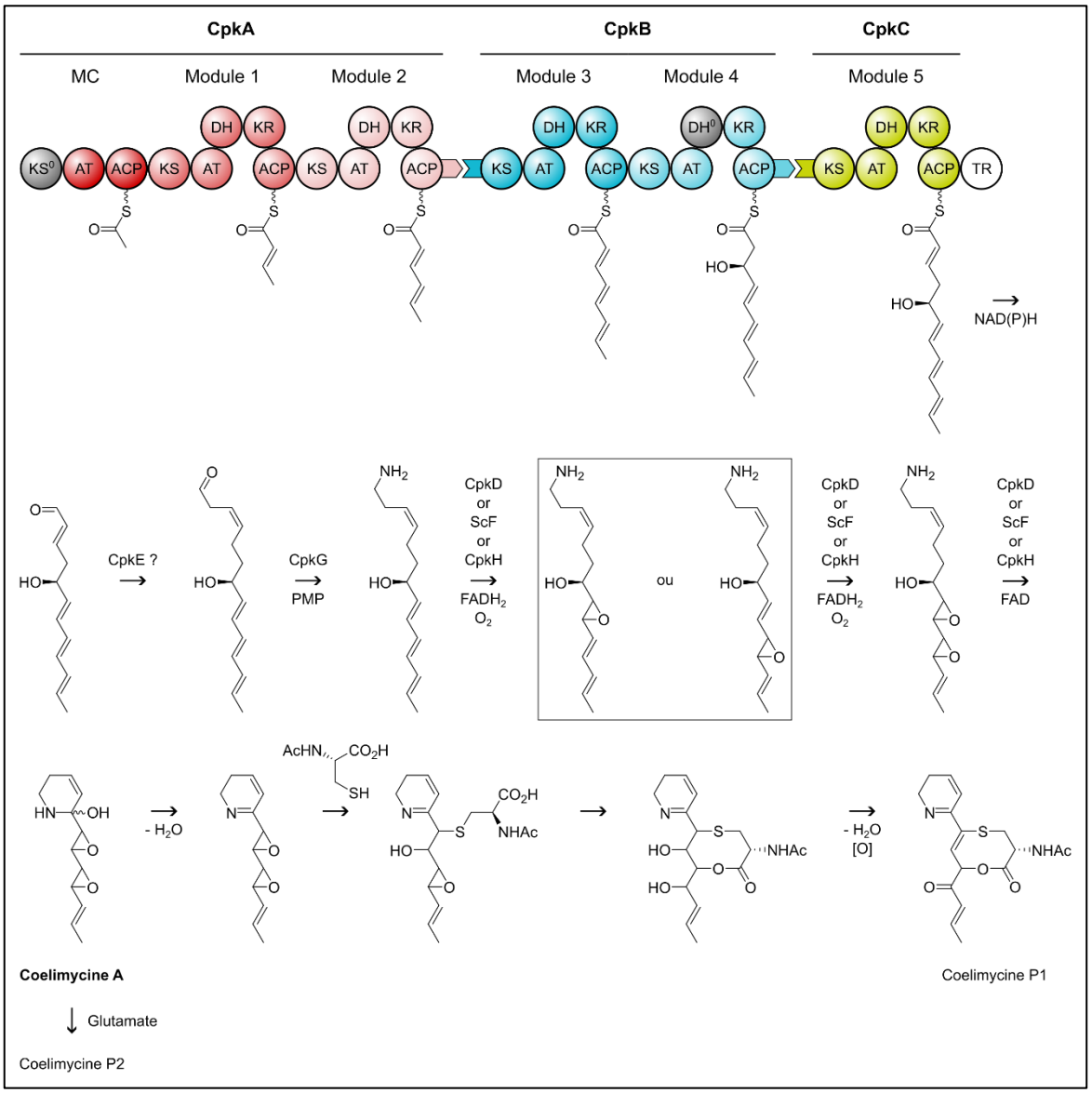


Figure 18. Voie de biosynthèse proposée de la coelimycine chez *Streptomyces coelicolor* (adapté de Gomez-Escribano et al. 2012). Les PKS de type I modulaires sont composées de plusieurs modules, catalysant chacun l'addition d'une unité acyl-CoA sur le polycétide en cours de synthèse. Le module de chargement (MC) recrute la première unité, ici un acétyl-CoA. Le domaine acyltransférase (AT) recrute la prochaine unité d'extension et le domaine cétosynthase (KS) catalyse sa liaison au polycétide en élongation. Ce dernier est rattaché à un domaine porteur d'acyle (ACP). D'autres domaines peuvent être présents, ayant pour activités de modifier les fonctions cétones : cétoréductase (KR), déshydratase (DH) et énoyl-réductase (ER). Enfin, un domaine thioester-réductase (TR) libère le polycétide, catalysant ainsi dans la majeure partie des cas la cyclisation du composé. Le composé produit est ensuite pris en charge par des enzymes post-PKS afin d'obtenir la coelimycine A, qui est un antibiotique. Ce produit est ensuite excrété dans le milieu, dans lequel il peut réagir de manière instantanée avec le glutamate ou la N-acétylcystéine afin de produire respectivement les coelimycines P2 et P1. La structure de la coelimycine P1 a été déterminée par spectrométrie de masse, tandis que celle de la coelimycine P2 reste à ce jour inconnue. Les domaines indiqué en gris (KS^0 et DH^0) sont des domaines catalytiquement inactifs.

La coelimycine A est ensuite exportée par le transporteur associé au système : CpkF. La coelimycine A peut alors réagir spontanément avec de la N-acétylcystéine ou de l'acide glutamique afin de former respectivement les coelimycines P1 et P2 (Gomez-Escribano et al., 2012; Gottelt et al., 2010).

La règle de colinéarité qui peut ainsi être établie entre le nombre, la composition et l'organisation des modules d'une PKS et la structure du polycétide qu'elle produit a inspiré des décennies d'ingénierie dans le but d'obtenir de nouvelles molécules.

II.C.2.b. Les PKS de type *trans*-AT

Les PKS *trans*-AT constituent la seconde classe de PKS de type I modulaires. Elles sont nommées ainsi car le domaine AT, présent au sein de la chaîne polypeptidique des PKS *cis*-AT, est ici extrinsèque et porté par une enzyme discrète. Elles sont remarquables notamment car elles possèdent une large gamme d'activités processives et contiennent souvent des modules NRPS. Ces derniers incorporent des acides aminés dans la chaîne en élongation (Helfrich & Piel, 2016).

La PKS qui synthétise la bacillaene chez *Bacillus subtilis* est l'une des premières PKS *trans*-AT hybride identifiée (Figure 19) (Butcher et al., 2007). Elle est composée de cinq sous-unités et de dix-huit modules d'élongation. A l'instar des PKS *cis*-AT, les modules catalysant l'ajout d'unités acylées sont *a minima* constitués de domaines KS et ACP. Cependant, l'activité AT est portée par une ou plusieurs enzymes discrètes, ici PksD et PksE. Cette dernière protéine porte également une activité oxydase (Ox) absente chez les PKS de type *cis*-AT. D'autres activités sont également retrouvées exclusivement chez les PKS *trans*-AT, telles que des domaines méthyltransférase (MT) et hydrolase (Hyd) (Butcher et al., 2007). Ceux-ci sont pointés d'une flèche rouge sur la Figure 19. Les acides aminés ajoutés par les modules NRPS sont ici une glycine et une alanine, respectivement représentés en orange et violet.

Certains modules ne catalysent pas l'ajout d'unité acylée ou peptidique. Ils peuvent cependant avoir d'autres rôles. Les modules KS^0–DH–ACP par exemple catalyse la déshydratation des fonctions hydroxyles (Butcher et al., 2007). Le module KS^0–DH*–ACP de PksR quant à lui entraîne la migration de la double liaison carbone-carbone (Butcher et al., 2007; Helfrich & Piel, 2016). Ce phénomène migratoire a lieu lorsqu'un domaine déhydratase-*like* DH* est présent au sein d'un module ne permettant pas d'ajouter de nouvelle unité acyl (notamment car le domaine KS est inactif) (Kusebauch et al., 2010). La liaison concernée ici par cette action est représentée dans un ovale (Figure 19). Une fois le produit hybride polycétide-NRP produit, celui-ci est libéré par un domaine TE et la bacillaene est ainsi obtenue (Butcher et al., 2007). Elle peut être modifiée par un cytochrome P450 (PksS), afin d'obtenir la bacillaene B (Helfrich & Piel, 2016).

Il est également possible que d'autres enzymes discrètes agissent sur le polycétide en cours d'élongation. C'est le cas ici des enzymes PksC et PksF – PksI. Elles font partie d'une cassette de β-méthylation et permettent l'ajout d'un groupement –CH3 au cours de l'activité de PksL (Butcher et al., 2007). Ce groupement est représenté en rouge dans la Figure 19.

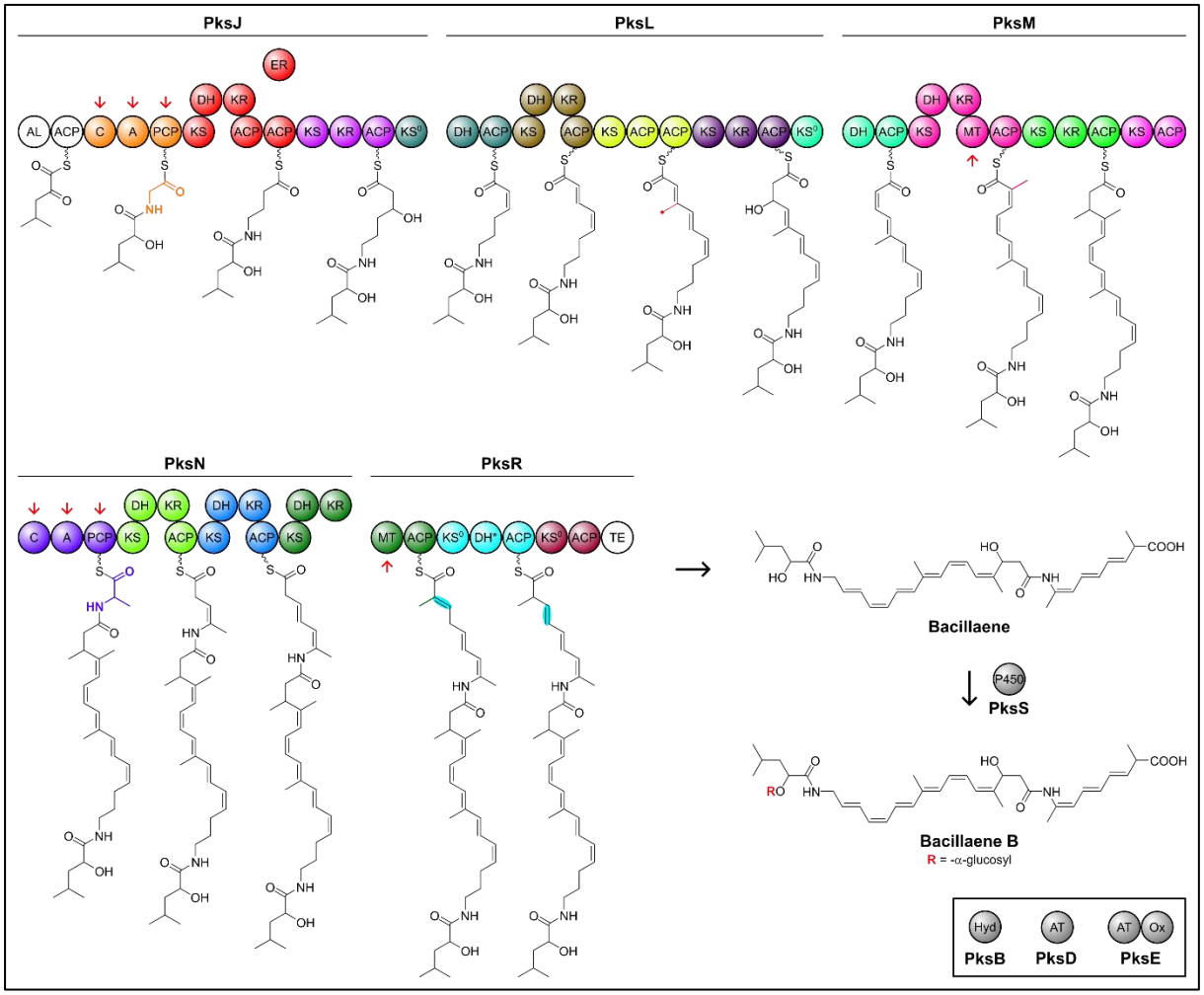


Figure 19. Voie de biosynthèse de la bacillaene par une PKS *trans*-AT retrouvée chez *Bacillus subtilis* (adapté de Butcher et al., 2007 et Helfrich & Piel, 2016). La PKS qui synthétise la bacillaene chez *B. subtilis* est composée de cinq sous-unités (PksJ, PksL–N et PksR). Elle catalyse l'assemblage de onze unités acylées et deux acides aminés. Comme les PKS *cis*-AT, les modules catalysant l'ajout d'unité acylée sont composées *a minima* d'un domaine cétosynthase (KS) et d'un porteur d'acyle (ACP). La différence ici est que l'activité acyltransférase (AT) est portée par les enzymes discrètes PksD et PksE. Des modules permettent l'ajout d'acides aminés et sont composés d'un domaine de condensation (C) analogue aux KS, d'adénylation (A) analogue aux AT et d'un domaine porteur de peptide (PCP) analogue aux ACP. Ces domaines sont pointés d'une flèche rouge et les acides aminés incorporés sont colorés en orange et violet. Les domaines méthyltransférase (MT) sont aussi pointés d'une flèche rouge car absent chez les PKS *cis*-AT, et les groupements méthyles ajoutés sont colorés en rose et vert foncé. Un groupement méthyle est également ajouté au polycétide en élongation au sein de la sous-unité PksL par une cassette de méthylation. Ce groupement est représenté en rouge. **Légende** – AL : acyl-ligase, DH : déshydratase, KR : cétoréductase, ER : énoyl-réductase, KS^0 : domaine KS inactif, MT : méthyltransférase, DH* : domaine de migration déshydratase-*like*, TE : thioestérase, P450 : cytochrome P450, Hyd : hydrolase, AT : acyltransférase, Ox : oxydase.

La cassette de β-méthylation est ici composée d'un domaine ACP (AcpK), d'un domaine AT (PksC), d'un domaine KS (PksF), d'une activité hydroxyméthylglutaryl-CoA synthase (HCS) et de deux enzymes énoyl-CoA hydratase / isomérase (ECH) (Figure 20) (Butcher et al., 2007; Nair et al., 2020). D'autres systèmes PKS *trans*-AT sont associés à des cassettes de méthylation, comme la PKS de la virginiamycine retrouvée chez *S. virginiae* (Pulsawat et al., 2007).

Le point commun de ces cassettes est leur champ d'activité, dans le sens où elles agissent au niveau de modules PKS contenant deux domaines ACP (Butcher et al., 2007; Pulsawat et al., 2007).

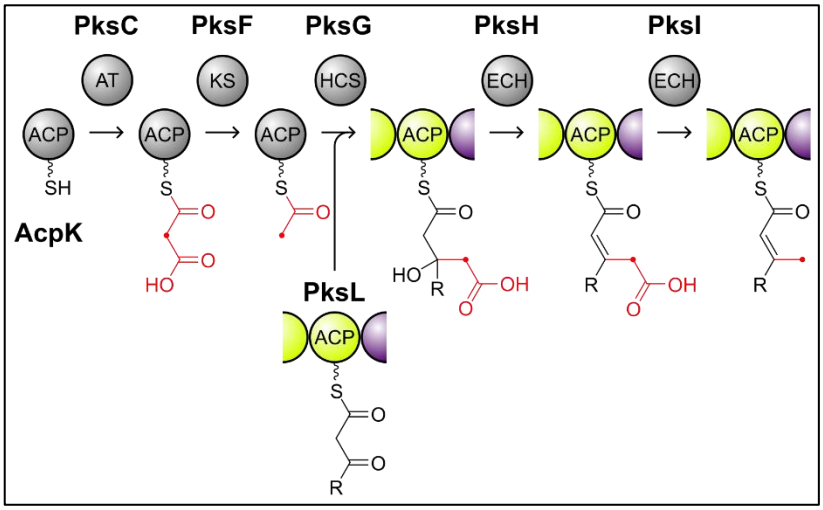


Figure 20. Cassette de β-méthylation de la bacillaene synthase (adapté de Butcher et al., 2007). La cassette de β-méthylation qui agit sur le polycétide porté par un domaine porteur d'acyle (ACP) de la sous-unité PksL est composée de cinq enzymes discrètes et d'un domaine ACP transitoire (AcpK). Un domaine acyltransférase (AT) porté par l'enzyme PksC recrute l'unité acyl sur le domaine AcpK. L'activité successive d'un domaine cétosynthase (KS) PksF, d'une hydroxyméthylglutaryl-CoA synthase (HCS) PksG et de deux enzymes énoyl-CoA hydratase / isomérase (ECH) PksH et PksI permettent l'ajout du groupement méthyle au polycétide en élongation porté par la sous-unité PksL.

II.C.3. Les modules itératifs au sein de PKS modulaires

La règle de colinéarité établie entre la composition en domaines et modules d'une PKS et la structure de son produit final connaît cependant quelques exceptions. Au sein des PKS *cis*-AT, cette exception touche notamment le système qui synthétise l'azalomycine F (Figure 21) (Luo et al., 2024; Zhai et al., 2023). Le cluster de gènes qui lui est associé contient huit gènes qui codent les huit sous-unités, parmi lesquelles sont répartis 20 modules d'élongation. En appliquant la règle de colinéarité inhérente aux PKS de type *cis*-AT, la structure de l'azalomycine F a pu être prédite, mais ne correspond cependant pas à la réalité du produit obtenu (Hong et al., 2016). L'analyse du produit réel obtenu semble montrer qu'un module d'élongation manquerait au début de la chaîne de biosynthèse.

Cela s'explique par le caractère itératif du premier module de l'azalomycine F synthase, qui ajoute les première et seconde unités acylées au polycétide en élongation. Ce premier module est par la suite nommé module 1/2, suivi du module 3, *etc.*

Au-delà de cette caractéristique, l'azalomycine F synthase présente une autre particularité. En effet, le module 1/2 contient un domaine énoyl-réductase, qui n'agit que lors de la seconde élongation catalysée. Le module 3 est quant à lui dépourvu de domaine ER, et pourtant la structure de l'azalomycine F montre que ce module devrait vraisemblablement contenir ce domaine or ce n'est pas le cas. La délétion du domaine ER du module 1/2 a permis d'identifier son rôle dans la modification de la fonction cétone prédite au module 3 (Zhai et al., 2020).

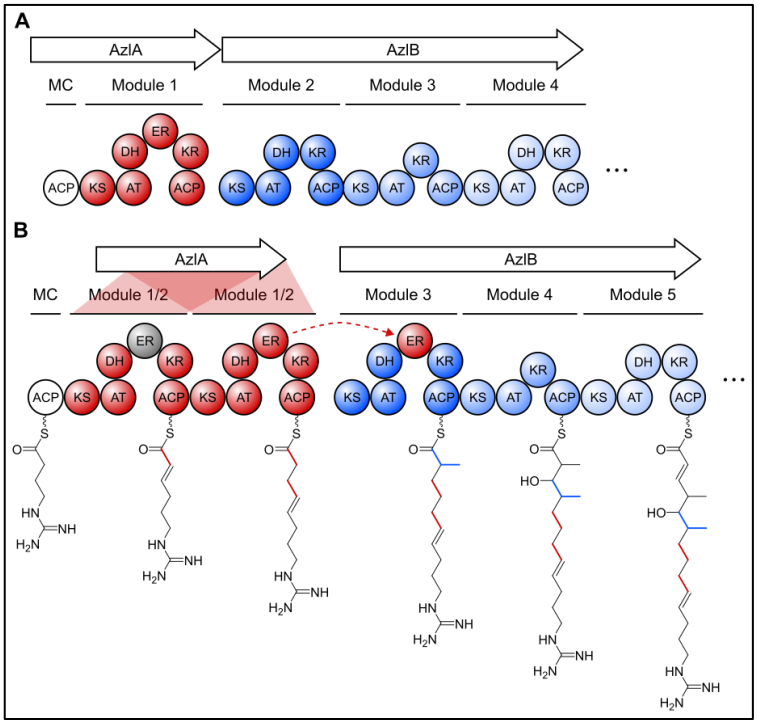


Figure 21. Représentation des deux premières sous-unités de l'azalomycine synthase (adapté de Zhai et al., 2020). **(A)** Prédiction de la composition en domaine des sous-unités à partir du cluster de gènes. **(B)** Composition réelle des sous-unités à partir de la structure de l'azalomycine F et des expériences de délétions. Le module 1 agit de manière itérative et est donc nommé module 1/2, le module 2 devient donc le 3, *etc* (Hong et al., 2016). La structure de l'azalomycine F permet de prédire la présence d'un domaine énoyl-réductase (ER) au sein du module 3, qui est cependant absent dans le gène qui code ce module. Des analyses bioinformatiques et la délétion du domaine ER du module 1/2 a permis d'établir son rôle dans la modification effectuée par le module 3 (Zhai et al., 2020). **Légende** – MC : module de chargement, ACP : domaine porteur d'acyle, KS : cétosynthase, AT : acyltransférase, DH : déshydratase, KR : cétoréductase.

Il existe également quelques exceptions au sein des PKS *trans*-AT. C'est le cas de la synthèse du polycétide associé à la nocardiose (NOCAP), maladie infectieuse causée par des bactéries du genre *Nocardia*. La synthèse de NOCAP implique une PKS hybride dont certains modules sont *cis*-AT et d'autres sont *trans*-AT (Guzman et al., 2021).

Cette PKS est composée d'un module de chargement et de huit modules d'élongation. Le module 5 de cette voie de biosynthèse se trouve scindé entre deux sous-unités. Le domaine KS se trouve en position C-terminale d'une sous-unité et les domaines DH–ACP–KR en N-terminal de la sous-unité suivante.

Ce module permet l'ajout de trois unités acylées au polycétide en élongation et non une seule (Guzman et al., 2021). Cela peut être dû à une faible affinité de substrat de la KS présente dans le module en question. En effet, l'équipe de Keasling *et coll.* a étudié la spécificité de substrat de la KS du module 5 de la borrelidine synthase, caractérisée comme catalysant trois cycles d'élongation également (Curran et al., 2018). Ils ont ainsi pu démontrer que ce domaine acceptait en effet une large gamme de substrats non natifs, pouvant donc expliquer sa capacité à accommoder un même polycétide en élongation durant trois cycles de suite.

III. Structure, mécanisme d'action et évolution des domaines de polycétide synthases

Chaque domaine au sein des PKS a une structure et une fonction spécifique et conservée entre les types de PKS. Comprendre la structure de chaque domaine permet de mieux appréhender leur mécanisme d'action et envisager des études d'ingénierie des PKS.

III.A. Le domaine porteur d'acyle (ACP)

Bien que dépourvu d'activité enzymatique, le domaine ACP est central dans la synthèse des polycétides (Figure 22).

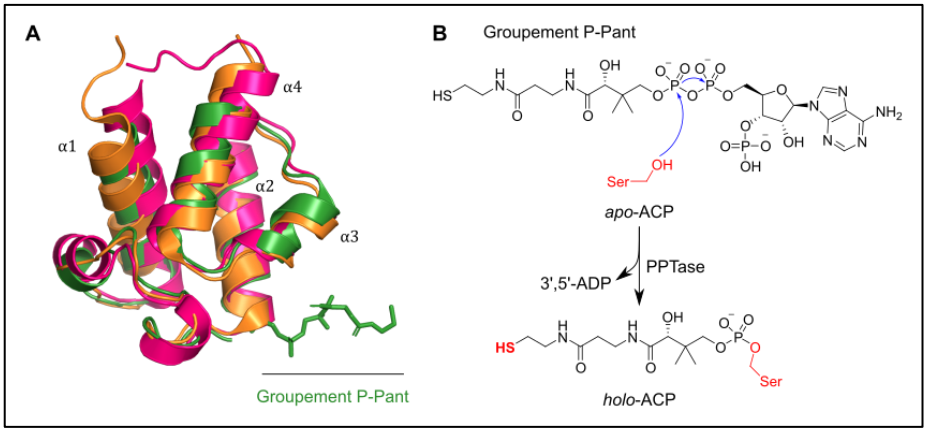


Figure 22. Alignement des structures de trois domaines porteurs d'acyl (ACP) et mécanisme de greffe du bras 4′-phosphopantéthéine (P-Pant) au domaine ACP. **(A)** Alignement de structures de domaines ACP. Les domaines ACP ont une structure reconnaissable constituée de quatre hélices α et possèdent une sérine conservée à laquelle est greffé un groupe 4′-phosphopantéthéine (P-Pant). La structure de la forme holo-ACP provient de l'organisme *Helicobacter pylori* P12 (vert ; PDB 5H9H), et les deux apo-ACP proviennent de *Thermotoga maritima* MSB8 (orange ; PBD 1VKU) et *Kuenenia stuttgartiensis* (rose ; PDB 7AUF) (Dietl & Barends, 2022). **(B)** Mécanisme de greffe du bras P-Pant au domaine ACP (adapté de Lambalot et al. 1996). Une enzyme, la phosphopantéthéinyl transférase (PPTase) permet le transfert du groupement P-Pant d'un groupement adénosine diphosphate (3′,5′-ADP) vers la sérine conservée du domaine ACP. C'est à ce bras P-Pant que le polycétide en élongation pourra se lier de manière covalente, *via* le groupement thiol ainsi ajouté.

Il est présent au sein des PKS de type I et II, et est intégré au sein d'un module ou porté par une protéine discrète respectivement. C'est ce domaine qui va porter les intermédiaires du polycétide en élongation et assure les interactions avec les domaines catalytiques dans le cas des PKS de type I. Ces domaines sont composés de quatre hélices α principales retrouvées au sein de toutes les structures d'ACP résolues (Figure 22A).

Le domaine ACP en lui-même n'est pas capable de porter le polycétide en élongation. Il doit subir la greffe du bras P-Pant, qui est ajouté par une phosphopantéthéine transférase (PPTase) (Figure 22B). Cette greffe se fait sur un résidu sérine conservé au sein des ACP, en position N-terminale de la seconde hélice α. Ainsi, l'extrémité du résidu devient un groupement thiol à la place d'un hydroxyle. Le polycétide peut alors être lié de manière covalente à l'atome de soufre ainsi présent.

III.B. Le domaine acyltransférase (AT)

Les domaines AT ont la spécificité de ne pas être présents au sein des chaînes polypeptidiques de toutes les sous-unités de PKS (Figure 23).

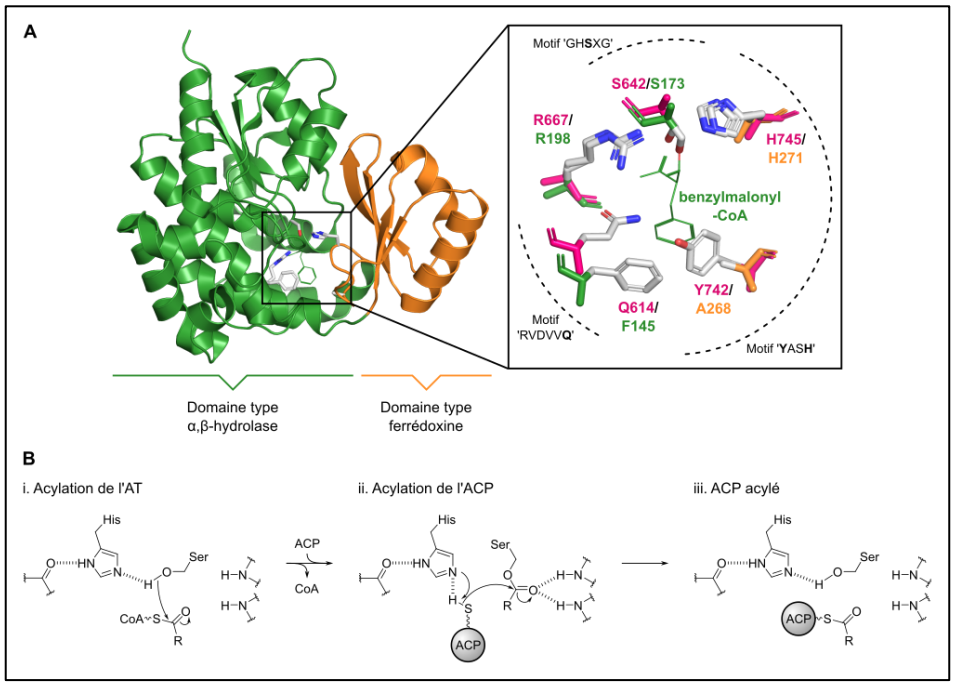


Figure 23. Structure d'un domaine acyltransférase (AT) de la splénocine synthase et mécanisme de transfert de l'unité acyl-coA du domaine AT au domaine porteur d'acyle. **(A)** Structure du domaine AT de la splénocine synthase (SpnD-AT) (Identifiant PDB : 5YDM) et comparaison avec le domaine AT5 de la 6-désoxyérythronolide B synthase (DEBS) (identifiant PDB : 2HG4). (*gauche*) Les domaines AT sont organisés en deux sous-domaines : un sous-domaine organisé de type α,β-hydrolase (vert) et un domaine de type ferrédoxine (orange). Le site actif des domaines AT est composé notamment d'une histidine et d'une sérine conservées. L'encombrement généré par les autres résidus présents dans la poche du site actif influe sur la taille des substrats accepté par le domaine. (*droite*) La comparaison entre les résidus présents dans les poches des domaines SpnD-AT et AT5 de DEBS (rose) montre que le premier contient des résidus moins encombrants, pouvant permettre l'accommodation de substrat plus important tel que le benzylmalonyl-coenzyme A (CoA) (*droite*). **(B)** Mécanisme d'action proposé pour les domaines AT (adapté de Tsai & Ames, 2009).

Ce domaine est le seul qui soit uniquement retrouvé chez les PKS de type I. Les domaines AT diffèrent dans leur spécificité selon leur structure, ce qui amène une grande diversité au sein des polycétides produits par la nature des unités acylées sélectionnées. L'étude cristallographique de ces domaines a permis d'identifier un repliement de la partie N-terminale de type α,β-hydrolase et de type ferrédoxine en C-terminal de l'enzyme (Y. Li et al., 2018; Tang et al., 2006) (Figure 23A).

Le site catalytique responsable de l'acylation du domaine ACP se trouve à l'interface de ces deux types de repliements (Figure 23A). Dans un premier temps, la sérine catalytique attaque le groupement carbonyle du substrat, ce qui entraîne l'acylation du domaine AT (Figure 23B). L'histidine catalytique effectue ensuite une attaque nucléophile du thiol porté par le bras P-Pant du domaine ACP, ce qui entraîne le transfert du groupement acylé de l'AT à l'ACP. L'analyse des séquences des domaines AT a permis d'établir que ces domaines sont très similaires au sein d'une même PKS et peuvent être classés en fonction de leur spécificité (Figure 24) (Ridley et al., 2008).

Les domaines AT présents au sein de la rapamycine synthase (RAPS) peuvent être regroupés en deux classes selon leur spécificité : huit domaines ont une spécificité pour les unités méthylmalonyl-CoA (mAT) et six pour le malonyl-CoA (AT) (Figure 24). Les deux classes de domaines AT de RAPS ont une faible identité de séquence (30 %) tandis que les régions linker de part et d'autre du domaine sont plus conservées : 80 % d'identité de séquence. Les séquences présentes de part et d'autre des domaines AT sont les régions les plus conservées au sein des PKS. Ces observations permettent d'établir l'hypothèse selon laquelle les domaines AT évoluent par recombinaison homologue au sein des clusters de gène de PKS (Figure 24).

Le domaine AT du module 14 cependant ressort du lot. En effet, ce domaine a une spécificité pour les unités malonyl-CoA et est pourtant retrouvé dans le clade des mAT. Cela peut s'expliquer par une perte de résidu au sein de la région C-terminale de la protéine au cours de l'évolution, entraînant une altération de sa spécificité. De plus, les auteur·rices se sont concentré·es principalement sur les domaines AT provenant de PKS *cis*-AT. Trois enzymes *trans*-AT ont cependant été étudiées – DszD, LmnG et PedD – et il a été montré que ces domaines sont éloignés phylogénétiquement des domaines présents au sein des PKS *cis*-AT (Figure 24).

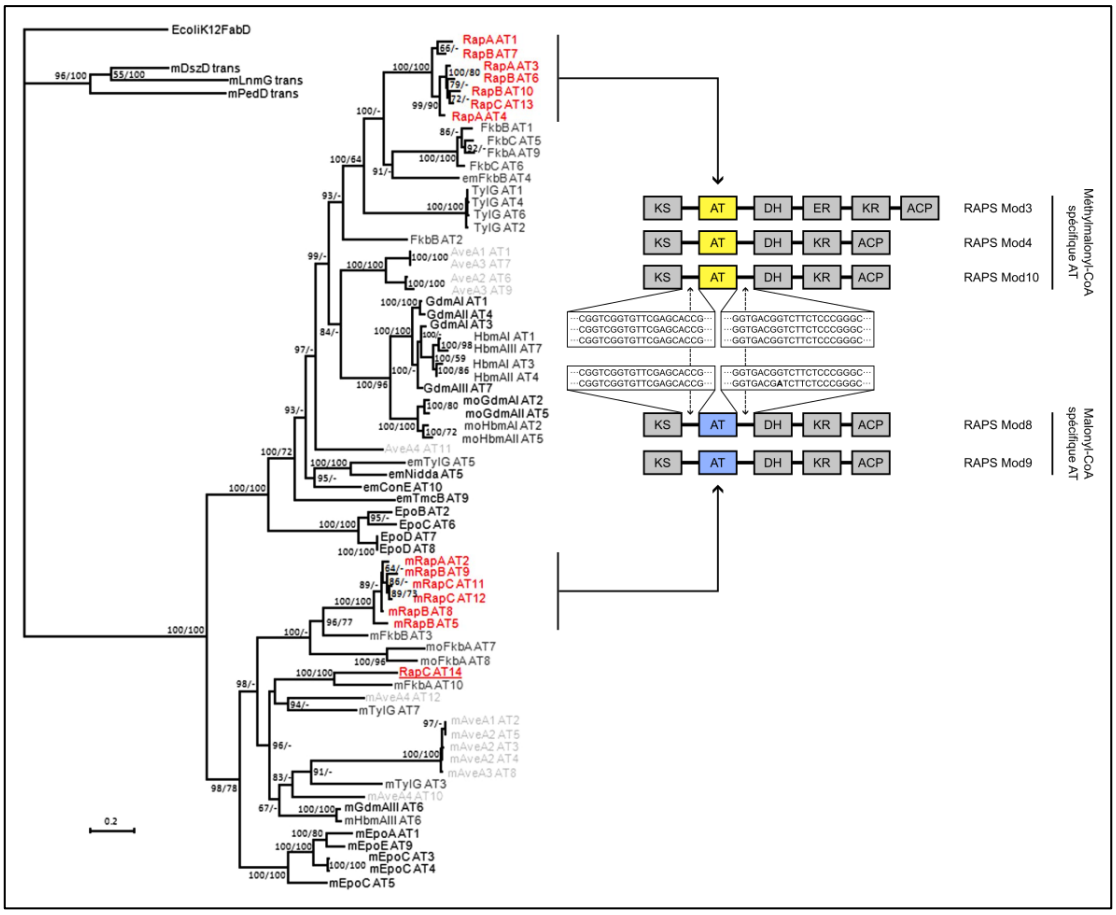


Figure 24. Arbre phylogénétique de domaines acyltransférases et représentation schématique des séquences pouvant permettre la recombinaison homologue de ces domaines pour la rapamycine synthase (adapté de Ridley et al. 2018). Les séquences AT étudiées à gauche ont différentes spécificités. Lorsque le nom AT est précédé de la lettre m, il est spécifique des méthylmalonyl-CoA, em correspond à l'éthylmalonyl-CoA et mo au méthoxymalonyl-CoA. L'absence d'indication signifie que ce sont des domaines spécifiques des unités malonyl-CoA. La probabilité de chaque clade a été calculée par la méthode de Bayesian et la technique des bootstrap. Les séquences AT retrouvées au sein de la rapamycine synthase sont indiquées en rouge. Une analyse des séquences des AT de RAPS à droite montre une identité de séquence forte entre les domaines linker KS–AT et AT–DH. Ces similarités peuvent être à l'origine d'événement de recombinaison homologue entraînant des duplications et échange de domaines AT au sein d'une même PKS.

III.C. Le domaine cétosynthase (KS)

Les domaines KS peuvent intervenir à différents niveaux au cours de la synthèse des polycétides, selon le type de PKS et leur composition en acides aminés. Les KS forment des homodimères et ont une structure de type thiolase caractérisée par un enchaînement d'hélices α et brins β (Figure 25) (Gay et al., 2014; Tang et al., 2006).

Au sein des PKS modulaires *cis*-AT, les domaines KS sont suivis des domaines AT au sein de la chaîne polypeptidique. La structure de la KS du module 5 de la 6-désoxyérythronolide B synthase (DEBS) a été résolue en présence du domaine AT adjacent. Les interactions entre ces deux domaines ont ainsi pu être caractérisées. Ces interactions se font entre le domaine AT et un domaine adaptateur situé en position C-terminale de la KS.

Au sein des PKS *trans*-AT, ce domaine adaptateur est également retrouvé malgré l'absence du domaine AT en aval. Il doit dans ce cas faciliter l'ancrage du domaine AT extrinsèque au cours de l'élongation du polycétide et est appelé AT docking (ATd). La structure du domaine KS du module 2 de la PKS *trans*-AT de la bacillaene a été résolue en présence du substrat natif. Les domaines KS possèdent une triade catalytique Cys/His/His. Le polycétide en élongation est porté lors de la condensation par le résidu cystéine. Afin de permettre la stabilité de l'intermédiaire KS lié au substrat, ce résidu a été muté en sérine pour la résolution de la structure présentée dans la Figure 25.

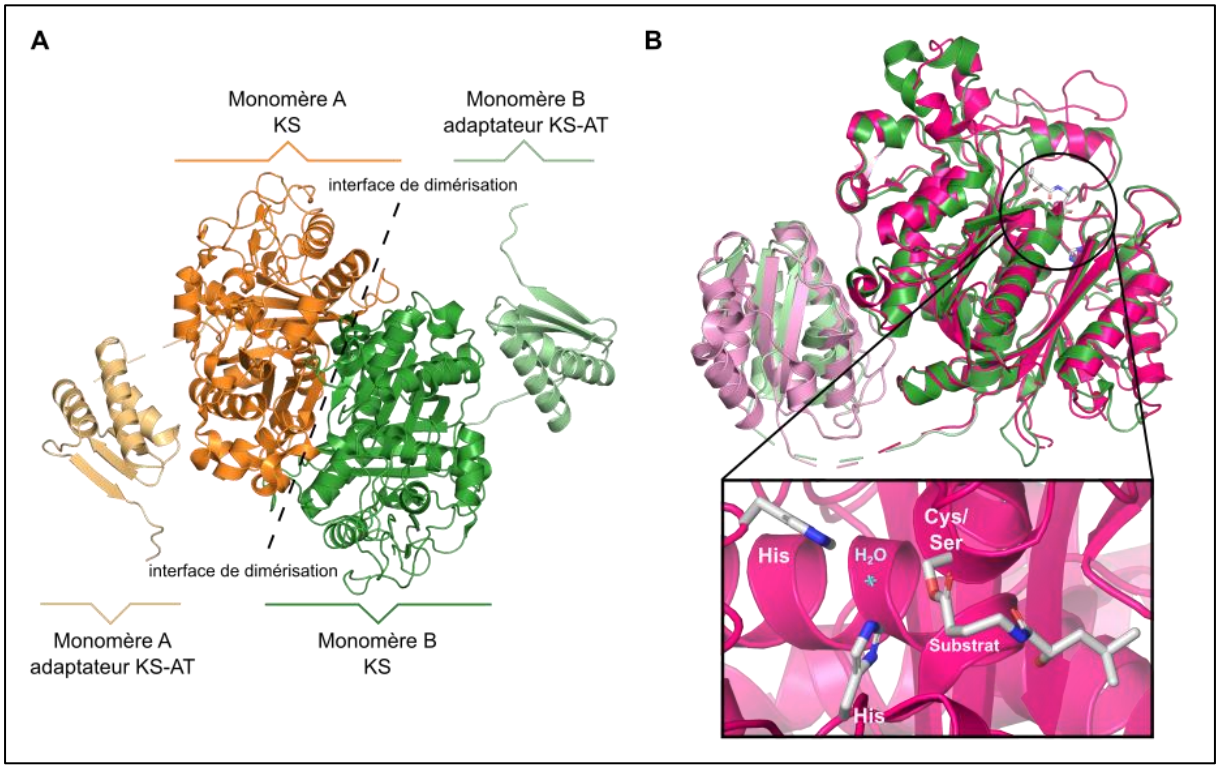


Figure 25. Structure de domaines cétosynthase issus de polycétide synthases. **(A)** Structure de l'homodimère de KS présent dans le module 5 de la 6-désoxyérythronolide B synthase (DEBS) (PDB 2HG4) (Tang et al., 2006). Les monomères représentés en vert et orange sont constitués d'un domaine principal en position N-terminale comportant l'activité catalytique, ainsi que de domaines adaptatifs du domaine acyltransférase en position C-terminale. **(B)** Alignement des structures de KS provenant de PKS *cis*-AT (vert, PDB 2HG4) (Tang et al., 2006) et *trans*-AT (rose, PBD 4NA2) (Gay et al., 2014). La structure de la KS du module 5 de DEBS est similaire à celle du module 2 de la PKS *trans*-AT synthétisant la bacillaene. La structure de ce dernier a été résolue avec le substrat natif du domaine. La triade catalytique est représentée en bas. Le substrat est initialement porté par une cystéine, mutée ici par une sérine pour permettre la stabilité de l'intermédiaire covalent. Les deux histidines catalytiques sont également visibles, ainsi que la molécule d'eau nécessaire à la décarboxylation de l'unité d'extension.

Le mécanisme catalytique des domaines KS s'effectue en trois étapes clefs (Figure 26). La première étape consiste au chargement du polycétide en élongation sur la cystéine catalytique. La cystéine effectue une attaque nucléophile sur le thioester porté par l'ACP du module précédent (ACP *n-1*) afin de le charger sur la KS. L'ACP *n-1* quitte ensuite le site actif et l'ACP *n* peut alors le pénétrer, portant l'unité d'extension apportée par le domaine AT.

Cette unité est ensuite décarboxylée puis liée de manière covalente au polycétide par l'action des deux histidines.

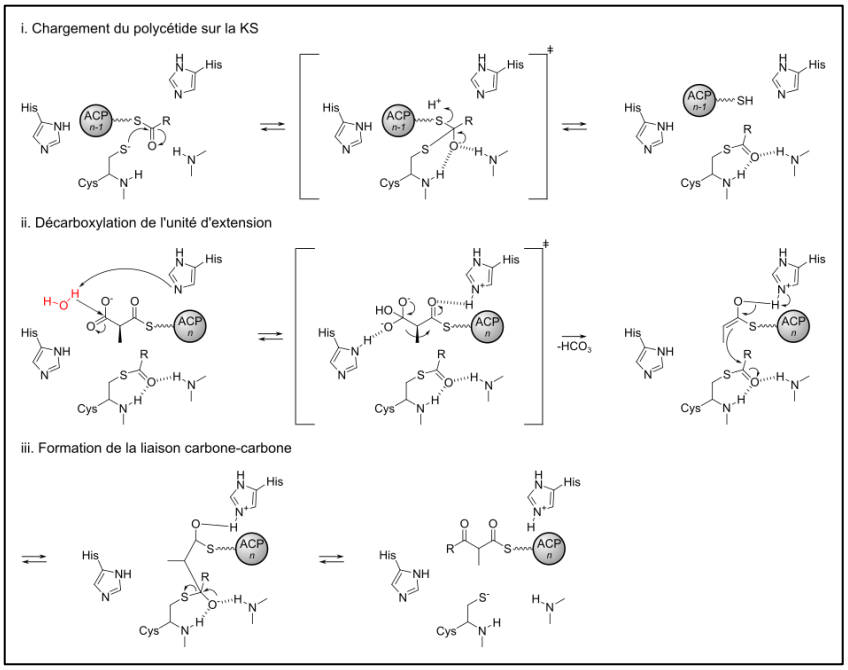


Figure 26. Mécanisme catalytique des domaines cétosynthase (adapté de Weissman, 2017). Dans un premier temps, la KS charge le polycétide en élongation par une attaque nucléophile de la cystéine sur le thioester porté par le domaine ACP du module précédent (i). Ensuite, les histidines permettent la décarboxylation de l'unité d'extension à l'aide d'une molécule d'eau (ii) et entraînent la formation de la liaison entre le polycétide et l'unité (iii).

III.D. Le domaine cétoréductase (KR)

Le domaine KR est un domaine non essentiel à la production des polycétides qui catalyse la première réaction de réduction de la fonction cétone. Ce domaine est composé de sous-domaines structuraux identifiés : un sous-domaine comportant le site actif et nommé « sous-domaine catalytique », et un sous-domaine uniquement impliqué dans le repliement du domaine et sa stabilisation et appelé « sous-domaine structural » (Liu et al., 2018). Ces deux domaines ont été identifiés comme étant similaires à des unités monomériques de la famille des déshydrogénases / réductases à chaîne courte (SDR). Le domaine KR est monomérique, et n'est donc pas impliqué dans les mécanismes de dimérisation et d'assemblage des PKS entre elles.

Ce domaine catalyse la réaction de réduction de la fonction cétone ajoutée par le module précédent en fonction hydroxyle, par le biais d'un cofacteur nicotinamide adénosine dinucléotide phosphate (NADPH). La structure d'un domaine KR retrouvé au sein de l'amphotéricine synthase a été résolue en présence du cofacteur NADPH, ainsi que d'un analogue de substrat mimant un polycétide en élongation, fusionné à un groupement N-acétylcystéamine (SNAC) mimant le bras P-Pant du domaine ACP (Figure 27).

Les résidus sérine, tyrosine et lysine représentés permettent de stabiliser le substrat et faciliter l'échange de protons avec le NADPH.

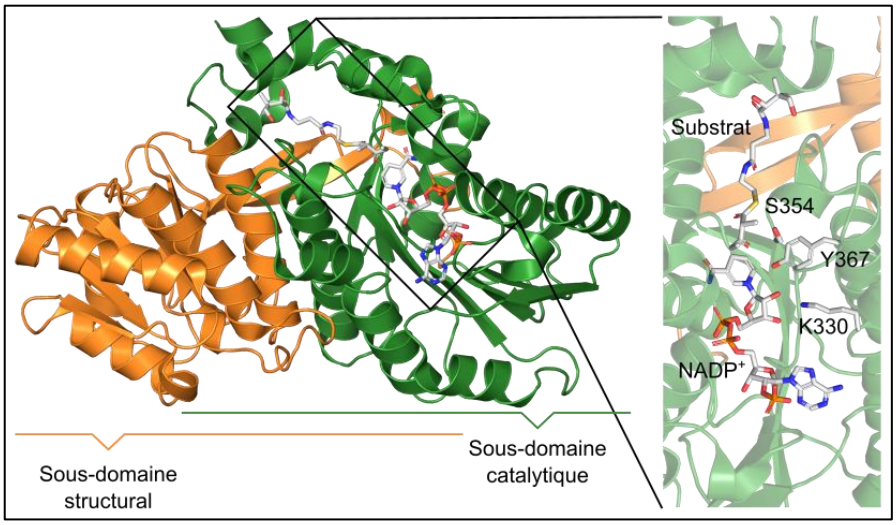


Figure 27. Structure d'un domaine cétoréductase de l'amphotéricine synthase (identifiant PDB : 5XWV) (adapté de C. Liu et al., 2018). Le domaine cétoréductase (KR) est composé de deux sous-domaines structural et catalytique, représentés respectivement en orange et vert. La structure résolue ici contient le cofacteur déprotoné $NADP^+$ nécessaire à la catalyse, ainsi qu'un analogue de substrat (SNAC). Les résidus sérine, tyrosine et histidine représentés stabilisent le substrat par le biais de liaisons hydrogène.

Les domaines KR peuvent être classés en deux types, selon la stéréochimie du produit obtenu. En effet, les domaines KR peuvent produire des fonctions hydroxyles au niveau du carbone 3 du polycétide en élongation qui sont de conformation *S* ou *R*, ce qui en font des KR de type A ou B respectivement (Figure 28A). L'alignement de séquences de KR de type A et B ont permis d'identifier les motifs récurrents permettant de discriminer ces classes en fonction de leur séquence peptidique (Caffrey, 2003).

A quelques exceptions près, les KR de type B contiennent un motif conservé composé d'une leucine suivie de deux acides aspartiques (Caffrey, 2003). Les KR de type A sont dépourvues de ce motif : aucun résidu aliphatique n'est retrouvé suivi de deux résidus chargés dans la même région. Au sein des KR de type B, des résidus proline et asparagine sont retrouvés une cinquantaine de résidus après le motif 'LDD', séparés par trois résidus. La majorité des KR de type A quant à elles contiennent un résidu tryptophane conservé dans la région où sont retrouvées la proline et l'asparagine, sans que ces dernières ne soient présentes. Ce sont donc ces motifs 'LDD' et 'W' qui permettent notamment de prédire la stéréochimie du produit obtenu (Caffrey, 2003).

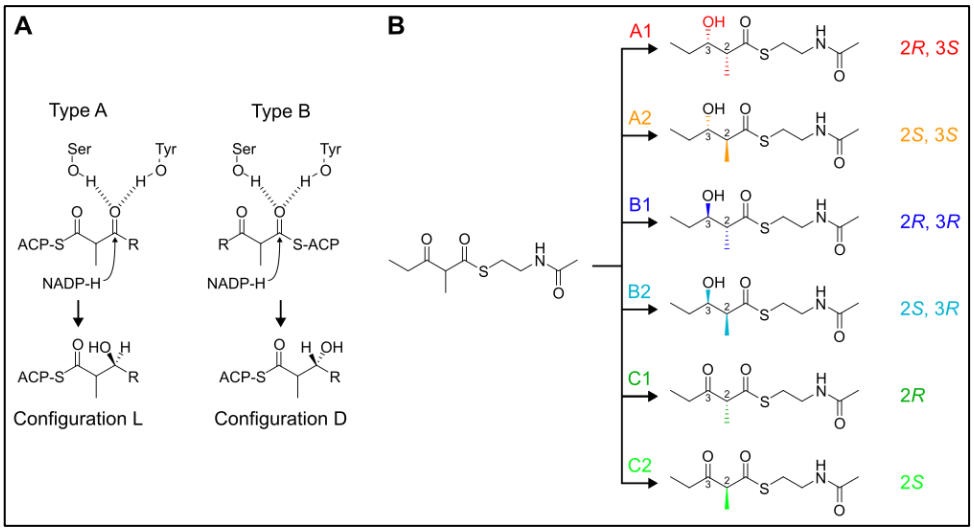


Figure 28. Mécanisme d'action des domaines cétoréductase. **(A)** Mécanisme de réduction de la fonction cétone de la chaîne polycétidique en fonction hydroxyle par les domaines cétoréductase (KR) (adapté de C. Liu et al., 2018). Les domaines KR contenant un motif 'W' sont de type A et leur produit montre une conformation L du groupement hydroxyle. A l'inverse, les domaines KR de type B contiennent un motif 'LDD' et produisent une molécule avec une conformation D du groupement hydroxyle. **(B)** Les classes de domaines KR et le résultat de leur catalyse sur les polycétides en élongation (adapté de Drufva et al., 2021 et Keatinge-Clay, 2007). Le type de KR 1 ne permet pas l'épimérisation à la position du C_2, induisant une conformation 2*R*, tandis que le type 2 possède une activité d'épimérisation et induit une conformation 2*S*. Le type C regroupe des KR qui ne possèdent pas d'activité réductrice, mais peuvent tout de même provoquer une épimérisation (C2) ou non (C1).

Ces résidus ne sont pas impliqués dans le mécanisme catalytique, par conséquent la différence de stéréochimie provient donc d'un autre paramètre. Cela peut s'expliquer par le sens dans lequel le substrat pénètre le site actif en premier lieu (Reid et al., 2003). Une autre hypothèse implique la manière dont les PKS interagissent entre elles. En effet, les PKS sont connues pour interagir entre elles et former des méga-complexes. Dans ce cas, le domaine ACP pourrait aussi bien interagir avec le domaine KR se trouvant sur la même chaîne polypeptidique que lui, ou sur une chaîne avoisinante. Ainsi, l'orientation de ces interactions est différente et peut expliquer la différence de stéréochimie obtenue après catalyse par le domaine KR (Caffrey, 2003).

Au-delà de la conformation induite au niveau du C_3, les domaines KR peuvent également porter une activité épimérase qui agit sur le C_2. Les KR ne portant pas cette caractéristique sont classées dans le type 1 et celles qui la possèdent dans le type 2 (Figure 28B) (Keatinge-Clay, 2007). Il existe une troisième famille de domaines KR, dépourvus de leur activité réductase et nommée type C (Keatinge-Clay, 2007). A l'instar des KR de type A et B, il existe des KR de type C1 et C2. Ces premières sont dépourvues du résidus tyrosine catalytique, et le substrat n'est dans ce cas pas modifié par le domaine. Les KR de type C2 quant à elles possèdent cette tyrosine, mais pas le résidu asparagine. Dans ce cas, une épimérisation a lieu au niveau du C_2 mais la fonction cétone n'est pas réduite en groupement hydroxyle.

III.E. Le domaine déshydratase (DH)

Les domaines DH prennent en charge le polycétide modifié préalablement par le domaine KR et catalysent la déshydratation de la fonction hydroxyle, ce qui amène à la formation d'une insaturation α,β. Les domaines DH ont une spécificité de substrat et ne sont pas capable d'accepter tous les isomères produits par les différents types de KR. En effet l'étude du domaine DH du module 4 de DEBS montre une spécificité de substrat pour les polycétides de conformation 2*R*, 3*R* (Valenzano et al., 2010). De même, le domaine DH retrouvé dans le module terminal de la rifamycine synthase possède une spécificité de substrat pour les polycétides de conformation 2*S*, 3*S* (Gay et al., 2013).

La structure du domaine DH retrouvé au sein du module 1 de la fluvirucin synthase montre une forme de double hot-dog, en ce sens que les brins β semblent représenter un pain contenant deux hélices α (les saucisses) (Figure 29A et B) (Barajas et al., 2019).

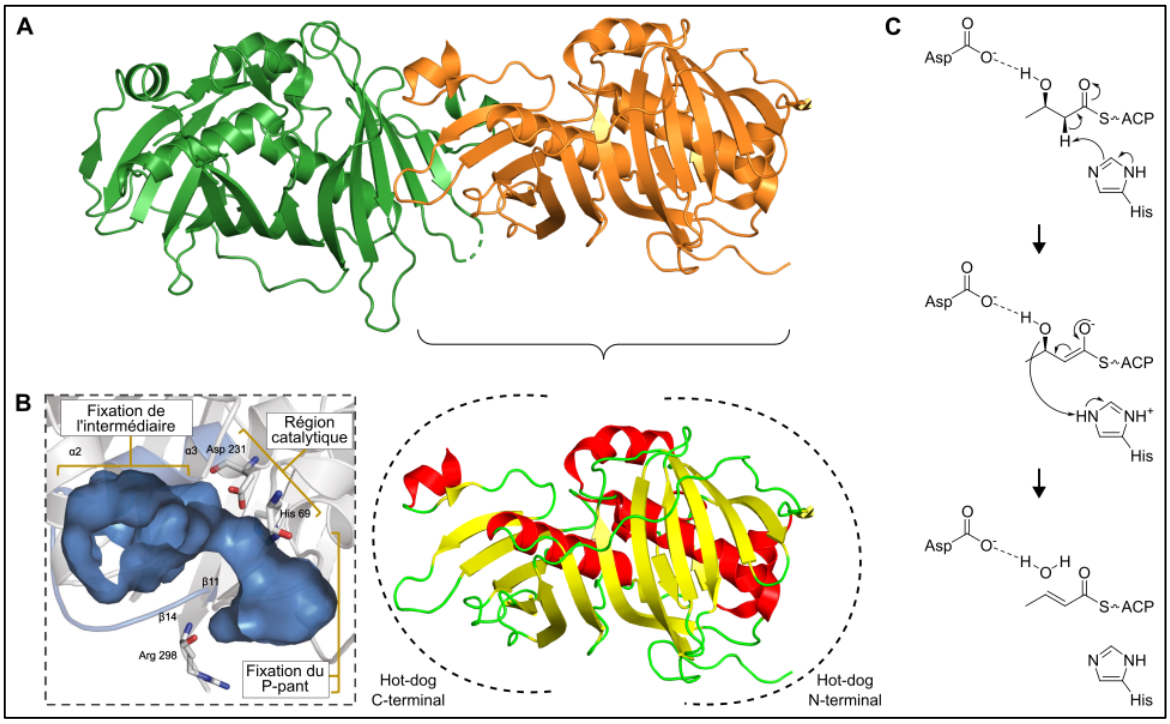


Figure 29. Structure et organisation du site actif du domaine déshydratase de la fluvirucin B1 synthase (adapté de Barajas et al., 2019). **(A)** Structure du domaine déshydratase (DH) de la fluvirucin B1 synthase (identifiant PDB : 6OBV). Ces domaines sont dimériques, et chaque monomère est représenté respectivement en vert et orange. Le domaine DH est organisé en structure double hot-dog, c'est-à-dire qu'il contient des brins β – représentés en jaune – entourant des hélices α – représentées en rouge. **(B)** Représentation du site actif du domaine DH de la fluvirucin B1 synthase. Le domaine DH possède au niveau de son site actif une histidine et un acide aspartique catalytiques. La résolution de la structure du domaine DH de la flurivurin B1 par diffraction aux rayons X a permis d'observer la présence de cavités au sein de la structure. Au niveau du site actif, trois régions ont ainsi été identifiées : celle permettant la fixation de l'intermédiaire acylé, la région catalytique et enfin celle où se fixe le bras 4′phosphopantéthéine (P-Pant). **(C)** Mécanisme catalytique proposé de déshydratation par le domaine DH (adapté de Xie & Cane, 2018). L'aspartate stabilise le substrat tandis que l'histidine déprotone la position C_2 et protone le groupement hydroxyle en position C_3 afin de catalyser sa libération sous forme H_2O. Cela permet la formation d'une insaturation C_2–C_3.

Le mécanisme proposé pour la déshydratation des substrats par le domaine DH repose sur l'action d'une base unique (Xie & Cane, 2018). Dans ce cas, cette base est représentée par le résidu histidine au sein du site actif (Figure 29C). Cette histidine permet la déprotonation du carbone C_2, puis la protonation du groupement hydroxyle porté par le carbone C_3. Cela entraîne la libération d'une molécule d'eau et la formation d'une insaturation C_2–C_3. L'aspartate présent au sein du site actif permet de stabiliser le substrat lors de la déshydratation.

III.F. Le domaine énoyl-réductase (ER)

Le domaine ER est le troisième domaine optionnel impliqué dans la modification de la fonction cétone des unités acylées. Il catalyse la réduction de l'insaturation obtenue à la suite de l'action des domaines KR puis DH, et entraîne alors la formation d'une liaison simple carbone-carbone. La résolution de la structure du bidomaine KR-DH du module 2 de la spinosyne synthase par cristallographie aux rayons X a permis d'élucider l'organisation de ces domaines (Figure 30) (Zheng et al., 2012).

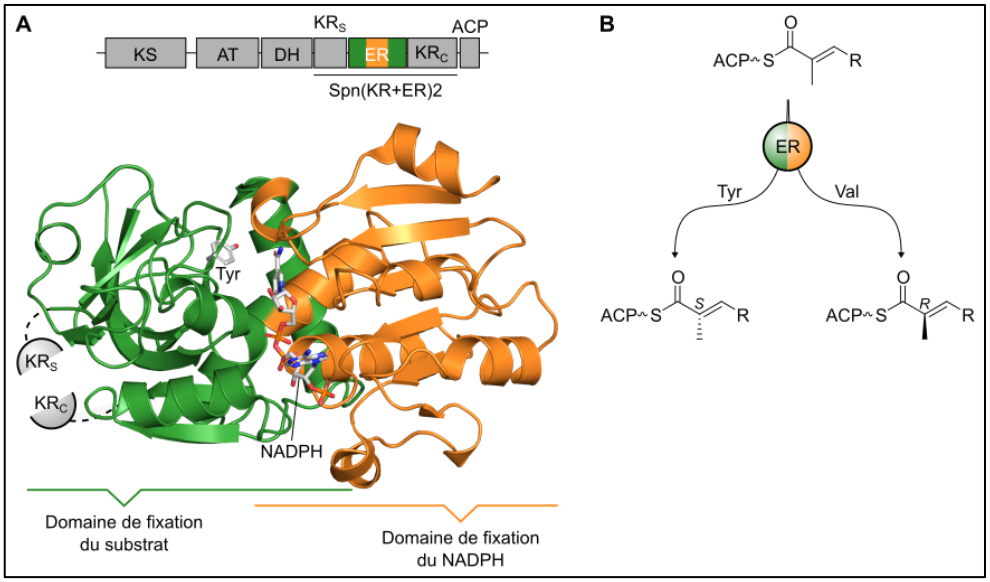


Figure 30. Structure d'un domaine énoyl-réductase et représentation schématique de la spécificité catalytique de ces domaines. **(A)** Structure du domaine énoyl-réductase (ER) du module 2 de la spinosyne synthase (Spn) (identifiant PDB : 3SLK) (adapté de Zheng et al., 2012). Une représentation schématique de l'organisation génétique du module 2 de la PKS Spn est montrée au-dessus de la structure. Afin d'élucider la structure du domaine ER, celui-ci a été produit et purifié en présence du domaine cétoréductase du module. Le bidomaine KR-ER y est noté Spn(KR+ER)2. La séquence codant le domaine KR est ici scindée par la séquence codant le domaine ER. En position 5′ de cette dernière se trouve la région codant le sous-domaine structural de la KR (KR_S) et en position 3′ celle codant le sous-domaine catalytique KR_C. Le domaine ER est composé de deux sous-domaines structuraux. L'un, représenté en orange, médie l'interaction avec le cofacteur nicotinamide adénosine dinucléotide phosphate (NADPH) et le second, en vert, permet la fixation du substrat. Le résidu tryptophane présent dans le domaine de fixation du substrat assure la spécificité du domaine. **(B)** Représentation schématique de la spécificité catalytique des domaines ER (adapté de Kwan et al., 2008 et Weissman, 2017). Les domaines ER contenant une tyrosine au sein d'un motif conservé 'ALGMX**Y**' permettent l'obtention d'un substrat de configuration 2*S*, tandis que la présence d'une valine à cette même position est corrélée avec une configuration 2*R* du produit.

La séquence codant le domaine ER se trouve entre les séquences codant les régions structurales du domaine KR (KR_S) d'un côté (en position 5′) et le domaine catalytique de la KR (KR_C) de l'autre (en position 3′). Le domaine KR complet se forme lors du repliement de ce bidomaine. Ce même phénomène a été observé lors de la résolution du tridomaine réducteur DH-ER-KR de la juvénimicine synthase (McCullough et al., 2023).

Deux sous-domaines sont identifiés au sein de la structure du domaine ER (Figure 30A). Le premier, constitué des extrémités C- et N-terminales du domaine ER, permet la fixation du substrat. En son sein, un motif 'ALGMXX' est retrouvé et duquel dépend la spécificité catalytique du domaine (Kwan et al., 2008). Lorsque le second 'X ' est une tyrosine, le produit de la catalyse a une conformation *S* (Figure 30B). Si ce résidu est une valine, alors la conformation obtenue est *R*. Le sous-domaine central quant à lui médie l'interaction avec le cofacteur NADPH nécessaire à la réaction catalytique.

III.G. Le domaine thioestérase (TE)

Il existe deux classes de domaines TE, selon leur rôle au sein de la biosynthèse des polycétides. Les domaines TE retrouvés à la fin des voies de biosynthèse des PKS modulaires, précédemment présentées, sont de type 1. Leur rôle est de libérer le polycétide produit, en catalysant dans certains cas sa cyclisation (Staunton & Weissman, 2001). Pour cela, ils sont retrouvés en position C-terminale du dernier module des PKS. La première structure d'un domaine TE de type 1 provenant d'une PKS modulaire à avoir été résolue est celle du domaine provenant de DEBS (Figure 31A) (Tsai et al., 2001). Celle-ci a permis d'observer la présence d'un sous-domaine de type α/β-hydrolase et d'un sous-domaine couvercle. La triade catalytique identifiée – composée d'une histidine, un acide aspartique et une sérine – se trouve à l'interface entre ces deux sous-domaines. La dimérisation de ce domaine est médiée par la présence de deux hélices α en position N-terminale.

Les TE de type 2 quant à eux sont des enzymes discrètes jouant un rôle de correcteur. En effet, ils permettent de libérer des substrats en cours d'élongation, notamment lorsque ceux-ci ne sont pas réactifs ou qu'ils ne correspondent pas au substrat attendu (Kotowska & Pawlik, 2014). La structure du domaine TE de type 2 provenant du système hybride NRPS/PKS de la rifamycine synthase montre une organisation en sous-domaines similaire à celle des TE de type 1 (Figure 31B). Cependant, elle ne contient pas les hélices α observée chez les domaines TE de type 1 comme permettant leur dimérisation.

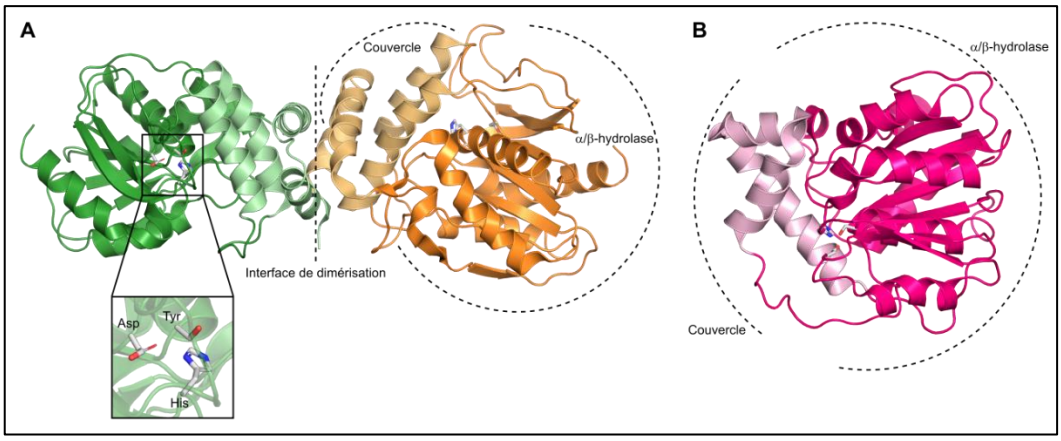


Figure 31. Structure de domaines thioestérase de polycétide synthases. **(A)** Structure du domaine thioestérase (TE) de type 1 de la 6-désoxyérythronolide B synthase (identifiant PDB : 1KEZ) (adapté de Tsai et al., 2001). Ce domaine est organisé en deux sous-domaines. Le premier est un domaine de type α/β-hydrolase et le second est un couvercle. Dans cette dernière région se trouvent deux hélices α assurant la dimérisation du domaine. La triade catalytique – composée d'une tyrosine, un acide aspartique et une histidine – se trouve à l'interface entre les deux sous-domaines. **(B)** Structure du domaine TE de type 2 de la rifamycine synthase (identifiant PDB : 3FLA) (adapté de Claxton et al., 2009). Ce domaine est organisé de la même manière que le TE de type 1 (A), à la différence que les hélices de dimérisation sont ici absentes. Par conséquent, ce domaine TE de type 2 est monomérique.

Le mécanisme catalytique des domaines TE des deux types se fait en deux grandes étapes majeures : le chargement du substrat sur le domaine, puis sa libération (Figure 32). Dans un premier temps, le polycétide substrat est chargé sur le résidu sérine de la triade, par le biais de l'action de l'histidine stabilisée par l'acide aspartique. Ensuite, la libération de ce substrat est médiée par la présence d'un nucléophile X. Celui-ci peut être une molécule d'eau, auquel cas le produit obtenu sera linéaire, mais il peut également être un des groupements hydroxyles du polycétide, et ainsi catalyser sa cyclisation.

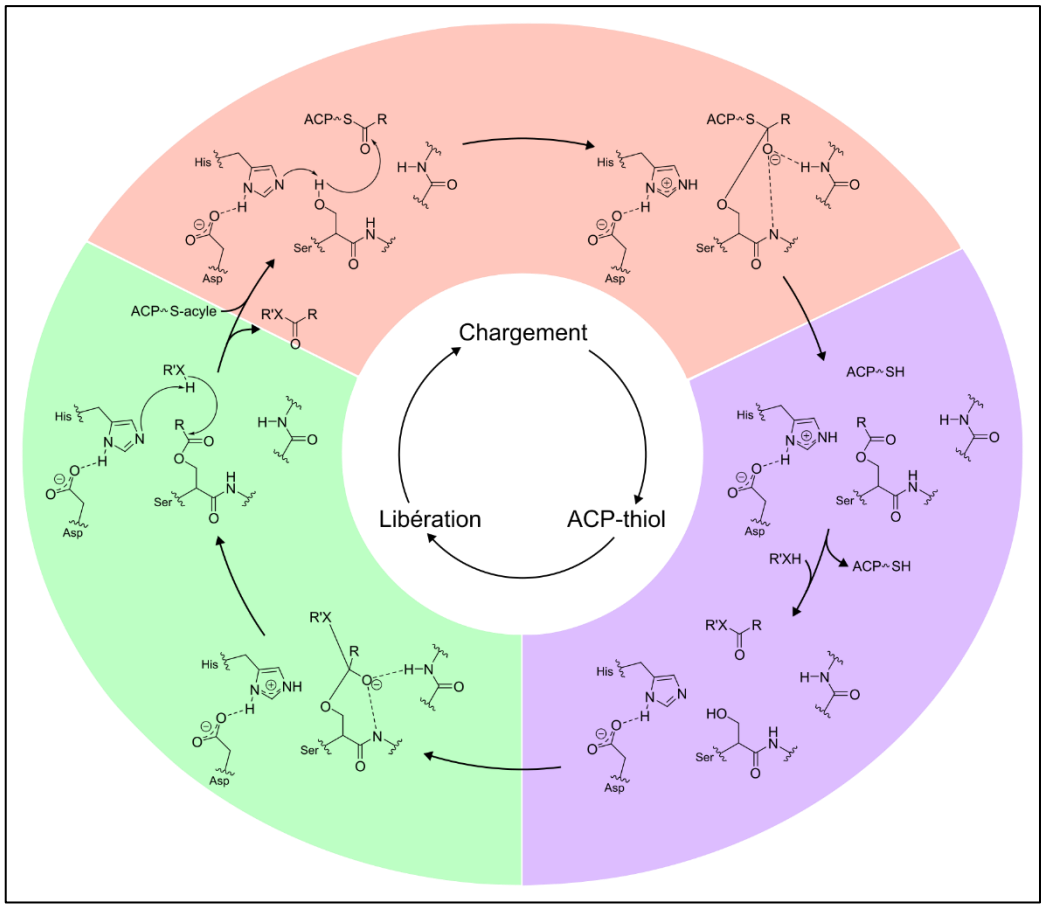


Figure 32. Mécanisme catalytique proposé des domaines thioestérase (adapté de Horsman et al., 2016). Le mécanisme d'action des domaines thioestérase (TE) se déroule en deux étapes majeures : le chargement du substrat et sa libération. Le substrat est chargé sur le résidu sérine de la triade catalytique par le biais de l'histidine, libérant ainsi le domaine porteur d'acyle (ACP). La présence d'un nucléophile X au niveau du site actif entraîne par la suite la libération du substrat préalablement chargé. Si ce nucléophile est interne au substrat, alors cela catalyse sa cyclisation également.

IV. Les interfaces d'interactions entre les sous-unités de polycétide synthases

Il existe différentes interactions permettant l'organisation des PKS, leur dimérisation et leur macro-complexation. La majeure partie des domaines constituant les PKS sont notamment capable de former des homodimères. De plus, la communication entre les modules est indispensable pour le bon fonctionnement des PKS, notamment entre les domaines ACP et KS. Cette notion sera discutée dans une prochaine partie. Parmi les éléments d'interactions entre modules sont également retrouvés les DD, présents aux jonctions entre les sous-unités de PKS et qui sont respectivement nommés ^{C}DD (domaine en position C-terminale) et ^{N}DD (domaine en position N-terminale). Les interactions entre les deux domaines sont non-covalentes et médiées par des interactions hydrophobes et parfois par la présence de résidus chargés (Smith et al., 2021). Ces interactions sont spécifiques, empêchant ainsi les liens entre deux sous-unités ne devant pas se suivre.

IV.A. Les interfaces d'interaction des PKS de type *cis*-AT

Au sein des PKS *cis*-AT, les jonctions entre les sous-unités sont retrouvées entre des domaines ACP d'un module (n) et KS du module suivant (n+1), soit à la jonction entre deux modules présents à la fin d'une sous-unité et au début de la sous-unité suivante. Deux types de DD ont à ce jour été caractérisés au sein des PKS *cis*-AT.

IV.A.1. Les domaines de docking de classe 1

Deux types de DD de classe 1 ont été caractérisés à ce jour : les types 1a et 1b. Les DD de classe 1a sont composés pour le ^{C}DD de trois hélices α, dont les deux premières assurent une homo-dimérisation des domaines et la troisième interagit avec le ^{N}DD. Le ^{N}DD quant à lui est composé d'une unique hélice α permettant à la fois la formation d'une structure en superhélice avec un second domaine ^{N}DD, et l'interaction avec le ^{C}DD partenaire (Broadhurst et al., 2003). Les premiers domaines de classe 1a à avoir été caractérisés proviennent de la PKS DEBS (Broadhurst et al., 2003) (Figure 33A).

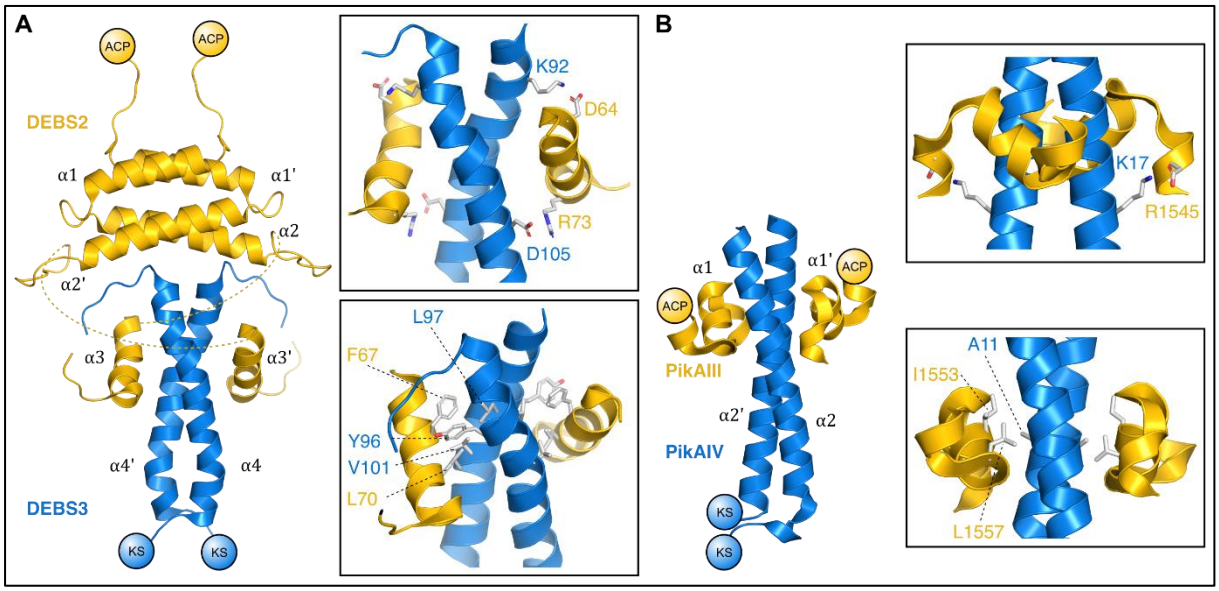


Figure 33. Structure de domaines de docking de type 1. **(A)** Structure des domaines retrouvés entre les sous-unités 2 et 3 de la 6-désoxyérythronolide B synthase (DEBS) (identifiant PDB : 1PZR et 1PZQ) (adapté de Broadhurst et al., 2003). Le domaine présent en C-terminal (^{C}DD) du domaine porteur d'acyle (ACP) est composé de deux hélices α1 et α2 qui assurent l'homodimérisation des DD, et une troisième hélice α3 qui interagit de manière spécifique avec le partenaire présent en position N-terminale (^{N}DD) du domaine cétosynthase (KS). Le ^{N}DD est composé d'une unique hélice α4. (*droite haut*) La spécificité d'interaction est médiée par des résidus chargés. (*droite bas*) Des résidus hydrophobes permettent de stabiliser l'interaction, de manière aspécifique (*droite bas*). **(B)** Structure des domaines retrouvés entre les sous-unités AIII et AIV de la pikromycine synthase (Pik) (identifiant PDB : 3F5H) (adapté de (Buchholz et al., 2009). Le domaine ^{C}DD ici ne contient pas les hélices de dimérisation car celles-ci n'ont pas pu être cristallisées. Il n'y a donc représenté ici que l'hélice α permettant l'interaction avec le domaine ^{N}DD, lui-même composée d'une unique hélice α formant un homodimère. Les interactions de charge sont ici médiées par la présence d'une lysine au sein du domaine ^{C}DD et d'une arginine au sein du domaine ^{N}DD (*droite haut*). Cette interaction est stabilisée par la présence de résidus hydrophobes : une leucine et une isoleucine au sein du ^{C}DD et une alanine dans le ^{N}DD (*droite bas*).

Les interactions entre les DD de classe 1a sont médiées par des résidus chargés et hydrophobes. La spécificité d'interaction provient de la nature des résidus chargés. Dans le cas présenté, les lysine 92 et aspartate 105 du domaine ^{N}DD interagissent respectivement avec les résidus aspartate 64 et arginine 73 du ^{C}DD. Des résidus hydrophobes permettent également de renforcer l'interaction entre les deux partenaires.

Les DD retrouvés entre les sous-unités PikAIII et PikAIV de la pikromycine synthase quant à eux présente notamment un profil d'interaction différent (Figure 33B) (Buchholz et al., 2009). Cette paire a donc été classée dans un nouveau type de DD : les DD de type 1b. Le domaine ^{C}DD dans ce cas n'est composé que d'une unique hélice α au sein de la structure résolue, car les hélices médiant la dimérisation du domaine n'ont pas pu être cristallisées. Pour ce qui est du ^{N}DD, celui-ci est composé d'une seule hélice α capable de s'homo-dimériser et d'interagir avec son partenaire ^{C}DD, à l'instar de ceux des DD de type 1a. Dans le cas de ces domaines, l'interaction est médiée entre une lysine présente au sein du ^{C}DD et une arginine du ^{N}DD. Ces résidus de plus ne sont pas situés à la même place au sein des domaines (Buchholz et al., 2009).

IV.A.2. Les domaines de docking de classe 2

Une nouvelle classe de DD a été identifiée par cristallographie aux rayons X en 2013 au sein de la curacine A synthase (Whicher et al., 2013) (Figure 34A). Ces domaines diffèrent des DD de classe 1 par leur organisation structurelle et la nature des interactions entre les partenaires.

Au sein de la première paire (Figure 34A), les ^{N}DD sont, à l'instar des ^{N}DD de classe 1, composés notamment d'une hélice α permettant l'homo-dimérisation des domaines et l'interaction avec le partenaire ^{C}DD (Whicher et al., 2013). Cependant, ici une seconde hélice constitue également le ^{N}DD et est impliquée uniquement dans l'interaction avec le partenaire. De l'autre côté, le domaine ^{C}DD est ici monomérique et composé de deux hélices α interagissant avec le ^{N}DD. La seconde paire est composé d'un fagot de quatre domaines – deux ^{C}DD et deux ^{N}DD – chacun comprenant une unique hélice α permettant à la fois l'homo-dimérisation et l'interaction entre les partenaires (Figure 34B).

Plus récemment, un autre système a été identifié comme contenant une paire de DD de classe 2 (Figure 34C) (Risser et al., 2020). Celle-ci est retrouvée entre les sous-unités Bamb_5925 et Bamb_5924 de l'énacyloxine synthase. Les partenaires de cette paire de DD sont structurés de la même manière que ceux de la paire retrouvée entre les sous-unités CurG et CurH de la curacine A synthase. Cela forme donc un fagot de huit hélices α.

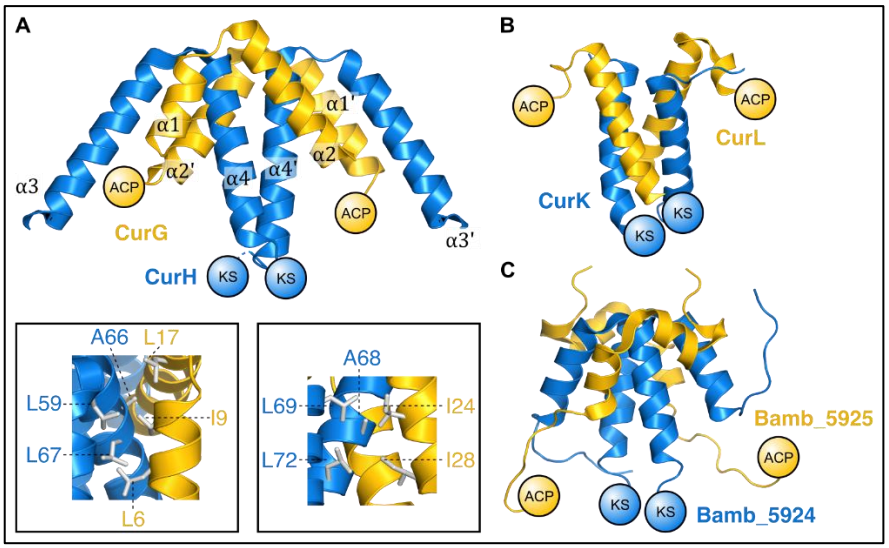


Figure 34. Structure de domaines de docking de type 2 (adapté de Smith et al., 2021). Les domaines de docking présents en position C-terminale (^{C}DD) des domaines porteurs d'acyl (ACP) sont représentés en jaune. Les DD en position N-terminale des domaines cétosynthase (KS) sont représentés en bleu. **(A)** Domaines de docking (DD) retrouvés entre les sous-unités CurG et CurH de la curacine A (identifiant PDB : 4MYY) (Whicher et al., 2013). Le domaine ^{N}DD est composé d'une hélice α assurant l'homo-dimérisation du domaine et d'une seconde interagissant avec le partenaire ^{C}DD. Ce dernier est ici monomérique et composé de deux hélices α qui interagissent avec le partenaire. Le partenaire ^{N}DD étant dimérique, deux domaines ^{C}DD peuvent interagir et ainsi former un fagot de huit hélices α. Dans les encadrés sont mis en avant les résidus assurant les interactions hydrophobes entre les deux partenaires. **(B)** DD de type 2 assurant l'interaction entre les sous-unités CurL et CurK de la curacine A synthase (identifiant PDB : 4MYZ) (Whicher et al., 2013). Les DD de cette paire sont structurés en une unique hélice α, assurant à la fois l'homo-dimérisation des domaines et les interactions entre les deux partenaires. Ces interactions permettent la formation d'un fagot de quatre hélices α. **(C)** DD de type 2 retrouvés entre les sous-unités Bamb_5925 et Bamb_5924 de l'énacyloxine synthase (identifiant PDB : 6TDN) (Risser et al., 2020). Ces domaines sont, à l'instar de (A), composés de deux hélices α entraînant l'homo-dimérisation et l'interaction entre les partenaires, afin de former un fagot de huit hélices α.

IV.B. Les interfaces d'interactions des PKS de type *trans*-AT

Les PKS *trans*-AT sont souvent scindées entre deux domaines intramodulaires – KS / KR, KR / DH et KS / KR par exemple – ce qui peut expliquer la plus grande diversité de DD observée au sein de cette classe (Smith et al., 2021). Un exemple typique est celui de la bacillaene synthase, dans laquelle certains modules sont scindés entre deux sous-unités. C'est le cas entre les sous-unités PksJ et PksL puisque la scission est présente entre le module KS^{0} et le bimodule DH–ACP et entre les sous-unités PksN et PksR, où la scission est entre les domaines KR et MT d'un module KS–DH–KR–MT–ACP (Butcher et al., 2007).

IV.B.1. Les domaines 4-helix bundle

La virginiamycine synthase est une PKS de type *trans*-AT, au sein de laquelle est retrouvé un nouveau type de DD entre un domaine ACP et un domaine KS. Les domaines constituant ce nouveau type sont composés chacun de deux hélices α de tailles similaires, plus courtes que celles caractérisant les DD des PKS de type *cis*-AT, et situées en position C-terminale du domaine ACP et en N-terminale du domaine KS (Figure 35).

Ces hélices permettent l'interaction des partenaires mais ne semblent pas permettre l'homo-dimérisation de ceux-ci. Cette interaction est médiée par des résidus chargés et elle est stabilisée par la présence de résidus hydrophobes.

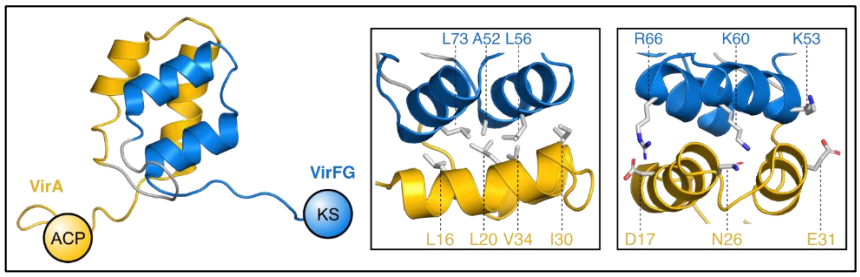


Figure 35. Structure des domaines de docking présents entre les sous-unités VirA et VirFG de la PKS *trans*-AT de la virginiamycine (identifiant PDB : 2N5D) (adapté de Dorival et al., 2016). Ces domaines sont composés chacun de deux courtes hélices α qui permettent l'interaction des deux partenaires (*gauche*). Les interactions entre le ^{C}DD et le ^{N}DD sont médiées par la présence de résidus hydrophobes (*milieu*) et chargés (*droite*).

IV.B.2. Les domaines déshydratase docking

Comme énoncé précédemment, les DD des PKS *trans*-AT ne sont pas nécessairement retrouvés entre des domaines ACP et KS. C'est le cas des domaines déshydratase docking (DHD), qui permettent l'interaction entre des domaines KS et DH portés par deux sous-unités différentes (Figure 36) (Jenner et al., 2018).

Cette interaction est notamment retrouvée entre les sous-unités GdnD4 et GdnD5 de la PKS *trans*-AT qui synthétise la gladioline. Ce domaine DHD est présent en position N-terminale d'un domaine KS, et interagit de manière directe avec le domaine DH présent dans la sous-unité voisine. Ainsi, il n'y a pas ici de paire de DD à proprement parlé, puisque le DHD interagit directement avec le domaine catalytique suivant. Des analyses bio-informatiques ont permis d'identifier qu'un domaine DHD serait également responsable de l'interaction entre les sous-unités PksJ et PksL de la bacillaene synthase, et a également été identifié au sein de dix-neuf autres PKS *trans*-AT caractérisées (Jenner et al., 2018).

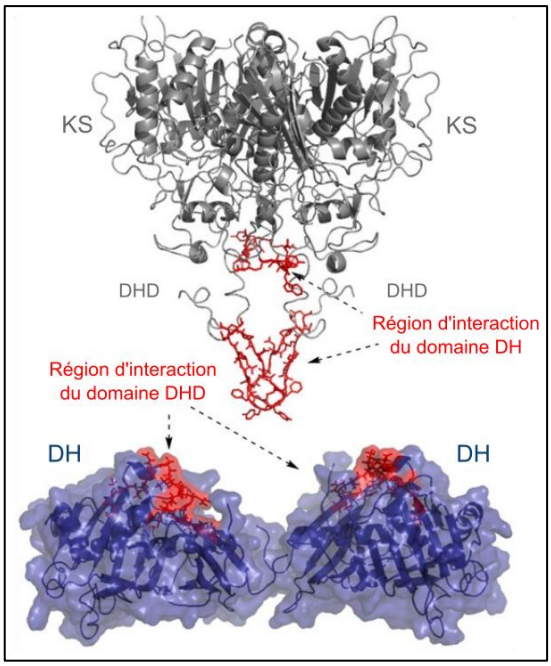


Figure 36. Modèle d'un bidomaine cétosynthase – déhydratase, interagissant par le biais d'un domaine déshydratase docking (adapté de Jenner et al., 2018 et Smith et al., 2021). Le domaine cétosynthase (KS) est représenté en gris, ainsi que le domaine déshydratase docking (DHD). Le domaine déshydratase (DH) est en bleu, et les régions d'interactions entre les domaines DHD et DH sont représentées en rouge. La région du domaine DHD qui interagit avec le domaine DH est ici déstructurée.

V. L'organisation des méga-enzymes au sein des cellules bactériennes

Les PKS sont des systèmes multi-enzymatiques complexes. L'organisation et l'harmonisation de tous les composants de ces systèmes constituent des informations importantes afin d'améliorer leur ingénierie et la production hétérologue de polycétides. L'organisation des PKS *trans*-AT a été étudiée par plusieurs équipes par microscopie à fluorescence et électronique.

V.A. L'organisation des PKS *trans*-AT

La PKS synthétisant la bacillaene a été utilisée comme modèle pour l'étude de l'organisation des PKS *trans*-AT. Elle est retrouvée chez *B. subtilis* et composée notamment de cinq sous-unités PKS – PksJ, PksL, PksM, PksN et PksR – et d'une enzyme discrète portant l'activité AT : PksE (Figure 19) (X.-H. Chen et al., 2006)

V.A.1. L'organisation de la bacillaene synthase au sein de son hôte bactérien

Dans un premier temps, l'équipe de Kolter *et coll.* a utilisé une stratégie de fusion des protéines de la bacillaene synthase à des protéines fluorescentes, couplée à des observations en microscopie à fluorescence. Dans ce cadre, la dernière sous-unité du système – PksR – et l'enzyme discrète portant l'activité AT – PksE – ont été fusionnées respectivement aux protéines fluorescentes cyan (CFP) et jaune (YFP) (Straight et al., 2007).

La localisation et l'organisation de la première sous-unité – PksJ – a également été étudiée par fusion avec la protéine CFP. Les souches modifiées ont ensuite été observées au microscope à fluorescence et par microscopie électronique à transmission (TEM).

Chacune des enzymes se retrouve au sein de foci dans les cellules, et ne sont pas distribuées de manière diffuse (Figure 37A). Sur la population totale des cellules de la souche exprimant la fusion PksR-YFP étudiées, 69 % ont montré un unique foci, tandis que 6 % en possédaient deux et 25 % aucun (Straight et al., 2007). Il y a donc une hétérogénéité dans l'expression des gènes de *B. subtilis*. Ces observations impliquent également la présence d'éléments au sein des séquences protéiques permettant cette macro-organisation au sein des cellules. Il a aussi été démontré que PksR et PksE colocalisent au sein de *B. subtilis* (Figure 37A), ce qui est en accord avec la nécessité pour ces enzymes d'être physiquement proches afin de permettre la synthèse du polycétide (Straight et al., 2007). L'abondance relative des protéines PksR et PksE au sein des bactéries a également été mesurée par Western Blot (Straight et al., 2007). L'enzyme portant l'activité AT (PksE) est présente en environ 100 fois plus de copies que la dernière sous-unité du système (PksR). Cette donnée est cohérente au vu du rôle de chacune de ces enzymes.

Dans cette étude, la localisation de la première sous-unité de la PKS – PksJ – a également été étudiée, en la fusionnant à la protéine fluorescente CFP (Straight et al., 2007). Une fois de plus, les images de microscopie montrent que cette enzyme forme des foci au sein des cellules (Figure 37B). De plus, la colocalisation entre PksJ et PksR a également pu être observée, ce qui est une nouvelle fois en accord avec une nécessité fonctionnelle. L'observation de ces différentes protéines au sein de l'hôte bactérien montre donc l'organisation d'une « usine » de PKS au sein des cellules.

Ce résultat est également en lien avec une autre observation de l'étude : l'organisation en foci du domaine *trans*-AT PksE est dépendante de la présence de la première sous-unité du système, PksJ (Straight et al., 2007). En effet, la délétion de cette dernière entraîne la perte de l'organisation de PksE, dont le signal, lié à la fusion avec la protéine CFP, devient diffus au sein des cellules (Figure 37C). Cela peut signifier que des éléments d'interactions existent entre la PKS et l'enzyme discrète, ou bien que le recrutement de cette dernière se fait lors de la synthèse du polycétide et donc que son organisation en foci soit dépendante de son activité.

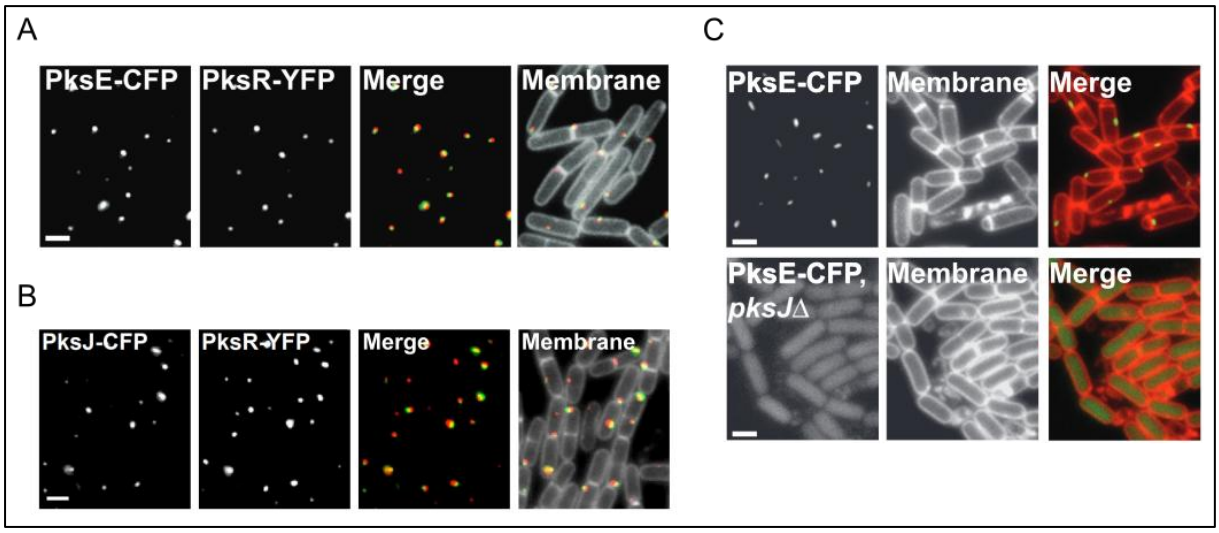


Figure 37. Localisation de la bacillaene synthase chez *Bacillus subtilis* par microscopie à fluorescence (adapté de Straight et al., 2007). La bacillaene synthase est composée de cinq sous-unités – PksJ–N et PksR – et d'un domaine acyltransférase (AT) discret : PksE. Les membranes sont colorées par l'utilisation de *p*-toluènesulfonate de 1-(4-triméthylammoniumphényl)-6-phényl-1,3,5-hexatriène **(A)** Observation simultanée de la localisation de la sous-unité PksR et du domaine *trans*-AT PksE. La protéine PksE a été marquée par la fusion d'une protéine fluorescente cyan (CFP) et PksR avec une protéine fluorescente jaune (YFP). **(B)** Observation de la localisation simultanée des sous-unités PksJ et PksR. L'enzyme PksJ a été fusionnée à la protéine CFP et la protéine YFP a été fusionnée à la sous-unité PksR. **(C)** Observation de la localisation du domaine *trans*-AT PksE en fonction de la présence ou non de la sous-unité PksJ. PksE a été fusionnée à la protéine CFP et le gène codant la sous-unité PksJ a été délété.

Cette observation est appuyée par des observations en microscopie électronique à transmission (Figure 38). Dans le cas de la souche sauvage, une masse est observée au niveau de la membrane bactérienne. Afin d'identifier si la PKS est présente ou non au niveau de cette masse, des particules d'or couplées à des anticorps anti-CFP et anti-YFP ont été utilisées (Straight et al., 2007). Dans les deux cas, PksE-CFP et PksR-YFP, ces particules s'agrègent au niveau des masses observées, près de la membrane. Cependant, cette masse ne semble pas être enclavée par une frontière bien définie, elle se retrouve donc au sein du cytosol de la cellule et non au sein de la membrane elle-même (Straight et al., 2007). L'hypothèse de cette localisation proche de la membrane est celle de la nécessité d'excréter la bacillaene, au risque qu'elle ne devienne toxique pour la cellule.

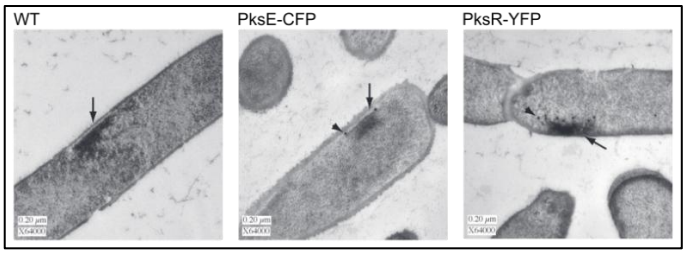


Figure 38. Observation par microscopie électronique à transmission des souches sauvage et mutées de *Bacillus subtilis* (adapté de Straight et al., 2007). Des anticorps couplés à des nanoparticules d'or ont été utilisés pour cibler les protéines fluorescentes cyan (CFP) et jaune (YFP) fusionnées respectivement au domaine *trans*-AT PksE et à la sous-unité PksR.

V.A.2. Les domaines LINKS

L'équipe de Keatinge-Clay *et coll*. a étudié les structures de différents domaines KS afin de déterminer le rôle des domaines ATd (Gay et al., 2016). En effet, ce domaine retrouvé en position C-terminale des domaines KS est considéré comme permettant l'ancrage du domaine *trans*-AT lors de la synthèse du polycétide. A la suite de ce domaine, cette équipe a montré la présence de domaines non caractérisés composés d'une quarantaine de résidus, qui ont été nommés *Laterally-INteracting Ketosynthase Sequences* (LINKS) (Gay et al., 2016).

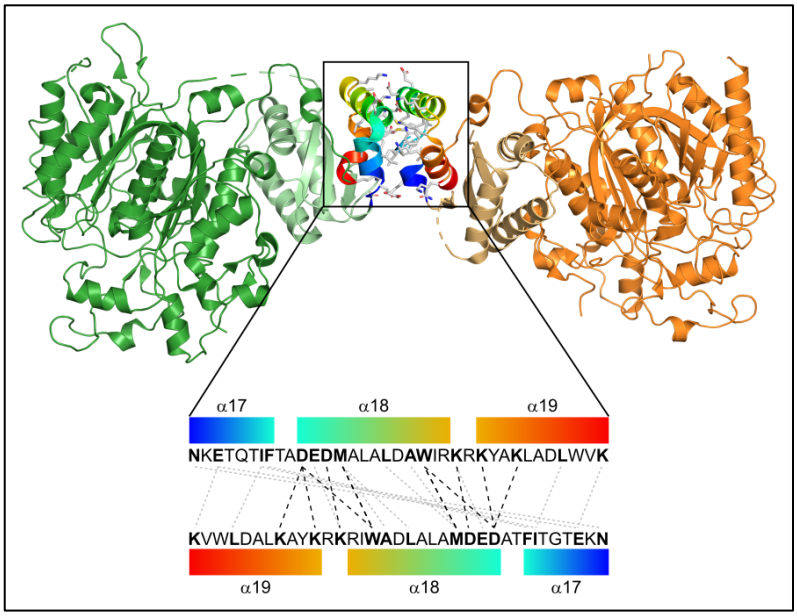


Figure 39. Analyse du domaine *Laterally-INteracting Ketosynthase Sequences* au sein d'un domaine cétosynthase de la bacillaene synthase (adapté de Gay et al., 2016). Identifiant PDB : 4NA1. Les domaines cétosynthase (KS) sont représentés en vert et orange foncé, tandis que les régions adaptatrices pour les domaines acyltransférase (AT) sont représentées en couleur plus claire. Les séquences *Laterally-INteracting Ketosynthase Sequences* (LINKS) sont représentées de couleur arc-en-ciel. Cette région est composée des hélices α17, α18 et α19 de la structure résolue. L'hélice α18 interagit avec son homonyme de manière anti-parallèle, tandis que l'hélice α17 interagit majoritairement avec l'hélice α19 du partenaire, et vice-versa. Les interactions entre des résidus éloignés d'entre 2,0 et 3,0 Å sont représentées en tiret noir, tandis que celles retrouvés entre deux résidus éloignés d'entre 3,1 et 4,0 Å sont représentées en pointillés gris clair.

Trois structures de domaine KS provenant de PKS *trans*-AT présentes dans la PDB, ainsi que sept nouvelles structures, ont été analysées (Gay et al., 2016). Le domaine KS du module 2 de la bacillaene synthase a fait partie de cette analyse (Figure 39). Trois hélices α – α17, α18 et α19 – ont ainsi été identifiées pouvant permettre la macro-complexation de ce système. L'hélice α18 interagit de manière antiparallèle avec l'hélice α18 homologue d'une seconde PKS, tandis que les hélices α17 et α19 interagissent respectivement avec les hélices α19 et α17 du partenaire (Gay et al., 2016). Les interactions à cette interface sont majoritairement de nature charge-charge, stabilisées par des interactions hydrophobes.

Une autre étude menée par Maier *et coll*. s'est penchée sur la structure et l'organisation de bimodules de PKS *trans*-AT (Tittes et al., 2022). Dans ce cadre, l'équipe s'est intéressée à une PKS-NRPS *trans*-AT orpheline de *Brevibacillus brevis* (Figure 40A).

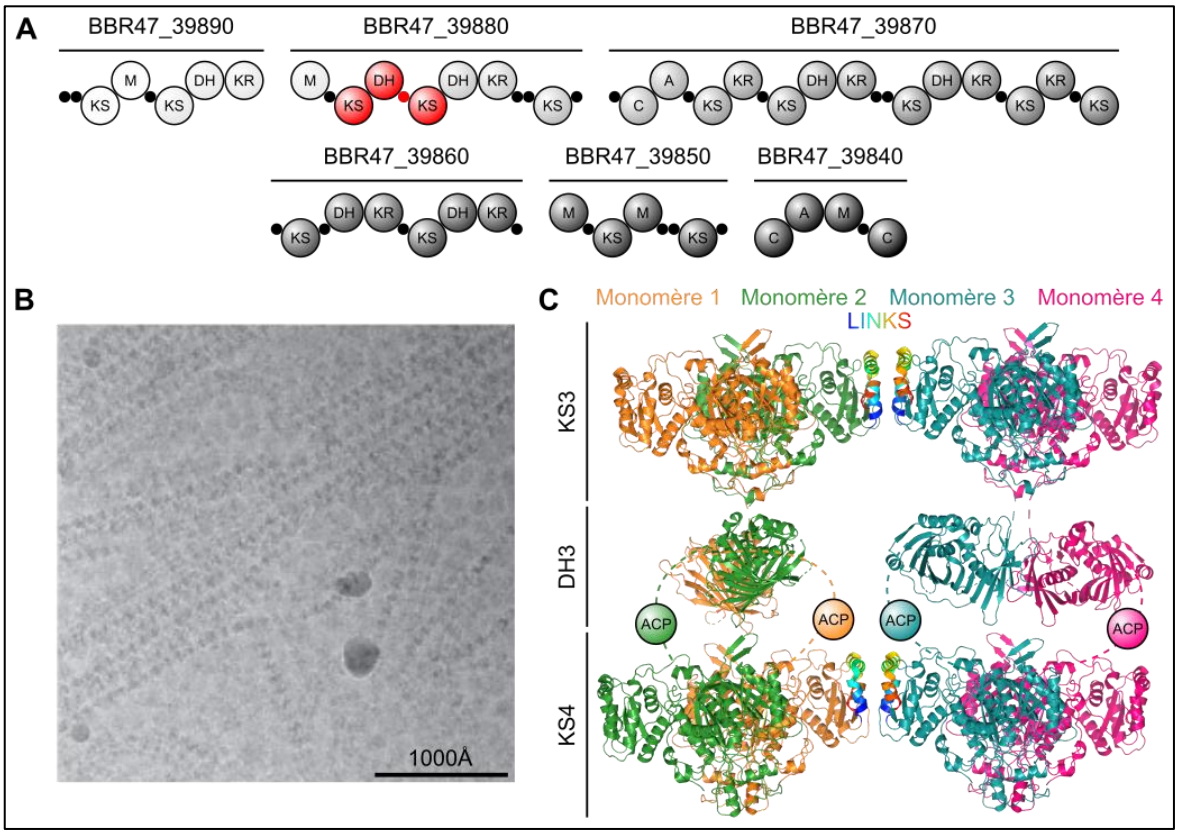


Figure 40. Macro-organisation d'un bi-module de polycétide synthase *trans*-acyltransférase (adapté de Tittes et al., 2022). **(A)** Représentation du système hybride polycétide synthase (PKS) – synthétase de peptides non ribosomaux (NRPS) *trans*-acyltransférase (AT) orphelin retrouvé chez *Brevibacillus brevis*. En rouge sont représentés les domaines étudiés – domaine cétosynthase (KS), déshydratase (DH) et domaine porteur d'acyle (ACP) du module 3 et KS du module 4 – formant le bi-module K3DAK4. Les cercles noirs représentent les domaines ACP. **(B)** Observation par cryo-microscopie électronique (cryo-EM) du bi-module K3DAK4. **(C)** Modèle de l'organisation des bi-modules K3DAK4 (identifiant PDB : 7ZSK). Ce modèle a été obtenu à partir des structures individuelles des domaines KS_3, DH_3 et KS_4 et de l'affinement des données de TEM (B). Chaque bi-module est représenté d'une couleur différente et les régions *Laterally-INteracting Ketosynthase Sequences* (LINKS) sont représentées de couleur arc-en-ciel. **Légende** – M : domaine de méthylation, KR : cétoréductase, C : domaine de condensation, A : domaine d'adénylation.

Cette PKS *trans*-AT est composé de dix-sept modules répartis en six sous-unités. L'étude s'est focalisée sur l'organisation d'un bi-module (ci-après nommé K3DAK4) composé des domaines KS, DH et ACP du module 3 et du domaine KS du module 4. Ce bi-module a été cloné, produit et purifié afin d'être observé par TEM (Figure 40B). Cette expérimentation a permis d'observer la formation de filaments par les bi-modules K3DAK4. En parallèle, les structures des domaines individuels KS_3, DH_3 et KS_4 ont été résolues par cristallographie aux rayons X. Couplées à l'affinement des données de cryo-EM, ces structures ont permis de

proposer un modèle de l'organisation des bi-modules K3DAK4 et des interactions qui la médient (Figure 40C).

V.B. L'organisation des PKS *cis*-AT

Ces domaines LINKS, bien que globalement conservés au sein des PKS *trans*-AT, semblent absents chez les PKS *cis*-AT. En effet, ces domaines se trouvent à la suite des domaines KS, là où se trouve le domaine AT chez ces dernières (Figure 41).

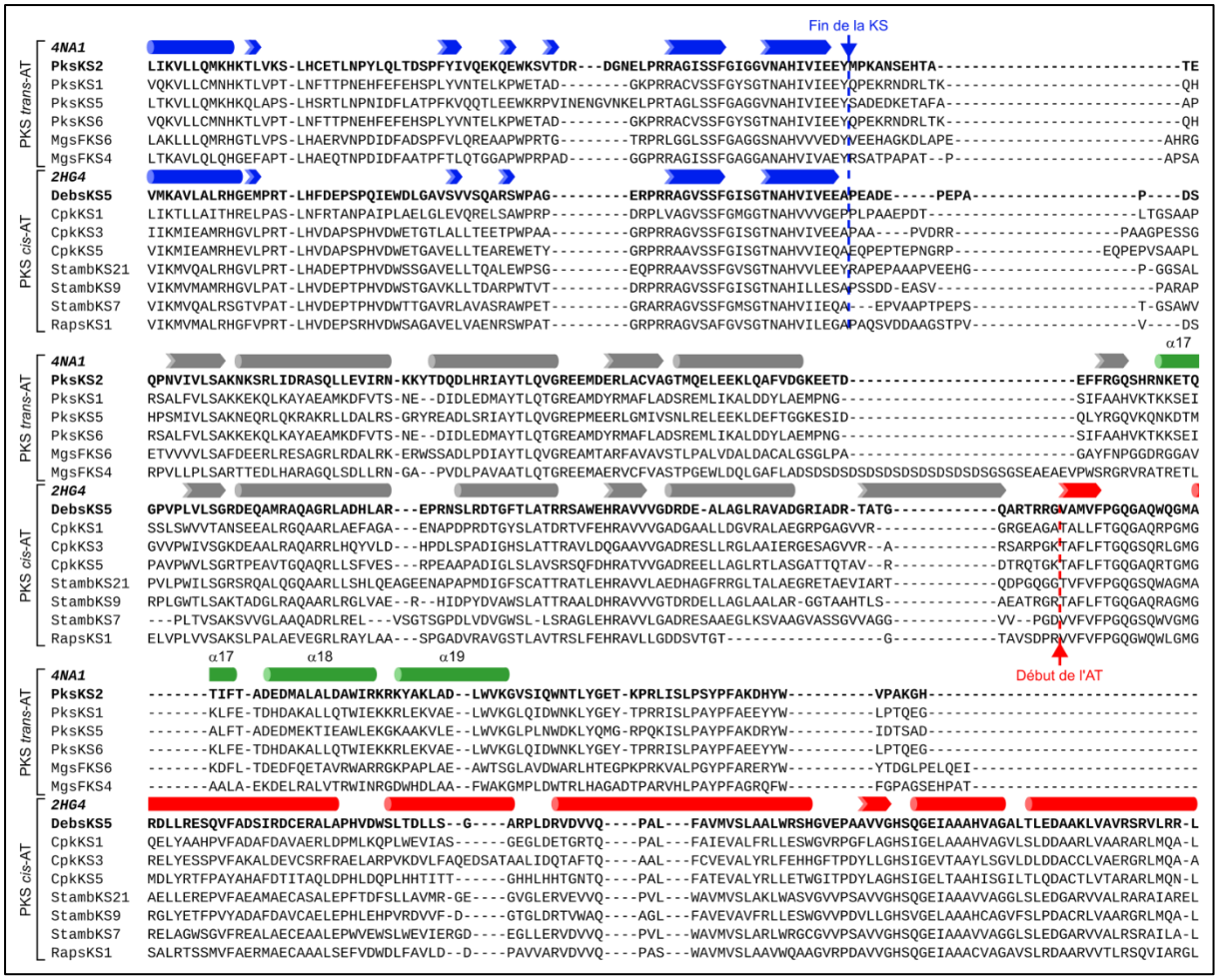


Figure 41. Alignements de séquences peptidiques et de structures secondaires de domaines cétosynthase et acyltransférase de polycétide synthase (données non publiées). Les séquences en gras sont celles dont la structure secondaire est représentée. La structure résolue du domaine cétosynthase (KS) suivi du domaine *Laterally-INteracting Ketosynthase Sequences* (LINKS) retrouvé au sein du module 2 de la PKS trans-acyltransférase (AT) bacillaene synthase (PksKS2) (identifiant PDB : 4NA1) a été utilisée dans le but de comparer les éléments de structures secondaires avec la structure du bidomaine KS-AT du module 5 de la PKS *cis*-AT 6-désoxyérythronolide B synthase (DEBS) (identifiant PDB : 2HG4). Les cylindres représentent des hélices α et les flèches des brins β et les séquences dont les éléments de structures sont présentés sont en gras. La structure des domaine KS est représentée en bleu, les régions linker en gris, le domaine LINKS en vert et le domaine AT en rouge. La fin domaine KS ainsi que le début du domaine AT sont représentées par des flèches de la couleur correspondante. **Légende** – MgsF : migrastatin synthase, Cpk : coelimycine synthase, Stamb : stambomycine synthase, Raps : rapamycine synthase.

L'alignement des régions KS-post-KS montre une conservation importante des séquences des domaines KS, ainsi que de leur structure, entre les PKS *cis*-AT et *trans*-AT. Les clades de KS *cis*-AT et *trans*-AT deviennent distinguables en termes de séquences au niveau de la région post-KS, en amont du domaine AT et des régions LINKS respectivement, malgré le fait que la composition en structure secondaire reste similaire. C'est au niveau des domaines AT ou LINKS qu'une différence de séquences et de structure est observée entre les PKS *cis*-AT et *trans*-AT respectivement. Ces alignements et comparaisons semblent montrer l'absence des régions LINKS au sein des PKS *cis*-AT.

V.C. La proximité entre les méga-complexes enzymatiques et la membrane bactérienne

L'entérobactine est un sidérophore produit par une NRPS retrouvée notamment chez *Escherichia. coli*. Ce système NRPS est composé des protéines EntB, EntE et EntF (Shi et al., 2020). EntB est un domaine PCP qui permet le recrutement du précurseur 2,3-dihydroxybenzoate par l'action du domaine d'acylation EntE. Ensuite, le module EntF, composé des domaines C–A–PCP–TE, permet d'ajouter une sérine au métabolite en production et de le libérer. Celui-ci va ensuite former un trimère capable de chélater le fer.

Ces enzymes responsables de la production de l'entérobactine ont été montrées comme étant associées à la membrane d'*E. coli* (Hantash & Earhart, 2000). Cette proximité avec la membrane peut s'expliquer par la nécessité d'excréter le composé. Il y aurait donc un lien ici entre la production et l'excrétion du métabolite.

La proximité de la bacillaene synthase avec la membrane chez *B. subtilis* n'est très certainement pas dû au hasard (Straight et al., 2007). Cette caractéristique a en effet été observé chez différents organismes et complexes enzymatiques, comme chez *Mycobacterium tuberculosis* (Jain & Cox, 2005). Le dimycocérosate de phthiocérol (PDIM) est un agent de virulence de *M. tuberculosis*, produit par l'assemblage de phthiocérol et d'acide mycocérosique. Ces deux molécules dérivent d'acides gras initiaux, pris en charge d'une part par les enzymes FadD26 et PpsA–E pour former le phthycérol et, d'autre part, par FadD28 et Mas pour obtenir l'acide mycocérosique. Ensuite, PapA5 catalyse l'assemblage de deux acides mycocérosiques sur les groupements hydroxyles du phthycérol. Enfin, ce composé est excrété notamment par MmpL7, un transporteur retrouvé dans la membrane des cellules bactériennes. L'objectif de cette étude était de mieux caractériser le transport du PDIM en identifiant l'interactome de MmpL7. Cela a permis d'identifier une interaction entre la dernière enzyme impliquée dans la synthèse du phthycérol – PpsE – et le transporteur. Cette proximité permet de coupler production et excrétion et l'adressage du métabolite produit.

VI. L'ingénierie des polycétide synthases pour la production de nouvelles molécules

L'une des voies pour la synthèse de nouveaux dérivés de polycétides et leur optimisation est l'ingénierie moléculaire des enzymes produisant ces composés (Figure 42).

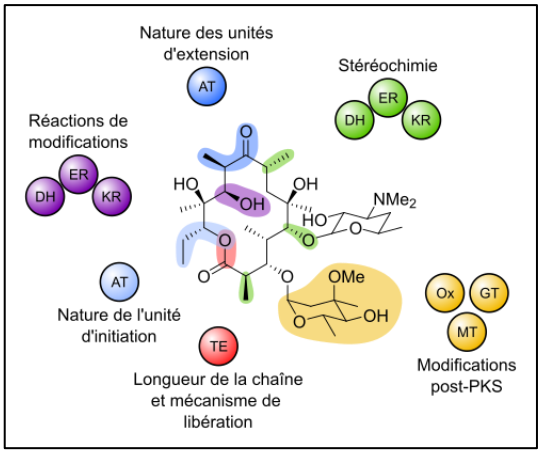


Figure 42. Potentielles cibles d'ingénierie au sein des PKS pour modifier la structure des polycétides (adapté de Weissman, 2016). L'exemple présenté est celui de l'érythromycine A. Les modifications sont colorées en accord avec les domaines ou les enzymes discrètes responsables. Tous ces domaines peuvent faire l'objet d'études d'ingénierie. **Légende** – AT : acyltransférase, DH : déshydratase, ER : énoyl-réductase, KR : cétoréductase, Ox : oxygénase, GT : glycosyltransférase, MT : méthyltransférase, TE : thioestérase.

L'ingénierie des PKS permet une *customisation* importante des composés en modifiant leur squelette polycétidique. La règle de colinéarité des PKS permet d'envisager et de mettre en place des stratégies d'ingénierie en inactivant, échangeant ou délétant certains domaines, modules ou sous-unités, tout en pouvant prédire le composé attendu produit par les PKS modifiées.

Pour des raisons de clarté, dans cette section les molécules d'intérêts seront numérotées de **1** à **31**, indépendamment des études dont elles sont issues et de la suite du manuscrit. Cette nomenclature n'est donc appliquée qu'au sein de cette partie et ne servira qu'à rendre le propos moins confus.

VI.A. La modification des groupements portés par les polycétides connus

VI.A.1. L'échange de domaine et modules pour influencer la nature des unités acylées recrutées

Lors de la synthèse des polycétides, les domaines AT sont responsables du recrutement de chaque unité acylée. La grande similarité des séquences nucléotidiques en position 5′ et 3′ des séquences codant ces domaines impliquent que l'évolution de ces derniers découle certainement d'événements de recombinaison homologue (Ridley et al., 2008). Cette caractéristique a conduit à de nombreuses études d'ingénierie des PKS par la substitution de

domaines AT afin de modifier la spécificité de ces complexes et ainsi obtenir de nouveaux composés.

VI.A.1.a. L'exemple de l'ingénierie de la 6-désoxyérythronolide B synthase

La PKS responsable de la production de l'érythromycine – DEBS – a fait l'objet de nombreuses tentatives d'ingénierie par le remplacement de ses domaines AT (Figure 43A). Cette PKS est composée de sept modules, répartis au sein de trois sous-unités (DEBS1–3) et est retrouvée dans le génome de *Sac. erythraea* (Cortes et al., 1990).

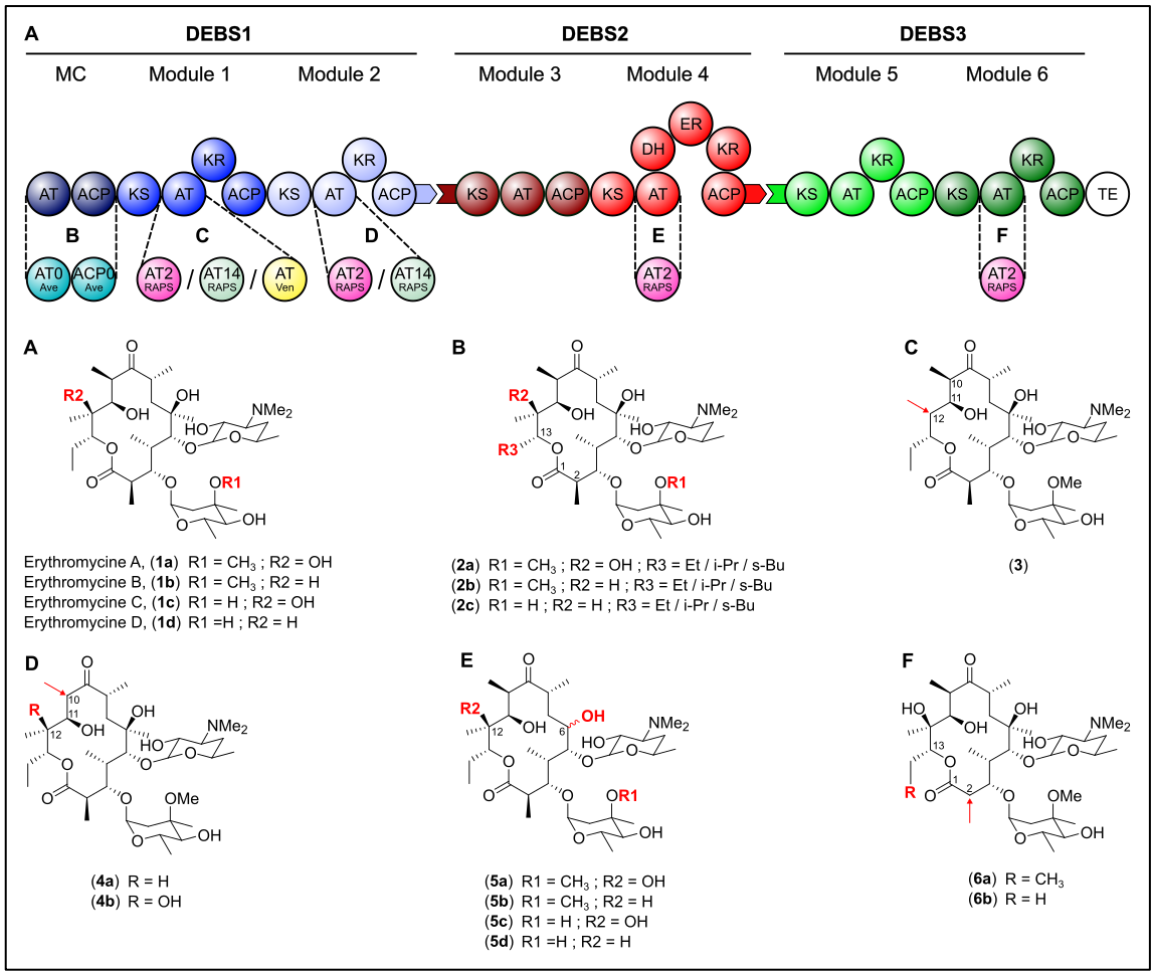


Figure 43. Ingénierie de la 6-désoxyérythronolide B synthase par modification des domaines acyltransférase. La 6-désoxyérythronolide B synthase (DEBS) est composée de six modules d'élongation et produit le 6-désoxyérythronolide B, précurseur des érythromycines A–D (**A**) (Cortes et al., 1990). Le module de chargement (MC) permet le recrutement d'une unité propionyl-CoA, et chaque module d'extension ajoute un méthylmalonyl-CoA. Le remplacement du module de chargement par celui présent au sein de l'avermectine synthase permet de diminuer la spécificité de substrat du domaine AT et d'obtenir de nouvelles molécules (**B**) (Marsden et al., 1998). Le remplacement du domaine AT des modules 1 ou 2 par le domaine AT du module 2 ou 14 de la rapamycine synthase (RAPS), ou de la vénémycine synthase, permet de recruter des unités malonyl-CoA et de nouveau obtenir de nouveaux composés (**C** et **D**) (Ruan et al., 1997). Enfin, le remplacement des domaines AT des modules des modules 4 et 6 par le domaine AT du module 2 de RAPS permet d'obtenir respectivement les composés présentés en **E** et **F** (Liu et al., 1997; Petkovic et al., 2003). **Légende** – AT : acyltransférase, ACP : domaine porteur d'acyle, KS : cétosynthase, KR : cétoréductase, DH : déshydratase, ER : énoyl-réductase, TE : thioestérase, Ave : avermectine synthase, RAPS : rapamycine synthase, Ven : vénémycine synthase.

Le module de chargement (MC) de DEBS a été remplacé *in vivo* par le MC provenant de l'avermectine synthase (Figure 43B) (Marsden et al., 1998). Le domaine AT du MC natif de DEBS permet le recrutement d'une unité propionyl-CoA, tandis que celui provenant de l'avermectine synthase permet de recruter des unités plus grandes : 2-méthylbutyryl-CoA ou isobutyryl-CoA (Yoon et al., 2004). Ce remplacement de MC a permis d'obtenir les composés **2a** à **2c**, qui sont des dérivés dont la taille du groupement porté par le carbone 13 du cycle (R3) diffère des érythromycines natives (**1a** à **1d**). Les composés **2a** et **2b** dont l'unité de départ a été le 2-méthylbutyryl-CoA montrent une activité antibactérienne contre *B. subtilis* semblable aux érythromycines A et B dont ils sont respectivement dérivés. Les rendements de production de ces composés ne sont cependant pas présentés par l'équipe de Leadlay *et coll*.

Un autre exemple de remplacement *in vivo* est celui du domaine AT du premier module d'élongation chez *Sac. erythraea* ER720 soit par le domaine AT du deuxième module d'élongation de la rapamycine synthase ($^{RAPS}AT_2$), soit par celui de son quatorzième module d'élongation ($^{RAPS}AT_{14}$), soit par le domaine AT provenant de la vénémycine synthase (^{Ven}AT) (Figure 43C) (Ruan et al., 1997). Ces trois domaines ont une spécificité de substrat pour le malonyl-CoA (Schwecke et al., 1995; Thanapipatsiri et al., 2016), tandis que le domaine AT natif de DEBS permet le recrutement de méthylmalonyl-CoA (Cortes et al., 1990). Ces modifications ont donc permis d'obtenir un dérivé 12-déméthyl-12-désoxyérythromycine A (**3**), ayant des propriétés antibactériennes contre *S. aureus* (Figure 43, 3). Basée sur cette analyse de l'activité du composé, il a été démontré que l'utilisation de domaine $^{RAPS}AT_2$ permet d'obtenir 2 à 4 fois plus de ce dérivé que les domaines $^{RAPS}AT_{14}$ et ^{Ven}AT. Cependant, les rendements de production des souches mutés n'ont ici pas été comparés à ceux de la souche sauvage.

Un remplacement similaire a également été effectué *in vivo* chez *Sac. erythraea* ER720 au sein du deuxième module d'élongation, dont le domaine AT natif a également une spécificité de substrat pour le méthylmalonyl-CoA (Figure 43D). L'utilisation ici des domaines $^{RAPS}AT_2$ ou $^{RAPS}AT_{14}$ a permis d'obtenir des versions 10-déméthylées des érythromycines B (**4a**) et C (**4b**) (Ruan et al., 1997). Ces composés sont également biologiquement actifs contre *S. aureus*.

Ensuite, le domaine AT du quatrième module d'élongation de DEBS, spécifique des unités méthylmalonyl-CoA, a été remplacé par le domaine $^{RAPS}AT_2$ (Figure 43E) (Petkovic et al., 2003). Cette modification a été apportée à la souche *Sac. erythraea* JC2, délétée du BGC lié à la production de l'érythromycine à l'exception de la séquence codant le domaine TE, utilisée pour l'insertion de matériel génétique par recombinaison homologue. Ainsi, les systèmes de DEBS modifiés ici ont été intégré au sein du génome de cette souche par recombinaison à partir d'un plasmide contenant les sous-unités de DEBS exprimées sous le

contrôle du promoteur *actI* provenant de *S. coelicolor*. Cela a permis d'obtenir des dérivés 6-désméthylés des érythromycines A – D (**5a** – **5d**). Le composé **5d** en particulier a montré une activité antibiotique semblable à l'érythromycine D native et a été obtenue à une concentration de 0,3 $mg \cdot L^{-1}$ de milieu de culture.

Enfin, la même substitution a été effectuée au niveau du dernier module d'élongation de DEBS, dont le domaine AT natif permet également le recrutement d'unités méthylmalonyl-CoA (Figure 43F) (Liu et al., 1997). Cette fois-ci, cette expérimentation a été menée par expression hétérologue chez *S. coelicolor* CH999, délétée du BGC lié à la production d'actinorhodine ainsi qu'une mutation inhibant la production d'undecylprodigiosine, et ne contenant aucun plasmide endogène (McDaniel et al., 1993). Un plasmide conjugatif contenant les gènes codant les sous-unités de DEBS a donc été utilisé. Ainsi, des dérivés 2-déméthylés de l'érythromycine A ont été obtenus (**6a** et **6b**).

Le goulot d'étranglement majeur des études d'ingénierie des PKS est le rendement de production. En effet, Ashley *et coll*. ont substitué les domaines AT des modules 2, 5 et 6 de DEBS par des domaines provenant d'autres PKS et ainsi construit 38 souches et identifié une trentaine de nouveaux composés (McDaniel et al., 1999). Cependant, les rendements de production de tous les composés (entre 0,7 et moins de 0,005 $mg \cdot L^{-1}$) sont bien inférieurs à la production de l'érythromycine (20 $mg \cdot L^{-1}$), et certains composés ne sont pas détectables.

Il n'est pas toujours nécessaire de modifier la PKS elle-même afin d'influencer la nature des unités acylées ajoutées au polycétide. En effet, une étude récente s'est intéressée au large spectre de spécificité du domaine AT du module 12 de la stambomycine synthase, capable de prendre en charge six substrats différents (Su et al., 2025). Dans ce cadre, le génome de *S. ambofaciens* a été modifié afin de diminuer la production des précurseurs natifs du domaine AT_{12}. Ensuite, des substrats non natifs ont été incorporés au milieu de culture afin d'obtenir de nouveaux dérivés de stambomycines, cette approche est appelée mutasynthèse.

Cependant, l'ingénierie des domaines de PKS a ses limites, en ce sens qu'intégrer ou déléter une partie de module peut drastiquement influencer leur activité et *in fine* la synthèse des polycétides. Cela peut s'expliquer par le fait que les tentatives d'ingénierie décrites jusqu'ici dans ce manuscrit ne s'intéressent en majorité qu'à la modification d'un domaine spécifique, sans en prendre en compte le contexte moléculaire. En effet, l'une des explications plausibles et qu'un domaine AT retrouvé au sein d'un module spécifique, au sein d'une PKS donnée, ne puisse être isolé de son environnement pour être placé dans un nouveau module sans conséquence sur l'efficacité de la communication entre des domaines non nativement retrouvés ensemble. Ainsi, une autre approche par modification de modules complets semble être une solution afin de garder en partie l'environnement natif des domaines modifiés.

VI.A.1.b. Le passage de l'échelle de l'échange de domaine à l'échange de modules : vers une nouvelle définition des modules pour l'ingénierie

Afin de faciliter les études menées en ingénierie des PKS, l'équipe de Leadlay et coll. a tout d'abord créé un système modèle en 1995, composé de la première sous-unité de DEBS – DEBS1, composée du module de chargement et de deux modules d'élongation – suivi du domaine TE du même système et ci-après nommé DEBS1-TE (Figure 44A) (Cortes et al., 1995). Ainsi, ils ont pu obtenir une PKS modèle produisant un tricétide lactone à raison de 8,9 $mg \cdot L^{-1}$. Ce système DEBS1-TE a également fait l'objet d'études d'ingénieries pour la modification des domaines AT.

En effet, l'équipe de Leadlay *et coll*. a pu, à partir d'un plasmide exprimant un système DEBS1-TE, remplacer le domaine AT du premier module d'élongation par celui du module 2 de RAPS, afin d'obtenir un tricétide lactone déméthylé (Figure 44B) (Oliynyk et al., 1996).

Ce composé a été obtenu par expression hétérologue du plasmide contenant la mutation d'intérêt au sein de la souche *S. coelicolor* CH999 précédemment citée. Cela a ensuite inspiré les tentatives d'ingénierie sur la PKS DEBS complète, citées précédemment (Figure 43, 3 et 4) (Ruan et al., 1997), puis d'échange de module complet entre DEBS et RAPS (Ranganathan et al., 1999). Dans ce cadre, le module 2 de DEBS a été échangé au sein du modèle DEBS1-TE avec le module 12 de RAPS (Figure 44C). Pour ce faire, les auteur·rices ont identifié des points de fusion au milieu des domaines $^{DEBS}ACP_1$ et $^{RAPS}ACP_{11}$ et des domaines $^{RAPS}ACP_{12}$ et $^{DEBS}ACP_2$. Cependant, cette PKS modifiée ne produit pas le composé attendu. Cette fois-ci, le système modifié a été exprimé par le biais de plasmide d'expression au sein de la souche *Sac. erythraea* JC2 et non pas *S. coelicolor* CH999 car cette dernière n'a pas permis d'obtenir de PKS hybrides fonctionnelles (Ranganathan et al., 1999).

Une seconde PKS hybride a été construite, cette fois-ci en échangeant les domaines AT-KR-ACP du module 2 de DEBS avec ceux du module 12 de RAPS (Figure 44C). Ainsi, le domaine $^{DESB}KS_2$ est conservé, tout comme les potentielles interactions avec le domaine $^{DEBS}ACP_1$ nécessaires à la synthèse des polycétides. Les composés attendus ont ainsi pu être obtenus et purifiés à un rendement de 1,0 et 2,4 $mg \cdot L^{-1}$ respectivement pour les composés **11** et **12**. Cette réussite lève un point important quant à la définition même d'un module au sein de ces enzymes et met en lumière l'importance des interactions entre un domaine KS et le domaine ACP du module précédent.

De la même manière, les modules 1 et 2 de DEBS1-TE ont été remplacés par les modules 11 et 12 de RAPS (Figure 42D). Ainsi, la nouvelle PKS obtenue est composée du module d'initiation de DEBS, puis des modules 11 et 12 de RAPS, et enfin du domaine TE de DEBS.

Une nouvelle fois, ce système ne permet pas la synthèse des tricétides lactones attendus. Une alternative a été construite, contenant cette fois-ci le module de chargement ainsi que le domaine KS_1 de DEBS, puis les modules 11 et 12 de RAPS à partir du domaine AT_{11} jusqu'au domaine ACP_{12} inclus. Ce système a permis d'obtenir les composés attendus à une quantité d'environ 1,0 $mg \cdot L^{-1}$ pour les molécules **9** et **10**, mettant une nouvelle fois en avant l'importance des interactions entre un domaine KS et le domaine ACP précédent. Ces résultats montrent que les interactions entre un domaine KS_n et le domaine ACP_{n-1} semblent primordiales pour la fonctionnalité des PKS.

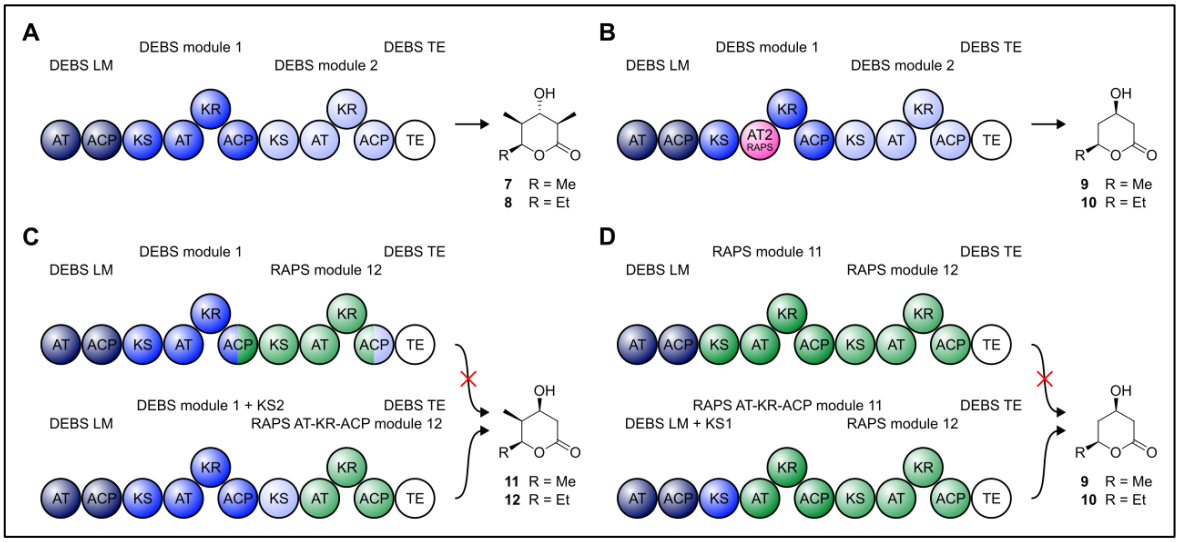


Figure 44. Ingénierie de la polycétide synthase minimale DEBS1-TE. **(A)** PKS minimale composée des deux premiers modules de la 6-désoxyérythronolide B synthase (DEBS), initialement portée par la première sous-unité DEBS1, suivi du domaine thioestérase (TE) de ce système (adapté de Cortes et al., 1995). Cette PKS a été nommée DEBS1-TE et permet la synthèse d'un polycétide de taille réduite. En effet, celui-ci est composé de trois blocs acylés, le composé est donc appelé tricétide lactone. **(B)** Ingénierie de DEBS1-TE par échange de domaines acyltransférase (AT) (adapté de Oliynyk et al., 1996). Le domaine AT du module 1 de DEBS1-TE a été remplacé par le domaine AT provenant du deuxième module de la rapamycine synthase (RAPS). Cela a pour but d'obtenir un tricétide déméthylé et mesurer l'efficacité d'une telle stratégie d'ingénierie. **(C)** Echange du module 2 de DEBS1-TE par le module 12 de RAPS (adapté de Ranganathan et al., 1999). La première stratégie consiste à intégrer le module 12 à partir de la moitié de l'ACP du module 11 de RAPS jusqu'au milieu de l'ACP du module 12. Cette construction ne permet pas la synthèse du composé attendu. La seconde stratégie est de garder le domaine KS du module 2 de DEBS et de remplacer les domaines AT, KR et ACP par ceux du module 12 de RAPS. Cette construction permet la synthèse du tricétide lactone attendu. **(D)** Echange des modules 1 et 2 de DEBS1-TE par les modules 11 et 12 de RAPS (adapté de Ranganathan et al., 1999). Le remplacement complet des modules n'a pas permis d'obtenir la molécule prédite. Le domaine KS du premier module de DEBS a donc été gardé, puis les domaines AT, KR et ACP remplacé par ceux du module 11 de RAPS, suivi du module 12 complet. Cela a permis d'obtenir le composé attendu.

La définition des modules au sein des PKS modulaires est une notion qui, de la définition donnée précédemment, tend à s'élargir au vu des tentatives d'ingénierie. Il existe aujourd'hui deux définitions pouvant s'appliquer à la répartition des domaines au sein des PKS modulaires. La première est celle auparavant énoncée : un module est composé *a minima* et dans l'ordre d'un domaine KS, puis d'un domaine AT et enfin d'un domaine ACP.

Cette définition a été utilisée dans le cadre d'échanges de modules entre PKS, cependant cela n'a pas donné l'effet escompté.

Des résultats similaires ont été observés dans le cadre de l'ingénierie des systèmes NRPS. Dans ce cadre, un nouveau terme d'« unité échangeable » a été proposé pour caractériser de nouvelles bornes de modules utilisables en ingénierie (Figure 45) (Bozhüyük et al., 2018). Ces bornes définissent un module comme composé *a minima* des domaines A-PCP-C et non pas C-A-PCP selon la définition classique des modules NRPS.

En parallèle, une nouvelle définition des modules au sein des PKS est également proposé, afin de borner les modules de manière à ce qu'ils soient composés dans l'ordre des domaines AT-ACP-KS et non plus KS-AT-ACP (Keatinge-Clay, 2017). Ainsi, le terme d'unité échangeable est également proposé afin de définir ces nouvelles bornes de modules PKS (Su et al., 2022). L'application de cette nouvelle définition sur les modules de DEBS est représentée dans la Figure 45.

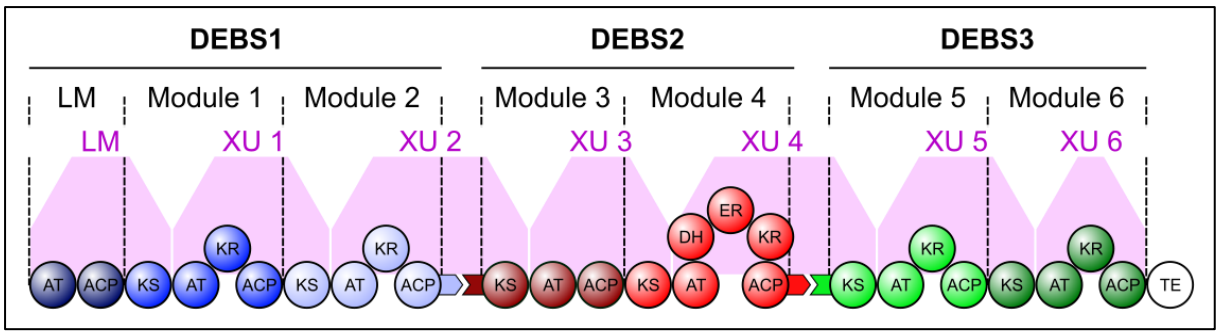


Figure 45. Représentation schématique de la répartition en modules et en unités échangeables de la 6-désoxyérythronolide B synthase (DEBS). La définition conventionnelle d'un module en défini les bornes de manière à contenir les domaines KS-AT-(DH)-(ER)-(KR)-ACP. Cependant, les tentatives d'ingénierie des PKS ont montré que les échanges de modules sont plus efficaces lorsque ceux-ci sont composés des domaines AT-(DH)-(ER)-(KR)-ACP-KS. Cela amène alors une seconde définition du module, appelée unité échangeable (XU).

VI.A.1.c. L'échange de modules pour l'ingénierie d'autres systèmes : le cas de la pikromycine synthase

Une étude plus récente menée par l'équipe de Keatinge-Clay *et coll*. a également permis d'étudier l'échange de modules dans le but d'obtenir de nouveaux dérivés de polycétides par le biais de domaines AT aux spécificités différentes. Cette étude a été menée cette fois-ci sur la PKS synthétisant la pikromycine (**13**) (Pik) (Wang et al., 2025). Cette PKS est composée de sept modules, répartis au sein de quatre sous-unités (PikAI–IV) (Figure 46A) (Xue et al., 1998). Dans ce cadre, la souche *S. venezuelae* ATCC15439 a été utilisée et son génome a été édité par l'utilisation du système *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR) associé à l'endonucléase Cas9.

La nouvelle définition des modules proposée a ainsi été utilisée, cependant pour des raisons de clarté et homogénéité, la nomenclature utilisée ici pour en discuter sera basée sur l'ancienne définition des modules.

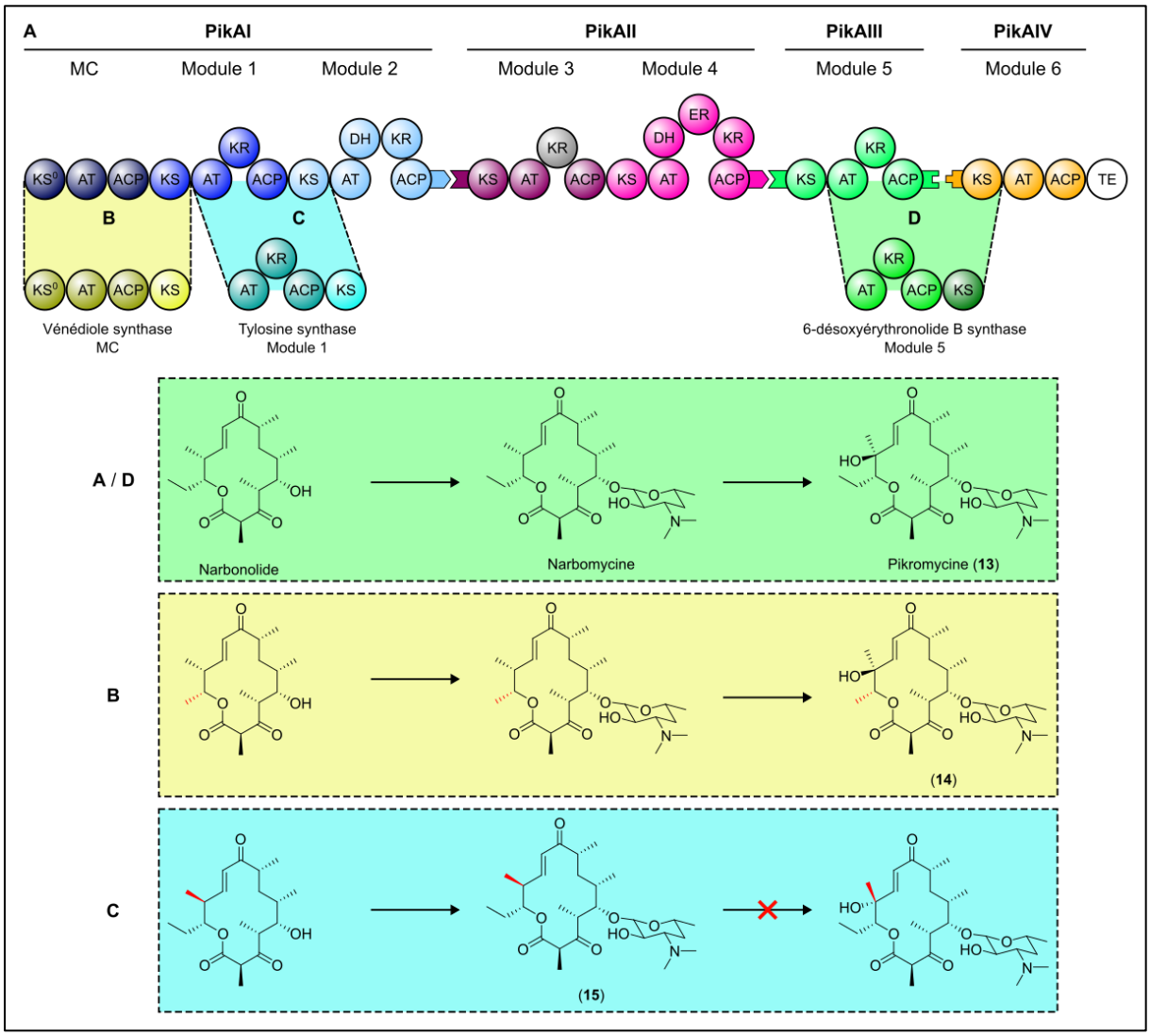


Figure 46. Ingénierie de la pikromycine synthase par échange de modules (adapté de Wang et al., 2025). **(A)** La pikromycine synthase (Pik) est composée de sept modules répartis au sein de quatre sous-unités (PikAI–IV) et produit le narbonolide, précurseur de la pikromycine (**13**) (Xue et al., 1998). **(B)** L'échange du module de chargement (MC) de la pikromycine synthase avec celui de la vénédiole synthase a permis d'obtenir un dérivé de la pikromycine portant un groupement méthyle et non plus éthyle au niveau de carbone 13 (**14**). **(C)** Lorsque le premier module d'élongation, composé des domaines AT_1-KS_1-ACP_1-KS_2, est échangé avec celui de la tylosine synthase, cela modifie l'activité du domaine cétoréductase (KR) et ainsi la stéréochimie du produit obtenu. Cela inhibe la réaction de 6-oxydation médiée par une enzyme post-PKS, et le produit obtenu est donc un dérivé de la narbomycine : la 12*R*-méthyl-6-désoxy-narbomycine (**15**). **(D)** Enfin, l'échange du module d'élongation 5 (AT_5-KR_5-ACP_5-KS_6) avec celui provenant de la 6-désoxyérythronolide B synthase (DEBS) ne change rien à la nature de l'unité acylée recrutée par le domaine AT, ni au mode d'action du domaine KR. Le produit obtenu dans ce cas est donc la pikromycine native (**13**). **Légende** – ACP : domaine porteur d'acyle, KS^0 : domaine cétosynthase inactif, KS : cétosynthase, AT : acyltransférase, KR : cétoréductase, DH : déshydratase, ER : énoyl-réductase, TE : thioestérase.

L'objectif de cette étude a été d'optimiser l'utilisation de CRISPR-Cas9 pour la délétion et l'insertion de grandes séquences nucléotidiques au sein de génomes riches en paire de bases G-C, ici *S. venezuelae*. Dans un premier temps l'équipe a échangé le module de chargement de la pikromycine synthase avec celui retrouvé au sein de la vénédiole synthase dont l'organisation est similaire (Figure 46B). Cet échange a concerné ainsi les domaines KS^0_{MC}-AT_{MC}-ACP_{MC}-KS_1. Cela a permis de modifier la spécificité de substrat du domaine AT, qui recrute une unité propionyl-CoA au sein de Pik et une unité acétyl-CoA au sein de la vénédiole synthase. Le composé ainsi obtenu est un dérivé de la pikromycine dont le groupement porté par le carbone 1 n'est plus un éthyle mais un méthyle (**14**). Ce composé a été obtenu à une concentration de 7,5 $mg \cdot L^{-1}$, ce qui représente 30 % des rendements de production en pikromycine de la PKS sauvage.

Ensuite, le premier module d'élongation (AT_1-KR_1-ACP_1-KS_2) a été échangé avec le premier module d'élongation de la tylosine synthase, dont la composition en domaine est similaire (Figure 46C). Cet échange de module a permis de modifier la nature du domaine KR et ainsi obtenir un composé à la stéréochimie modifiée. Cela n'a pas empêché le précurseur de subir une 3*O*-glycosylation, cependant le produit obtenu n'a pas pu être méthylé en position 12. Le produit obtenu est donc un analogue 12*R*-méthyl-6-désoxy-narbomycine (**15**). Une nouvelle fois cependant, les rendements de production de ce système étaient plus faibles que ceux de la PKS sauvage : 0,15 $mg \cdot L^{-1}$.

Enfin, le cinquième module d'élongation de Pik a été échangé avec celui provenant de DEBS (Figure 46D). Les deux systèmes étant très similaires et ces deux modules identiques en termes de spécificité de substrat, cela ne change rien au composé final obtenu qui est alors la pikromycine (**13**). Les rendements de production de ce système modifié comprenant le module 5 de DEBS représentaient 80 % (20 $mg \cdot L^{-1}$) des rendements de la PKS native.

Ces résultats montrent que, malgré la prise en compte du contexte moléculaire des domaines d'intérêts lors des tentatives d'ingénierie, les rendements de production ne sont toujours pas optimisés (entre 20 et 80 % des rendements de production de la PKS sauvage). Dans le cas de la modification de la nature du domaine KR, cela peut s'expliquer par le fait que les domaines en aval de celui-ci aient une affinité plus faible pour un substrat dont la stéréochimie n'est pas celle du substrat natif (Wang et al., 2025). Et pour ce qui est de la modification de la nature du domaine AT, cela peut s'expliquer par le même phénomène, mais également par la disponibilité en substrat au sein de la cellule.

Toutefois, des études d'ingénierie des PKS dans le but de modifier la stéréochimie des polycétides produits ont déjà été menées. L'objectif de ces études a été de modifier la nature des domaines KR présents au sein de PKS, par l'échange de domaines seuls ou de modules complets.

VI.A.2. L'échange de domaines et modules pour modifier la stéréochimie des polycétides

La stéréochimie des polycétides dépend de la nature des domaines KR présents au sein de la PKS. Une manière d'obtenir de nouvelles molécules consiste donc à l'échange de domaines KR afin de produire des composés à la stéréochimie modifiée. C'est ce qui a été fait par l'équipe de Weissman *et coll.* dont l'étude s'est tout d'abord portée sur la modification d'un domaine KR au sein de la PKS DEBS1-TE afin d'en changer la spécificité (Annaval et al., 2015). Le domaine KR retrouvé au sein du module 2 est de type B1, et ici l'objectif a été de le remplacer par des domaines de type A2 et B2. Cependant, les rendements de production ainsi obtenus sont faibles, lorsque les composés attendus sont observés (Tableau 1). Afin de mesurer l'impact de l'environnement moléculaire des domaines KR sur les rendements de production, le module complet a été échangé avec des modules provenant d'autres PKS et contenant des domaines KR de type A2 ou B2 (Massicard et al., 2020).

Quatre stratégies de remplacement du module 2 de DEBS1-TE ont été étudiées, différant par les bornes sélectionnées au niveau des domaines KS et ACP (Figure 47) (Massicard et al., 2020). La borne sélectionnée en position N-terminale permet soit de conserver le domaine KS du module 2 de DEBS, soit de le changer par le domaine du module introduit en suivant la définition des XU. En position C-terminale, le domaine ACP est soit entièrement remplacé par celui du module introduit, soit en partie conservé par la création d'un domaine hybride afin de conserver les interactions potentielles avec le domaine KS précédent. Ainsi, les stratégies A et B sont de garder intact le domaine ACP introduit, en ayant respectivement en position N-terminale le domaine KS du module introduit ou de DEBS. Les stratégies C et D quant à elles sont d'utiliser un domaine ACP hybride entre le domaine du module introduit et celui de DEBS, tout en ayant en position N-terminale le domaine KS introduit ou natif respectivement.

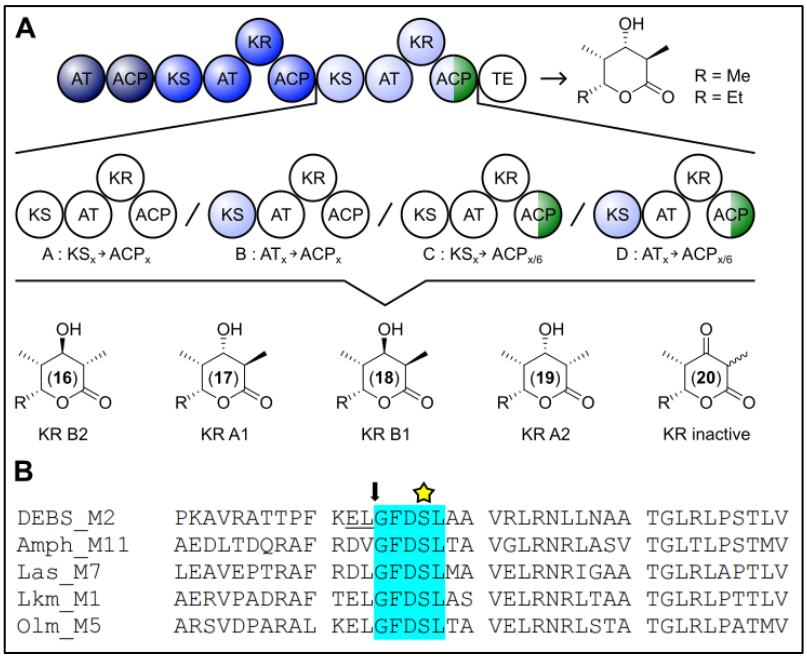


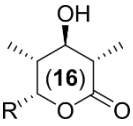
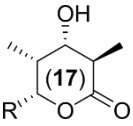
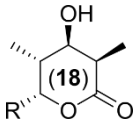
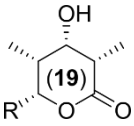
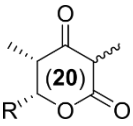
Figure 47. Stratégies d'ingénierie d'une polycétide synthase pour modifier la stéréochimie de la molécule produite (adapté de Massicard et al., 2020). **(A)** Représentation schématique des stratégies envisagées. Les quatre stratégies diffèrent dans les bornes utilisées pour introduire le module exogène au sein de la PKS modèle DEBS1-TE Le domaine cétosynthase (KS) est soit remplacé par celui du module introduit (stratégies A et C), soit celui du module 2 de DEBS1-TE est conservé (stratégies B et D). De l'autre côté, soit le domaine porteur d'acyle (ACP) est celui du module introduit (stratégies A et B), soit un hybride construit entre le domaine natif et introduit (stratégies C et D). **(B)** Alignement de séquences pour identifier le point de jonction au sein du domaine ACP hybride. La flèche indique le point de fusion entre le domaine ACP introduit et natif, en amont du motif conservé GFDS contenant la sérine modifiée par le bras phosphopantéthéine, marquée d'une étoile. Les résidus soulignés correspondent au site de reconnaissance de l'enzyme de restriction *Sac*I utilisée lors de la construction originelle de DEBS1-TE. **Légende** – AT : acyltransférase, KR : cétoréductase, Amph : amphotéricine synthase, Las : lasalocide synthase, Lkm : lankamycine synthase, Olm : oligomycine synthase.

Dans cette étude, huit modules de PKS contenant des domaines KR de type A2 ont été utilisés, ainsi que trois modules comprenant des domaines KR de type B2. Parmi ces modules, seuls quatre KR de type A2 ont permis d'observer la production des composés recherchés, contre une seule pour les types B2 (Tableau 1). Dans chaque cas, les rendements de production des composés majeurs attendus ainsi que des versions non épimérisées et non réduites du tricétide lactone ont été mesurés par spectrométrie de masse.

Les modules utilisés ayant permis d'obtenir les composés attendus montrent que les stratégies C et D, dans lesquelles le domaine ACP est un hybride entre le domaine natif et celui introduit, ne semblent pas fonctionnelles. Dans le cas des stratégies A et B, les résultats sont différents selon que le domaine KR est de type A ou B. En effet, dans le cas de l'insertion de domaines KR A2, la stratégie A intégrant le domaine KS exogène en entier montre de meilleurs rendements que la B, tandis que l'inverse est observé dans le cas des KR B2. La production de tricétide lactone non réduit a également été mesurée.

Cela a permis d'observer que malgré l'absence des composés d'intérêt dans le cas de l'utilisation du module 5 de l'oligomycine synthase (Olm M5), celui-ci n'interrompt pas pour autant la voie de biosynthèse puisque le composé non réduit est retrouvé.

Tableau 1. Rendements de production de lactones des systèmes DEBS1-TE modifiées par introduction de domaines KR de type A2 et B2 (adapté de Annaval et al., 2015; Massicard et al., 2020).

		Rendements de production (mg/L)				
		 (16)	 (17)	 (18)	 (19)	 (20)
Module introduit	**Stratégie**					
Amph M11 (A2)	KR seule	traces			3,75	5,5
	A	6,1 ± 0,5			36 ± 1	6 ±1
	B	6,8 ± 0,7			20,2 ± 0,8	8,7 ± 0,4
	C	nd			nd	nd
	D	nd			nd	nd
Nys M1 (A2)	A	traces			9,4 ± 0,2	≈ 2
	B	traces			≈ 1,2	≈ 1,6
Nys M11 (A2)	A	≈ 2			18,9 ± 0,8	≈ 2,6
	B	≈ 4			9 ± 2	≈ 3,8
Olm M5 (A2)	KR seule	nd			nd	52
	A	nd			nd	14,6 ± 0,6
	B	nd			nd	6 ± 2
	C	nd			nd	nd
	D	nd			nd	nd
Lkm M1 (B2)	KR seule	traces	nd			5
	A	≈ 3,5	nd			≈ 2,7
	B	14 ± 1	traces			≈ 4,7
	C	nd	nd			nd
	D	nd	nd			nd

Pour chaque construction, les rendements de production du composé principal attendu sont surlignés en bleu. Les stratégies « KR seul » correspondent au remplacement des domaines seuls (Annaval et al., 2015) et les stratégies A–D au remplacement de modules (Massicard et al., 2020). **Légende** – nd : non détecté.

Ces résultats permettent d'illustrer l'importance du contexte moléculaire sur l'activité et la fonctionnalité des domaines individuels. En effet, le remplacement seul des domaines KR au sein de DEBS1-TE montre des rendements de production des composés attendus plus faibles que dans le cas du remplacement des modules complets.

VI.B. La modification de la taille des polycétides

L'une des voies pour produire de nouveaux dérivés de polycétides est la synthèse chimique de ces dérivés, comme précédemment présenté avec l'exemple de l'éribuline (Figure 12). Simplifier la structure de polycétides d'intérêts est dans ce cas nécessaire dans le but de faciliter leur synthèse chimique.

Ce chapitre se concentrera sur les efforts fait dans le but de simplifier la rapamycine et les stambomycines par ingénierie des PKS les produisant, ces deux systèmes ayant fait l'objet d'études lors de cette thèse.

VI.B.1. La synthèse de stambomycines de taille réduite

Un exemple récent d'ingénierie des PKS, effectuée au sein des laboratoires DynAMic et IMoPA par Dr. Li Su durant sa thèse, concerne la modification des stambomycines (**21**) (Su, 2021 ; Su et al., 2022). Les stambomycines sont des macrolides produits par la bactérie *S. ambofaciens.* Il s'agit d'un complexe de macrolides (stambomycine A à F) formés d'un cycle macrolactone de 51 atomes de carbone, et qui montrent des propriétés antiprolifératives et antibactériennes (Laureti et al., 2011). Le cluster de gènes responsable de la synthèse de ces composés est constitué de 25 gènes, dont neuf codent des sous-unités de PKS (Figure 48A). Ces neuf sous-unités contiennent un total de 24 modules d'élongation (Figure 48B).

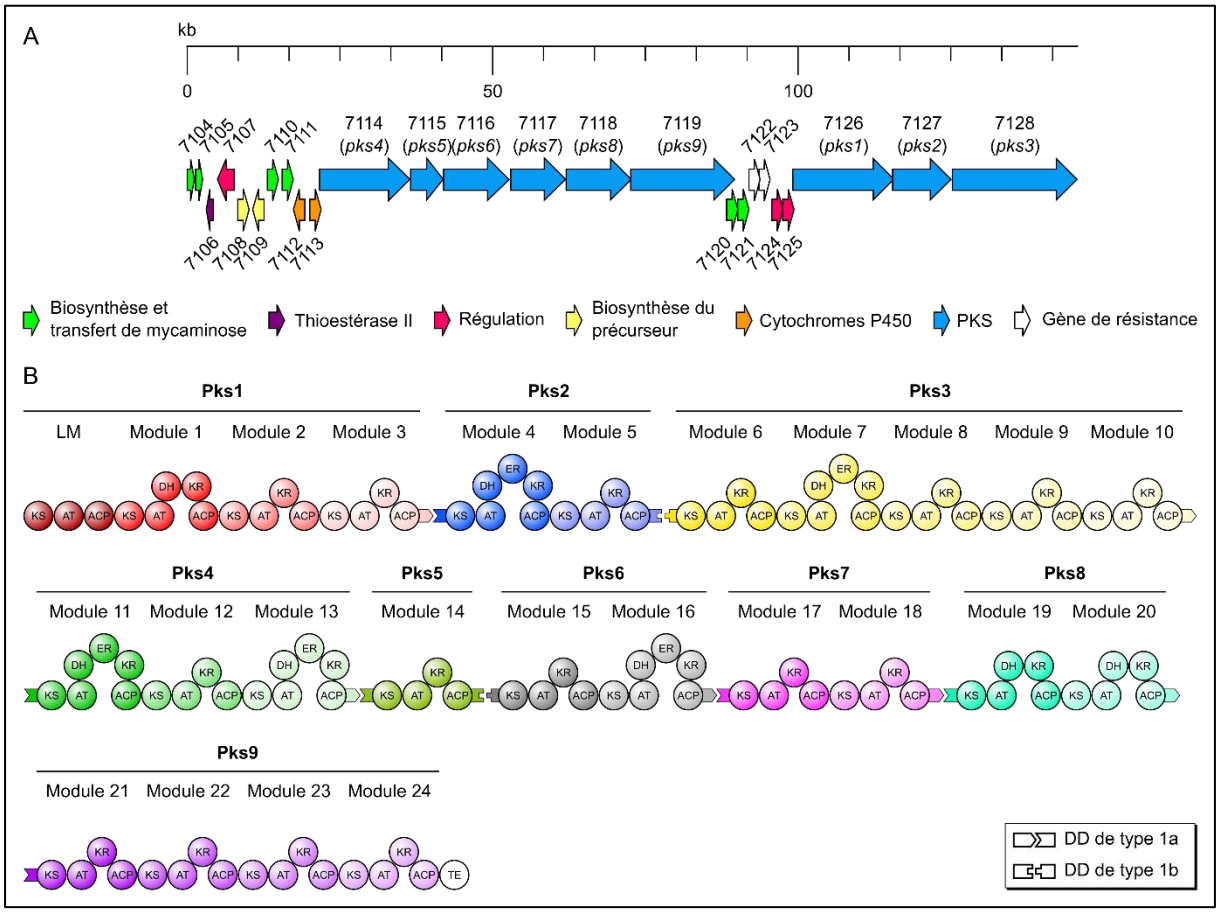


Figure 48. Représentation schématique de la PKS des stambomycines (adapté de Laureti et al., 2011). **(A)** Cluster de gènes codant la stambomycines synthase chez *Streptomyces ambofaciens*. **(B)** Organisation en domaines des neuf sous-unités de la PKS des stambomycines. **Légende** – DD : domaines de docking, LM : module de chargement, KS : cétosynthase, AT : acyltransférase, ACP : domaine porteur d'acyle, DH : déshydratase, KR : cétoréductase, ER : énoyl-réductase, TE : thioestérase.

Par ingénierie génétique, Dr Li Su a durant sa thèse pu introduire une nouvelle interface d'interactions entre les sous-unités 4 (Pks4) et 9 (Pks9) du système, tout en délétant les gènes codant les sous-unités Pks5 à Pks8. Cela a permis de produire des molécules composées d'un cycle de 37 carbones et appelées mini-stambomycines (**22** et **23**) (Figure 49) (Su et al., 2022). Deux familles de mini-stambomycines ont été caractérisées, selon l'état d'hydroxylation du carbone en position 14 du cycle. Celle-ci est médiée par une enzyme post-PKS, un cytochrome P450 codé par le gène *sam23877-7113*. L'interface entre les sous-unités Pks4 et Pks9 a été générée par ingénierie génétique en utilisant le système d'édition CRSPR-Cas9. Les rendements de production du système natif a été déterminé par spectrométrie de masse comme étant de 30 ± 2 $mg \cdot L^{-1}$.

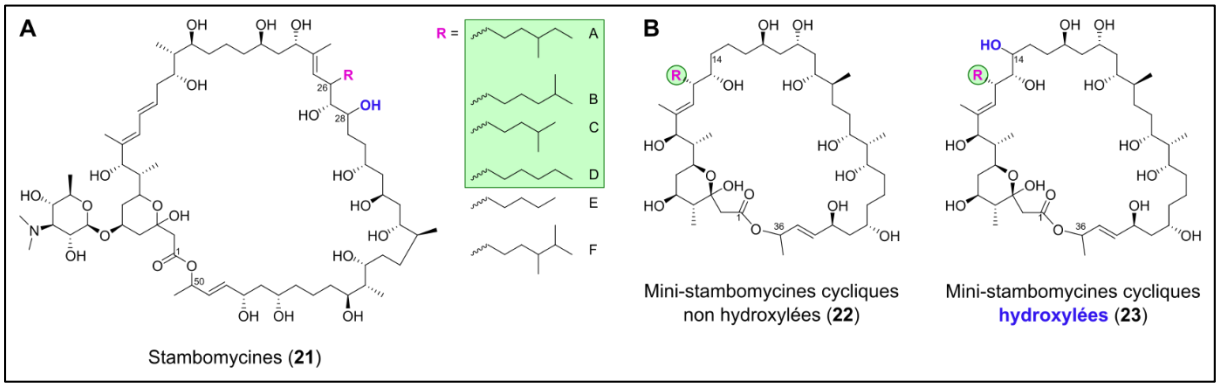


Figure 49. Structure des stambomycines et des mini-stambomycines obtenues par ingénierie génétique (adapté de Su et al., 2022). **(A)** Structure des stambomycines, produites par une PKS présente chez *Streptomyces ambofaciens* (Laureti et al., 2011). **(B)** Structure des mini-stambomycines obtenues par ingénierie de la stambomycines synthase (Su et al., 2022). Seules les mini-stambomycines A à D sont retrouvées (indiquées en vert). La fonction hydroxyle indiquée en bleu est ajoutée par une enzyme post-PKS codée par le gène *sam23877-7113*.

Afin de créer *in vivo* une nouvelle interface entre les sous-unités Pks4 et Pks9, les DD présents entre les sous-unités Pks4 / Pks5 et Pks8 / Pks9 ont été utilisés et seront après nommés respectivement $^{C}DD_4$ / $^{N}DD_5$ et $^{C}DD_8$ / $^{N}DD_9$. Dans un premier temps, Dr Li Su a gardé intact le $^{N}DD_9$ et modifié le $^{C}DD_4$ (Su et al., 2022). Les résidus de l'hélice α responsables de l'interaction avec le partenaire ont été modifiés afin de correspondre à ceux présents dans le $^{C}DD_8$ et ainsi reconstituer l'interaction initiale. Cependant, ces modifications n'ont pas permis d'obtenir les composés attendus.

Il est à noter que les interactions entre les sous-unités sont médiées par des DD (Broadhurst et al., 2003), mais également par des interactions entre le domaine ACP en amont et KS en aval de ces DD. Les six premiers résidus du domaine ACP_{13} (Pks4) ont donc été mutés afin de correspondre au début de l'ACP_{20} (Pks8) et ainsi renforcer la reconnaissance entre les sous-unités Pks4 et Pks9 (Figure 50A) (Su et al., 2022). Ainsi, les mini-stambomycines **22**A et **22**B ont pu être obtenues (Figure 49).

Cependant, comme cela est souvent observé en ingénierie des PKS, les rendements de production des souches modifiées sont plus faibles que ceux de la souche sauvage (Figure 50A) (Su et al., 2022).

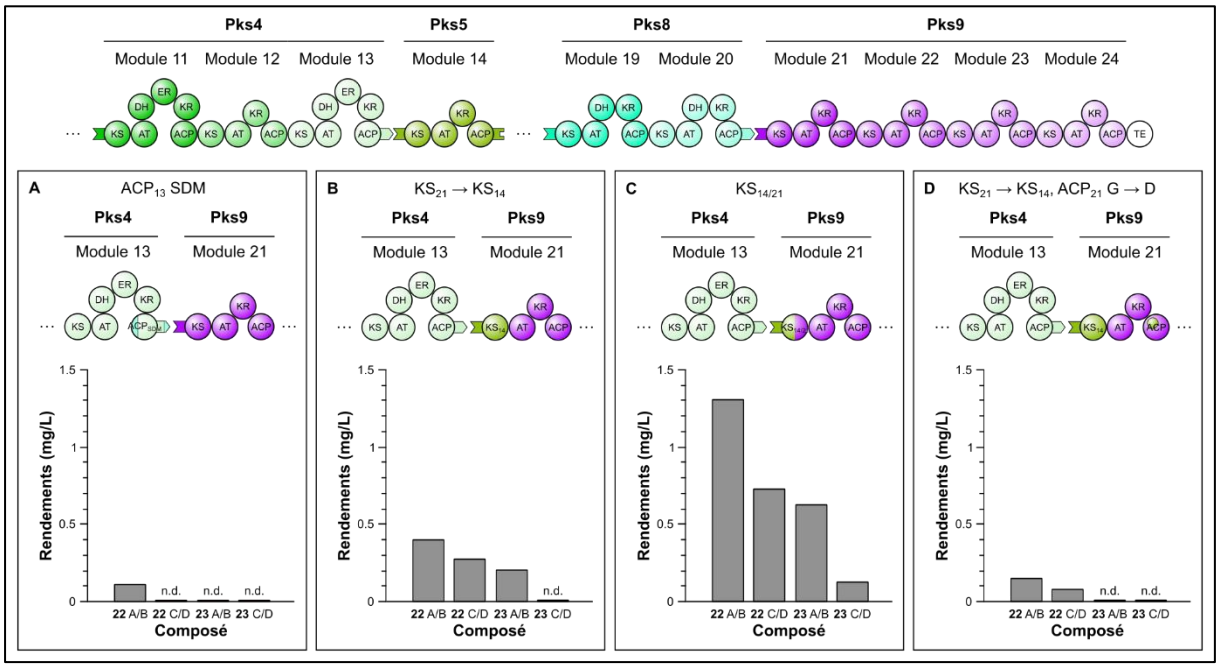


Figure 50. Représentation des interfaces générées entre les sous-unités 4 et 9 de la PKS des stambomycines et rendements de production des différentes mini-stambomycines (adapté de Su et al., 2022). Les sous-unités de la PKS originale sont représentées en haut de la figure. **(A)** Le domaine porteur d'acyle du module 13 (ACP_{13}) est muté sur les six premiers résidus pour correspondre au début du domaine ACP_{20} et le domaine de docking en C-terminal de Pks4 ($^{C}DD_4$) comporte l'hélice α d'interaction avec $^{N}DD_9$, qui provient de $^{C}DD_8$. **(B)** Le module 21 a été modifié pour contenir le domaine cétosynthase du module 14 (KS_{14}) ainsi que $^{N}DD_5$ qui interagit avec $^{C}DD_4$. **(C)** Le ^{N}DD provient de la sous-unité Pks5 et le domaine KS qui suit est un hybride entre ceux des modules 14 et 21. **(D)** Le ^{N}DD et le domaine KS utilisé dans le module 21 proviennent du module 14 et le domaine ACP_{21} a été muté pour contenir le motif d'interaction avec KS_{14}. **Légende** – Pks : sous-unité de la PKS des stambomycines, AT : acyltransférase, DH : déshydratase, ER : énoyl-réductase, KR : cétoréductase, n.d. : non déterminé.Une autre stratégie a été investiguée, consistant à l'utilisation du couple de DD intacts $^{C}DD_4$ / $^{N}DD_5$ au lieu du couple $^{C}DD_{4/8}$ / $^{N}DD_9$ préalablement étudié. Dans le même temps, le domaine KS_{21} (Pks9) a été remplacé par le domaine KS_{14} (Pks5, $KS_{21} \rightarrow KS_{14}$) afin d'assurer une continuité des interactions avec le module 13 (Figure 50B). Cela a permis alors d'augmenter les rendements de production des métabolites **22**A et **22**B en passant de 0,10 à 0,39 $mg \cdot L^{-1}$, mais également de voir l'apparition de **22**C / **22**D (0,27 $mg \cdot L^{-1}$) et **23**A / **23**B (0,20 $mg \cdot L^{-1}$), soit dans ce dernier cas des composés hydroxylés (Figure 50B).

Cependant, une étude publiée par Gregory *et coll.* en 2017 menée sur RAPS a mis en évidence l'importance des domaines KS dans l'interaction entre les modules (Wlodek et al., 2017). Par évolution accélérée, cette équipe a obtenu des PKS modifiées de manière aléatoire et elle a ainsi pu montrer que la majorité des nouvelles combinaisons de modules ont leur jonction au sein des domaines KS.

Cette étude sera présentée plus en détails par la suite, et a donné lieu à une nouvelle stratégie dans la miniaturisation de la stambomycine synthase, qui a consisté à créer un domaine hybride $KS_{14/21}$ au sein du module 21, tout en utilisant les domaines $^{C}DD_{Pks5}$ et $^{N}DD_{Pks8}$ (Figure 50C). Ainsi, les rendements de production des composés déjà obtenus ont pu être augmentés et les produits **23**C / **23**D (0,12 mg·L^{-1}) ont pu être observés (Figure 50C). Cela semble mettre en évidence que les interactions entre les domaines KS et les modules en amont sont nécessaires pour le bon fonctionnement des PKS.

Enfin, le domaine ACP_{21} a également été modifié afin de contenir le motif d'interaction de l'ACP_{14} avec le domaine KS_{14}, en mutant une glycine en acide aspartique (KS_{21} → KS_{14}, ACP_{21} G → D) (Figure 50D). Ainsi, l'interaction entre les domaines KS et ACP du module 21 modifié a été reconstituée. Cependant dans ce cas, seuls les mini-stambomycines non hydroxylées ont été obtenues, et les rendements de production étaient semblables à ceux obtenus avec la construction ACP_{13} SDM (Figure 50D).

Ces différentes expériences d'ingénierie semblent montrer que les interactions entre le domaine KS d'un module *n* et les domaines ACP des modules *n-1* et *n* sont les plus critiques dans le transfert de la chaîne acylée. En effet, malgré un rendement largement inférieur à celui des stambomycines qui est de 30 ± 2 mg·L^{-1}, la construction d'un domaine KS hybride entre celui du module 14 et 21 semble permettre un transfert optimisé de la chaîne entre les modules 13 et 21, comparé aux autres tentatives d'ingénierie.

VI.B.2. La synthèse de dérivés de la rapamycine par évolution accélérée

D'autres études d'ingénierie des PKS ont été effectuées, comme celles concernant la rapamycine synthase (RAPS). La rapamycine (**24**) est un composé ayant une activité inhibitrice de la formation du complexe de la sérine / thréonine kinase chez les mammifères (mTOR). RAPS est composée de quatorze modules d'élongation, répartis en trois sous-unités (Figure 51).

Par évolution accélérée, l'équipe de Gregory *et coll.* a pu obtenir des systèmes modifiés de RAPS, produisant des dérivés de la rapamycine (Wlodek et al., 2017). Ces systèmes ont été obtenus par sérendipité, puisque l'objectif initial était de remplacer les domaines DH, ER et KR inactifs du module 3 par les domaines actifs des modules 11 et 13. Les expérimentations menées ont permis d'obtenir des souches produisant des dérivés de la rapamycine. Sept analogues nommés rapalogues ont été caractérisés, et les génomes des souches les produisant séquencés.

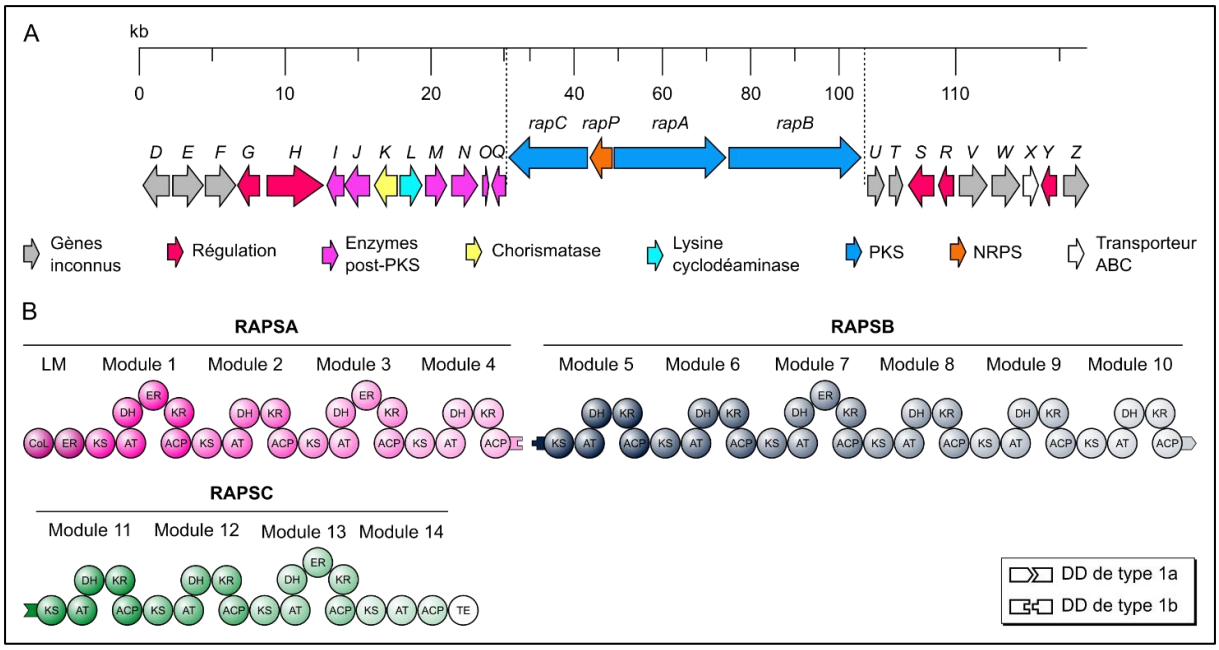


Figure 51. Représentation schématique de la rapamycine synthase (adapté de Aparicio et al., 1996). **(A)** Cluster de gènes codant la rapamycine synthase (RAPS) chez *Streptomyces hygroscopicus*. **(B)** Organisation en domaines des trois sous-unités de RAPS. **Légende** – LM : module de chargement, KS : cétosynthase, AT : acyltransférase, ACP : domaine porteur d'acyle, DH : déshydratase, KR : cétoréductase, ER : énoyl-réductase, TE : thioestérase, DD : domaines de docking.

Parmi ces rapalogues, six sont produits par des PKS dépourvues de certains modules, et un par un système comprenant un module supplémentaire. Les systèmes délétés du plus grand nombre de modules produisent un composé linéaire. En moyenne, la production de rapalogue est diminuée d'un facteur 2 par rapport à la souche sauvage (Wlodek et al., 2017). Cinq de ces rapalogues ont été analysés (Figure 52A). Leur production par les souches modifiées ainsi que leur activité biologique ont été mesurées (Figure 52B). Cette activité a été évaluée par la constante d'inhibition de chaque composé face au récepteur au tacrolimus (FKBP12), une cible connue de la rapamycine. Il a été observé que plus le cycle est petit, et donc éloigné en structure de la 39-déméthoxy-rapamycine produite par le complexe catalytique sauvage, moins son activité biologique inhibitrice de FKBP12 était importante (Wlodek et al., 2017).

Les composés **28** et **29** ont ainsi une activité biologique qui se rapproche de celle de la rapamycine (**24)** (Figure 52B) (Wlodek et al., 2017). Cependant, les rendements de production des souches concernées sont réduits par rapport à la souche ISOM-4010 qui produit le composé **24** (193 µM), notamment ceux de la souche ISOM-4309 qui produit le rapalogue **29** (9 µM), et de manière moins importante pour la molécule **28** (121 µM) (Figure 52B) (Wlodek et al., 2017). Ces résultats soulignent que les rendements de production des systèmes PKS modifiés sont un frein à la caractérisation et l'exploitation de nouvelles biomolécules d'intérêt.

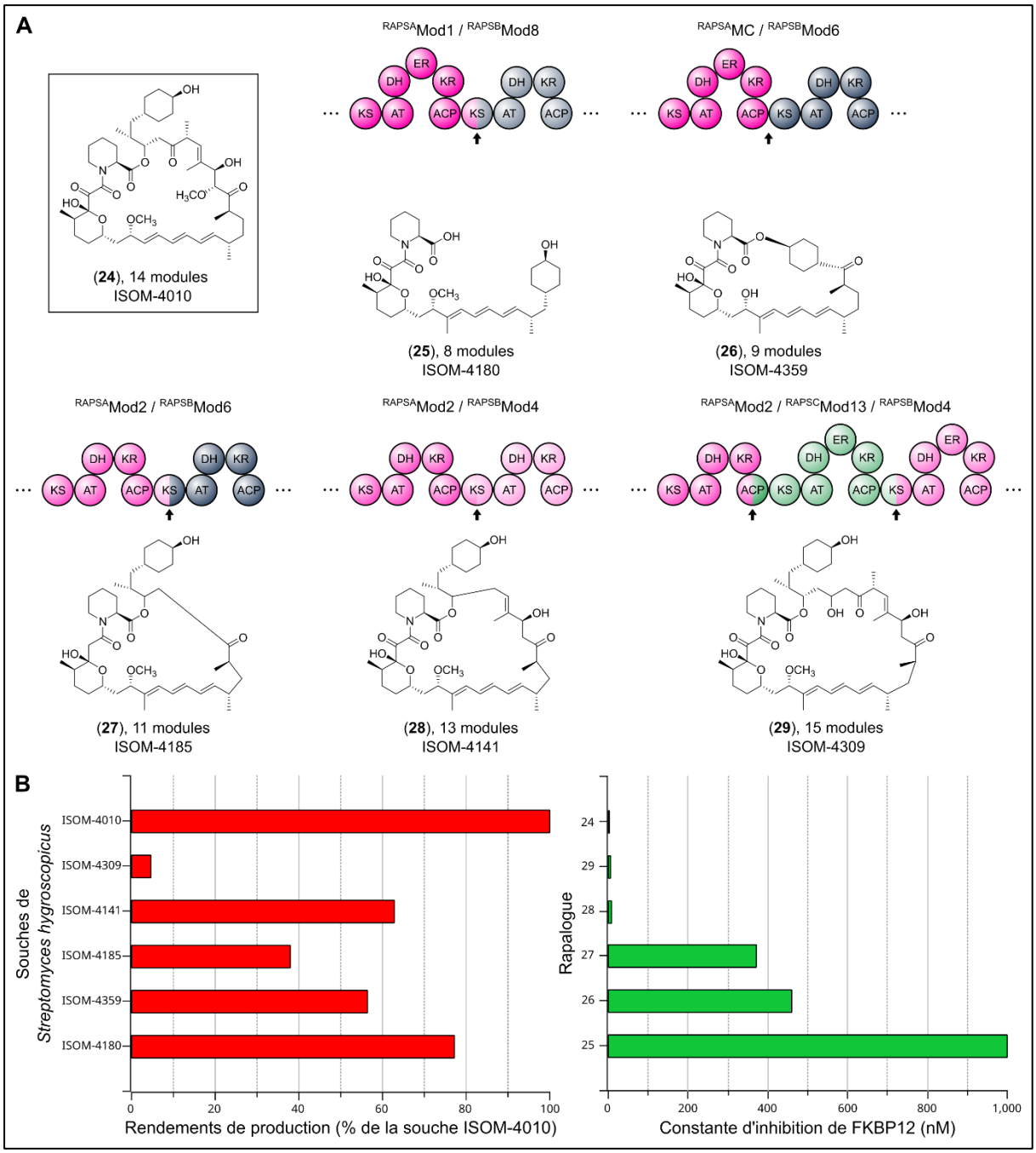


Figure 52. Analyse des souches de *Streptomyces hygroscopicus* mutées pour le cluster de gènes codant la rapamycine synthase (adapté de Wlodek et al., 2017). **(A)** Formules semi-développées des cinq rapalogues étudiés par la suite (**25**, **26**, **27**, **28**, **29**) et de la région modifiée de la rapamycine synthase (RAPS). Le nombre de module qui constitue la PKS est indiquée en-dessous de chaque composé, ainsi que le nom de la souche productrice. Les points de fusion entre différents modules de la PKS sont représentés par des flèches noires. **(B)** Production de rapamycine par la souche au génotype parental (ISOM-4010) et 5 souches mutées (*gauche*) et activité biologique des rapalogues associés mesurée par leur inhibition du récepteur au tacrolimus (FKBP12) (*droite*). Chaque rapalogue est placé à côté des rendements de production de la souche qui le synthétise.

La majorité des systèmes modifiés obtenus par mutations aléatoires et dont l'activité biologique du rapalogue a été étudiée montre une fusion de modules au milieu d'un domaine KS. Par exemple la souche ISOM-4180 produisant le rapalogue **25** exprime une rapamycine synthase délétée des modules 2 à 7, et le domaine KS du module 8 est un hybride entre ceux des modules 2 et 8. La seule souche ne montrant pas de KS hybride lors de la délétion de modules est la souche ISOM-4359 produisant le rapalogue **26**.

VI.C. L'ingénierie des polycétide synthases pour l'optimisation des rendements de production

Une étude menée sur 500 sous-unités de PKS répertoriées au sein de la base de données *Minimal Information about a Biosynthetic Gene cluster* (MIBiG) a montré que celles-ci sont majoritairement composées d'un nombre de modules restreint, allant d'un à cinq par sous-unité (Figure 53). Posséder un nombre restreint de modules semble donc être un avantage évolutif pour les sous-unités de PKS. De plus, dans le contexte des tentatives d'ingénierie des PKS, la taille de ces sous-unités peut représenter un frein à leur expression en système hétérologue. En effet, exprimer au sein d'un hôte non natif des systèmes de cette ampleur peut impacter le cycle de vie de l'hôte sélectionné et sa capacité à produire le composé d'intérêt.

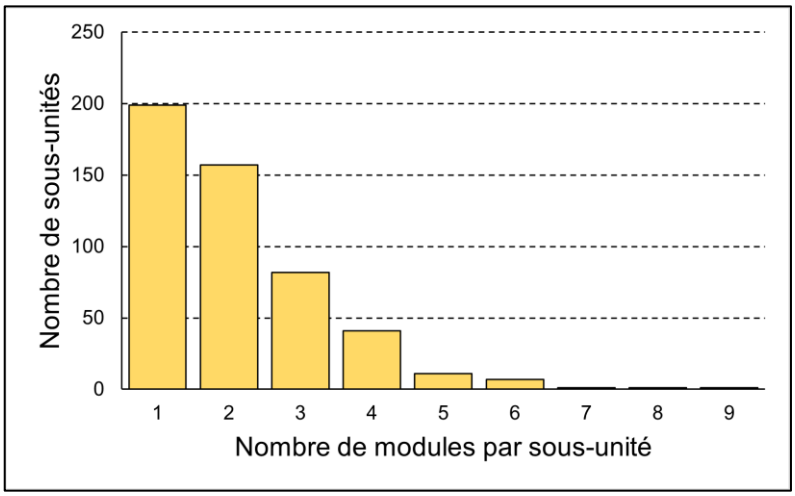


Figure 53. Nombre de modules au sein de sous-unités de polycétide synthases répertoriées au sein de la base de données MIBiG (données non publiées). Cette étude a été menée sur 102 systèmes, regroupant 1044 modules répartis au sein de 500 sous-unités, référencées au sein de la base de données *Minimal Information about a Biosynthetic Gene cluster* (MIBiG). Parmi ces systèmes, 97 sont des PKS *cis*-acyltransférase (AT) et seulement 5 sont des PKS *trans*-AT. La composition moyenne des sous-unités est de 2,6 modules. Parmi les 500 sous-unités étudiées, 98 % (n=490) sont composées d'entre un et cinq modules, tandis que les 2 % (n=10) restants sont des sous-unités de six à neuf modules.

VI.C.1. L'utilisation de domaines de docking : étude portant sur la pikromycine synthase

La pikromycine synthase a été utilisée comme modèle d'étude pour l'ingénierie de domaines KR (Figure 54A). Dans ce cadre, l'équipe de Keatinge-Clay *et coll*. a développé des PKS de tailles réduites, ne contenant que trois modules (Figure 54B) (Miyazawa et al., 2021). Ainsi, le module de chargement de la pikromycine, suivi des modules d'élongation 1 et 6 et du domaine TE ont permis d'obtenir une PKS miniature produisant un tricétide lactone (**30**) et ci-après nommée Pik016. Le module 1 a été sélectionné car il porte un domaine KR, tandis que le module 6 a été choisi car il est suivi du domaine TE. Ce système a été construite de deux manière différentes, selon les deux définitions pouvant être attribuées à un module de PKS. Alors, une première version de Pik016 contient le domaine KS_2 en amont du module 6 (Pik016a) selon la définition adaptée des modules, et une seconde contient à cette place le domaine KS_6 (Pik016b) construite selon la définition classique (Figure 54B).

Ces systèmes ont été produits par *Escherichia coli* K207-3, qui contient le gène codant la PPTase Sfp provenant de *B. subtilis*, et des étiquettes histidines ont été placées aux extrémités C-terminale et N-terminale afin de faciliter purification. La synthèse du tricétide lactone attendue a ensuite été effectué *in vitro* et la productivité des différents systèmes modifiés a été analysée par la quantité de tricétide obtenu, mesurée par spectrométrie de masse.

Ainsi, les rendements de production de la PKS miniature Pik016a construite à partir de la définition actualisée des modules ont été plus importants d'un facteur d'environ 10 que ceux obtenus à partir de la définition classique (Pik016b). Afin d'optimiser les rendements de production de Pik016, une piste d'amélioration a été proposée : scinder le système obtenu par l'intégration de domaines de docking (DD). Pour cela, la paire de DD retrouvée nativement entre les sous-unités PikAIII et PikAIV a été sélectionnée. Cette paire a donc été intégrée entre les domaines ACP_1 et KS_2 au sein de Pik016a et entre les domaines ACP_1 et KS_6 au sein de Pik016b. Cela a permis de multiplier les rendements par un facteur 22 dans le cas de Pik016a et d'un facteur d'environ 70 dans le cas de Pik016b (Figure 54B). Cependant, les rendements de production de la version scindée de Pik016a restent supérieurs à ceux de Pik016b scindée d'un facteur d'environ 3 (Miyazawa et al., 2021).

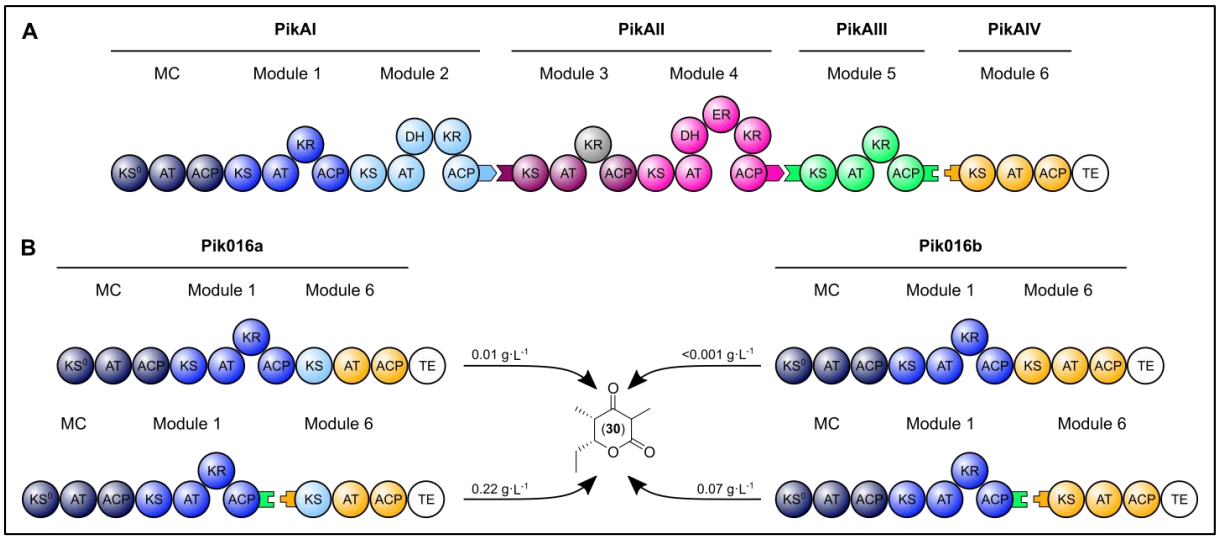


Figure 54. Ingénierie du domaine cétoréductase de la pikromycine synthase. **(A)** Composition en modules et sous-unités de la pikromycine synthase (adapté de Xue et al., 1998). **(B)** Systèmes de la pikromycine synthase modifiés pour l'étude du domaine cétoréductase (adapté de Miyazawa et al., 2021). Deux systèmes miniatures de la pikromycine synthase ont été construits, composés du module de chargement et des modules d'élongation 1 et 6, ainsi que du domaine thioestérase (TE). Le premier a été construit selon la définition des unités échangeables (Pik016a), et le second selon la définition initiale des modules (Pik016b). Ces systèmes permettent de produire un tricétide lactone (**30**). L'utilisation des unités échangeables a permis d'obtenir des rendements de production plus importants que la définition initiale des modules. De plus, la scission du système par l'utilisation des domaines de docking présents entre les modules 5 et 6 de la pikromycine synthase a également permis d'augmenter les rendements de production des tricétides lactones. **Légende** – MC : module de chargement, KS^0 : domaine cétosynthase inactif, KS : cétosynthase AT : acyltransférase, ACP : domaine porteur d'acyle, KR : cétoréductase.

VI.C.2. L'utilisation de domaines d'interactions (semi-)synthétiques

Au-delà des domaines DD présents nativement au sein des PKS et pouvant être utilisés dans le cadre de tentatives d'ingénieries, il existe également des domaines d'interaction synthétiques et semi-synthétiques pouvant être utilisé à des fins d'ingénierie.

VI.C.2.a. Les domaines synthétiques *Synthetic Zippers*

Les *Synthetic Zippers* (SynZips) sont un exemple de domaines d'interaction synthétiques (Figure 55) (Thompson et al., 2012). Ces domaines ont été conçues dans le but de fournir une boîte à outil pour le développement de la biologie synthétique. Vingt-trois domaines SynZip ont été ainsi obtenus et leurs interactions caractérisées. Ils sont structurés en superhélice et sont constitués de répétitions de sept résidus $(abcdefg)_n$. Des résidus hydrophobes y occupent les positions *a* et *d*, tandis que les positions *e* et *g* sont occupées par des résidus qui permettent des interactions électrostatiques. Les SynZips ont une constante de dissociation plus faible que les DD, de l'ordre du nM. Cela en fait des outils d'ingénierie des PKS intéressants, en ce sens où leur interaction se trouve entre celle plus flexible des DD et celle plus rigide des liaisons covalentes.

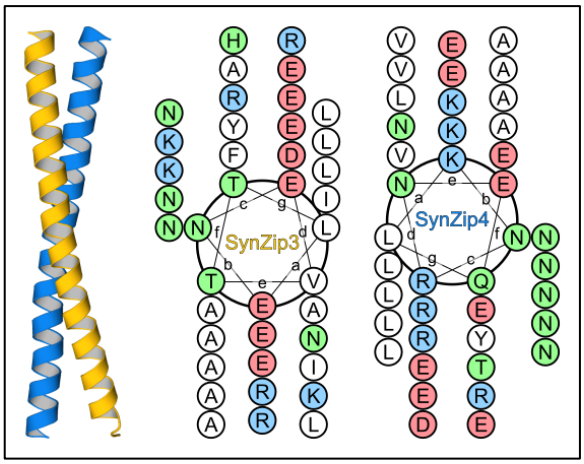


Figure 55. Structure et diagramme en hélice de la paire de *Synthetic Zippers* 3:4 (identifiant PDB : 3HE5) (adapté de Reinke et al., 2010). Les deux partenaires sont chacun composés d'une unique hélice α et sont composés de répétitions de sept résidus (*abcdefg*)*n*. Les résidus *a* et *d* médient des interactions hydrophobes et *e* et *g* permettent des interactions électrostatiques. Les résidus chargés positivement sont représentés en rouge et ceux chargés négativement en bleu. Les résidus en vert sont polaires et ceux en blanc sont hydrophobes.

Ces vingt-trois domaines sont capables de former 25 paires qui interagissent pour former des dimères. Parmi elles, 24 forment des interactions parallèles et une seule paire permet de former une interaction antiparallèle (Thompson et al., 2012).

VI.C.2.b. Les domaines semi-synthétiques SpyCatcher:SpyTag

Il existe également des domaines semi-synthétiques qui interagissent de manière covalente, dérivés du domaine adhésine de collagène de type immunoglobuline 2 (CnaB2) présents dans la protéine de liaison à la fibronectine de *Streptococcus pyogenes* (Zakeri et al., 2012). Ce domaine forme de manière spontanée des liaisons isopeptidiques entre un résidu aspartate et une lysine (Hagan et al., 2010). Ce domaine a été scindé en deux protéines distinctes nommées SpyCatcher et SpyTag (Figure 56), afin de permettre par la suite leur fusion à des protéines d'intérêts et leur interaction *via* la formation d'une liaison entre la lysine présente dans la moitié SpyCatcher et l'aspartate présente chez le partenaire SpyTag.

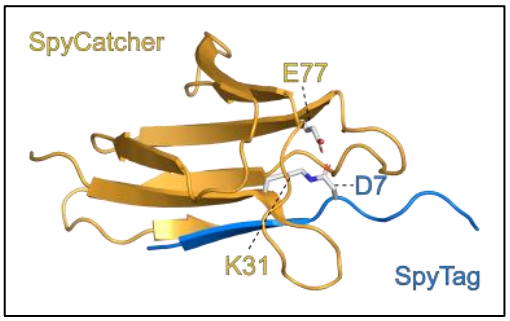


Figure 56. Structure du couple SpyCatcher:SpyTag (identifiant PDB : 4MLI) (adapté de Li et al., 2014). Les deux partenaires, organisés en brins β, interagissent par la formation d'une liaison covalente isopeptidique entre la lysine 31 de SpyCatcher et l'aspartate 7 de SpyTag. Cette liaison est stabilisée par la présence d'un acide glutamique.

VI.C.2.c. L'utilisation de ces domaines pour l'ingénierie de la vénémycine synthase

L'utilisation de ces différents domaines d'interaction a pu être comparée dans l'interaction entre les deux sous-unités de la vénémycine synthase (VEMS) (Buyachuihan et al., 2023), qui est une PKS modulaire retrouvée chez *S. venezuelae* (Thanapipatsiri et al., 2016). Cette PKS est composée de trois modules, répartis au sein des sous-unités VemG et VemH. La première sous-unité VemG contient le module de chargement et un module d'élongation, tandis que VemH contient le second module d'élongation (Figure 57, 1). Les DD retrouvés entre ces deux sous-unités sont de type 1a, et sont donc formés de deux hélices α de dimérisation et d'une hélice α d'interaction avec le partenaire d'une part (^{C}DD) et d'une unique hélice α de l'autre (^{N}DD), comme précédemment décrit (voir paragraphe IV.A.1 page 44).

L'objectif ici a été d'évaluer l'efficacité de DD hétérologues, de SynZips (Thompson et al., 2012), ainsi que du complexe SpyCatcher:SpyTag (Zakeri et al., 2012), dans le maintien de l'interaction entre VemG et VemH (Buyachuihan et al., 2023). Les DD natifs de VEMS ont soit été délétés, soit remplacés par les DD présents entre les sous-unités 1 et 2 de DEBS, par la paire de SynZips 3:4, ou par le complexe SpyCatcher:SpyTag (Figure 57).

La délétion des DD natifs entraîne une perte d'activité totale de la PKS. Cela peut s'expliquer par le fait que la délétion du domaine ^{C}DD aboli l'interaction avec la sous-unité suivante et la dimérisation de la sous-unité VemG (Figure 57, 2) (Buyachuihan et al., 2023). Ensuite, l'utilisation des domaines $^{C}DD_{DEBS1}$ et $^{N}DD_{DEBS2}$ a été étudiée pour rétablir l'activité de la PKS. Deux stratégies ont ici été utilisées concernant le remplacement du domaine ^{C}DD, dans ces deux cas l'hélice α formant le ^{N}DD ayant été entièrement remplacée. Dans le premier cas, les trois hélices α du $^{C}DD_{VemG}$ ont été remplacées par celles du domaine $^{C}DD_{DEBS1}$ (3, Figure 57). Cela a permis de rétablir l'activité de la PKS, malgré une efficacité réduite d'un facteur d'environ 2 comparativement à la PKS native. Cette construction a également pu rétablir la dimérisation de la sous-unité VemG.

Dans un second temps, seule l'hélice α permettant l'interaction avec le ^{N}DD a été remplacée, et cela a permis d'obtenir une PKS modifiée légèrement plus efficace que la PKS native et une sous-unité VemG capable de dimérisation (Figure 57, 4). Cette dimérisation joue donc un rôle dans l'efficacité de synthèse des polycétides mais n'en est pas le seul critère. Les éléments d'interaction SynZips 3:4 ont également permis de rétablir la productivité de la PKS, avec cependant une diminution d'un facteur 2 environ à l'instar des domaines $^{C}DD_{DEBS1}$ et $^{N}DD_{DEBS2}$ (Figure 57, 5). Cela peut être expliqué par le fait que les SynZips aient une affinité plus forte que les DD, ce qui réduit la flexibilité des sous-unités.

Enfin, le système SpyCatcher:SpyTag a également été utilisé afin de remplacer l'interface d'interaction native entre VemG et VemH. Dans ce cas, la fusion du domaine SpyCatcher en position C-terminale du second domaine ACP de VemG a été effectuée à la suite des deux hélices α impliquées dans la dimérisation des sous-untiés (Figure 57, 6). Cependant, l'utilisation de ces domaines a permis d'obtenir un système modifié ayant une activité correspondant à 39 % de celle de la PKS native.

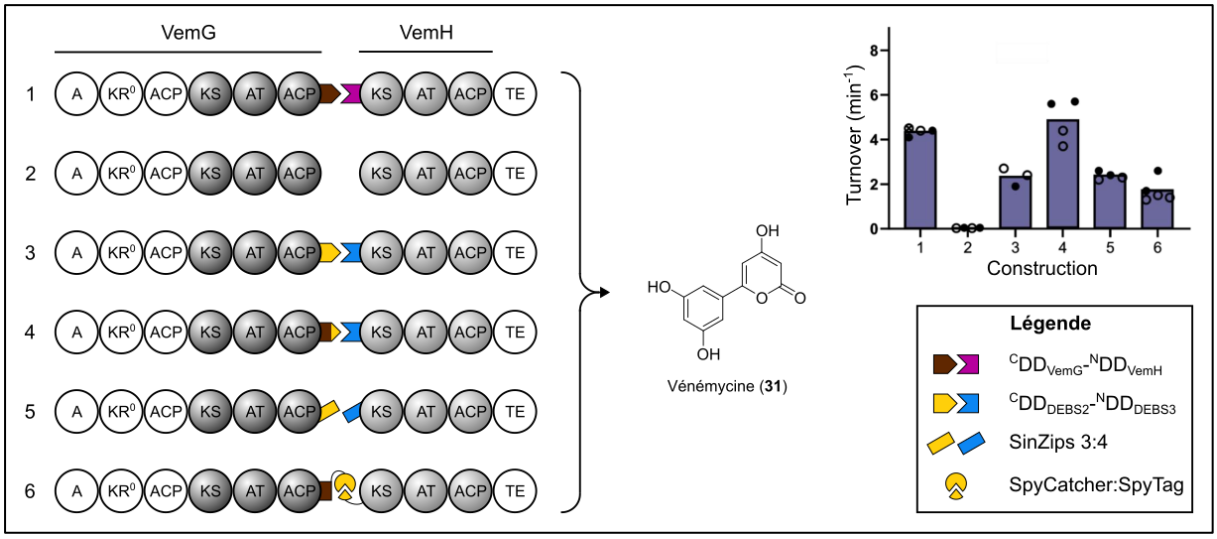


Figure 57. Représentation schématique des interfaces générées entre VemG et VemH (adapté de Buyachuihan et al., 2023). L'interface d'interaction native retrouvée entre les sous-unités VemG et VemH est constituée de domaines de docking (DD) de type 1. Ceux-ci ont soit été délétés, soit remplacés par des DD provenant d'un autre système – la 6-désoxyérythronolide B synthase (DEBS) –, une paire de domaines synthétiques appelés *Synthetic Zippers* (SynZip) ou par le système SpyCatcher:SpyTag. L'activité de chaque système a été mesurée par le turnover de chacune des PKS modifiées. La délétion des DD natifs a entraîné l'abolition de toute activité du système (2). Le remplacement des DD natifs par ceux présents entre les sous-unités 2 et 3 de DEBS a permis de recouvrir environ 50 % de l'activité de la PKS native (3), et l'activité native du système a pu être rétablie par la formation d'un domaine ^{C}DD hybride composé des hélices α de dimérisation native et de l'hélice α d'interaction avec le partenaire provenant de DEBS (4). L'utilisation de la paire de SynZip 3:4 quant à elle a également permis d'obtenir un système dont l'activité est environ 2 fois inférieur à celle de la vénémycine synthase (5), et l'utilisation du système SpyCatcher:SpyTag, formant une liaison covalente entre les sous-unités, n'a permis de retrouver que 39 % de l'activité du système natif (6). **Légende** – A : domaine d'adénylation, KR^0 : cétoréductase inactive, ACP : domaine porteur d'acyle, KS : cétosynthase, AT : acyltransférase, TE : thioestérase.

Enfin, l'équipe s'est intéressée à l'effet de la scission de VemG sur la production de chaque sous-unité en expression hétérologue chez *E. colir* et l'impact de l'utilisation de plus petites sous-unités sur l'efficacité de production de la vénémycine (Buyachuihan et al., 2023). Dans le cas de la scission entre les deux modules de VemG au sein du système contenant des DD entre VemG et VemH, cela n'influence pas l'activité du système, mais permet une expression plus importante de ces sous-unités en hôte hétérologue (Figure 58, 7 et 8). En effet, lorsqu'elle n'est pas modifiée, la sous-unité VemG est exprimée de manière hétérologue chez *E. coli* a raison de 2,5 g·L^{-1}.

Lorsqu'elle est scindée, le rendement de production du premier module est de 5,4 $g \cdot L^{-1}$ et de 9,4 $g \cdot L^{-1}$ pour le second, soit une augmentation du rendement de production d'un facteur d'environ 2 et 4 respectivement. L'expression de plus petites sous-unités chez un hôte hétérologue facilite donc leur production. La scission de VemG dans le cas où les sous-unités VemG et VemH interagissent par le biais de SynZips permet d'améliorer le turnover du système (Figure 58, 9 et 10). Cela peut être expliqué par le fait que l'introduction à nouveau de domaines de docking permette de restaurer une dimérisation des sous-unités, ainsi que d'améliorer la flexibilité du système.

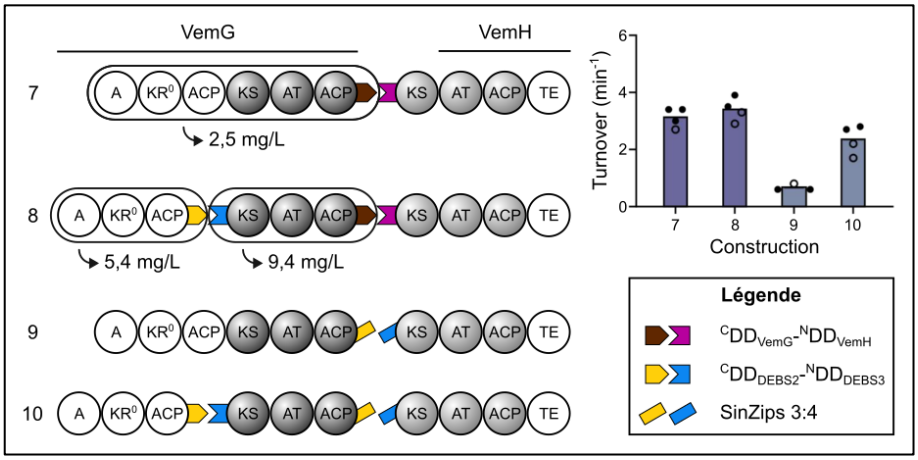


Figure 58. Scission de la sous-unité VemG de la vénémycine synthase (adapté de Buyachuihan et al., 2023). La sous-unité VemG a été scindée par l'utilisation des domaines de docking (DD) présents entre les sous-unités 2 et 3 de la 6-désoxyérythronolide B synthase (DEBS). Cette scission a été faite dans le système natif est dans celui contenant les domaines d'interaction SynZips 3:4 à la place des DD natifs présents entre VemG et VemH, précédemment décrits (5, Figure 57). L'impact sur la production de la sous-unité VemG entière et scindée chez *Escherichia coli* a été mesuré. La sous-unité complète VemG a pu être purifiée à une concentration de 2,5 $mg \cdot L^{-1}$, tandis que les modules scindés exprimés séparément chez *E. coli* ont permis d'obtenir après purification des concentrations respectives de 5,4 $mg \cdot L^{-1}$ et 9,4 $mg \cdot L^{-1}$. La productivité des systèmes modifiés a également été mesurée en calculant le turnover *in vitro* pour la production de vénémycine. **Légende** – A : domaine d'adénylation, KR^0 : cétoréductase inactive, ACP : domaine porteur d'acyle, KS : cétosynthase, AT : acyltransférase, TE : thioestérase.

Produire ainsi de plus petites sous-unités chez un hôte hétérologue – ici *E. coli* – modifie notamment la taille des protéines produites, ce qui facilite sûrement le travail des cellules hétérologues. De plus, cela impact également la taille de l'ARN messager (ARNm) transcrit et diminue les probabilités d'obtenir des ARNm tronqués au sein des cellules. Cela a en effet été montré comme étant un frein à la productivité des PKS et de leur dérivés modifiés par ingénierie (Liu et al., 2025).

VI.C.3. La modification de la taille des ARN messagers : étude de la butényl-spinosyne synthase

L'un des freins de l'efficacité de la production de dérivés polycétidiques par les PKS modifiées et la production de sous-unités PKS elles-mêmes. L'un des facteurs influençant cette efficacité est la présence au sein des cellules d'ARNm tronqués, sur lesquelles les ribosomes vont également se fixer et commencer la synthèse protéique, mais seront bloqués à la localisation de la troncature de l'ARNm. Un système de sauvetage des ribosomes est alors enclenché, permettant de récupérer ces enzymes et de pouvoir les utiliser pour la traduction des ARNm complets. La présence d'ARNm tronqués au sein des cellules ralentit donc la synthèse de protéines complètes et fonctionnelles. Ce souci est rencontré au sein de cellules bactériennes, dans lesquelles la transcription et la traduction sont couplées, ce qui ne permet pas de discriminer entre des ARNm qui seront complets ou tronqués en amont de la traduction.

En s'intéressant au cas de la PKS qui produit la butényle-spinosyne (Bus) chez, l'équipe de Wang *et coll.* a permis d'augmenter le rendement de production de ce métabolite en modifiant la taille des ARNm codant la première sous-unité (Figure 59) (Liu et al., 2025). Cette étude a été menée en exprimant le cluster lié à la production de la butényle-spinosyne au sein d'un hôte hétérologue : *S. albus* J1074.

Plus un ARNm est long, plus le risque d'erreur durant sa synthèse est grand, et plus la probabilité de retrouver des ARNm tronqués est grande. Alors, minimiser la taille de ces ARNm permettrait d'améliorer la qualité de transcription et *in fine* de traduction. C'est donc cette hypothèse que l'équipe de Wang *et coll.* a voulu tester sur la PKS Bus. Ce système est composé de douze modules, répartis au sein de cinq sous-unités (Hahn et al., 2006) et produit un polycétide qui a des propriétés insecticides (Hahn et al., 2002).

Cette étude s'est penchée sur la première sous-unité BusA du système, composée du module de chargement et de deux modules d'élongation (Figure 59A) (Hahn et al., 2006). Dans un premier temps, la protéine BusA a été scindée, sans modifier la taille de l'ARNm (Figure 59B) (Liu et al., 2025). Ici, un codon stop, un site de liaison au ribosome (RBS) et un codon start ont été introduits entre le module de chargement et le module 1 de BusA, et / ou entre les modules 1 et 2. Les domaines de docking choisis pour la scission de BusA proviennent de la salinomycine synthase (SlnA), car les sous-unités sont codées par des gènes présents au sein d'un opéron et non à son extrémité, prévenant ainsi le risque d'insérer une région promotrice ou terminatrice avec les DD (Y. Liu et al., 2025).

La scission entre le MC et les deux modules d'élongation a permis d'augmenter les rendements de production de 5,9 %, tandis que la scission entre les deux modules d'élongation a permis une augmentation de 5,3 %. Enfin, la double scission permettant d'obtenir trois modules séparés par des DD a augmenté la production de buténYl-spinosyne de 13,3 %. Cependant, aucune étude des interactions potentielles entre les DD natifs et introduits ne permet ici d'assurer l'assemblage correct de la PKS modifiée.

L'étude s'est ensuite penchée sur l'utilisation d'un promoteur constitutif fort kasOp* en amont de la séquence codant le domaine $^{N}DD_{SlnA2}$ au sein des trois constructions précédemment citées (Figure 59B) (Liu et al., 2025). L'introduction du promoteur entre les codons start de la séquence $^{C}DD_{SlnA1}$ et le codon stop de $^{N}DD_{SlnA2}$ permet d'obtenir deux types d'ARNm. Dans le cas, par exemple, de la scission permettant d'obtenir les sous-unités BusA-1-1 et BusA-1-2, un ARNm codant les deux sous-unités sera obtenu, ainsi qu'un ARNm ne codant que la sous-unité BusA-1-2. Cela a permis dans ce cas d'augmenter les rendements de production de 2,6 % par rapport à la PKS scindée en absence du promoteur kasOp* et donc de 13,8 % par rapport à la PKS sauvage. Dans le cas des scissions BusA-2 et BusA-3, les rendements de production ont ainsi été augmentés de 11,2 % et 18,6 % par rapport à la PKS sauvage en présence du promoteur en amont de $^{N}DD_{SlnA2}$.

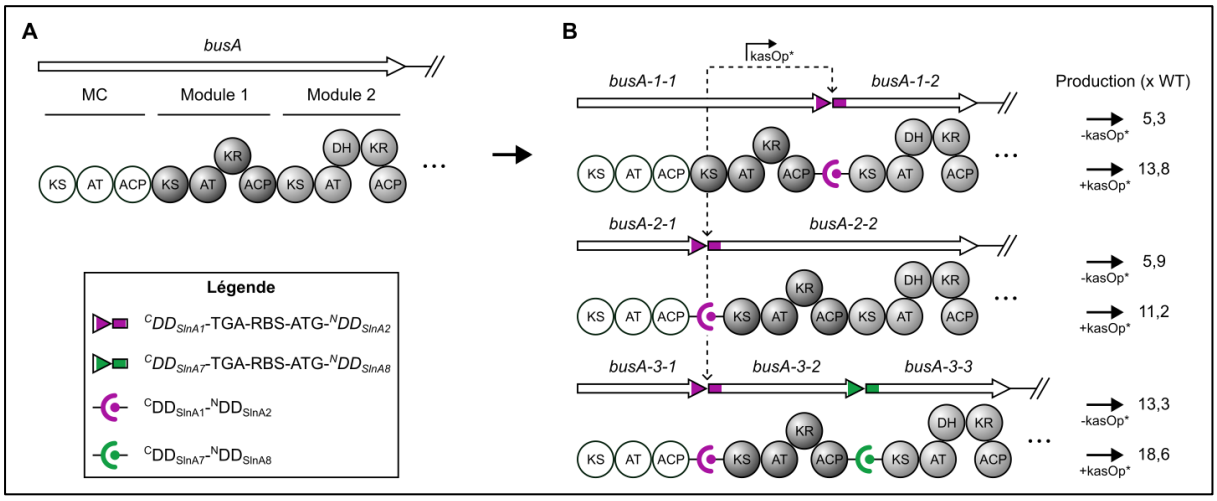


Figure 59. Scission de la buténYl-spinosyne synthase (Bus) par ingénierie moléculaire et génétique (adapté de Y. Liu et al., 2025). **(A)** Représentation schématique de la première sous-unité de Bus et de la stratégie de scission. **(B)** Représentation schématique des premières sous-unités scindées au sein des systèmes Bus modifiés. Les domaines de docking (DD) retrouvés entre les sous-unités 1 et 2 de la salinomycine synthase (SlnA) ont été utilisés pour les scissions uniques, et la paire de DD retrouvée entre les sous-unités 7 et 8 du même système Sln a été ajoutée pour la double scission. Les rendements de production des polycétide synthases (PKS) modifiées ont été mesurés en absence puis en présence du promoteur fort kas*O*p* en amont de la séquence codant le domaine $^{N}DD_{SlnA2}$. **Légende** – KS : cétosynthase, AT : acyltransférase, ACP : domaine porteur d'acyle, KR : cétoréductase, DH : déshydratase, RBS : site de liaison du ribosome, WT : souche sauvage.

Les études présentées dans ce chapitre permettent d'envisager la scission de systèmes PKS comme un levier pour l'optimisation de leur productivité. De plus, les tentatives d'ingénierie effectuées le sont très souvent en négligeant l'impact que cela pourrait avoir sur la macro-organisation des PKS et l'impact sur l'environnement moléculaire natif du système qui se trouve modifié, pouvant expliquer les rendements de production faibles observés dans la majorité des cas.

Objectifs

Les produits naturels sont une source précieuse de molécules de grand intérêt en médecine et en agro-alimentaire. Parmi eux, les polycétides sont des composés particulièrement intéressants dans la manière dont ils sont produits par les polycétide synthases. Les PKS modulaires ont en effet la particularité de présenter un mode de fonctionnement linéaire, en ce sens que la séquence des gènes et l'organisation en domaines des PKS permet de déterminer la structure du composé produit. Cette caractéristique fait de ces enzymes de parfaits sujets pour l'ingénierie moléculaire afin d'obtenir de nouvelles molécules aux propriétés thérapeutiques optimisées.

Cependant, à ce jour les tentatives d'ingénierie ont permis d'obtenir, dans la majeure partie des cas, des systèmes modifiés présentant des rendements de production bien plus faible que les systèmes natifs. Cette baisse de productivité peut être expliquée par différents paramètres, notamment les interactions entre les domaines qui sont impactées lors de modification des PKS et le fait que ces enzymes forment des méga-complexes enzymatiques, ce qui rend compliqué leur expression hétérologue.

Afin de contribuer à l'amélioration des tentatives d'ingénierie des PKS, cette thèse a été menée autour de deux objectifs principaux. Les bactéries du genre *Streptomyces* étant les plus grandes productrices de polycétides, elles ont également fait l'objet des études menées lors de cette thèse. Le premier objectif repose sur les tentatives précédentes de scission des PKS modifiées ayant permis d'augmenter parfois les rendements de production. L'objectif de cette thèse a, dans ce cadre, été de s'intéresser à la scission de la rapamycine et de la stambomycine synthases afin d'évaluer l'impact sur les rendements de production des composés d'intérêts. L'objectif final étant d'appliquer cette stratégie au système PKS modifié produisant les mini-stambomycines. Le second objectif a été de s'intéresser à l'organisation cellulaire de la PKS *cis*-AT de la coelimycine synthase, dans le but d'approfondir et compléter les connaissances quant à la macro-complexation de ces systèmes.

Investiguer la scission des PKS dans l'objectif d'augmenter les rendements de production.

Le premier objectif de ce projet de thèse a donc été d'étudier cette stratégie de scission sur deux modèles d'étude. Les cas des rapamycine et stambomycine synthases ont été en particulier investigués. La PKS RAPS a été utilisée comme preuve de concept. L'objectif ici a été d'intégrer au sein de RAPS une paire de DD provenant d'une autre PKS : la 6-désoxyérythronolide B synthase (DEBS). Cela permet de scinder la PKS sans modifier profondément la nature des interactions entre sous-unités, comme lors de l'utilisation de SynZips ou du couple SpyCatcher:SpyTag.

Ce premier objectif contient deux axes : l'ingénierie génétique de *S. hygroscopicus* et la caractérisation des interactions entre les DD natifs et introduits. En effet, il est nécessaire de s'assurer que la paire de DD introduite n'interagit pas avec ceux présents chez RAPS, afin de ne pas interférer avec l'assemblage de la PKS.

Cette stratégie de scission a également été appliquée à la stambomycine synthase. Ici, la scission a été effectuée à l'aide de domaines SynZips, afin d'assurer aucune interaction avec les huit couples de DD natifs de la PKS. L'utilisation de ces domaines a permis d'éliminer l'axe de caractérisation des interactions, et ainsi de focaliser l'étude sur l'ingénierie du génome de *S. ambofaciens*. Ici, l'objectif est d'appliquer cette stratégie de scission à la PKS miniaturisée des stambomycines créée par la Dr. Li Su durant sa thèse.

Au-delà de l'objectif d'augmenter les rendements de production en modifiant les PKS, comprendre l'organisation et l'assemblage de ces systèmes est un élément clef pour une meilleure appréhension des tentatives d'ingénierie et de reroutage des voies de biosynthèse. Dans ce cadre, beaucoup d'études ont été menées sur les structures de chaque domaine, ainsi que sur la structure générale de modules de PKS modulaire. Cependant, la localisation de ces systèmes au sein d'un hôte bactérien n'a été étudiée que sur une PKS *trans*-AT. De plus, des domaines d'interaction nommés LINKS semblent avoir été identifiés en leur sein, médiant leur macro-organisation. Une telle étude n'a jamais été menée sur les systèmes PKS *cis*-AT, et cela constitue donc le second objectif de cette thèse.

Etudier la localisation et l'organisation d'une PKS *cis*-AT au sein de son hôte bactérien.

Dans le cas de l'étude de la localisation d'une PKS *cis*-AT, le système d'étude sélectionné est celui de la coelimycine synthase de *S. coelicolor*. L'objectif a été d'étudier la localisation et l'organisation de la première et dernière sous-unité de la PKS, ainsi que du transporteur associé. Ce projet a été initié durant la thèse du Dr. Jean-Malo Massicard. Il a ainsi respectivement fusionné les protéines fluorescentes mCherry et eGFP aux sous-unités CpkA et CpkC. Cela lui a alors permis d'observer la formation de foci au sein des hyphes bactériens, mais sans pouvoir observer de colocalisation des sous-unités. L'objectif a donc été d'utiliser de nouvelles protéines fluorescentes – mYPet et mTurquoise2 – afin de valider ou invalider ces premières observations. L'objectif de cette partie a également été d'étudier la localisation du transporteur de la coelimycine – CpkF – fusionné à la protéine fluorescente mCherry, ainsi que la localisation conjointe de ce transporteur et des sous-unités CpkA et CpkC. Enfin, l'étude de l'impact de la délétion d'un des partenaires sur la localisation et l'organisation des autres composants marqués du système Cpk a également fait l'objet des objectifs de cette partie.

Matériels et méthodes

I. Souches bactériennes, conditions de culture et plasmides

I.A. Souches bactériennes

Les souches bactériennes utilisées dans le cadre des études menées durant cette thèse sont présentées dans le Tableau 2.

Tableau 2. Liste des souches bactériennes utilisées au cours de cette thèse.

Souche	Génotype	Utilisation	Référence
Escherichia coli BL21	F-, *ompT*, *gal*, *dcm*, *lon*, *hsdSB* (rB- mB-), λ(DE3)	Cette souche est utilisée pour la production de protéines recombinantes et est notamment caractérisé par la présence dans son génome de la séquence codant l'ARN polymérase ADN dépendante T7 sous le contrôle d'un promoteur – *lacUV5* – inductible à l'IPTG.	(Studier & Moffatt, 1986)
Escherichia coli DH5α	*supE*44, Δ*lacU*169 (φ80*lacZΔM15*), *hsdR*17, *rec*A1, *end*A1, *gyrA*96, *thj*-1, *relA*1	Cette souche est utilisée pour la propagation des plasmides et est caractérisée notamment par un taux élevé de transformation.	(Grant et al., 1990)
Escherichia coli ET12567(pUZ8002)	*F⁻*, *dam-13*::*Tn9*, *dcm-6*, *hsdM*, *hsdS*	Cette souche a été utilisée pour la conjugaison de *Streptomyces* spp. Elle possède le plasmide pUZ8002 contenant le gène de résistance à la kanamycine et l'opéron *tra* qui code des protéines de transfert d'ADN plasmidique.	(Larcombe et al., 2024; MacNeil et al., 1992)
Saccharopolyspora erythraea	–	Cette souche contient le cluster de gènes de DEBS, la PKS de type I modulaire responsable de la synthèse de l'érythromycine.	
Streptomyces ambofaciens ATCC23877	–	Cette souche contient le cluster de gènes de la PKS des stambomycines, qui est de type I modulaire et a été utilisée pour des expériences d'ingénierie génétique.	(Pinnert-Sindico, 1954)
Streptomyces ambofaciens ATCC/hy59_S2	Δ*sam23877-7112*, ^{Pks9}KSmid$_{14/21}$	Cette souche contient le cluster de gènes de la PKS des stambomycines modifié pour la production de mini-stambomycines (Figure 50C).	(Su et al., 2022)

Tableau 2. Liste des souches bactériennes utilisées au cours de cette thèse.

Souche	Génotype	Utilisation	Référence
Streptomyces coelicolor FM145	–	Cette souche contient le cluster de gènes de la coelimycine synthase. Elle dérive de la souche *S. coelicolor* M145 et présente une auto-fluorescence réduite, favorisant la visualisation des protéines fluorescentes présentes en faible abondance. Cette souche a été fournie par le Pr. Van Wezel (University of Leiden, Netherlands) et a été utilisée pour des expériences d'ingénierie génétique	(Willemse & van Wezel, 2009)
Streptomyces hygroscopicus DSM41530	–	Cette souche contient le cluster de gènes de la rapamycine synthase, une PKS de type I modulaire. Elle a été utilisée pour des expériences d'ingénierie génétique.	(Baranasic et al., 2013; Kumar & Goodfellow, 2008)
Streptomyces hygroscopicus ISOM-3410	–	Cette souche de *S. hygroscopicus* a été mutée afin de surproduire la rapamycine et a été fournie par Isomerase (Chesterford Research Park, Cambridge CB10 1XL, Royaume-Uni). Elle a été utilisée pour des expériences d'ingénierie génétique.	

I.B. Milieux de cultures

Tous les milieux de culture sont stérilisés par autoclavage durant 20 minutes à 1 bar, et les pH sont ajustés par l'ajout de NaOH ou de HCl. Afin d'assurer une pression de sélection, les milieux ont été supplémentés en apramycine (50 µg·L^{-1}), kanamycine (50 µg·L^{-1}), chloramphénicol (25 µg·L^{-1}) ou acide nalidixique (30 µg·L^{-1}), selon les besoins. La composition des milieux utilisés durant cette thèse est présentée dans le Tableau 3.

Tableau 3. Liste des milieux de culture utilisés au cours de cette thèse.

Nom	Composition pour 1L	pH	Utilisation
Difco™ Nutrient Agar (DNA)	23 g Difco™ Nutrient Agar	6,8	Il est supplémenté en acide L-glutamique (50 mM) pour la production de coelimycine P2 par *S. coelicolor*.
ISP medium n°2 (ISP2)	10 g extrait de malt 4 g extrait de levure 4 g glucose	7,5	Ce milieu est utilisé pour la culture liquide de *S. hygroscopicus*.
Lysogeny Broth (LB)	10 g tryptone 5 g extrait de levure 5 g chlorure de sodium	7,5	Ce milieu est utilisé pour la culture des souches d'*E. coli* et est supplémenté en agar (17 g·L^{-1}) pour la culture solide.
Milieu B (MB)	5 g extrait de malt 4 g saccharose 2 g extrait de levure 1 g lait écrémé en poudre 20 g agar	7,1	Ce milieu est utilisé pour la culture en milieu solide de *S. hygroscopicus*.
M-Isp4	5 g farine de soja torréfiée 5 g mannitol 5 g amidon de maïs 2 g tryptone 2 g sulfate d'ammonium 1 g extrait de levure 1 g chlorure de sodium 1 g carbonate de calcium 1 g hydrogénophosphate de potassium 20 g agar	7,1	Ce milieu est utilisé pour la conjugaison de *S. hygroscopicus*.
Soya Flour Mannitol (SFM)	20 g farine de soja torréfiée 20 g mannitol 20 g agar	7,2	Il est supplémenté en chlorure de magnésium (10 mM) pour la conjugaison de *S. coelicolor*.
Tryptone Soya Broth (TSB)	17 g tryptone 5 g chlorure de sodium 3 g soytone 2,5 g glucose 2,5 g hydrogénophosphate de potassium		Ce milieu est supplémenté en agar (20 g·L^{-1}) pour la culture solide de *S. ambofaciens*. Il ne permet pas la sporulation des souches et a été utilisé pour le criblage du mycélium des clones mutés.

I.C. Vecteurs plasmidiques

pCRISPomyces-2 : Ce plasmide conjugatif est apramycine résistant et a été utilisé pour la modification génétique de *S. coelicolor* par l'outil d'ingénierie génétique *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR)-Cas9 (Figure 60) (Cobb et al., 2015).

pKCCpf1 : Ce plasmide conjugatif est apramycine résistant et a été utilisé pour la modification génétique de *S. coelicolor* par l'outil d'ingénierie CRISPR-Cpf1 (Figure 61) (L. Li et al., 2018).

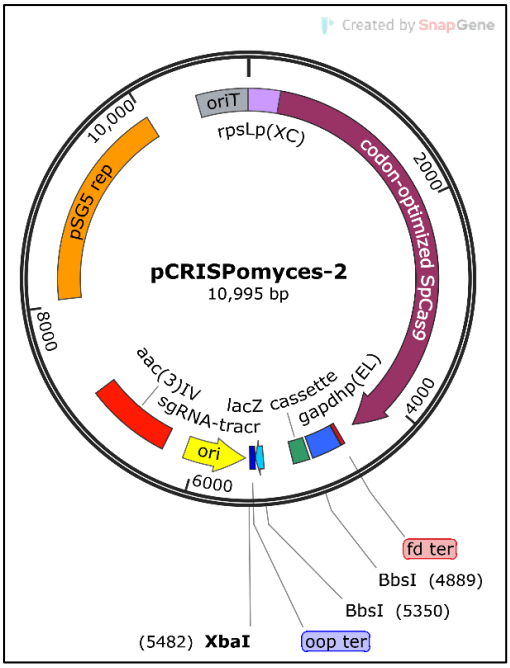


Figure 60. Carte du plasmide pCRISPomyces-2 qui permet l'expression du système CRISPR-Cas9 chez *Streptomyces* spp (Cobb et al., 2015). Ce vecteur contient le gène codant l'endonucléase Cas9, dont la séquence a été optimisée pour son expression chez *Streptomyces* spp, sous le contrôle du promoteur *rpsLp*(XC). La cassette *LacZ* est sous dépendance du promoteur *Pgapdh* et permet la sélection des clones lors du clonage de l'ARN guide entre les sites de restriction *Bbs*I. Les matrices d'éditions sont clonées au site de restriction *Xba*I. Un gène de résistance à l'apramycine (*aac(3)IV*) permet la sélection des clones de *Streptomyces* spp. ayant intégré le vecteur après conjugaison. Enfin, l'origine de réplication thermosensible pSG5 permet la perte du plasmide lors de la culture des souches transformées à 37°C.

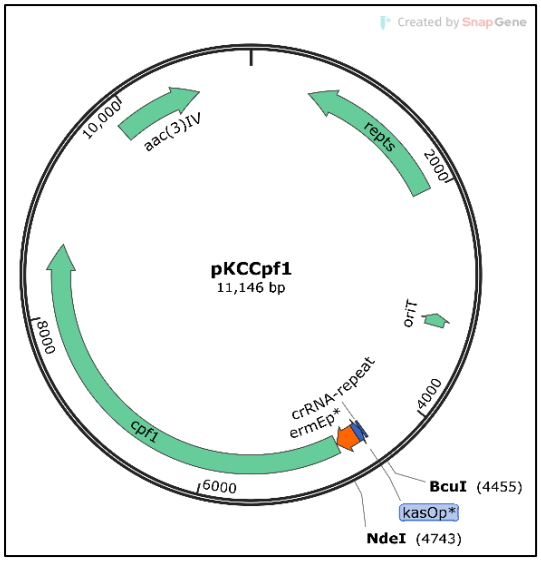


Figure 61. Carte du plasmide pKCCpf1 qui permet l'expression du système CRISPR-Cpf1 chez *Streptomyces* spp (L. Li et al., 2018). Ce vecteur contient le gène codant l'endonucléase Cpf1 (Cas12a), sous le contrôle du promoteur ermEp*. La séquence « crRNA-repeat », suivi lors de la construction des matrices par l'ARN guide, est placée sous le contrôle du promoteur kasOp* et permet le recrutement du système CRISPR sur l'ADN bactérien ciblé. Un gène de résistance à l'apramycine (*aac(3)IV*) permet la sélection des clones de *Streptomyces* spp. ayant intégré le vecteur après conjugaison. Enfin, l'origine de réplication thermosensible *repts* permet la perte du plasmide lors de la culture des souches transformées à 37°C.

pBG102 : Ce vecteur a été utilisé dans le cadre de la production et purification des protéines d'intérêt. Il contient notamment un gène de résistance à la kanamycine et les séquences codant les étiquettes 6xHis et *Small Ubiquitin-like MOdifier* (SUMO) pour la purification des protéines produites (Figure 62). L'expression de la protéine recombinante est sous le contrôle d'un promoteur T7 (Center for Structural Biology, Vanderbilt University).

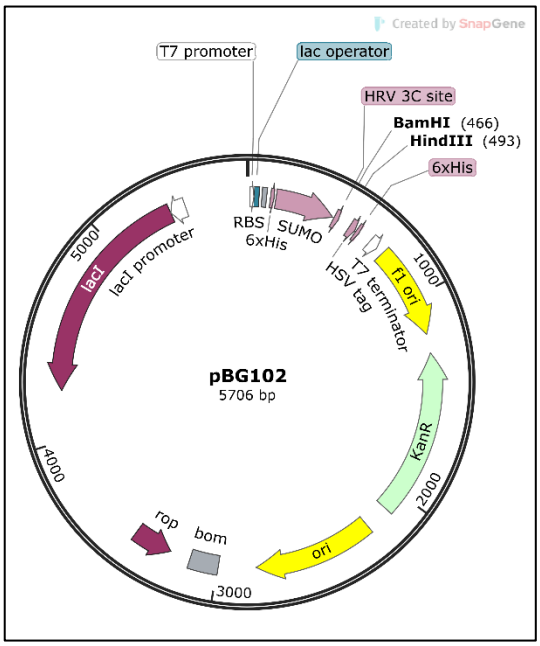


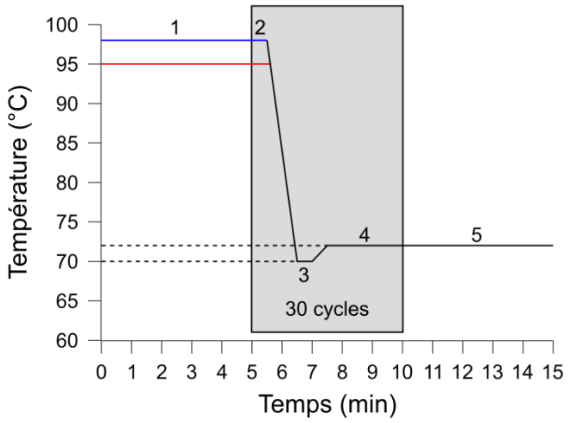
Figure 62. Carte du vecteur d'expression pBG102 (Center for Structural Biology, Vanderbilt University). Ce vecteur contient notamment les séquences codant une étiquette 6xHis et SUMO en position N-terminale des protéines d'intérêt produites. La synthèse de la protéine de fusion d'intérêt est placée sous le contrôle du promoteur T7. Les sites de restrictions *Bam*HI et *Hind*III sont utilisés pour le clonage des séquences d'intérêt au sein du vecteur. Enfin, un gène de résistance à la kanamycine (KanR) permet la sélection des clones bactériens contenant ce plasmide.

I.D. Enzymes

Deux acide désoxyribonucléique (ADN) polymérases ADN dépendantes ont été utilisées durant cette thèse : la Phusion HotStart II et la DreamTaq. Cette première a été utilisée pour l'amplification de séquences dans le cadre des clonages et du contrôle de l'intégrité des modifications génétiques, et la seconde a été utilisée pour le crible de clones bactériens. La composition des mélanges réactionnels ainsi que les programmes d'amplification utilisés sont décrits dans la Figure 63. Pour des questions de simplicité, toutes les amorces conçues dans le cadre de cette thèse l'ont été de sorte à avoir une température d'hybridation aux alentours de 70°C. Cela permet d'utiliser toutes les amorces en combinaison.

Dans le cas du criblage de bactéries Gram positives – ici les *Streptomyces* –, une lyse cellulaire a dû être effectuée avant l'amplification par PCR afin de libérer le matériel génétique. Les colonies ont été placées dans 10 µL de DMSO puis incubées durant 10 minutes à 98°C.

Ensuite, une partie de ce mélange est placée dans le mélange réactionnel de PCR pour atteindre une concentration finale de DMSO de 7 %.



Composant	Quantité
Matrice	*x*
Désoxyribonucléotides	1,2 mM
Amorce directe	0,2 mM
Amorce indirecte	0,2 mM
DMSO	10 %
Tampon	1 X
Polymérase	0,5 U / 1 U
H_2O	qsp 50 µL / 25 µL

Figure 63. Programme d'amplification par réaction en chaîne par polymérase et composition du mélange réactionnel des différentes polymérases utilisées lors de cette thèse. A droite est représenté le programme d'amplification utilisé, et à gauche la composition du mélange réactionnel. En bleu sont représentées les conditions d'utilisation spécifiques à la polymérase Phusion HotStartII (HSII) et en rouge celles à la DreamTaq. Une dénaturation initiale de 5 minutes à 98°C ou 95°C selon la polymérase (**1**) initie l'amplification, puis 30 cycles de 30 secondes de dénaturation à 98°C ou 95°C (**2**), puis 30 secondes d'hybridation des amorces à 70°C (**3**) et d'élongation à 72°C (**4**) sont répétés. Le temps d'élongation est de 30 sec/kb dans le cas de la polymérase Phusion HSII et de 1 min/kb pour la DreamTaq. Enfin, une étape d'élongation finale est effectuée à 72°C durant 5 minutes (**5**). La quantité *x* de matrice utilisée dépend des expérimentations menées.

La protéase du rhinovirus humain 3C (HRV 3C) utilisée pour le clivage des protéines recombinantes produites est exprimée chez *E. coli* BL21(DE3), fusionnée à une étiquette composée de six histidines. Sa purification est faite en une unique étape de chromatographie par chélation de métaux immobilisés (IMAC). Les enzymes de restrictions et l'ADN ligase utilisées pour le clonage des différentes matrices proviennent de chez ThermoFisher Scientific. Elles ont été utilisées en suivant les protocoles fournis par le fournisseur.

II. Extraction de matériel génétique

II.A. Extraction d'ADN génomique bactérien

Les ADN génomiques de *S. coelicolor*, *S. hygroscopicus*, *S. ambofaciens* et *Sac. erythraea* ont été extraits afin d'amplifier les éléments nécessaires à l'édition génomique des souches. Ces extractions ont été faites avec le kit Wizard® Genomic DNA Purification (Promega). La quantité d'ADN a été mesurée par l'absorbance de l'échantillon à 260 nm et sa pureté a été analysée à l'aide du rapport d'absorbance 280/260 nm. Ensuite, l'échantillon a été conservé à -20°C.

II.B. Mini-préparation de matériel plasmidique

Les préparations d'ADN plasmidique ont été réalisées en utilisant le kit NucleoSpin® (low copy) Plasmid (Macherey-Nagel, Allemagne) à partir de cultures bactériennes de 10 mL. Les échantillons finaux ont été élués dans 50 µL d'eau. La concentration de plasmide purifié a été mesurée par l'absorbance de l'échantillon à 260 nm, la pureté a été analysée par le rapport d'absorbance 280/260 nm et l'échantillon a été conservé à -20°C.

II.C. Extraction de matériel génétique sur gel d'agarose

Les extractions de matériel génétique sur gel d'agarose ont été effectuées à l'aide du kit NucleoSpin Gel and PCR clean-up (Macherey-Nagel). Le matériel génétique a été élué par l'ajout de 30 µL d'eau. La concentration des échantillons obtenus a été mesurée par leur absorbance 260 nm et leur pureté analysée par le rapport d'absorbance 280/260 nm, puis ils ont été conservés à -20°C.

III. Construction des plasmides

III.A. Conception, amplification et assemblage des éléments nécessaires à l'édition des génomes de *Streptomyces* spp.

Dans le cadre de l'ingénierie génétique des différentes espèces de *Streptomyces*, deux systèmes d'édition *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR) ont été utilisés : celui associé à l'endonucléase Cas9 (Jinek et al., 2012) et celui lié à l'endonucléase Cpf1 (Cas12a) (Zetsche et al., 2015). Les différences majeures entre les systèmes CRISPR-Cas9 et CRISPR-Cpf1 sont les séquences PAM reconnues et le clivage de l'ADN double brin (L. Li et al., 2018; Yamano et al., 2016). Le système CRISPR-Cas9, communément utilisé et provenant de l'organisme *Streptococcus pyogenes*, reconnaît une séquence PAM « NGG » (N = A, T, C ou G), tandis que le système Cpf1 provient de *Francisella novicida* et reconnaît la séquence « TTV » (V = A, C ou G) (L. Li et al., 2018). Pour ce qui est du clivage, CRISPR-Cas9 opère une coupure donnant naissance à des bouts francs, tandis que CRISPR-Cpf1 laisse des extrémités cohésives de 5 nucléotides (Yamano et al., 2016; Zetsche et al., 2015).

III.A.1. Identification des ARN guides

Il existe différents outils en ligne permettant d'identifier des séquences cibles au sein de génomes pour l'utilisation de systèmes CRISPR. Les outils CRISPy-Web et CHOPCHOP en particulier ont été utilisés au cours de cette thèse (Blin et al., 2016; Labun et al., 2019).

III.A.2. Construction des matrices d'édition génétique

III.A.2.a. Construction des matrices dans le cas de la scission de polycétide synthase

Chaque élément des matrices a été amplifié par PCR en utilisant la polymérase Phusion HSII. Les amorces utilisées sont présentées dans le Tableau 4.

Les régions d'homologie (RH1 et RH2) ont été amplifiées à partir de l'ADN génomique extrait de *S. hygroscopicus* dans le cas de la scission de RAPS, et de *S. ambofaciens* pour la scission de la stambomycines synthase. Dans le cas de la scission de la stambomycines synthase, les séquences codant les SynZips sont également retrouvées au sein du gène synthétique (GS).

Tableau 4. Amorces oligonucléotidiques utilisées pour la construction des cassettes d'édition dans le cadre de la scission de polycétide synthases.

Amplicon	Nom ESCOLA	Sens	Séquence 5' → 3'
Matrices d'édition de la rapamycine synthase			
RH1	***I952***	***direct***	***atatcgacgcACTAGTgaccggctgcgggagatcatc***
	H019	indirect	ccgggccagccactgcaccgacgag
GS	H020	direct	ctcgtcggtgcagtggctggcccggttggcgccggtggagcgggagaagg
	H021	indirect 1	ccaccgccggtgaggccgcgaacagcaactcccccagccgggctgccagg
	H022	indirect 2	ccagccggtcgccgatgtccaccgcaggagtttcaccgacgaacaactcc
Séquence codant les DD	H023	direct 1	ctgttcgcggcctcaccggcggtgg
	H024	direct 2	gcggtggacatcggcgaccggctgg
	H025	**indirect**	**accggcccggtgctcgacttcacgc**
RH2	**H026**	**direct**	**gcgtgaagtcgagcaccgggccggtgagccgttggcgattgtcggcatgg**
	J021	indirect	gcagcccggctgggggagttgttcgcgtatgacccagattcgacttcacc
ARNcr	J006	direct	cgaacaactcccccagccgggctgcatctacaacagtagaaatttgg
	I410	***indirect***	***ggtagatggaCATATGtggatcctaccaaccggca***
Matrices d'édition de la stambomycines synthase			
RH1	***I945***	***direct***	***atatcgacgcACTAGTccggcacctggtcaccgaaca***
	I341	indirect	tcgaagacgagggtggcggggaggc
GS	I325	direct	gcctccccgccaccctcgtcttcgaccatccgaccccgaccgtgctcgcc
	I658	**indirect 1**	**tccagaaccccctcccgagccacctttttttgtgcgccagacggttgcgcag**
	I657	**indirect 2**	**tccagaaccccctcccgagccaccgaagtacgcttcttcacgctccagttt**
RH2	**I469**	**direct**	**ggtggctcgggaggggggttctggaggcggcgagcccgcccccac**
	J020	indirect	ggcgagcacggtcggggtcggatgggccgggtgccggtaccgtgcgcctc
ARNcr	J007	direct	ccatccgaccccgaccgtgctcgccatctacaacagtagaaatttgg
	I410	***indirect***	***ggtagatggaCATATGtggatcctaccaaccggca***

Les amorces indiquées en gras ont été utilisées pour le sous-assemblage des fragments de PCR, et celles indiquées en gras et italique pour l'assemblage finale des matrices d'édition. Les sites de restrictions sensibles aux enzymes *Spe*I et *Nde*I sont représentés en majuscule, respectivement au sein des amorces I945 et I410. **Légende –** RH : région d'homologie, DD : domaines de docking, GS : gène synthétique, ARNcr : acide ribonucléique CRISPR.

Dans le cas de RAPS, la séquence codant les DD de DEBS (DD) a été amplifiée à partir de l'ADN génomique extrait de *Sac. erythraea*. Enfin, la séquence de l'ARNg a été amplifiée à partir du vecteur pCRISPomyces-2. Les gènes synthétiques (GS) nécessaires aux différentes constructions ont été commandés chez Integrated DNA Technologies ™.

Ensuite, les fragments RH1, GS et DD d'une part, et RH2 et ARNcr de l'autre, ont été assemblés par PCR afin de former les sous-fragments RH1-GS-DD et RH2-ARNcr. Cet assemblage a été rendu possible par l'utilisation d'amorces chevauchantes. Enfin, une dernière étape de PCR a permis d'assembler les deux sous-fragments afin d'obtenir la matrice d'édition complète.

III.A.2.b. Construction des matrices dans le cadre de l'étude de la localisation de la coelimycine synthase

La majeure partie des constructions concernant cette partie ont été réalisées par le Dr. Jean-Malo Massicard durant sa thèse. La liste des constructions nécessaires à l'étude de la localisation et l'organisation de la coelimycine synthase Cpk chez *S. coelicolor* est présentée dans le Tableau 5.

Tableau 5. Liste des constructions plasmidiques nécessaires à l'étude de l'organisation de la coelimycine synthase chez *Streptomyces coelicolor*.

Construction	Description	Source
pCRISPomyces-2-*mYPet-CpkA*	Expression de la protéine fluorescente mYPet fusionnée à la sous-unité CpkA en position N-terminale.	Ce travail
pCRISPomyces-2-*cpkC-mTurquoise2*	Expression de la protéine fluorescente mTurquoise2 fusionnée à la sous-unité CpkC en position C-terminale.	(Massicard, 2020)
pCRISPomyces-2-*cpkF-mCherry*	Expression de la protéine fluorescente mCherry fusionnée au transporteur CpkF en position C-terminale.	(Massicard, 2020)
pCRISPomyces-2-Δ*cpkA*	Délétion du gène codant la sous-unité CpkA.	(Massicard, 2020)
pCRISPomyces-2-Δ*cpkC*	Délétion du gène codant la sous-unité CpkC.	(Massicard, 2020)
pCRISPomyces-2-Δ*cpkF*	Délétion du gène codant le transporteur CpkF.	(Massicard, 2020)

L'amplification et l'assemblage de la matrice ont été faits par PCR en utilisant la polymérase Phusion HSII. Les régions d'homologie ont été amplifiées à partir de l'ADN génomique de *S. coelicolor* et le gène codant la protéine fluorescente mYPet à partir du plasmide pSS149 (Schlimpert et al., 2017), qui a été gracieusement donné à l'équipe par la Dr. Susan Schlimpert. Une fois ces éléments amplifiés, les fragments RH1 et GS d'une part, et les fragments mYPet et RH2 d'autre part, ont été assemblés par PCR comme décrit précédemment. Ces sous-fragments ont ensuite été assemblés afin d'obtenir la matrice finale. Toutes les amorces utilisées pour la construction de cette matrice sont décrites dans le Tableau 6.

Tableau 6. Amorces oligonucléotidiques utilisées pour la construction de la matrice d'édition *mYPet-cpkA*.

Amplicon	Nom ESCOLA	Sens	Séquence 5' → 3'
RH1	***F634***	***direct***	***gccgggcgtttttttatctagacggtgaacagcagggcgg***
	F635	indirect	gtgtcaatctgaacctcctcgccgagaacgccactacg
GS	F636	direct	tccgtagtggcgttctcggcgaggaggttcagattgacaccg
	F640	**indirect**	**gcatgaacgagctgtacaagggcggcggcggcggcggcgtcaccagtg aaaataactcggttgct**
mYPet	**F639**	**direct**	**gagttattttcactggtgacgccgccgccgccgccgccgcccttgtacagctcgtt catgccc**
	F638	indirect	ccaggaaagggcagccgagcatggtctccaagggcgagg
RH2	F637	direct	tccgtagtggcgttctcggcgaggaggttcagattgacaccg
	F641	***indirect***	**ttacggttcctggcctctagacgtgttgatggagatgttgccct**

Les amorces indiquées en gras ont été utilisées pour le sous-assemblage des fragments de PCR, et celles indiquées en gras et italique pour l'assemblage final de la matrice d'édition. **Légende –** RH : région d'homologie, GS : gène synthétique, ARNcr : acide ribonucléique CRISPR.

III.B. Clonage des matrices et ARN guides au sein des vecteurs d'intérêts pour l'ingénierie génétique de *Streptomyces* spp.

Dans le cas de l'utilisation du vecteur pCRISPomyces-2, la séquence de l'ARNg n'est pas introduite avec la matrice d'édition mais clonée directement au sein du vecteur. Il a donc été nécessaire de cloner la séquence ARNg avant de procéder au clonage de la matrice.

III.B.1. Clonage de la séquence ARN guide au sein du plasmide pCRISPomyces-2

Le vecteur pCRISPomyces-2 a été construit de telle sorte à pouvoir cloner la séquence de l'ARNg au niveau d'un site de restriction *Bbs*I. Celle-ci est une enzyme de restriction de type IIS, ce qui signifie qu'elle reconnaît une séquence ADN asymétrique et qu'elle clive en dehors de son site de reconnaissance. Le clonage de l'ARNg dans le plasmide a été fait par un assemblage de type Golden Gate (Figure 64). Pour ce faire, des extensions de 4 nucléotides sont présentes à l'extrémité des amorces qui reconnaissent spécifiquement les extrémités de 4 nucléotides provoquées par le clivage de *Bbs*I dans le vecteur.

Les oligonucléotides contenant la séquence de l'ARNg (Tableau 7) ont dû être hybridées ensemble avant de cloner l'ADN double brin obtenu. Pour ce faire, 0,5 µg de chaque amorce sont mélangés dans du tampon contenant de l'acide 4-(2-hydroxyéthyl)-1-pipérazineéthanesulfonique (HEPES) à une concentration de 30 mM, pH 7,8. Le mélange a ensuite été dénaturé par une étape de chauffe à 95°C pendant 5 minutes, puis la température a été diminuée de 0,1°C·s^{-1} jusqu'à atteindre une température de 4°C afin de permettre l'hybridation des amorces. Ensuite, cet ADN double brin est mis en contact avec le vecteur pCRISPomyces-2, l'enzyme de restriction *Bbs*I et l'ADN ligase T4 afin d'effectuer l'assemblage de type Golden Gate tel que décrit dans le Tableau 7.

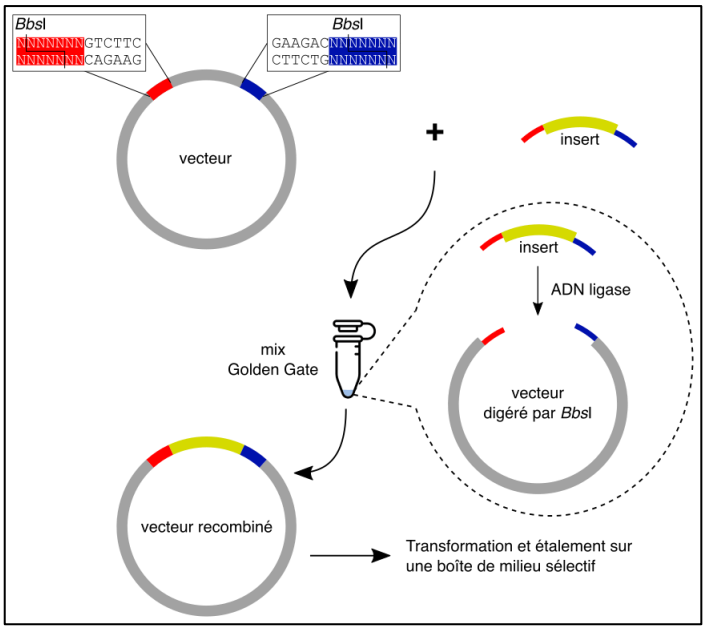


Figure 64. Principe de la méthode d'assemblage de type Golden Gate (adapté de Engler et al., 2008. Le vecteur d'intérêt – ici le pCRISPomyces-2 – ainsi que la séquence à intégrer sont placés dans le mélange réactionnel contenant l'enzyme de restriction *Bbs*I et l'ADN ligase T4. L'enzyme *Bbs*I permet dans un premier temps de cliver le vecteur et de créer des extrémités de 4 nucléotides non compatibles. L'insert est conçu de manière à avoir 4 nucléotides en positions 3′ et 5′ qui soient compatibles avec les extrémités obtenues dans le plasmide. Une fois les séquences hybridées, l'ADN ligase T4 permet de ligaturer l'insert dans le vecteur. Ensuite, le mélange réactionnel est transformé dans des bactéries compétentes, puis ces bactéries sont sélectionnées sur un milieu supplémenté en antibiotique (ici l'apramycine).

Tableau 7. Amorces oligonucléotidiques, mélange réactionnel et programme utilisés pour le clonage de l'ARN guide au sein du vecteur pCRISPomyces-2 par Golden Gate.

Amorce	Nom ESCOLA	Séquence 5′ → 3′
cpkA-ARNg direct	D455	ACGCtgctacggaattgttctca
cpkA-ARNg indirect	D454	AAACgtgagaacaattccgtagca

Composant	Quantité
pCRISPomyces-2	100 ng
Amorces hybridées	0,03 µL
Tampon ADN ligase T4	2 µL
*Bbs*I	1 µL
ADN ligase T4	1 µL
H_2O	qsp 20 µL

Durée	Température	Cycles
10 min	37°C	10
10 min	16 °C	
5 min	50 °C	1
20 min	65°C	1

Les extensions de 4 nucléotides permettant l'insertion de l'ARN guide (ARNg) au sein du vecteur pCRISPomyces-2 au site de clivage *Bbs*I sont indiquées en majuscules.

III.B.2. Clonage des matrices d'édition au sein des vecteurs d'intérêt

Les matrices d'édition ont été clonées au sein des vecteurs pKCCpf1 pour le cas de l'édition des génomes de *S. hygroscopicus* et de *S. ambofaciens*, et pCRISPomyces-2 pour l'édition du génome de *S. coelicolor*. Cela a été fait par assemblage de type Gibson (Figure 65) (Gibson et al., 2009).

Pour cela, le vecteur d'intérêt a préalablement été digéré – par *Bcu*I et *Nde*I dans le cas du pKCCpf1 et par *Xba*I dans le cas du pCRISPomyces-2 – et déphosphorylé pour éviter tout événement de ligature du plasmide sur lui-même. Une quantité de 50 ng de vecteur digéré est incubé avec la matrice d'édition à un ratio de 2:1 et 0,5 µL du mélange Gibson commercial In-Fusion Cloning (Takara). Ce mélange contient une 5′-exonucléase qui permet de digérer les extrémités du vecteur linéarisé et de la matrice d'édition entraînant leur hybridation, une ADN polymérase qui reconstitue les nucléotides digérés des deux séquences et d'une ADN ligase qui joint le vecteur à la matrice. L'incubation est faite durant 15 minutes à 55°C puis le mélange est transformé au sein de la souche *E. coli* DH5α.

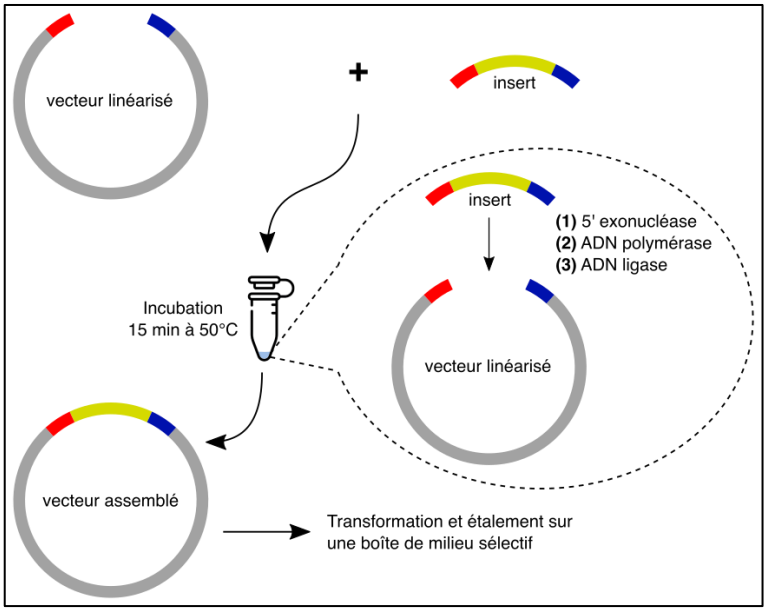


Figure 65. Principe de la méthode d'assemblage de type Gibson (adapté de Gibson et al., 2009). Cette méthode repose sur l'utilisation de trois activités enzymatiques. Une 5′-exonucléase va dans un premier temps digérer l'extrémité 5′ du vecteur linéarisé et du / des inserts. Les extrémités sortantes, conçues pour être complémentaires, vont ensuite s'hybrider et une ADN polymérase ADN dépendante viendra combler les espaces laissés par l'exonucléase. Enfin, une ADN ligase va permettre la formation de liaisons phosphodiester entre le(s) fragment(s) et le vecteur.

III.C. Conception et clonage des séquences codant les domaines de docking

Afin d'étudier les interactions entre les DD natifs de RAPS et ceux introduits provenant de DEBS, les séquences les codant ont tout d'abord été clonées dans le vecteur d'expression pBG102. Les séquences codant les DD provenant de DEBS ont été amplifiées à partir de l'ADN génomique de *Sac. erythraea*. Les domaines $^{C}DD_{RAPSA}$ et $^{C}DD_{RAPSB}$ ont été amplifiés à partir de l'ADN génomique de *S. hygroscopicus*. L'amplification des gènes codant les domaines $^{N}DD_{RAPSB}$ et $^{N}DD_{RAPSC}$ ayant été infructueuse à partir de l'ADN génomique, ceux-ci ont été reconstitués par PCR en utilisant des amorces chevauchantes. Les amplifications ont été faites en utilisant la polymérase Phusion HSII et les amorces décrites dans le Tableau 8.

Tableau 8. Amorces oligonucléotidiques utilisées pour l'amplification des séquences codant les domaines de docking.

Amplicon	Nom ESCOLA	Sens	Séquence 5′ → 3′
$^{C}DD_{DEBS2}$	H510	direct	aaaaGGATCCttcgcggcctcaccggcggtg
	H509	indirect	aaaaAAGCTTctacaggtcctctcccccgccgaa
$^{N}DD_{DEBS3}$	H508	direct	aaaaGGATCCatgagcggtgacaacggcatg
	H507	indirect	aaaaAAGCTTtcactcaccggcccggtgctcgac
$^{C}DD_{RAPSA}$	H506	direct	aaaaGGATCCtatgctccggcggagccggaacct
	H505	indirect	aaaaAAGCTTtcatcgttcttcctccagcacgtg
$^{N}DD_{RAPSB}$	I063	direct	GGATCCatgcctgagcaagacaaagtggtcgaatatcttcgctg**ggcgactgcggagttgcacaccacccgcgc**
	I062	indirect	AAGCTTtcaggtgttggcggcggcaagtgcttccagtttg**gcgcgggtggtgtgcaactccgcagtcgcc**
$^{C}DD_{RAPSB}$	H502	direct	aaaaGGATCCcatgaacccgagggttcagtcg
	H501	indirect	aaaaAAGCTTtcatcaggccgtgggatcaccgaa
$^{N}DD_{RAPSC}$	I061	direct	GGATCCctatatgagggaagatcagcttctcgacgcattgcgcaagt**cggtcaaggagaacgcgcgg**
	I060	indirect	AAGCTTtcagtccatggctgcccgcaacgacgtgttcgccttgcggag**ccgcgcgttctccttgaccg**

Les séquences indiquées en gras sont les extrémités complémentaires des amorces, permettant l'assemblage des gènes par chevauchement. Les sites de restrictions reconnus par *Bam*HI et *Hind*III, utilisés pour le clonage des séquences au sein du plasmide pBG102, sont représentés en majuscule et respectivement présents au sein des amorces directes et indirectes.

Les séquences ainsi obtenues ont ensuite été clonées au sein du vecteur d'expression pBG102 entre les sites de restrictions *Bam*HI et *Hind*III. Pour cela, le vecteur a préalablement été linéarisé puis purifié sur gel d'agarose. Le clonage a été fait à l'aide de la ligase T4 puis 4 µL des mélanges réactionnels ont été transformés chez *E. coli* DH5α afin d'obtenir des mini-préparations de ces constructions. L'intégrité des vecteurs obtenus a été validée par séquençage.

IV. Préparation et transformation de cellules bactériennes

IV.A. Préparation et transformation des souches d'*Escherichia coli* chimio-compétentes

Les cellules électro-compétentes ont été préparées à partir d'une préculture de bactérie dans 50 mL de milieu LB et placée à 37°C sous agitation à 190 rpm durant 18 à 20 heures. Ensuite, une culture de 100 mL de milieu LB est inoculée afin d'obtenir une DO_{600} finale de 0,1, puis placée à 37°C sous agitation à 190 rpm jusqu'à atteindre une DO_{600} aux alentours de 0,5. La culture a ensuite été centrifugée à 1780 *g* durant 15 minutes à 4°C. Le culot a été repris dans 30 mL de solution TfbI (Tableau 9), supplémenté en chlorure de manganèse (10 $g \cdot L^{-1}$), puis incubé à 4°C durant 10 minutes et centrifugé à nouveau. Le culot final a été repris dans 2 mL de solution TfbII puis des fractions aliquotes de 50–100 µL ont été conservées à -80°C jusqu'à utilisation.

Tableau 9. Solutions tampons utilisées pour la préparation des bactéries *Escherichia coli* chimio-compétentes.

Tfb I		Tfb II	
Composants	**Concentrations**	**Composants**	**Concentrations**
acétate de potassium	30 mM	acide 3-morpholino-1-propanesulfonique (MOPS)	30 mM
chlorure de potassium	100 mM	chlorure de potassium	100 mM
chlorure de calcium	10 mM	chlorure de calcium	10 mM
glycérol > 99 %	15 %	glycérol > 99 %	15 %

Pour la transformation des bactéries *E. coli*, 50 µL de cellules compétentes ont été utilisés et mis en contact avec 50–100 ng de plasmide. Le mélange a été incubé à 4°C durant 15 à 30 minutes, puis 30 secondes à 42°C et 2 minutes à 4°C à nouveau. Ensuite 100 µL de milieu LB ont été ajoutés et les cellules ont été incubées à 37°C sous agitation durant 1h à 1h30. Enfin, les cellules ont été étalées sur milieu LB solide supplémenté avec le(s) antibiotique(s) adéquat(s) pour la sélection des clones transformés.

IV.B. Préparation et transformation des souches d'*Escherichia coli* électro-compétentes

Les cellules électro-compétentes ont été préparées à partir d'une préculture de bactérie dans 50 mL de milieu LB et placée à 37°C sous agitation à 190 rpm durant 18 à 20 heures. Ensuite, une culture de 100 mL de milieu LB est inoculée afin d'obtenir une DO_{600} finale de 0,1, puis placée à 37°C sous agitation à 190 rpm jusqu'à atteindre une DO_{600} aux alentours de 0,5. Les cultures sont ensuite centrifugées durant 15 minutes à 1780 *g* à 4°C. Le surnageant a été éliminé et le culot suspendu dans 40 mL de glycérol 10 % glacé, puis le mélange est à nouveau centrifugé comme précédemment. Cette étape est répétée une seconde fois afin de laver les cellules. Ensuite, le culot est suspendu dans 1 mL de glycérol 10 % glacé puis des fractions aliquotes de 50 µL sont stockées à -80°C jusqu'à utilisation.

La transformation des souches d'*E. coli* électro-compétentes a été réalisée en mélangeant 50 µL de bactéries avec 100 ng du vecteur d'intérêt. Le mélange est introduit dans une cuve d'électroporation puis placé dans un appareil Eporator (Eppendorf ®). Une tension de 1,7 kV a été appliqué sur l'échantillon, puis 100 µL de milieu LB ont rapidement été ajoutés et le tout a été placé à 37°C sous agitation à 190 rpm pendant 1 heure. Ensuite, le mélange a été étalé sur une boîte de milieu LB supplémentée avec le(s) antibiotique(s) adéquat(s).

IV.C. Conjugaison intergénérique des bactéries du genre *Streptomyces*

IV.C.1. Récolte des spores de Streptomyces *spp*.

Des boîtes de milieu de sporulation – SFM pour les espèces *S. coelicolor* et *S. ambofaciens* et MB dans le cas de *S. hygroscopicus* – sont ensemencées à raison de 10^8 spores de *Streptomyces*, puis placées à 30°C durant 7 à 10 jours. Une fois que les souches ont sporulé, 5 mL de milieu d'eau stérile sont déposés sur la boîte et les spores sont grattées à l'aide d'un coton-tige stérile. Le mélange est ensuite filtré à l'aide d'une seringue qui contient du coton cardé stérile. Ensuite, les solutions sont centrifugées à 1780 *g* durant 10 minutes et le culot est suspendu dans 500 µL d'eau.

IV.C.2. Préparation des cellules d'Escherichia coli

Les cellules d'*E. coli* ET12567(pUZ8002) ont été rendues électro-compétentes puis transformées avec les constructions d'intérêt. Une pré-culture de 5 mL de LB supplémenté avec les antibiotiques appropriés a été inoculée de ces bactéries puis placée à 37°C sous agitation durant 16 à 18 heures. Une culture de 25 mL de LB, également supplémentée en antibiotiques, a ensuite été inoculée à une DO_{600} de 0,1 et placée à 37°C sous agitation jusqu'à atteindre une DO_{600} comprise entre 0,45 et 0,6. La culture a été ensuite centrifugée à 1780 *g* durant 12 minutes. Le culot a été repris dans 25 mL de milieu LB puis centrifugé à nouveau. Cette étape de lavage a été répétée deux fois. Le culot final a été repris dans 2 mL de milieu LB.

IV.C.3. Transformation par conjugaison des espèces de Streptomyces

Des fractions aliquotes de 100 µL de spores ont été chauffées ensuite à 50°C durant 10 minutes afin d'induire leur germination. A chaque fraction aliquote de spores de *Streptomyces* spp. a été ajouté 500 µL de cellules d'*E. coli* contenant le vecteur d'intérêt. Le mélange a ensuite été étalé sur le milieu approprié et les boîtes ont été placées à 30°C durant 16 à 18 heures. De l'acide nalidixique a ensuite été ajouté afin d'inhiber la croissance d'*E. coli*, ainsi que de l'apramycine pour sélectionner les clones de *Streptomyces* spp. ayant intégré le plasmide. Les boîtes sont ensuite replacées à 30°C durant 5 à 7 jours.

IV.C.4. Génotypage des exconjuguants par PCR sur colonies

Une fois les clones obtenus, il est nécessaire de s'assurer que le génome a bien été modifié. Dans un premier temps, un criblage par PCR sur colonie est effectué avec la polymérase Dream Taq. Les amorces utilisées dans ce cas sont présentées dans le Tableau 10. Une fois les clones recombinants identifiés, une nouvelle amplification a été faite avec les mêmes amorces, en utilisant comme enzyme la polymérase Phusion HSII afin de limiter les erreurs d'amplification, et l'amplicon a ensuite été séquencé afin de valider l'intégrité de la modification génétique.

Tableau 10. Amorces oligonucléotidiques utilisées pour le criblage par PCR sur colonies et le génotypage des exconjuguants de *Streptomyces* spp.

Mutation ciblée	Criblage	Nomenclature ESCOLA	Sens	Séquence 5' → 3'
Δ*cpkA*	+ / -	E848	direct	accacgtgcgcgttggtgccg
	+	H718	indirect	actatcgcgcggccttcgcgc
	-	H719	indirect	ccctcctcgacgacctcggct
Δ*cpkC*	+ / -	F439	direct	tcgcgcggactgaccgtcgac
	+	F437	indirect	cggcgaactgatggccctcgac
	-	F436	indirect	tccggtgctgacgcgggagct
Δ*cpkF*	+	F428	direct	acttcccggcaggcctgctgg
	+	F427	indirect	agaaacggctcggcgaagccgtc
	-	F424	direct	acttcccggcaggcctgctggtca
	-	F423	indirect	agctccaccgcgggagcggtg
mYPet-cpkA	+	F646	direct	atgaagccctcggggtcggtg
	+	F647	indirect	agaacttgtggccgttgacgtcg
cpkC-mTurquoise2	+	F651	direct	ggagacggagaacttgtggccgt
	-	F653	direct	tgtacgccgtcgaggtcgtcc
	+ / -	F650	indirect	cggaactgtcggcggacggtt
rapsB(3-3)-DD	-	H635	direct	ggacaccactgacagccgggg
	+	H636	direct	aggcgctcgaagccctgtccg
	+ / -	H634	indirect	accgcgctgcagcgccatcac
pks3(2-3)-SZ	+	K009	direct	tactggacctcgtccgaggcaacgc
	+	K010	indirect	actcgaacgcctcccaggcggtttc

IV.C.5. Sauvegarde des souches de Streptomyces *spp.*

Une fois l'intégrité de la modification génétique vérifiée, il a été nécessaire de leur faire perdre le plasmide d'édition pCRISPomyces-2 ou pKCCpf1 selon la stratégie. Pour cela, les clones sont cultivés à 37°C jusqu'à ce qu'ils deviennent sensibles à l'apramycine. Ensuite, les clones sont mis en culture de manière à obtenir des spores, et ces spores sont ensuite récoltées et sauvegardées dans du glycérol 20 % à -20°C.

V. Production et purification de protéines recombinantes

Toutes les solutions tampons utilisées pour la production et purification des protéines recombinantes sont présentées dans le Tableau 11.

Tableau 11. Solutions tampons utilisées pour la purification des protéines d'intérêts.

Solution	Composition	pH
Tampon A	Tris-HCl 50 mM, NaCl 500 mM	8,0
Tampon B	Tris-HCl 50 mM, NaCl 500 mM, imidazole 500 mM	8,0
Tampon phosphate	NaP 100 mM	8,0

V.A. Production des protéines recombinantes

Les bactéries *E. coli* BL21(DE3) chimio-compétentes ont été transformées avec les constructions d'intérêt. Les clones sélectionnés ont été placés dans une préculture de 50 mL de milieu LB supplémenté en chloramphénicol et kanamycine. Ces précultures ont été placées à 37°C sous agitation durant 16 à 18 heures. Les 2 litres de milieu de culture, supplémentés en chloramphénicol et kanamycine, ont ensuite été ensemencés afin d'atteindre une DO_{600} de 0,1 et placés à 37°C sous agitation jusqu'à atteindre une DO_{600} de 0,7. La production a alors été induite par l'ajout d'isopropyl β-D-1 thiogalactopyranoside (IPTG) (0,5 mM) et les cultures ont été placées à 20°C sous agitation durant 24h. Par la suite, les milieux de culture ont été centrifugés durant 40 minutes à 4780 *g* à 4°C. Les culots ont été repris dans 10 mL de tampon A (Tableau 11), puis à nouveau centrifugés à 4780 *g* à 4°C. Après centrifugation, le surnageant a été éliminé et les culots ont été conservés à -20°C.

V.B. Purification des protéines recombinantes

La purification des domaines de docking a été effectuée en quatre étapes : une chromatographie IMAC, une dialyse et une protéolyse, une chromatographie IMAC inverse et une chromatographie d'exclusion de taille par gel filtration (SEC-GF). Toutes les colonnes de chromatographie ont été couplées à un système ÄKTA pure™ (cytiva). A la fin de chaque étape, une fraction aliquote de 20 µL est récoltée afin d'évaluer la pureté des échantillons par migration sur un gel SDS-PAGE.

V.B.1. Lyse cellulaire

Les culots cellulaires décongelés ont été repris dans 30 mL de tampon A par litre de culture. La suspension cellulaire a ensuite subi trois cycles de sonication de deux minutes à 40 kHz. Le lysat a ensuite été centrifugé à 18000 *g* durant 40 minutes à 4°C. Le surnageant contenant la protéine d'intérêt a ensuite été filtré à l'aide d'une seringue et d'une membrane de polyéthersulfone (PES) d'une porosité de 0,22 µm.

V.B.2. Chromatographie par chélation de métaux immobilisés

Afin d'éviter les interactions aspécifiques entre les protéines d'*E. coli* et la colonne, du tampon B est ajouté à l'échantillon afin d'obtenir une concentration finale en imidazole de 30 mM. Celui-ci a ensuite été chargé sur une colonne Nickel-Sépharose (cytiva), préalablement équilibrée à l'aide des tampons A et B (Tableau 11) afin d'obtenir une concentration en imidazole de 30 mM. Alors, l'étiquette histidine permet à la protéine recombinante de fixer les ions nickel de la colonne et ainsi d'être retenue. La colonne a ensuite été lavée avec une concentration en imidazole de 30 mM (6 % tampon B) afin d'éliminer les protéines qui n'interagissent pas avec la colonne. Puis les protéines ayant interagit avec la colonne sont éluées avec une concentration en imidazole de 300 mM (60 % tampon B).

Enfin, toutes les protéines toujours retenues sur la colonne à cette étape sont éliminées par un lavage de la colonne à 500 mM d'imidazole (100 % tampon B). Tous les détails du programme utilisé pour la purification des protéines d'intérêt par chromatographie par chélation de métaux immobilisés sont présentés dans le Tableau 12.

Tableau 12. Programme utilisé pour la chromatographie par chélation de métaux immobilisés.

Etape	Tampon A (%)	Tampon B (%)	Débit	Volume colonne
Equilibration	94	6	5 mL·min^{-1}	3
Injection	–	–	2 mL·min^{-1}	–
Lavage	94	6	5 mL·min^{-1}	15
Elution	40	60	5 mL·min^{-1}	7
Lavage	0	100	5 mL·min^{-1}	4

Une fois la purification terminée, la colonne est lavée à l'eau puis à l'éthanol 70 % avant d'être stockée à 4°C. La concentration en protéines au sein des fractions récoltées durant la purification est suivie par leur absorbance à 280 nm. Ainsi, les fractions obtenues lors de l'élution et contenant la plus haute concentration en protéines sont récoltées et utilisées par la suite.

V.B.3. Dialyse et protéolyse

Afin de permettre le clivage entre l'étiquette et la protéine d'intérêt ainsi que l'élimination de l'imidazole de l'échantillon, une dialyse et une protéolyse ont été effectuées sur l'échantillon d'intérêt. Pour cela, un bain de dialyse de 4 L de tampon A a été utilisé (Tableau 11). Aux fractions précédentes contenant la protéine d'intérêt sont ajoutés 150 µg de protéase HRV 3C et le tout est placé dans un boudin de dialyse d'une porosité de 3 kDa. Le boudin est ensuite placé dans le bain de dialyse sous agitation à 4°C durant 18 à 20 heures.

V.B.4. Chromatographie par chélation de métaux immobilisés inverse

Après dialyse et clivage des étiquettes His-SUMO de la protéine d'intérêt, l'échantillon a de nouveau été chargé sur une colonne Nickel-Sépharose (cytiva). Cette colonne a préalablement été équilibrée avec 30 mM d'imidazole (6 % tampon B) comme précédemment. Une fois l'échantillon chargé, la colonne est à nouveau lavée avec 30 mM d'imidazole (6 % tampon B). Durant cette étape, la protéine d'intérêt est éluée tandis que les étiquettes sont retenues sur la colonne. Les étiquettes sont éluées par la suite avec une étape de lavage de la colonne avec 500 mM d'imidazole (100 % tampon B). Les étapes de cette purification sont décrites dans le Tableau 13.

Tableau 13. Programme utilisé pour la chromatographie par chélation de métaux immobilisés inverse.

Etape	Tampon A (%)	Tampon B (%)	Débit	Volume colonne
Equilibration	94	6	5 mL·min^{-1}	3
Injection	–	–	2 mL·min^{-1}	–
Lavage	94	6	5 mL·min^{-1}	6
Elution	0	100	5 mL·min^{-1}	3

La présence de protéines dans chaque fraction a été observée par leur absorbance à 280 nm. Les fractions obtenues durant l'étape de lavage et contenant la plus grande quantité de protéines ont été retenues et rassemblées pour la suite des étapes de purification.

V.B.5. Chromatographie d'exclusion de taille

L'échantillon obtenu après chromatographie IMAC inverse a été, au besoin, concentré afin d'obtenir un volume final compris entre 3 et 5 mL à l'aide d'un concentrateur Amicon (Merck Millipore) avec une porosité de 3 kDa. L'échantillon chargé sur le concentrateur a été centrifugé à 4000 *g* jusqu'à atteindre le volume attendu. Une colonne Superdex 200 26/60 prep grade (cytiva) a quant à elle été équilibrée avec le tampon phosphate (Tableau 11) avant injection de l'échantillon concentré. L'échantillon est chargé sur la colonne *via* une boucle de 10 mL. Une fois l'échantillon chargé, la colonne est lavée et les protéines éluées avec le même tampon phosphate. Les fractions contenant la protéine d'intérêt, identifiées par leur absorbance à 280 nm, ont ensuite été rassemblées et l'échantillon final a été concentré de la même manière que précédemment jusqu'à atteindre un volume final compris entre 1 et 2 mL. Une fois la purification terminée, la colonne a été lavée avec 130 mL d'eau à un débit de 5 mL·min^{-1}.

V.B.6. Dosage et stockage des protéines

La concentration en protéines des échantillons obtenus a été mesurée par leur absorbance à 280 nm et calculée à l'aide de la loi de Beer-Lambert $A = \varepsilon lc$, où A est l'absorbance de l'échantillon, ε le coefficient d'extinction molaire de la protéine d'intérêt, l la longueur du trajet optique parcouru (égal à 1 cm ici) et c la concentration de l'échantillon en protéine d'intérêt. Le coefficient ε a été calculé à l'aide de l'outil en ligne ProtParam (Expasy, Swiss Bioinformatics Resource Portal) (Gasteiger et al., 2005). Les échantillons sont ensuite aliquotés puis congelés dans l'azote liquide et placés à -80°C.

VI. Caractérisation des interactions entre les domaines de docking

VI.A. Analyse de l'intégrité des domaines purifiés

VI.A.1. Mesures de l'agrégation par diffusion dynamique de la lumière

Afin de s'assurer que les domaines ne s'agrègent pas, les échantillons ont été analysés par diffusion dynamique de la lumière (DLS). Cela permet d'estimer le rayon hydrodynamique des particules en solution en mesurant la variation d'intensité lumineuse qu'elles diffusent au cours du temps. Lorsqu'une particule est petite, cette variation sera plus rapide que lorsqu'une particule est plus grosse. Dans un premier temps, l'échantillon est centrifugé à 15000 *g* durant 15 minutes à 4°C afin d'éliminer les poussières et les agrégats protéiques. La concentration de l'échantillon a été ajustée aux alentours de 5 $mg \cdot L^{-1}$, puis placé dans la cuve. Les mesures ont été prises par un appareil Zetasizer Nano S (Malvern Panalytical), par une dizaine d'acquisition de 1 seconde à 25°C. Cette méthode permet d'obtenir notamment l'indice de polydispersité (PdI) de l'échantillon. Un échantillon est considéré monodisperse lorsque cette valeur est inférieure à 0,20. Il est primordial pour la suite des expérimentations que les protéines d'intérêts soient monodisperses.

VI.A.2. Analyse des séquences et poids moléculaires des domaines par spectrométrie de masse

Les échantillons ont été équilibrés avec 0,5 % d'acide trifluoroacétique (TFA) et 2 % d'acétonitrile (ACN) puis dilués si nécessaire à une concentration de 1 µM. Ensuite, 5 µL de l'échantillon ont été fractionnés par nano-HPLC couplé à un système Ultimate 3000 équipé d'une colonne de dessalage pepMap 100 C18 et d'une colonne de fractionnement PepMap RSLC C18 de 15 cm (ThermoFisher Scientific). L'échantillon a été élué avec un gradient de 2 % à 45 % d'ACN pendant 30 minutes à un débit de 300 $nL \cdot min^{-1}$. L'absorbance à 214 nm des échantillons élués permet de suivre la qualité de l'échantillon ainsi que sa concentration en protéines. Les fractions ont été collectées à l'aide d'un Proteineer FcII puis mélangées à de l'acide α-cyano-4-hydroxycinnamique sur un échantillon cible MTP-1536 TF. Les échantillons correspondant aux éluas contenant la protéine d'intérêt, selon l'absorbance à 214 nm, ont été étudiés par spectrométrie de masse par désorption-ionisation laser assistée par matrice – temps de vol (MALDI-TOF) et MALDI-TOF/TOF (mode LIFT) à l'aide d'un Autoflex IV (Bruker) utilisant un étalonnage externe. Les données de MS et MS/MS ont été utilisées pour l'identification des molécules par l'utilisation de la base de données protéique non-redondante du Centre National pour les Informations Biotechnologiques Etats-Unien (NCBI), assisté par un serveur Mascot, avec les paramètres suivants : une tolérance de 50 ppm en MS, une tolérance de 0,8 Da en MS/MS, pas de protéolyse et un score peptidique minimum supérieur à 20.

Ces analyses ont été menées par la plateforme de protéomique IBSLor, Université de Lorraine, F-54000 Nancy, France, que nous remercions pour ces précieuses analyses.

VI.A.3. Caractérisation des structures secondaires par dichroïsme circulaire

Afin de s'assurer que les domaines soient correctement structurés en solution, des analyses par dichroïsme circulaire (DC) ont été menées. Cette méthode permet l'étude des structures secondaires protéiques en fonction de leurs propriétés d'absorbance (Miles et al., 2021). Les protéines sont composées d'acides aminés chiraux, qui absorbent différemment la lumière selon que leur polarisation soit circulaire droite ou gauche.

Les spectres ont été enregistrés avec un appareil Chirascan CD (IBS-Lor UMS 2008 Plateforme de Biophysique et Biologie Structurale) dans l'UV lointain entre 190 et 280 nm. La vitesse de balayage était de 73 secondes·nm^{-1} à 20°C. Les échantillons ont été préparés à une concentration de 50 µM dans un tampon NaP 100 mM (pH 8,0). Ces analyses ont été faites dans deux cuves de quartz possédant un trajet optique de 0,1 cm et contenant 30 µL d'échantillon chacune. Afin d'analyser la structure secondaire des protéines d'intérêt seules, les deux cuves ont été remplies du même échantillon à une concentration protéique de 50 µM. Afin d'évaluer un potentiel changement de structure secondaire lorsque deux protéines sont ensemble, induit par une interaction entre les partenaires, deux mesures ont été prises. Pour la première, les protéines sont chacune placées dans une cuve à une concentration de 50 µM. Cela permet d'obtenir le spectre moyen correspondant au repliement des deux partenaires. La seconde mesure est faite en plaçant un mélange des deux protéines dans les deux cuves, à une concentration de 50 µM. La courbe obtenue correspond alors au repliement des protéines en présence d'un potentiel partenaire. La comparaison de ces deux courbes permet alors d'évaluer la présence ou non d'interaction entre les deux protéines.

Pour chaque analyse, trois spectres ont été enregistrés puis moyennés et ajustés en soustrayant le spectre du tampon seul. La courbe a ensuite été lissée à l'aide d'un filtre Savitzky-Golay, avec une taille de fenêtre comprise entre 1 et 3 points (Savitzky & Golay, 1964). Enfin, les données brutes obtenues en milli-degrés ont été transformées en valeurs d'ellipticité moyenne par résidu (EMR) (deg·cm^2·$dmol^{-1}$) à l'aide du logiciel Pro-Data Viewer (Applied Photophysics Ltd, Angleterre).

VI.B. Etude des interactions entre les domaines de docking par titrage calorimétrique isotherme

Les interactions entre les domaines de docking ont été analysées par titrage calorimétrique isotherme (ITC). Cette méthode repose sur l'analyse du dégagement de chaleur induit par l'interaction entre deux partenaires. Elle permet notamment de déterminer la stœchiométrie et la constante de dissociation (K_D) de ces interactions.

Le partenaire titré a été placé dans la cuve à une concentration de 80 µM et le partenaire titrant à une concentration de 800 µM dans la seringue. Une première injection de 0,5 µL de la solution titrante dans la cuve a été suivie de dix-neuf injections de 2 µL, sous agitation à 750 rpm à une température de 25°C. La tangente de la courbe obtenue a permis de calculer la constante de dissociation (K_D) de l'interaction et le point d'inflexion en a indiqué la stœchiométrie.

VII. Localisation de la coelimycine synthase chez *S. coelicolor* par microscopie confocale

VII.A. Préparation des échantillons

Afin de visualiser les protéines fluorescentes exprimées par *S. coelicolor* et fusionnées aux enzymes du complexe Cpk, les souches mutées ont été cultivées sur un milieu DNA. Afin de récupérer le mycélium aérien pour les observations, deux techniques ont été utilisées. La première consiste en la culture des bactéries sur une membrane de cellophane déposée préalablement sur le milieu. Ainsi, il est possible de récupérer uniquement le mycélium aérien présent sur cette membrane. La seconde méthode consiste à placer dans la gélose, après inoculation, des lamelles de verre afin que le mycélium croisse dessus et de pouvoir ainsi le récupérer et l'observer au microscope. Une quantité de 10^7 spores a été inoculée dans chaque boîte, et plusieurs temps de culture à 30°C ont été testés : 24, 48 et 72 heures.

VII.B. Acquisition des images par microscopie

Les protéines fluorescentes mYPet et mTurquoise2 ont des spectres d'excitation qui ont la particularité de ne pas se chevaucher (Figure 66). Cela permet leur utilisation conjointe au sein des souches bactériennes. Pour la détection de mTurquoise2, la longueur d'onde d'excitation utilisée a été de 405 nm et la gamme d'émission entre 470 et 525 nm (Figure 66A). Pour l'analyse de mYPet, l'excitation s'est faite à 517 nm et l'émission a été mesurée entre 525 et 600 nm (Figure 66B).

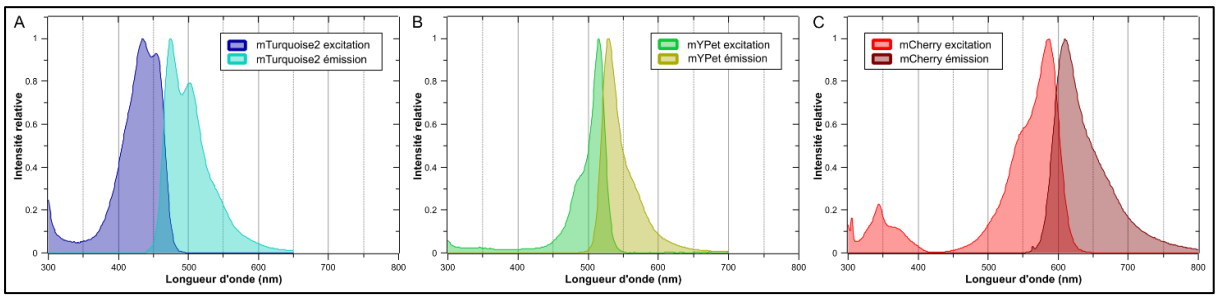


Figure 66. Spectres d'excitation et d'émission des protéines fluorescentes utilisées lors de cette étude. Spectres relatifs à la protéine fluorescente turquoise monomérique (mTurquoise2) (Goedhart et al., 2012). **(B)** Spectres relatifs à la protéine fluorescente jaune monomérique (mYPet) (Nguyen & Daugherty, 2005). **(C)** Spectres relatifs à la protéine fluorescente Cherry monomérique (mCherry) (Shaner et al., 2004).

Enfin, l'analyse de mCherry a été faite avec une excitation à 580 nm et l'émission a été recueillie entre 600 et 750 nm (Figure 66C). Les observations ont été faites à l'aide d'un microscope confocal Leica TCS SP8, utilisant des détecteurs hybrides.

VII.C. Traitement des clichés obtenus par microscopie confocale

Les images obtenues par microscopie ont été analysées et traitées avec le logiciel Fiji (Fiji is just ImageJ). Certains clichés – précisé le cas échéant – ont également été déconvolus. Cette déconvolution a été faite avec le logiciel Huygens ™ Deconvolution par la Dr. Marta Alonso Guervós (Université d'Oviedo, Espagne).

Résultats

I. Effet de la scission de polycétide synthases : focus sur la rapamycine synthase et la stambomycine synthase

Les polycétides sont des métabolites spécialisés dont les propriétés biologiques présentent un intérêt majeur en médecine et agro-alimentaire notamment. L'ingénierie des enzymes responsables de leur production – les PKS – est une voie prometteuse pour la synthèse de nouvelles molécules. Cependant, à ce jour les tentatives d'ingénierie se traduisent par une baisse des rendements de production des systèmes modifiés (Ranganathan et al., 1999; Su et al., 2022; Wlodek et al., 2017). Certaines études ont montré que la taille des sous-unités enzymatiques des PKS peut jouer un rôle dans leur productivité. Ainsi réduire leur taille en introduisant des domaines d'interactions peut permettre d'améliorer cette productivité (Buyachuihan et al., 2023; Liu et al., 2025; Miyazawa et al., 2021). Le premier objectif de cette thèse a ainsi été de s'intéresser à deux systèmes d'étude : la rapamycine synthase (RAPS) et la stambomycines synthase. L'objectif est de scinder une sous-unité de ces systèmes afin d'observer le résultat sur les rendements de production de rapamycine et de stambomycines, respectivement.

I.A. Etude de la rapamycine synthase : constructions plasmidiques et caractérisation biophysique des domaines de docking

Le premier objectif de ces travaux de thèse a été de scinder la plus grande sous-unité de RAPS, RAPSB, composée de six modules. La scission a été effectuée entre les troisième et quatrième modules de la sous-unité, afin d'obtenir deux chaînes polypeptidiques composées chacune de trois modules.

I.A.1. Choix des domaines de docking et du point de fusion

L'analyse de RAPS a montré que l'interface entre RAPSA et RAPSB est composée de DD de type 1b, tandis que celle entre RAPSB et RAPSC est constituée de DD de type 1a (Figure 67). La distinction entre ces classes de DD se fait notamment par l'analyse des résidus impliqués dans l'interaction entre les partenaires, indiqués par des étoiles jaunes dans la Figure 67. Ces DD seront après nommés respectivement $^{C}DD_{RAPSA}$ / $^{N}DD_{RAPSB}$ et $^{C}DD_{RAPSB}$ / $^{N}DD_{RAPSC}$. Les DD choisis afin de scinder RAPSB proviennent du complexe enzymatique de la 6-désoxyérythronolide B synthase (DEBS), retrouvée chez *Saccharopolyspora erythraea* (Cortes et al., 1990). DEBS est composée de trois sous-unités, contenant chacune de deux modules d'élongation, qui interagissent par le biais de DD de type 1a.

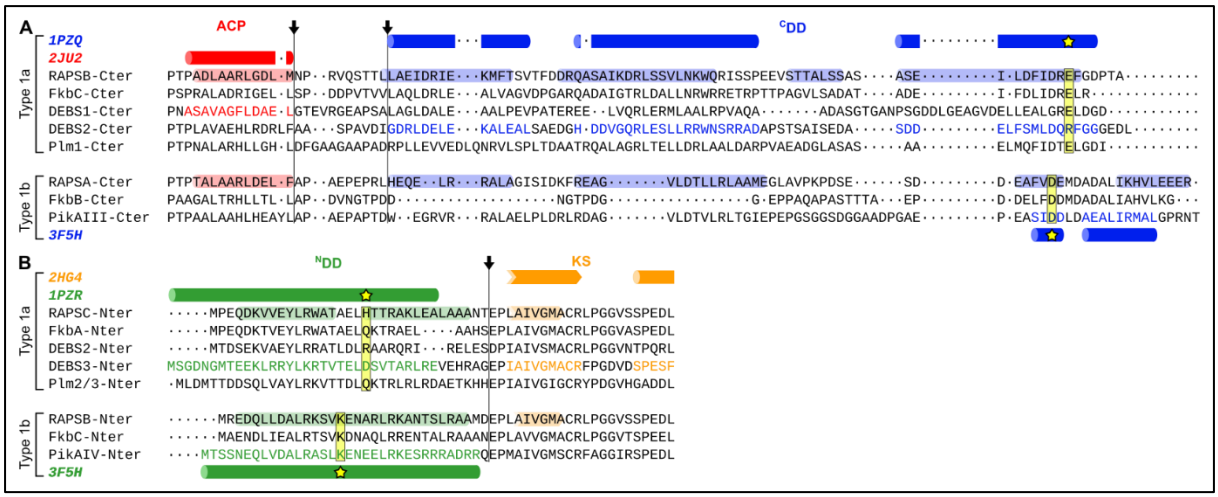


Figure 67. Alignement de séquences et structures secondaires des domaines de docking d'intérêt, de domaines porteurs d'acyl et de domaines cétosynthase. Les éléments de structure secondaire sont représentés par des cylindres (hélices α) et des flèches (brins β) de couleur. Les éléments pleins et les résidus associés indiqués avec la même couleur représentent les éléments structurels résolus, tandis que les éléments retrouvés en clair au sein des séquences protéiques proviennent de prédictions faites à l'aide de l'outil en ligne PsiPred. Les résidus impliqués dans les interactions de charge spécifiques de chaque paire de DD sont marqués par une étoile jaune et surlignés en jaune. **(A)** Alignement de séquences et structures secondaires des régions C-terminales de sous-unités de polycétide synthases. Les structures résolues du domaine porteur d'acyle (ACP) du deuxième module de la 6-désoxyérythronolide B synthase (DEBS) (identifiant PDB : 2JU2) et du domaine de docking (DD) présent en position C-terminale de la sous-unité DEBS2 (identifiant PDB : 1PZQ) sont représentés respectivement en rouge et bleu. La fin du domaine ACP et le début du domaine ^{C}DD sont représentés par des flèches noires. **(B)** Alignement de séquences et structures secondaires des régions N-terminales de polycétide synthases. Les structures résolues du DD présent en position N-terminale de la sous-unité DEBS3 (identifiant PDB : 1PZR) est représenté en vert et en orange est représenté le domaine cétosynthase (KS) du module 5 de DEBS (identifiant PDB : 2HG4). La transition entre les domaines ^{N}DD et KS est représenté par une flèche noire. **Légende** – RAPS : rapamycine synthase, Fkb : tacrolimus synthase, Plm : phoslactomycine synthase, Pik : pikromycine synthase.

Le domaine $^{C}DD_{RAPSB}$ contient un acide glutamique (E) identifié comme impliqué dans les interactions de charge par alignement de séquence, interagissant avec une histidine (H) du domaine $^{N}DD_{RAPSC}$. Les DD de DEBS quant à eux sont tous de type 1a. Le profil d'interaction du couple $^{C}DD_{DEBS1}$ / $^{N}DD_{DEBS2}$ est composé des résidus glutamate (E) et arginine (R) respectivement, ce qui en fait un couple similaire à $^{C}DD_{RAPSB}$ / $^{N}DD_{RAPSC}$. La seconde paire de DD de DEBS contient une arginine (R) au sein du domaine $^{C}DD_{DEBS2}$ et un acide aspartique (D) au niveau du domaine $^{N}DD_{DEBS3}$. C'est donc ce couple de DD qui a été sélectionné afin de scinder la sous-unité RAPSB, car son profil d'interaction est différent de celui des DD natifs de RAPS de même type (Figure 68). Cette première analyse *in silico* a par la suite été complétée par une caractérisation biophysique des interactions entre les partenaires (voir ci-dessous).

Une étude récente, publiée peu après l'analyse *in silico* des interactions entre les domaines effectuée au cours de cette thèse, s'est intéressée aux interactions possibles entre les DD de DEBS et ceux de RAPS.

L'objectif de cette étude était de favoriser le contact entre deux partenaires enzymatiques impliqués dans une même voie métabolique, afin d'en améliorer l'efficacité (Sun et al., 2022). Dans ce cadre, les auteurs ont étudié les interactions potentielles entre certains DD de ces deux systèmes. Ils ont ainsi pu montrer notamment l'absence d'interactions entre $^{C}DD_{RAPSA}$ et $^{N}DD_{DEBS2}$, ce qui renforce le choix qui a été fait.

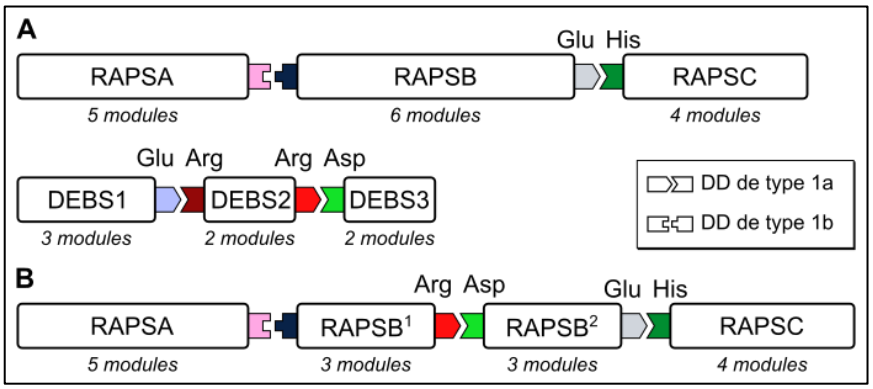


Figure 68. Analyse des profils d'interaction de charge des domaines de docking de type 1a chez DEBS et RAPS ainsi que du système RAPS scindé (Weissman, non publié). **(A)** Analyse des interactions de charge entre les domaines de docking des systèmes de la rapamycine synthase (RAPS) et de la 6-désoxyérythronolide B synthase (DEBS). RAPS est dotée d'une paire de domaines de docking (DD) de type 1a. Les interactions de charge sont médiées par un acide glutamique au sein du domaine C-terminal et d'une histidine en N-terminal. La PKS DEBS quant à elle contient deux paires de DD de type 1a. Le premier a un profil d'interaction similaire à la paire de RAPS, avec un acide glutamique en position C-terminale et une arginine en N-terminal. La seconde paire montre un profil inversé, avec une arginine au sein du domaine en C-terminal et un acide aspartique en N-terminal. **(B)** Représentation de la PKS RAPS scindée avec la paire de DD $^{C}DD_{DEBS2}$ / $^{N}DD_{DEBS3}$. Le profil d'interaction de la paire de DD de type 1a de RAPS est incompatible avec celui des DD introduits.

L'objectif a donc été d'introduire la paire $^{C}DD_{DEBS2}$ / $^{N}DD_{DEBS3}$ entre les modules 7 et 8 de la plus grande sous-unité de RAPS : RASPB (6 modules). Afin de déterminer le point de jonction entre le domaine ACP_7 du module 7 de RAPSB et le domaine $^{C}DD_{DEBS2}$, des alignements de séquences des ACP de DEBS et de RAPS ont été effectués (Figure 69). Les informations issues des structures du domaine ACP_2 de DEBS (Alekseyev et al., 2007), ainsi que celles des domaines $^{C}DD_{DEBS2}$ et $^{N}DD_{DEBS3}$ (Broadhurst et al., 2003) ont été utilisées afin de guider cette analyse.

Cette analyse a permis d'établir les bornes de chaque domaine et de déterminer les points de jonction possibles pour l'insertion des DD hétérologues au sein de RAPS. Deux stratégies ont été envisagées à partir de ces alignements, selon la séquence linker gardée entre le domaine ACP_7 et $^{C}DD_{DEBS2}$. La première consiste à intégrer les DD dès la fin de l'ACP_7 en gardant la séquence linker présente chez DEBS : « AASPAVDI ». La phénylalanine située nativement en première position du linker, en amont de la première alanine, n'est ici pas conservée pour éviter la présence de deux phénylalanines – qui sont des acides aminés ayant une chaîne latérale encombrante – lors de la fusion avec la fin du domaine $^{RAPSB}ACP_7$.

La seconde stratégie consiste en la reconstitution d'un linker hybride entre les résidus qui suivent le domaine ACP_7 et ceux du linker présent chez DEBS : « VGETP / AVDI ».

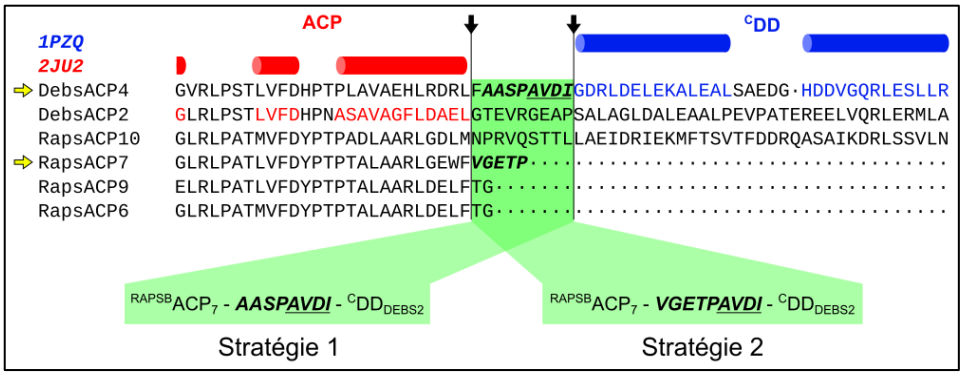


Figure 69. Alignement de séquences et d'éléments de structure secondaire de domaines porteurs d'acyle et de domaine de docking de polycétide synthase. Les structures résolues des domaines $^{DEBS1}ACP_2$ (identifiant PDB : 2JU2) en rouge (Alekseyev et al., 2007) et $^{C}DD_{DEBS2}$ (identifiant PDB : 1PZQ) en bleu (Broadhurst et al., 2003) ont été utilisées afin de définir les bornes de chaque domaine. Les cylindres représentent des hélices α. La fin du domaine $^{DEBS1}ACP_2$ et le début du domaine $^{C}DD_{DEBS2}$ sont indiqués par des flèches noires et permettent d'identifier la région linker entre les domaines ACP et ^{C}DD surlignée en vert. Les résidus indiqués en gras sont les séquences linker utilisées pour la scission de RAPSB, et la séquence soulignée est celle qui est retrouvée dans les deux stratégies. Les séquences impliquées dans la fusion du domaine $^{RAPSB}ACP_7$ et $^{C}DD_{DEBS2}$ sont indiquées par des flèches jaunes. La stratégie 1 consiste en l'utilisation du linker natif présent entre les domaines $^{DEBS2}ACP_4$ et $^{C}DD_{DEBS2}$, tandis que la stratégie 2 consiste en la construction d'un linker hybride entre les résidus retrouvés à la suite du domaine $^{RAPSB}ACP_7$ et ceux retrouvés en amont du domaine $^{C}DD_{DEBS2}$.

I.A.2. Préparation du matériel génétique pour l'ingénierie génétique de Streptomyces hygroscopicus

Une fois le point de scission établi, il a fallu concevoir et créer les vecteurs plasmidiques permettant l'édition du génome de *S. hygroscopicus*. Pour cela, le vecteur pKCCpf1, optimisé pour l'expression du système CRISPR-Cpf1 chez *Streptomyces* spp., a été utilisé (L. Li et al., 2018). Dans un premier temps, la séquence du protospacer ciblé au sein du génome a été identifiée, puis la matrice contenant les éléments nécessaires à la modification à apporter a été construite puis clonée au sein du plasmide pKCCpf1.

I.A.2.a. Identification du protospacer

Afin d'identifier le protospacer à cibler au sein du génome de *S. hygroscopicus*, la première étape a été d'investiguer des outils en ligne tels que CRISPy-Web et CHOPCHOP (Blin et al., 2016; Labun et al., 2019). Cependant, le génome de cette espèce n'est pas enregistré au sein de leur base de données. Un script python a été écrit afin de rechercher des séquences protospacer dans le cas de l'ingénierie par CRISPR-Cas9 ou CRISPR-Cpf1, dans un génome donné. Les étapes suivies par ce programme sont décrites dans la Figure 70.

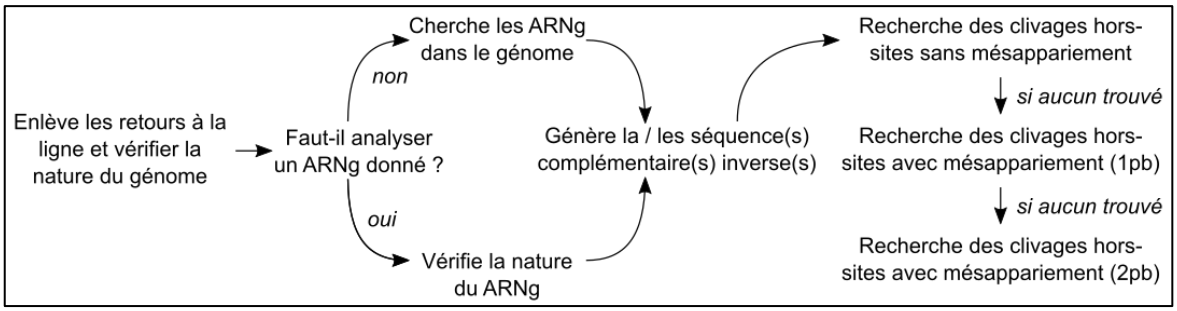


Figure 70. Représentation schématique du programme développé pour trouver des ARN guides potentiels pour l'utilisation de CRISPR-Cpf1 ou CRISPR-Cas9 dans un génome donné. La première étape du programme est de créer un fichier texte contenant le génome sans retour à la ligne et sans caractère ne correspondant pas à des nucléotides. Une fois cette étape franchie, il est possible de rechercher des sites d'hybridation de l'ARN guide (ARNg) hors-cible dans le génome. Il est également possible de rechercher des ARNg dans une portion du génome, en indiquant les positions entre lesquelles les rechercher. Ensuite, la ou les séquence(s) complémentaire(s) inverse(s) sont générées afin d'étudier par la suite la possibilité pour l'ARNg de reconnaître un autre site dans le génome, sans mésappariement puis avec 1 et 2 bases non appariées.

Ce programme a été validé par son utilisation sur le génome de *S. ambofaciens* en parallèle de l'outil en ligne CHOPCHOP (Labun et al., 2019). Il a permis d'identifier plusieurs séquences cibles dans la région de la fin du domaine ACP_7 ± 100 bases (Tableau 14). La séquence protospacer sélectionnée ici a été la suivante : « 5′-GCAGCCCGGCTGGGGGAGTTGTTCG-3′ ». Ce protospacer a été choisi car sa séquence n'est retrouvée nulle part ailleurs dans le génome, même à 1 ou 2 bases près. De plus, il est situé au plus proche de la fin du domaine ACP_7.

Tableau 14. Séquences des protospacers potentiels identifiés dans le génome de *S. hygroscopicus* pour la scission de la rapamycine synthase.

Séquence	PAM	Distance avec ACP_7	Hybridation hors cible		
			0 pb	*1 pb*	*2 pb*
CCCGCCACGCTGGTGTTCGACTACC	TTG	41 pb	0	1	1
GACTACCCCACCCCGACGGCCTTGG	TTC	23 pb	0	1	1
ACCGACGAACAACTCCCCCAGCCGG	TTC	19 pb	0	1	0
GCAGCCCGGCTGGGGGAGTTGTTCG	TTG	-1 pb	0	0	0
TTCGTCGGTGAAACTCCTGTTCCGG	TTG	-22 pb	0	0	0
GTCGGTGAAACTCCTGTTCCGGTAC	TTC	-25 pb	0	0	0
CGGTACGTACGTCGGTGTCGGTGGT	TTC	-44 pb	0	0	0
GCGATTGTCGGCATGGCGTGTCGCT	TTG	-88 pb	0	0	0
TCGGCATGGCGTGTCGCTTGCCGGG	TTG	-95 pb	0	0	0

La distance avec la fin du domaine ACP_7 est calculée à partir de la position du dernier nucléotide de la séquence du protospacer. La séquence sélectionnée est surlignée en jaune.

I.A.2.b. Conception et assemblage des matrices d'édition

Une fois le protospacer identifié, les autres éléments constituant les matrices d'édition ont pu être déterminés. Tout d'abord, les régions d'homologie, qui sont composées des environ 1000 paires de bases en amont du point de coupure induit par le système CRISPR-Cpf1 dans le cas de la région RH1 et en aval de ce point pour la région RH2. Le système CRISPR-Cpf1 clive la sous-unité RAPSB en amont de la fin du domaine ACP_7.

Il a donc fallu reconstruire la région de la séquence codant le domaine ACP_7 et retrouvée en aval du site de coupure, et intégrer une mutation dans la région PAM afin d'éviter que le vecteur ne puisse être la cible du clivage. Ces deux points ont été réglés par l'utilisation d'un gène synthétique (GS), qui se trouve en position 3′ de la région RH1. Entre cette région GS et la région RH2 se trouvent les séquences codant les DD à intégrer au sein du génome de *S. hygroscopicus*. La matrice d'édition contient alors dans l'ordre les séquences suivantes : RH1-GS-DD-RH2 (Figure 71). La matrice contient également une partie du plasmide pKCCpf1 puisqu'entre les sites de restriction *Spe*I et *Nde*I se trouvent la séquence codant l'ARN transactivateur, le promoteur kasOp* permettant l'expression de l'ARNcr et le promoteur ermEp* contrôlant l'expression du gène de l'endonucléase Cpf1.

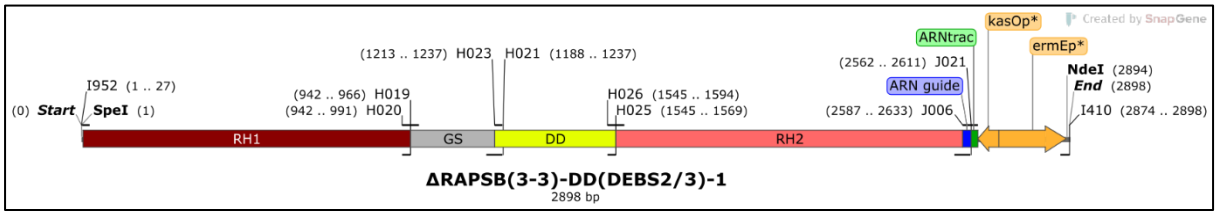


Figure 71. Carte de la matrice d'édition pour l'insertion des DD de DEBS au sein de la PKS RAPS chez *S. hygroscopicus* pour la stratégie 1 par CRISPR-Cpf1. Cette matrice d'édition contient les régions d'homologies (RH) permettant la réparation de l'ADN à la suite du clivage par CRISPR-Cpf1, le gène synthétique (GS) qui correspond à la région du protospacer mutée pour éviter des événements multiples de clivage, la région codant les domaines de docking (DD) de la 6-désoxyérythronolide B synthase (DEBS) à intégrer au génome et la séquence de l'ARN guide. Des éléments du plasmide pKCCpf1 sont également présents afin de ne pas les éliminer lors du clonage entre les sites *Spe*I et *Nde*I. Ces éléments sont la séquence codant l'ARN transactivateur (ARNtracr), le promoteur kasOp* qui permet la synthèse de l'ARN guide et de l'ARNtracr et le promoteur de la région codant l'endonucléase Cpf1, ermEp*. Les amorces utilisées pour l'amplification et l'assemblage de la matrice sont représentées par leur nomenclature ESCOLA propre au laboratoire (voir Tableau 4).

Chaque élément de cette matrice a été amplifié puis assemblé par PCR à l'aide d'amorces chevauchantes. Les profils de migration de chaque élément et étape d'assemblage sur gel d'agarose sont représentés dans la Figure 72.

I.A.2.c. Clonage des matrices d'édition

Une fois obtenues, les matrices d'édition ont été clonées dans le vecteur pKCCpf1 permettant l'expression du système d'édition génétique CRISPR-Cpf1 chez *Streptomyces* spp (L. Li et al., 2018). Pour se faire, les sites de restrictions reconnus par les enzymes *Nde*I et *Spe*I ont été utilisés. Le vecteur a préalablement été linéarisé puis les matrices ont été clonées par la méthode d'assemblage de type Gibson (Figure 65, page 104). La présence des matrices au sein du vecteur a été vérifiée par digestion des constructions avec les mêmes enzymes ayant permis le clonage, puis leur intégrité a été validée par séquençage.

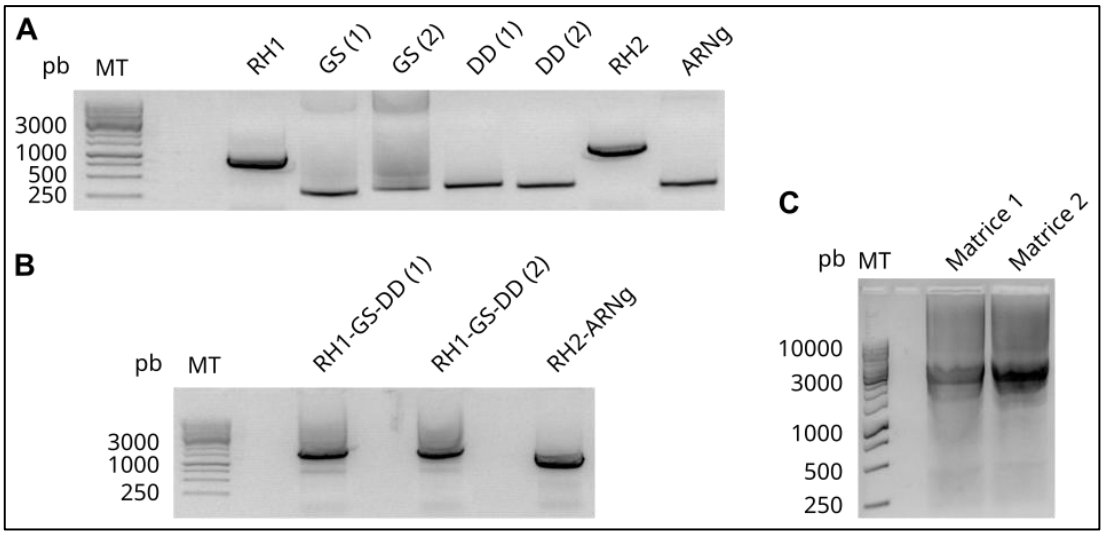


Figure 72. Profil de migration sur gel d'agarose (0,9 %) des étapes d'amplification et assemblage des éléments des matrices d'édition pour la scission de la rapamycine synthase. Les éléments spécifiques à chaque stratégie sont indiqués par un (1) pour la première stratégie et un (2) pour la seconde. **(A)** Amplification des éléments constitutifs des matrices. Les régions d'homologie RH1 et RH2 ont des tailles respectives d'environ 960 et 1015 pb. Les gènes synthétiques ont une taille aux alentours de 245 pb pour la première stratégie – GS(1) – et d'environ 265 pb dans le cas de la seconde stratégie : GS(2). Les séquences codant les domaines de docking (DD) provenant de la 6-désoxyérythronolide B synthase (DEBS) à intégrer au génome de *S. hygroscopicus* ont une taille d'environ 360 pb pour la première stratégie et 340 pb pour la seconde. La différence de taille de ces fragments s'explique par la différence du nombre de résidus gardés pour la constitution du linker dans les deux stratégies. Enfin, la séquence contenant l'ARN guide (ARNg) est d'environ 320 pb. **(B)** Sous-assemblage des éléments constitutifs des matrices d'édition. Les régions RH1, GS et DD ont été assemblées d'une part et les fragments RH2 et ARNg d'une autre. La taille des fragments obtenus pour le sous-assemblage RH1-GS-DD est d'environ 1565 pb pour les deux stratégies, et celle du fragment RH2-ARNg est d'environ 1335 pb. **(C)** Assemblage des matrices d'édition. Les sous-fragments obtenus en (B) ont été assemblés par PCR. La taille des fragments attendus est aux alentours de 2900 pb. **Légende** – pb : paires de bases, MT : marqueur de taille.

I.A.3. Etude des interactions in vitro entre les domaines de docking

Afin de s'assurer de l'absence d'interactions entre les domaines de docking hétérologues de DEBS et natifs de RAPS, chaque DD doit être obtenu et les interactions entre chaque couple caractérisées par titrage calorimétrique isotherme (ITC).

I.A.3.a. Clonage des séquences codant les domaines de docking d'intérêts

Les bornes des séquences des DD d'intérêts ont été identifiées par alignement de séquences, l'utilisation de structures de domaines ACP, ^{C}DD, ^{N}DD et KS résolues et la prédiction des structures secondaires des protéines d'intérêts n'ayant pas été résolues (Figure 67, page 120). Les séquences codant les ^{C}DD de DEBS2, RAPSA et RAPSB ont clonées à partir de la fin du domaine ACP identifié en amont, et comprennent donc les régions linker (Figure 67A, page 120). Les domaines ^{N}DD de DEBS3, RAPSB et RAPSC quant à eux ont été clonés à partir de la méthionine du début de la sous-unité, jusqu'au début du domaine KS identifié (Figure 67B, page 120).

Les séquences codant les DD étudiés lors de cette thèse ont été amplifiés à partir d'ADN génomique de *Sac. erythraea* pour ce qui est du couple $^{C}DD_{DES2}$ / $^{N}DD_{DEBS3}$ et de *S. hygroscopicus* pour les domaines $^{C}DD_{RAPSA}$ et $^{C}DD_{RAPSB}$ ou reconstruites par amorces chevauchantes dans le cas des domaines $^{N}DD_{RAPSB}$ et $^{N}DD_{RAPSC}$. Lorsque les séquences natives ne contiennent pas de résidus absorbant à 260 nm permettant de suivre la présence de protéines dans les échantillons, une tyrosine a été ajoutée en position N-terminale du domaine. Ces séquences ont ensuite été clonées au sein du vecteur pBG102 à l'aide des enzymes de restrictions *Bam*HI et *Hind*III.

I.A.3.b. Production et purification des domaines de docking

Les domaines dont la séquence a été clonée dans le vecteur pBG102 ont été produits chez *E. coli* BL21(DE3). Les protéines recombinantes obtenues après production et avant purification sont donc composées d'une étiquette 6xHis en position N-terminale, suivi d'une étiquette SUMO, du site de reconnaissance pour la protéase HRV 3C puis du DD d'intérêt.

L'étiquette SUMO permet de favoriser la solubilité de la protéine recombinante, et l'étiquette 6xHis est utilisée pour la première étape de purification de la protéine de fusion par IMAC. Ensuite, une protéolyse par la protéase HRV 3C couplée à une dialyse permet de cliver les étiquettes de la protéine d'intérêt et d'éliminer l'imidazole présent dans le tampon d'élution. Une seconde étape IMAC a permis de séparer l'étiquette 6xHis, fusionnée à l'étiquette SUMO, de la protéine d'intérêt, et enfin une dernière étape de chromatographie par exclusion de taille a permis d'obtenir une protéine pure et utilisable pour les prochaines étapes.

Chaque étape de purification est suivie par la mesure de l'absorbance à 280 nm de l'échantillon à la sortie de la colonne. Ainsi, les chromatographes obtenus permettent d'obtenir des informations quant à la qualité de l'échantillon purifié. Dans le cas de la purification du domaine $^{C}DD_{DEBS2}$, chaque chromatographe montre des pics uniques correspondant à la protéine d'intérêt purifiée (Figure 73A–C).

A chaque étape de la purification du domaine, des fractions aliquotes ont été prélevées afin d'observer la pureté des échantillons par une électrophorèse SDS-PAGE. Dans le cas de la purification du domaine $^{C}DD_{DEBS2}$, la taille de la protéine tagguée est de 21,0 kDa et une fois clivée, les étiquettes et le domaine $^{C}DD_{DEBS2}$ ont une taille respective de 12,1 kDa et 8,9 kDa. Ce gel montre que la protéine d'intérêt est bien soluble dans les conditions utilisées puisqu'une bande est observée entre 20 et 25 kDa dans le surnageant de sonication (Figure 73D). La première étape de purification par IMAC a permis d'obtenir un échantillon plus concentré en éliminant les autres protéines de l'échantillon, mais quelques bandes parasites sont encore présentes. La protéolyse a permis de cliver les étiquettes de la protéine d'intérêt, puisque deux bandes sont observées à des tailles attendues à la suite de cette étape.

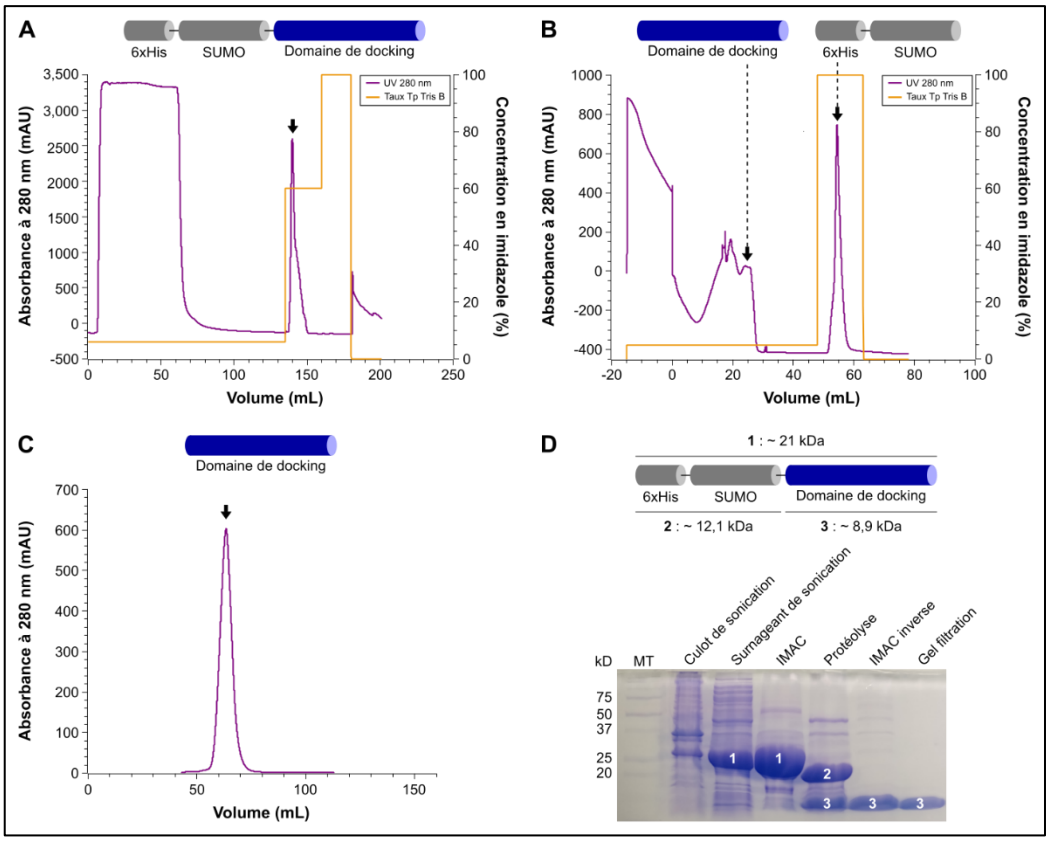


Figure 73. Chromatographes et gel SDS-PAGE des différentes étapes de purification du domaine de docking présent en position C-terminale de la sous-unité 2 de la 6-désoxyérythronolide B synthase. **(A–C)** Chromatographes obtenus lors de la purification de la protéine recombinante par chélation de métaux immobilisés (IMAC) (A), puis de séparation par IMAC inverse de la protéine d'intérêt des étiquettes protéolysées (B) et de la chromatographie d'exclusion de taille (C). La présence de protéines dans les fractions éluées à chaque étape a été suivie par l'absorbance à 280 nm et représentée en violet. Dans le cas des chromatographies IMAC, la concentration en imidazole en pourcentage du tampon B est indiquée en orange. **(D)** Migration sur gel de polyacrylamide (12,5 %) en condition dénaturante des fractions aliquotes récoltées lors des étapes de purification du domaine. La protéine recombinante (1) est bien retrouvée dans le surnageant de sonication et donc soluble dans le tampon utilisé par la suite pour la première étape de purification par IMAC. Après protéolyse, deux bandes aux tailles attendues d'environ 12,1 kDa pour les étiquettes (2) et 8,9 kDa pour le domaine $^{C}DD_{DEBS2}$ (3) sont observées, indiquant que la protéolyse a bien été effectuée. L'étape d'IMAC inverse a bien permis d'éliminer les étiquettes et de purifier le domaine seul, et l'étape de gel filtration a permis d'obtenir un échantillon plus pur en protéine d'intérêt.

Ensuite, l'étape d'IMAC inverse permettant d'éliminer le tag de l'échantillon a également été un succès puisqu'une bande unique, correspondant au domaine $^{C}DD_{DEBS2,}$ est observée. Les contaminants observés à la suite de la première étape IMAC sont toujours présents, l'échantillon n'est donc pas tout à fait pur. La dernière étape de chromatographie par exclusion de taille, quant à elle, a permis d'éliminer ces contaminants : une seule bande correspondant à $^{C}DD_{DEBS2}$ est observée. L'échantillon est ensuite concentré par centrifugation jusqu'à obtenir un volume entre 1 mL et 500 µL. La concentration finale en protéines purifiées a été dans ce cas de 6,2 mg par litre de production. Une fois les protéines d'intérêts produites et purifiées, les séquences des DD obtenues contiennent une séquence résiduelle « GPGS » en position N-terminale liée au clivage par la protéase HRV 3C (Tableau 15).

Tableau 15. Séquences protéiques des domaines de docking produits à partir des constructions plasmidiques dérivées du vecteur pBG102.

Domaine	Séquence	ε	MM
$^{C}DD_{DEBS2}$	**GPGS**FAASPAVDIGDRLDELEKALEALSAEDGHDDVGQRLESLLR RWNSRRADAPSTSAISEDASDDELFSMLDQRFGGGEDL	5500	8906,55
$^{N}DD_{DEBS3}$	**GPGS**MSGDNGMTEEKLRRYLKRTVTELDSVTARLREVEHRAGE	1490	4833,39
$^{C}DD_{RAPSA}$	**GPGS**Y**Y**APAEPEPRLHEQELRRALAGISIDKFREAGVLDTLLRLAAM EGLAVPKPDSESDDEAFVDEMDADALIKHVLEEER	1490	8898,93
$^{N}DD_{RAPSB}$	**GPGS**Y**Y**MREDQLLDALRKSVKENARLRKANTSLRAAMD	1490	4162,76
$^{C}DD_{RAPSB}$	**GPGS**MNPRVQSTTLLAEIDRIEKMFTSVTFDDRQASAIKDRLSSVL NKWQRISSPEEVSTTALSSASASEILDFIDREFGDPTA	5500	9221,25
$^{N}DD_{RAPSC}$	**GPGS**MPEQDKVVEYLRWATAELHTTRAKLEALAAANT	1490	4026,54

La séquence « GPGS » indiquée en gras est un artefact de clonage du vecteur pBG102 et la tyrosine en rouge a été ajoutée pour suivre la concentration en protéines des échantillons par leur absorbance à 280 nm. Les coefficients d'extinction molaire (ε) et la masse moléculaire (MM) en Dalton ont été calculées à l'aide de l'outil en ligne ProtParam d'Expasy.

Les domaines $^{C}DD_{DEBS2}$, $^{N}DD_{DEBS3}$, $^{C}DD_{RAPSA}$, $^{N}DD_{RAPSB}$ et $^{C}DD_{RAPSB}$ ont pu être obtenus. Le domaine $^{N}DD_{RAPSC}$ n'a pas pu être produit ni purifié dans ces conditions, un peptide synthétique a donc été commandé. Cependant, ce peptide n'est pas soluble dans le tampon utilisé (Tableau 11, page 108). A ce jour, ce problème de solubilité n'a pas pu être résolu.

I.A.3.c. Analyse de l'intégrité des domaines purifiés

L'intégrité des DD produits et purifiés a été validée par diffusion dynamique de la lumière (DLS), dichroïsme circulaire (DC) et spectrométrie de masse (MS). Le DLS a permis d'évaluer l'état d'oligomérisation du domaine et la pureté de l'échantillon, le DC a donné des informations sur la structure secondaire des domaines et la spectrométrie de masse a permis d'obtenir la masse et de séquencer chaque protéine d'intérêt.

Dans le cas du domaine $^{C}DD_{DEBS2}$, l'analyse par DLS a permis d'obtenir un pic correspondant à une protéine d'un diamètre de 5,935 nm (Figure 74A). L'indice de polydispersité (PdI) de l'échantillon est de 0,202. Cette valeur est légèrement plus élevée que la valeur maximale acceptée pour un échantillon pur – qui est de 0,200 – ce qui peut être expliqué par la présence d'une bosse entre 3000 et 6000 d. nm (diamètre en nanomètre). Le profil de ce pic résiduel est caractéristique de la présence de poussière intégrée dans l'échantillon au moment de sa préparation pour la mesure. Ainsi, cette impureté est ignorée et la protéine d'intérêt est considérée comme pure.

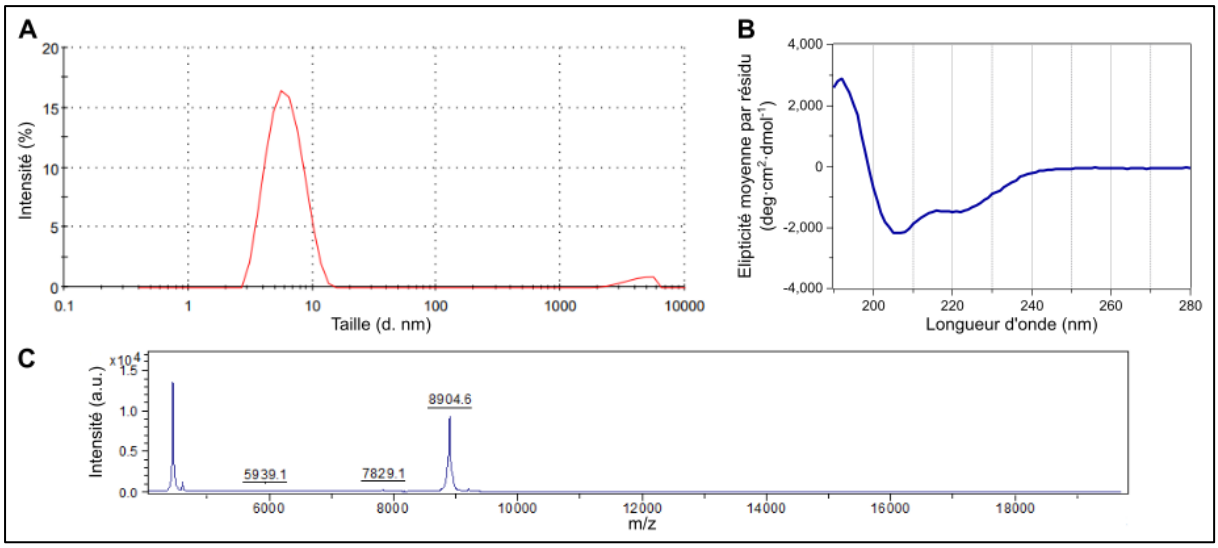


Figure 74. Analyse de l'intégrité du domaine de docking présent en position C-terminale de la sous-unité 2 de la 6-désoxyérythronolide B synthase. Le domaine $^{C}DD_{DEBS2}$ a été produit chez *E. coli* puis purifié en quatre étapes telles que présentées plus haut. **(A)** Chromatographe de l'analyse du domaine $^{C}DD_{DEBS2}$ par diffusion dynamique de la lumière. **(B)** Ellipticité moyenne par résidu du domaine $^{C}DD_{DEBS2}$ mesurée par dichroïsme circulaire. **(C)** Chromatographe de l'analyse du domaine $^{C}DD_{DEBS2}$ par spectrométrie de masse.

L'analyse par DC a permis d'accéder aux informations concernant la structuration secondaire globale du domaine (Figure 74B). Le spectre obtenu montre un maximum aux alentours de 192 nm, et des minima aux alentours de 205 et 222 nm. Ce profil est caractéristique d'une organisation majoritaire en hélices α du domaine, ce qui est en accord avec la littérature (Broadhurst et al., 2003).

Enfin, l'analyse de l'échantillon par MS en mode positif a permis d'obtenir la masse exacte des composants de chaque échantillon (Figure 74C). Dans le cas de l'analyse de $^{C}DD_{DEBS2}$, le spectre obtenu montre que le composant majeur de l'échantillon a une masse de 8904,6 Da. La différence entre cette valeur et la valeur déterminée théoriquement par le logiciel ProtParam (Tableau 15) peut provenir de l'état de protonation de la protéine. En effet, la différence de masse observée entre la valeur théorique de 8906,55 Da calculée et la valeur réelle observée de 8904,6 Da correspond à celle de deux hydrogènes. Les échantillons ont également été analysés par MS / MS afin de valider leur intégrité en termes de séquence. La séquence du domaine $^{C}DD_{DEBS2}$ correspond bien à celle de la protéine majoritaire de l'échantillon, dont la masse est de 8904,6 Da.

De cette manière, les peptides $^{C}DD_{DEBS2}$, $^{N}DD_{DEBS3}$ et $^{N}DD_{RAPSB}$ ont pu être validés. Le domaine $^{C}DD_{RAPSA}$ montre quant à lui un profil en DLS anormal et caractéristique d'une agrégation du domaine par la présence de deux courbes (Figure 75A). Le PdI obtenu pour ce domaine a été de 0,247, ce qui est supérieur à 0,2 et ne permet pas d'attester non plus de l'état monomérique et soluble du domaine.

L'analyse de la structure secondaire du domaine par DC montre bien une organisation majoritaire en hélices α (Figure 75B) comme pour le domaine précédent, et en adéquation avec la littérature (Buchholz et al., 2009). Cependant l'analyse du domaine par MS n'a pas permis d'observer la masse moléculaire attendue pour le domaine $^{C}DD_{RAPSA}$, qui est de 8898,93 Da (Figure 75C). Cela peut s'expliquer par le fait que le domaine précipite dans les solutions tampons utilisées.

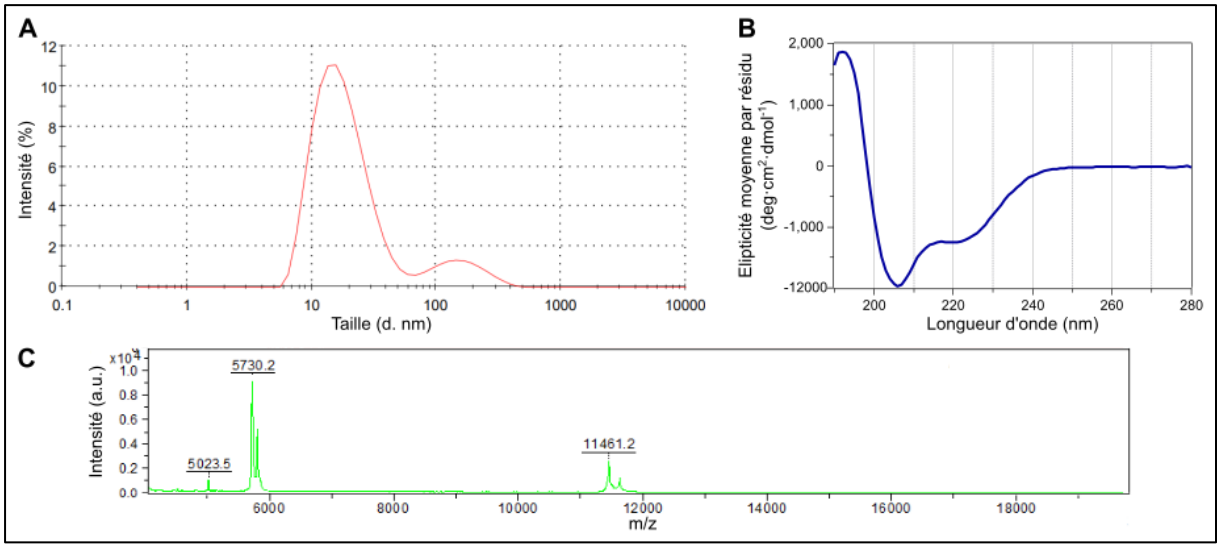


Figure 75. Analyse de l'intégrité du domaine de docking présent en position C-terminale de la sous-unité A de la rapamycine synthase. Le domaine $^{C}DD_{RAPSA}$ a été produit chez *E. coli* puis purifié en quatre étapes telles que présentées plus haut. **(A)** Chromatographe de l'analyse du domaine $^{C}DD_{RAPSA}$ par diffusion dynamique de la lumière. **(B)** Ellipticité moyenne par résidu du domaine $^{C}DD_{RAPSA}$ mesurée par dichroïsme circulaire. **(C)** Chromatographe de l'analyse du domaine $^{C}DD_{RAPSA}$ par spectrométrie de masse.

Le domaine $^{C}DD_{RAPSB}$ quant à lui n'a pour le moment pas pu être validé par ces analyses biophysiques. En effet, le profil observé en spectrométrie de masse correspondait pour cet échantillon à celui de $^{N}DD_{RAPSB}$, correspondant donc à une erreur lors de la production et purification des protéines d'intérêt. Par manque de temps cependant, ces étapes n'ont pas pu être reprises. Enfin, le domaine $^{N}DD_{RAPSC}$ n'a pas pu être solubilisé lors de cette thèse et n'a donc pas pu être étudié.

I.A.3.d. Etude des interactions entre les domaines de docking

Après la production, purification et validation des domaines, leurs interactions ont été analysées par titrage calorimétrique isotherme (ITC). Dans un premier temps, l'interaction entre les DD de DEBS a été caractérisée. Pour se faire, l'échantillon titré a été placé dans la cuve de l'appareil à une concentration de 60 µM, et le partenaire titrant dans la seringue à une concentration dix fois plus importante. Lorsque le domaine $^{N}DD_{DEBS3}$ est titré par le domaine $^{C}DD_{DEBS2}$, un unique événement est observé, lié à l'interaction entre les deux partenaires, dont la constante de dissociation calculée est de 22,17 ± 15,67 µM (Figure 76A).

La stœchiométrie de cette interaction est de 1,42 ± 0,272. Lorsque le domaine $^{C}DD_{DEBS2}$ est titré par le domaine $^{N}DD_{DEBS3}$, deux événements sont observés, caractérisés par la présence d'une pente avant la courbe caractéristique de l'interaction observée dans le premier cas (Figure 76B).

Cependant, les données obtenues dans ce cas n'ont pas permis de calculer les constantes de ces événements.

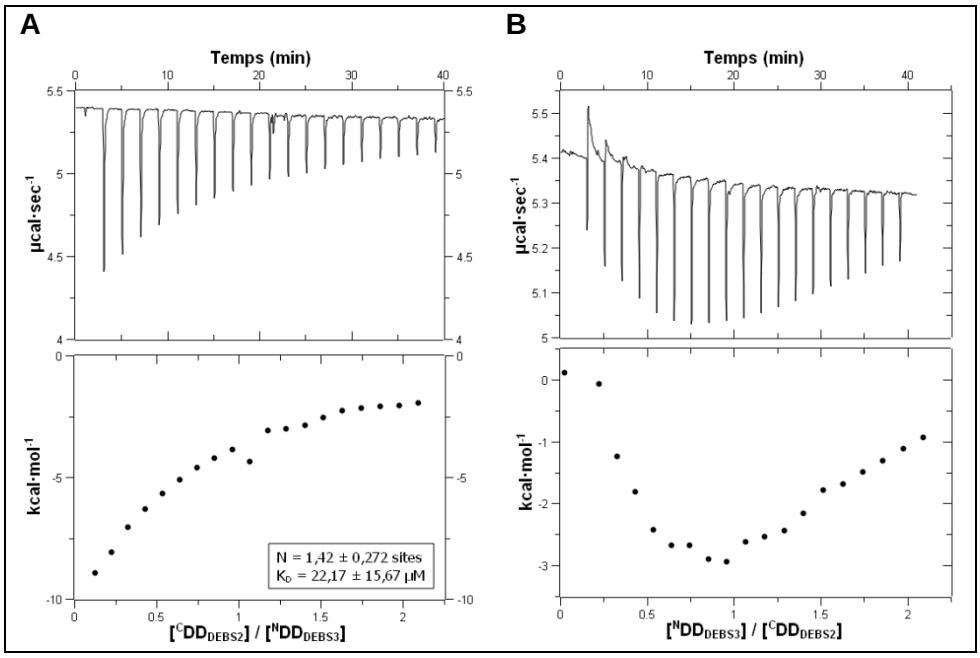


Figure 76. Spectres obtenus par titrage calorimétrique isotherme du couple de domaines de docking présents entre les sous-unités 2 et 3 de la 6-désoxyérythronolide B synthase. **(A)** Spectre obtenu lors du titrage du domaine $^{N}DD_{DEBS3}$ par le partenaire $^{C}DD_{DEBS2}$. Un unique événement est observé, dont la constante de dissociation est de 22,17 ± 15,67 µM et la stœchiométrie de 1,42 ± 0,272. **(B)** Spectre obtenu lors du titrage du domaine $^{C}DD_{DEBS2}$ par le partenaire $^{N}DD_{DEBS3}$. Deux événements sont observés lors du titrage.

L'étude des interactions par DC a également été menée. L'interaction entre deux partenaires protéiques peut entraîner une structuration d'un des partenaires qui serait nativement plus désordonné. Dans ce cas, il est alors possible d'observer ce changement par DC. Les expériences menées sur le couple $^{C}DD_{DEBS2}$ / $^{N}DD_{DEBS3}$ n'ont pas permis d'observer de tels changements de conformation (Figure 77). Les résultats obtenus montrent que les deux partenaires sont ordonnés majoritairement en hélices α (Figure 77A), et qu'aucun ne semble plus désordonné. L'analyse des deux partenaires placés chacun dans une cuve ou mélangés dans chacune ne montre aucune différence entre les courbes obtenues (Figure 77B). Cela signifie que l'interaction entre les deux domaines – observée en ITC – n'induit pas de changement conformationnel des DD.

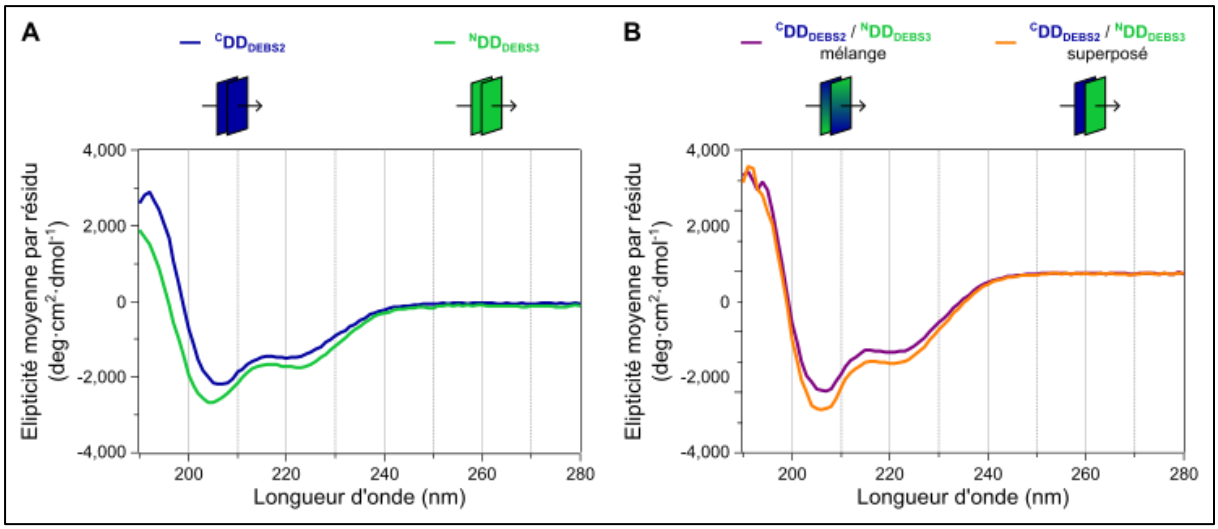


Figure 77. Spectres obtenus par dichroïsme circulaire pour le couple de domaines de docking présents entre les sous-unités 2 et 3 de la 6-désoxyérythronolide B synthase. La légende des courbes et la composition des deux cuves sont représentées en haut de chaque graphique. **(A)** Spectres des domaines individuels. Les deux spectres ont un profil caractéristique d'une organisation majoritaire en hélices α des domaines. **(B)** Spectres des domaines superposés ou mélangés. Le profil des spectres des domaines superposés ou mélangés ne montre pas de différence dans l'organisation en hélices α. En effet, il n'y a pas de décalage des minima observés aux alentours de 205 nm et 222 nm dans les deux cas.

Seul le domaine $^{C}DD_{RAPSB}$ ayant pu être obtenu, aucun couple de DD n'a pu être étudié. En l'absence de contrôle, l'utilisation de ce domaine pour évaluer les potentielles interactions avec le domaine $^{N}DD_{DEBS3}$ n'a pas été possible pour le moment, les résultats ne pouvant être correctement interprétés.

L'étude des interactions entre les domaines de RAPS n'a pas pu être menée dû au manque de protéines disponibles. En effet, seul le domaine $^{N}DD_{RAPSB}$ est disponible à ce jour. La caractérisation de la présence d'une interaction entre les paires canoniques de DD est indispensable à l'étude d'interactions potentielles non canoniques. C'est pour cela que l'analyse du domaine $^{C}DD_{DEBS2}$ en contact avec le domaine $^{N}DD_{RAPSB}$ n'a, pour le moment, pas été effectuée.

I.B. Etude de la stambomycine synthase : construction des souches mutées

La stambomycine synthase est une PKS de type I modulaire *cis*-AT composée de 24 modules d'élongation, répartis au sein de neuf sous-unités (Figure 48) (Laureti et al., 2011). Lors de précédentes expérimentations d'ingénierie génétique, une souche de *S. ambofaciens* a été obtenue exprimant cette PKS modifiée, au sein de laquelle les modules 14 à 20 (Δ*pks5*–*pks8*) ont été délétés et une nouvelle interface d'interaction a été créée entre les sous-unités Pks4 et Pks9. Ce système modifié a permis d'obtenir de nouveaux composés appelés mini-stambomycines, avec cependant un rendement de production 20 fois plus faible que celui des stambomycines natives chez la souche sauvage (Su et al., 2022). L'objectif de cette thèse a été de scinder une sous-unité de la PKS afin de mesurer les effets sur les rendements. Dans ce cas, la sous-unité Pks3 a été choisie car c'est la plus grande du système – composée de 5 modules – et elle est retrouvée au sein du complexe de PKS sauvage et muté.

I.B.1. Choix des séquences Synthetic Zippers *et du point de fusion*

La sous-unité Pks3 est composée des modules d'élongation 6 à 10. Lors de cette thèse, la scission a été effectuée entre les modules 7 et 8, soit entre les domaines ACP_7 et KS_8, afin d'obtenir des sous-unités dont les tailles sont les plus similaires possible. Les domaines KS sont des domaines hautement conservés en séquences et qui débutent par un motif 'EPIA'. Ainsi, le point de fusion entre le domaine SynZip en position N-terminale et le domaine KS a pu facilement être établi. Pour ce qui est du domaine ACP_7, une analyse a été menée afin de déterminer la fin de ce domaine et donc le point de fusion en position C-terminale. L'alignement de la séquence de l'ACP7 avec des séquences d'autres domaines ACP, ainsi que la prédiction de la structure secondaire de ce domaine par PsiPred, Prabi et AlphaFold a permis d'établir le point de fusion à la suite d'un résidu leucine (Figure 78).

Ici, la scission a été faite par l'insertion de domaines SynZip. Ce choix a été fait pour assurer qu'aucune interaction non canonique ne soit possible entre les domaines introduits et les DD natifs de la PKS. Pour ce qui est du choix des paires de SynZips à introduire, la paire 3:4 a été sélectionnée car elle est a déjà pu être utilisée lors d'expériences d'ingénierie des PKS (Buyachuihan et al., 2023). Cette paire forme une interaction parallèle dont la constante de dissociation est inférieure à 30 nM (Thompson et al., 2012). Une seconde paire a été choisie, qui est la paire 17:18. Celle-ci forme une interaction antiparallèle dont la constante de dissociation est inférieure à 10 nM (Thompson et al., 2012).

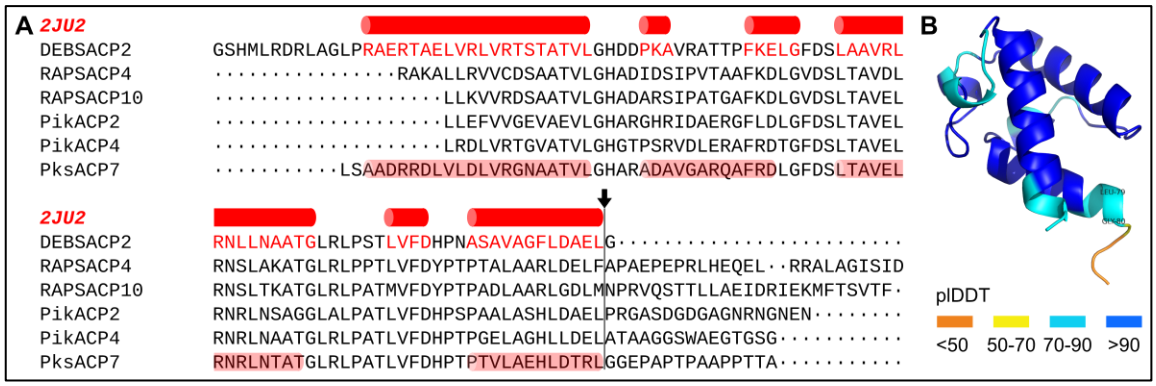


Figure 78. Analyses du domaine porteur d'acyle du module 7 de la stambomycine synthase. **(A)** Alignement de séquences entre plusieurs domaines porteur d'acyle (ACP) provenant de différentes polycétide synthases (PKS) de type *cis*-acyltransférase (AT) et prédiction de la structure secondaire du domaine Pks-ACP7. Le domaine Pks-ACP7 est le domaine ACP retrouvé au sein de la sous-unité 3 de la stambomycine synthase et après lequel les domaines d'interactions seront insérés. L'alignement de séquences a été effectué à l'aide de l'outil en ligne Clustal Omega, et les prédictions de structures ont été faites par l'outil en ligne PsiPred. La structure résolue du domaine $^{DEBS1}ACP_2$ (identifiant PDB : 2JU2) a été utilisée afin de déterminer la fin des domaines ACP et donc le point de fusion, indiquée par la flèche noire. Les hélices α de ce domaine sont représentées par des cylindres rouges, et les résidus qui y sont impliqués sont également en rouge. Le domaine $^{Pks3}ACP_7$ est prédit comme composé d'hélices α, représentées par des cylindres rouges en filigrane. **Légende** – DEBS : 6-désoxyérythronolide B synthase, RAPS : rapamycine syntase, Pik : pikromycine synthase. **(B)** Modèle de la structure du domaine Pks3-ACP7 construit par AlphaFold3. L'indice de confiance du modèle est ici exprimé par le test des différences de distances locales (plDDT), qui est un score allant de 0 à 100. Plus la valeur est proche de 100, plus la confiance est élevée. Les derniers résidus impliqués dans une hélice α ici sont les résidus leucine et glycine identifiés en (A) comme étant la fin du domaine.

I.B.2. Conception, assemblage et clonage des matrices d'édition

Le système d'édition CRISPR-Cpf1 a également été sélectionné pour la modification du génome de *S. ambofaciens*. En effet, le système CRISPR-Cas9 présente des ARNg ayant un trop haut risque de clivage hors site pour la scission de Pks3, ou qui sont trop éloignés de la fin du domaine ACP_7 (plus de 1000 pb). Afin de déterminer la séquence du protospacer, l'outil en ligne CHOPCHOP a été utilisé (Labun et al., 2019). Les cinq premières séquences proposées sont présentées dans le Tableau 16. La séquence sélectionnée a été celle remplissant dans l'ordre les critères suivants : 1) ne pas s'hybrider ailleurs dans le génome, à 0 pb, 1 pb ou 2 pb près et 2) être le plus proche possible de la fin du domaine ACP_7. La séquence ainsi identifiée est la suivante : « 5′-GGCGAGCACGGTCGGGGTCGGATGG-3′ »

Tableau 16. Séquences des protospacer potentiels identifiés dans le génome de *Streptomyces ambofaciens* pour la scission de la stambomycine synthase.

Séquence	PAM	Distance avec ACP_7	Hybridation hors cible		
			0 pb	*1 pb*	*2 pb*
GGCGAGCACGGTCGGGGTCGGATGG	TTC	24 pb	0	0	0
GACCATCCGACCCCGACCGTGCTCG	TTC	26 pb	0	0	0
CGCGACCTCGGCTTCGACTCGCTCA	TTC	122 pb	0	2	5
GACCGCGGTGAGCGAGTCGAAGCCG	TTC	111 pb	0	0	2
GACTCGCTCACCGCGGTCGAACTCC	TTC	104 pb	0	0	2

Les séquences protospacer potentielles ont été identifiées à l'aide de l'outil en ligne CHOPCHOP (Labun et al., 2019). La distance avec la fin du domaine ACP_7 est calculée à partir de la position du dernier nucléotide de la séquence du protospacer. La séquence sélectionnée est surlignée en jaune.

Comme précédemment présenté pour la rapamycine synthase, les matrices d'édition doivent contenir les régions d'homologie permettant la réparation du génome après le clivage par Cpf1, un gène synthétique contenant à la fois la région du protospacer mutée et les séquences des SynZips, et la séquence de l'ARNcr. La carte de la matrice d'édition dans le cas de l'utilisation des domaines SynZips 3:4 est présentée dans la Figure 79.

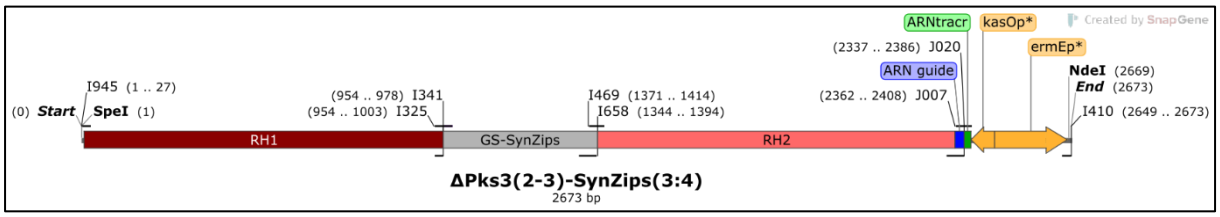


Figure 79. Carte de la matrice d'édition pour l'insertion de la paire de SynZip 3:4 au sein de la stambomycine synthase chez *S. ambofaciens* par CRISPR-Cpf1. Cette matrice contient les régions d'homologie (RH) nécessaires à la réparation de l'ADN à la suite de l'action de CRISPR-Cpf1, un gène synthétique permettant à la fois de muter la région du protospacer et d'insérer les séquences codant la paire de *synthetic zippers* (SynZips) 3:4, ainsi que la séquence de l'ARN guide. Des éléments du plasmide pKCCpf1 sont également présents afin de ne pas les éliminer lors du clonage entre les sites *Spe*I et *Nde*I. Ces éléments sont la séquence répétée qui suit l'ARN guide, le promoteur kasOp* qui permet la synthèse de l'ARN guide et le promoteur ermEp* en amont de la région codant l'endonucléase Cpf1. Les amorces utilisées pour l'amplification et l'assemblage de la matrice sont représentées par leur nomenclature ESCOLA (voir Tableau 4).

Chaque élément des matrices d'édition a été amplifié par PCR puis assemblé en deux étapes de PCR à l'aide d'amorces chevauchantes, comme précédemment présenté dans le cas de la scission de la rapamycine synthase. Les matrices – contenant soit la paire SynZip 3:4 soit la paire SynZip 17:18 – ont ensuite été clonées au sein du plasmide pKCCpf1 puis l'intégrité des vecteurs recombinants a été validée par un séquençage complet des plasmides.

I.B.3. Construction des souches de S. ambofaciens

Les plasmides préalablement obtenus ont ensuite été transférés par conjugaison chez *S. ambofaciens*. Les souches ATCC23877 et ATCC/hy59_S2 (Su et al., 2022) de *S. ambofaciens* ont été utilisées.

La souche ATCC/hy59_S2 est délétée des sous-unités Pks5 à Pks8, et la nouvelle interface d'interaction entre les sous-unités Pks4 et Pks9 est constituée des DD des sous-unités Pks4 et Pks5. Le premier domaine KS de la sous-unité Pks9 a également été modifié afin d'obtenir un hybride entre celui-ci est le premier domaine KS de la sous-unité Pks5. Cela dans le but d'améliorer l'interaction entre le domaine ACP en amont du domaine $^{C}DD_{Pks4}$ et la sous-unité Pks9 (Figure 50C, page 69).

Les exconjuguants obtenus après conjugaison des deux souches de *S. ambofaciens* avec les vecteurs d'intérêts ont été criblés par PCR (Figure 80). Les amorces K009 et K010 utilisées pour ce criblage ont été conçues de manière à s'hybrider de part et d'autre du site d'insertion, afin d'obtenir un fragment d'une taille d'environ 515 pb dans le cas d'un génome non modifié et d'environ 930 pb dans le cas de l'insertion des SynZip. Dans le cas de l'insertion des domaines SynZip 3:4, quatre clones ont été identifiés comme modifiés à partir de la souche ATCC23877 et deux à partir de la souche ATCC/hy59_S2. Dans le cas de l'insertion des domaines SynZip 17:18, deux clones de chaque souche ont été identifiés.

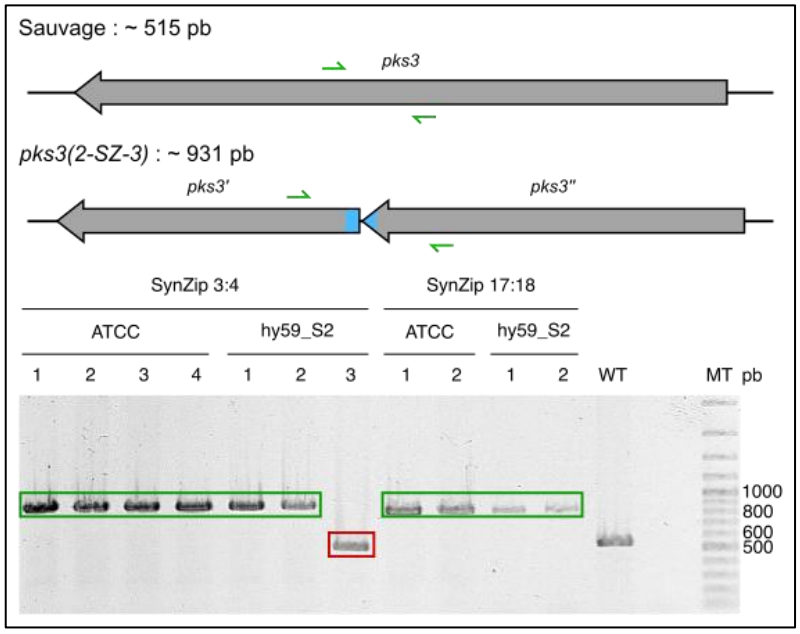


Figure 80. Criblage positif par PCR des exconjuguants de *Streptomyces ambofaciens* pour l'insertion des domaines SynZip au sein de la stambomycine synthase. Les amorces K009 et K010 ont été conçues de manière à s'hybrider de part et d'autre du point de fusion. Les domaines SynZip sont représentés en bleu. La taille des fragments attendus dans la souche non mutée (WT) est d'environ 515 pb. Lorsque les domaines SynZip ont correctement été intégré, la taille attendue du fragment est d'environ 930 pb. Les souches ATCC23877 et ATCC/hy59_S2 de *S. ambofaciens* ont été mutées. Les fragments correspondant à la taille attendue après insertion sont encadrés en vert, et le fragment obtenu à la taille de la souche non mutée est encadré en rouge. **Légende** – WT : sauvage, MT : marqueur de taille, pb : paire de bases.

Les fragments obtenus à une taille d'environ 900 pb ont été purifiés et séquencés pour valider l'intégrité de la modification génétique. La séquence de la région modifiée par l'insertion de la paire SynZip 3:4 est présentée en exemple dans la Figure 81.

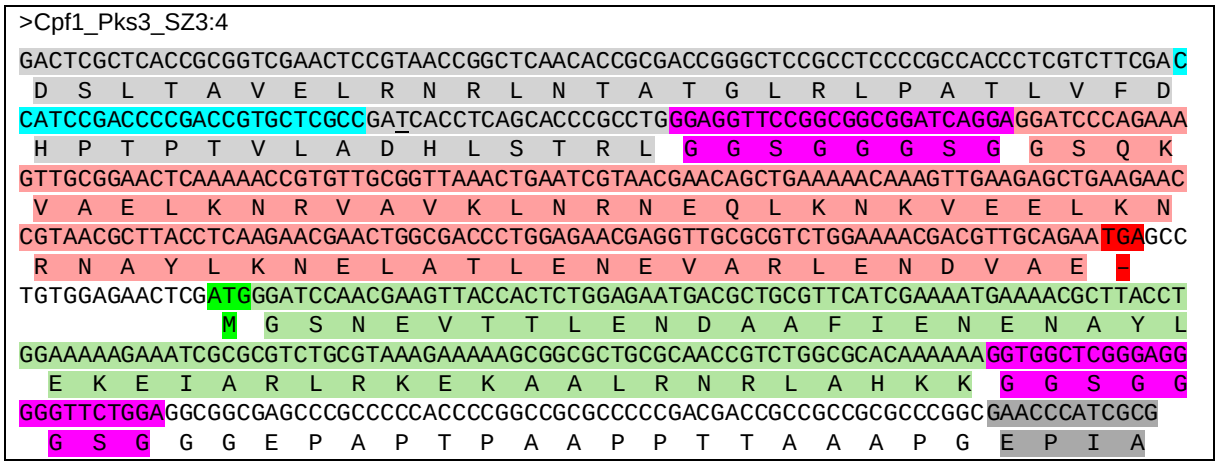


Figure 81. Séquence nucléotidique et protéique résultant de la mutation apportée au génome de *S. ambofaciens* . La fin du domaine porteur d'acyle (ACP) du deuxième module de Pks3 est représenté en gris clair. La séquence du protospacer est en bleu au sein de la séquence nucléotidique et le nucléotide muté au sein de la séquence adjacente au protospacer (PAM), afin d'éviter que le système CRISPR ne clive le vecteur d'intérêt contenant cette région ainsi que d'éviter des événements multiples de clivage et recombinaison, est souligné. Les régions codant les linker « GGSG » x2 sont en rose et les SynZip 3 et 4 sont respectivement représenté en rouge et vert. Le codon stop présent à la fin de la région codant le SynZip 3 et le codon start au début de la séquence codant le SynZip 4 sont représentés en plus foncés. Enfin, le domaine cétosynthase du troisième module de Pks3 est représenté en gris foncé.

La même démarche a été suivie afin de séquencer les régions mutées de chacun des clones des souches ATCC23877 et ATCC/hy59_S2, contenant soit les SynZips 3:4 soit la paire de SynZip 17:18. Les souches obtenues au cours de cette thèse sont présentées dans le Tableau 17.

Tableau 17. Liste des souches conçues dans le cadre de la scission de la stambomycine synthase durant ce travail de thèse.

Souche	Mutation	Description
ATCC23877-MT-S1	*pks3(2-SZ3:4-3)*	Insertion des SynZips 3:4 entre le deuxième et le troisième module de la sous-unité Pks3.
ATCC23877-MT-S2	*pks3(2-SZ17:18-3)*	Insertion des SynZips 17:18 entre le deuxième et le troisième module de la sous-unité Pks3.
ATCC/hy59_S2-MT-S3	*pks3(2-SZ3:4-3)*	Insertion des SynZips 3:4 entre le deuxième et le troisième module de la sous-unité Pks3.
ATCC/hy59_S2-MT-S4	*pks3(2-SZ17:18-3)*	Insertion des SynZips 17:18 entre le deuxième et le troisième module de la sous-unité Pks3.

L'étude de la production de stambomycines et mini-stambomycines par les souches obtenues n'a cependant pas encore pu être menées. Tout d'abord, les souches doivent être transformées à nouveau afin de contenir le vecteur pOE-0484, contenant le gène *sam23877-7107* codant l'activateur transcriptionnel du cluster sous le contrôle du promoteur constitutif fort ermE* (Laureti et al., 2011). Une fois ces souches obtenues, les rendements de production pourront être étudiés par l'extraction des métabolites produits et leur analyse par spectrométrie de masse.

II. Localisation *in situ* de la PKS *cis*-AT coelimycine synthase

Le second objectif de cette thèse porte sur un autre aspect, largement négligé lors des tentatives d'ingénierie des PKS, qui est leur macro-organisation au sein de leur organisme hôte. Des études menées sur la bacillaene synthase de *B. subtilis*, une PKS modulaire *trans*-AT, ont montré que celle-ci s'organise en méga-complexes localisés au niveau de la membrane bactérienne (Straight et al., 2007). Des hélices α pouvant médier cette macro-organisation ont été identifiées au niveau des domaines KS qui forment une région nommée LINKS (Gay et al., 2016). Cependant, ces domaines ne sont pas présents au sein des PKS *cis*-AT (Figure 41, page 53), et aucune analyse de leur organisation au sein de leur hôte n'a, à ce jour, été menée. Dans ce cadre, la PKS *cis*-AT produisant la coelimycine chez *S. coelicolor* a été choisie comme modèle d'étude. L'organisation des sous-unités de ce système, ainsi que de son transporteur, a été analysée par l'utilisation de protéines fluorescentes couplées à de la microscopie confocale.

II.A. Contexte et résultats antérieurs à ce projet de thèse

Ce projet a été initié par le Dr. Jean-Malo Massicard durant sa thèse (Massicard, 2020). La souche sauvage utilisée dans cette étude est la souche FM145 de *S. coelicolor* A3(2) (Willemse & van Wezel, 2009). Le choix de cette souche repose sur le fait qu'elle présente une diminution de l'auto-fluorescence dans le canal vert d'un facteur 2,9, et d'un facteur 7 dans le canal rouge comparée à la souche M145. Afin d'étudier la localisation et l'organisation de la coelimycine synthase (Cpk), la première sous-unité (CpkA) a été fusionnée à la protéine fluorescente mCherry en position N-terminale, et la dernière sous-unité CpkC a été fusionnée à la protéine eGFP en position C-terminale. Les fusions ont été faites de manière à ne pas interférer avec les interactions entre les sous-unités médiées par les DD présents en position C-terminale de CpkA et N-terminale de CpkC. De plus, la localisation du transporteur CpkF a également été étudiée en lui fusionnant la protéine mCherry en position C-terminale.

Les souches nécessaires à ces observations ont été générées par ingénierie génétique à l'aide du système d'édition CRISPR-Cas9. Dans ce cadre, le plasmide pCRISPomyces-2 a été modifié afin de contenir les éléments nécessaires à la fusion des protéines fluorescentes aux protéines du système Cpk ciblées. Les souches obtenues et étudiées lors de ces premières expérimentations sont présentées dans le Tableau 18.

Tableau 18. Liste des souches conçues et étudiées dans le cadre de la localisation de la coelimycine synthase par le Dr. Jean-Malo Massicard (adapté de Massicard, 2020).

Souche	Génotype	Description	Source
FM145-JMM1	*mCherry-cpkA*	Insertion de la séquence codant mCherry à l'extrémité 5′ de *cpkA*	(Massicard, 2020)
FM145-JMM2	*cpkC-eGFP*	Insertion de la séquence codant eGFP à l'extrémité 3′ de *cpkC*	(Massicard, 2020)
FM145-JMM3	*mCherry-cpkA, cpkC-eGFP*	Insertion de la séquence codant mCherry à l'extrémité 5′ de *cpkA* et insertion de la séquence codant eGFP à l'extrémité 3′ de *cpkC*	(Massicard, 2020)
FM145-JMM4	*cpkF-mCherry*	Insertion de la séquence codant mCherry à l'extrémité 3′ de *cpkF*	(Massicard, 2020)

Les souches ont été observées en microscopie confocale en collaboration avec l'équipe de Pr. van Wezel (Université de Leiden, Pays-Bas). Les souches ont été cultivées sur des lamelles dans un milieu DNA et observées après 27 heures d'incubation à 30°C. Les résultats obtenus ont montré que les sous-unités CpkA et CpkC s'organisent en foci au sein des hyphes bactériens, mais ne colocalisent pas (Figure 82). De plus, leur expression n'est pas observée dans toute la colonie, mais seulement dans quelques hyphes et des îlots d'expression ont pu être identifiés.

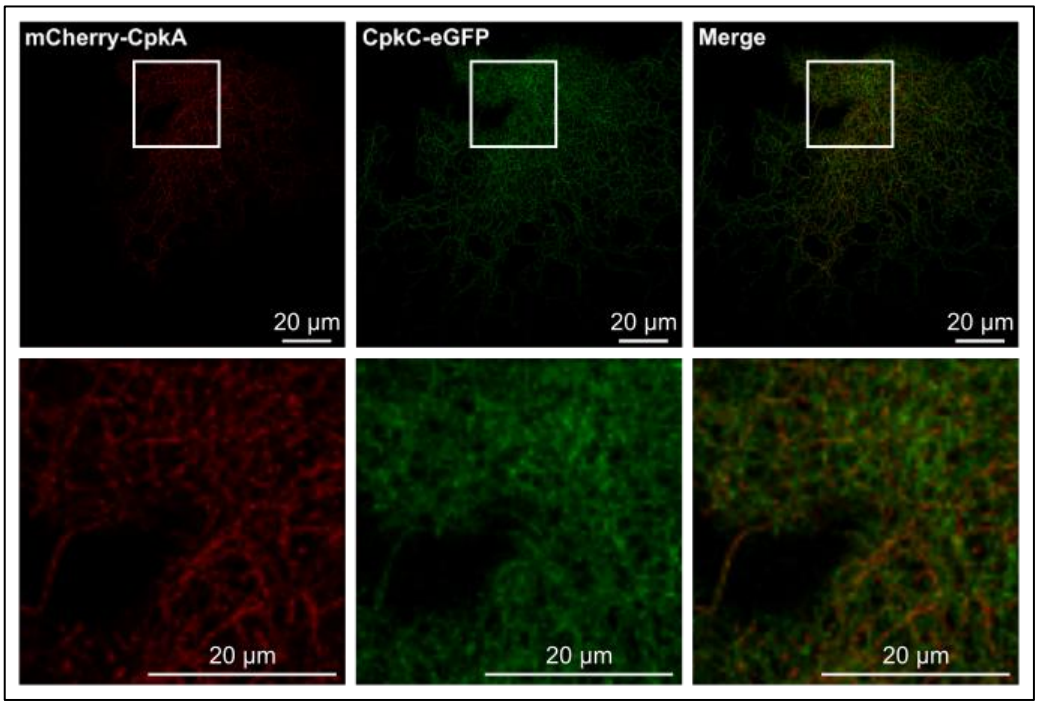


Figure 82. Observation au microscope confocal de la souche FM145-JMM3. La fusion mCherry-CpkA a été observée en excitant mCherry à une longueur d'onde de 522 nm et en récoltant l'émission à 586–789 nm (*gauche*). Celle-ci s'organise en foci au sein des hyphes bactériens. Cependant, cette observation n'a pas été faite sur tous les échantillons étudiés, seuls quelques îlots d'hyphes au sein des colonies expriment la coelimycine synthase. La fusion CpkC-eGFP a été excitée à 488 nm et l'émission a été récoltée à 500–548 nm (*milieu*). Elle forme également des foci au sein d'îlots d'hyphes retrouvés dans les colonies. Enfin, la superposition de ces deux clichés n'a pas permis d'observer une colocalisation de ces deux protéines fusionnées (*droite*).

L'absence de colocalisation ne permet pas de conclure quant à l'organisation de la PKS, car les protéines fluorescentes utilisées montrent des caractéristiques pouvant influencer ce résultat. D'une part, le signal lié à mCherry est difficilement détectable à travers l'auto-fluorescence de la souche, et d'autre part la protéine eGFP est capable de s'oligomériser, ce qui peut interférer avec l'organisation propre à Cpk. Afin de valider ou non les observations faites, la stratégie a alors été d'observer l'organisation de ces sous-unités fusionnées à d'autres protéines fluorescentes monomériques : mYPet et mTurquoise2. C'est donc dans cet objectif que s'intègrent les résultats présentés dans ce chapitre.

II.B. Ingénierie génétique de *Streptomyces coelicolor*

II.B.1. Conception, assemblage et clonage des matrices d'édition

Les constructions nécessaires à la fusion des protéines fluorescentes mYPet et mTurquoise2 aux sous-unités de Cpk ont été conçues par le Dr. Jean-Malo Massicard durant sa thèse. Le vecteur pCRISPomyces-2-cpkC-mTurquoise2 avait également été construit précédemment, et seul le vecteur pCRISPomyces-2-*mYPet-cpkA* n'avait pas encore été construit (Tableau 19).

Tableau 19. Liste des plasmides utilisés pour l'édition génétique par CRISPR-Cas9 de *Streptomyces coelicolor* au cours de cette thèse.

Construction	Description	Source
pCRISPomyces-2-*mYPet-cpkA*	Insertion du gène codant mYPet en position 5′ de *cpkA*	Ce travail
pCRISPomyces-2-*cpkC-mTurquoise2*	Insertion du gène codant mTurquoise2 en position 3′ de *cpkC*	(Massicard, 2020)
pCRISPomyces-2-Δ*cpkA*	Délétion du gène *cpkA*	(Massicard, 2020)
pCRISPomyces-2-Δ*cpkC*	Délétion du gène *cpkC*	(Massicard, 2020)
pCRISPomyces-2-Δ*cpkF*	Délétion du gène *cpkF*	(Massicard, 2020)

Dans un premier temps le plasmide pCRISPomyces-2-*mYPet-cpkA* a été construit. La matrice d'édition permettant d'insérer la séquence codant la protéine fluorescente mYPet en position 5′ de l'ORF du gène *cpkA* est présentée dans la Figure 83. Elle contient les régions d'homologie, un gène synthétique permettant de reconstituer le début de la sous-unité CpkA tronquée lors du clivage – puisque le protospacer est retrouvé en amont de la fin du gène – et d'apporter des mutations au sein du protospacer afin d'éviter que le vecteur puisse être la cible du système CRISPR-Cas9, et le gène codant mYPet.

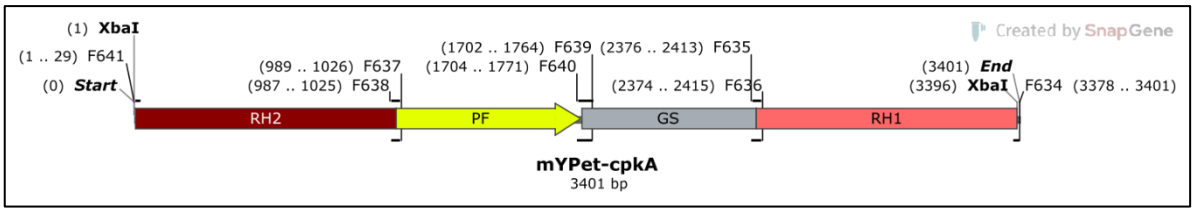


Figure 83. Carte de la matrice d'édition utilisée pour l'insertion de la protéine fluorescente mYPet fusionnée à la première sous-unité de la coelimycine synthase chez *Streptomyces coelicolor*. La matrice est composée des régions d'homologie (RH) permettant la réparation du génome par recombinaison homologue à la suite de la coupure double-brin générée par la caspase, un gène synthétique afin de reconstruire la partie de l'ORF *cpkA* clivée par CRISPR-Cas9 et intégrer des mutations au sein du protospacer et le gène codant la protéine fluorescente (PF) à intégrer en 5' de *cpkA*. Les amorces sont indiquées par leur nomenclature ESCOLA (voir Tableau 6, page 102).

Les éléments de cette matrice ont été amplifiés puis assemblés par PCR à l'aide d'amorces chevauchantes, comme précédemment présenté lors de la construction des matrices pour la scission de la rapamycine synthase. La matrice a ensuite été clonée dans le vecteur pCRISPomyces-2 au niveau du site de restriction *Xba*I par un assemblage de type Gibson. Par la suite, la séquence codant l'ARN guide a été clonée dans le vecteur par la méthode de Golden Gate au niveau du site de restriction *Bbs*I. La construction obtenue a été validée par séquençage avant d'être utilisée pour conjuguer la souche FM145 de *S. coelicolor*.

II.B.2. Construction des souches de S. coelicolor *FM145*

Afin de construire les souches mutées de *S. coelicolor* FM145, la première étape a été de transformer la souche FM145 par conjugaison intergénérique avec les vecteurs d'édition génétique. Ensuite, les exconjuguants ont été criblés par PCR et les fragments séquencés afin de valider l'intégrité des modifications effectuées. Enfin, le plasmide d'intérêt utilisé pour transformer *S. coelicolor* a été éliminé des exconjuguants mutés afin de pouvoir l'utiliser à nouveau pour de nouvelles modifications génétiques au sein des souches mutées.

II.B.2.a. Criblage et séquençage des exconjuguants dans le cas de l'insertion des gènes codant les protéines fluorescentes

Les constructions pCRISPomyces-2-*mYPet-cpkA* et pCRISPomyces-2-*cpkC-mTurquoise2* ont été utilisées afin de transformer la souche FM145 de *S. coelicolor*, dans le but d'intégrer les protéines fluorescentes mYPet et mTurquoise2 fusionnées respectivement aux sous-unités CpkA et CpkC. Une fois les exconjuguants obtenus, ils ont été criblés par PCR afin de déterminer lesquels portaient la mutation attendue. L'exemple du crible des exconjuguants obtenus après conjugaison pour l'insertion de *mTurquoise2* en position 3′ de *cpkC* est présenté dans la Figure 84.

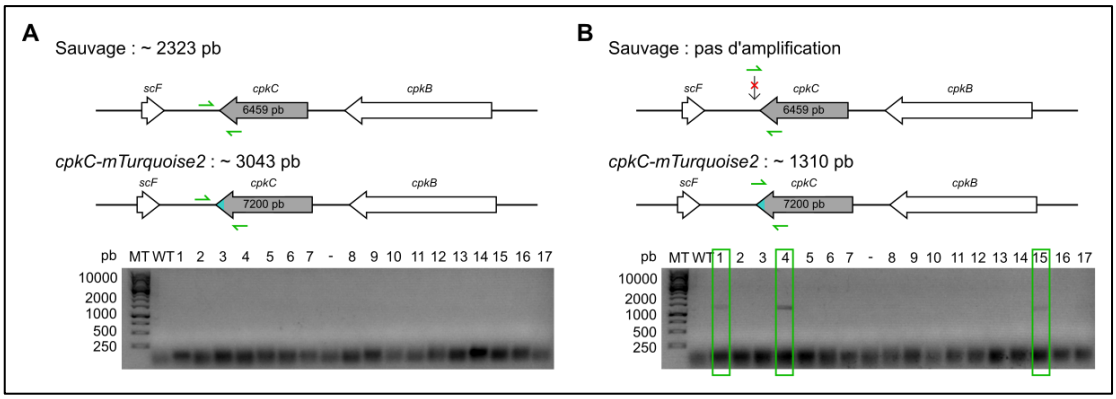


Figure 84. Criblage par PCR des exconjuguants de *Streptomyces coelicolor* pour l'insertion de la séquence codant la protéine mTurquoise2 en position 3′ de *cpkC*. (A) Premier criblage positif des exconjuguants. Les amorces F650 et F653 conçues pour ce criblage s'hybrident de part et d'autre du point de fusion, pour obtenir des fragments d'environ 2320 pb à partir d'un génome non modifié et de 3040 pb après mutation. Aucun fragment n'a été observé à la suite de la PCR. **(B)** Second criblage positif des exconjuguants. Les amorces oligonucléotidiques F650 et F651 ont été conçues de manière que l'une s'hybride au sein de la séquence insérée. Ainsi, aucune amplification n'est possible à partir d'un génome sauvage, et un fragment d'environ 1310 pb est attendu à partir des exconjuguants mutés. Ainsi, les clones 1, 4 et 15 ont été identifiés. Le clone 1 a été sélectionné pour la suite des expérimentations.

Deux criblages positifs ont été effectués. Pour le premier criblage, les amorces utilisées – F650 et F653 – ont été conçues de manière à s'hybrider de part et d'autre du point de fusion. Ainsi, le fragment obtenu à partir du génome muté est plus long (~3040 pb) que celui obtenu à partir du génome sauvage (~2320 pb). Ce criblage cependant n'a pas fonctionné puisqu'aucun fragment n'a été observé à la suite de la PCR, autant à partir de la souche sauvage FM145 qu'à partir des exconjuguants (Figure 84A). Cela peut s'expliquer par le fait que les amorces conçues ne puissent pas s'hybrider sur la matrice d'ADNg dans les conditions utilisées. Il aurait donc fallu essayer plusieurs températures d'hybridation, ainsi que plusieurs concentrations de DMSO dans le mix réactionnel.

Le second criblage a été effectué en utilisant un couple d'amorces – F650 et F651 – dont l'une s'hybride dans le gène *cpkC*, et la seconde s'hybride au sein de *mTurquoise2*. Un signal sera donc obtenu dès lors que la mutation aura été introduite. Ici, la taille du fragment obtenu à partir d'un génome muté est d'environ 1310 pb. Ainsi, trois exconjuguants ont pu être identifiés, et le clone n°1 a été sélectionné pour la suite des expérimentations (Figure 84B). Son génome a été extrait et utilisé comme matrice pour amplifier à nouveau ce fragment, avec la polymérase Phusion HSII. Le fragment obtenu a été extrait puis séquencé afin de valider l'intégrité de la région modifiée. La séquence nucléotidique de la région modifiée par l'introduction de la séquence codant *mTurquoise2* en position 3′ de *cpkC* est présentée dans la Figure 85.

La même méthode a été utilisée afin de construire toutes les souches contenant un ou plusieurs éléments du système Cpk fusionné(s) à des protéines fluorescentes.

>Cas9_CpkC_mTurquoise2

```
TTCAGCCACCGTCTCGGCATCCGCAAGCCGTAGCCGTAGCCACAGCCACAGCCGTGGAAAGAAGCCCCCGCCGGCCTCGGG
  F  S  H  R  L  G  I  R  K  P  –
CCGGCGGGGGCTTTCGCGTGCGGTGGGTCACTTGTACAGCTCGTCCATGCCCAGGGTGATGCCGGCGGCGGTGACGAACTC
                              –  K  Y  L  E  D  M  G  L  T  I  G  A  A  T  V  F  E
CAGCAGGACCATGTGGTCGCGCTTCTCGTTCGGGTCCTTGGACAGCTTGGACTGGGTGGACAGGTAGTGGTTGTCCGGCAG
  L  L  V  M  H  D  R  K  E  N  P  D  K  S  L  K  S  Q  T  S  L  Y  H  N  D  P  L
CAGGACCGGGCCGTCGCCGATCGGGGTGTTCTGCTGGTAGTGGTCGGCCAGCTGGACGCCGCCGTCCTCGATGTTGTGGCG
  L  V  P  G  D  G  I  P  T  N  Q  Q  Y  H  D  A  L  Q  V  G  G  D  E  I  N  H  R
GATCTTGAAGTTGGCCTTGATGCCGTTCTTCTGCTTGTCGGCGGTGATGTAGACGTTGTCGGAGAAGTAGTTGTACTCCAG
  I  K  F  N  A  K  I  G  N  K  Q  K  D  A  T  I  Y  V  N  D  S  F  Y  N  Y  E  L
CTTGTGGCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTGACCAGGGTGTCGCCCTCGAA
  K  H  G  L  I  N  G  D  E  K  F  D  I  G  K  L  E  I  R  N  V  L  T  D  G  E  F
CTTGACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCCTCCGGCATGGC
  K  V  E  A  R  T  K  Y  N  G  D  D  K  F  F  I  T  R  E  Q  V  Y  G  E  P  M  A
GGACTTGAAGAAGTCGTGCTGCTTCATGTGGTCCGGGTAGCGGGCGAAGCACTGGACGCCCCAGGACAGGGTGGTGACCAG
  S  K  R  R  D  H  Q  K  M  H  D  P  Y  R  A  F  C  Q  V  G  W  S  L  T  T  V  L
GGTCGGCCACGGGACCGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCGTCGCCCTCGCC
  T  P  W  P  V  P  L  K  G  T  T  C  I  F  K  L  T  L  K  G  Y  T  A  D  G  E  G
CTCGCCGGAGACGGAGAACTTGTGGCCGTTGACGTCGCCGTCCAGCTCGACCAGGATCGGGACGACGCCGGTGAACAGCTC
  E  G  S  V  S  F  K  H  G  N  V  D  G  D  L  E  V  L  I  P  V  V  G  T  F  L  E
CTCGCCCTTGGAGACGCCGCCGCCGCCGCCGCCCCCGGCCGCCGGGAAGTGTCCCTCCTCGACGAAGAAGGCCACGTA
  E  G  K  S  V     G  G  G  G  G  G  G     G  A  A  P  F  H  G  E  E  V  F  F  A  V  Y
TTTTTCAAAGAGTTCCCGCGTCAGCACCGGTATGCCTATGCCCGTGCCCTCCAGCGCCGCCACCGTCTCAGCGGTGTCGAT
  K  E  F  L  E  R  T  L  V  P  I  G  I  G  T  G  E  L  A  A  V  T  E  A  T  D  I
```

Figure 85. Séquence nucléotidique et protéique résultant de l'insertion de la séquence codant mTurquoise2 en position 3′ de l'ORF *cpkC* chez *Streptomyces coelicolor*. Le gène *cpkC* est représenté en gris foncé, et le gène qui se trouve en position 3′ de *cpkC* au sein du cluster (*cpkD*) est représenté en gris clair. La séquence du protospacer est en bleu au sein de la séquence nucléotidique, et les nucléotides mutés qui y sont apportés afin d'éviter des événements multiples de clivage et recombinaisons sont soulignés. La région en rose correspond au linker de sept glycine introduit entre la sous-unité CpkC et la protéine fluorescente mTurquoise2, dont la séquence nucléotidique est représentée en jaune.

II.B.2.b. Criblage et séquençage des exconjuguants dans le cas de la délétion d'ORF

Des souches dérivées de FM145 et délétées des gènes *cpkA*, *cpkC* ou *cpkF* ont également été construites afin de déterminer l'impact de la délétion d'un partenaire sur l'organisation des autres éléments de Cpk. Pour cela, les vecteurs pCRISPomyces-2-Δ*cpkA*, pCRISPomyces-2-Δ*cpkC* et pCRISPomyces-2-Δ*cpkF* construits par le Dr. Jean-Malo Massicard durant sa thèse ont été utilisés. Comme précédemment, les exconjuguants obtenus après conjugaison ont été criblés par PCR et les fragments obtenus séquencés afin de valider l'intégrité de la délétion. Le criblage des exconjuguants obtenus après transformation de *S. coelicolor* avec le plasmide pCRISPomyces-2-Δ*cpkC* permettant de déléter *cpkC* est présenté en exemple dans la Figure 86.

Pour ce qui est du criblage négatif (Figure 86A), les amorces F436 et F439 ont été conçues de manière que l'une d'elle s'hybride au sein du gène délété. Ainsi, aucune amplification n'est possible à partir des clones mutés, mais une amplification d'environ 1225 pb est obtenue à partir du génome sauvage. Dans le cas du criblage positif (Figure 86B), les amorces F437 et F439 s'hybrident de part et d'autre du gène délété.

Ainsi, l'amplification à partir du génome sauvage est d'environ 7670 pb et d'environ 1210 pb dans le cas du génome délété de *cpkC*. Dans les conditions de PCR utilisées, l'amplification du fragment à partir du génome sauvage n'est pas possible car il est trop long pour le temps d'élongation utilisé.

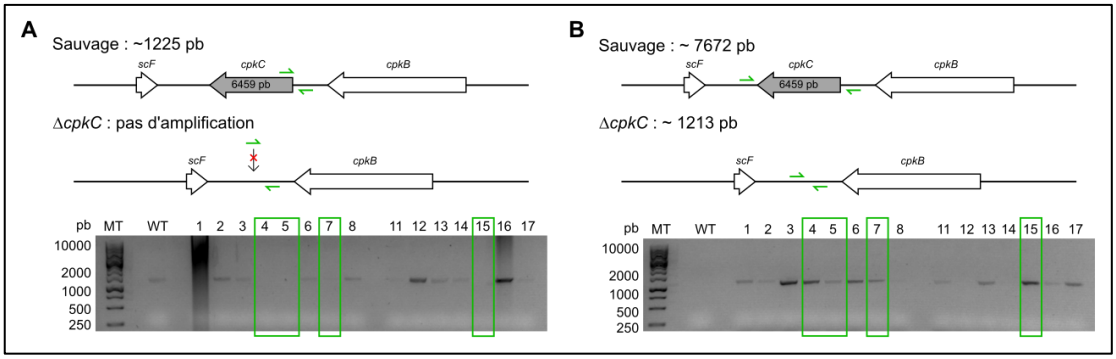


Figure 86. Criblage par PCR des exconjuguants de *Streptomyces coelicolor* pour la délétion de *cpkC*. (A) Criblage négatif des exconjuguants Δ*cpkC*. Les amorces F436 et F439 conçues pour ce criblage s'hybrident l'une dans le gène délété et la seconde en dehors. Ainsi, un fragment d'environ 1220 pb est attendu à partir du génome sauvage et aucune amplification n'est possible à partir des exconjuguants mutés. Les clones à partir desquels aucune amplification n'est observée ont été retenus pour la suite et sont entourés en vert. **(B)** Criblage positif des exconjuguants Δ*cpkC*. Les amorces F437 et F439 ont été conçues de manière à s'hybrider de part et d'autre du gène délété. Ainsi, une amplification d'environ 7670 pb est attendue à partir du génome sauvage et un fragment d'environ 1210 pb à partir d'un génome muté. Les conditions expérimentales utilisées ne permettent pas d'observer une amplification d'environ 7670 pb. Les clones pour lesquels une amplification est observée aux alentours de 1200 paires de bases ont été sélectionnés pour la suite et sont encadrés en vert.

Ainsi, les clones sélectionnés par la suite sont ceux dont le criblage négatif (Figure 86A) ne permet pas d'obtenir d'amplification, tandis que le criblage positif donne un signal correspondant à la taille attendue (Figure 86B). Le clone n°4 a été sélectionné pour la suite des expérimentations et une nouvelle PCR sur colonie est effectuée, cette fois-ci avec la polymérase Phusion HSII, afin de purifier le fragment obtenu et de le séquencer pour valider l'intégrité de la région mutée. L'exemple de la séquence nucléotidique obtenue pour la région mutée dans le cas de la délétion de *cpkC* est présenté dans la Figure 87.

>Cas9_ΔCpkC

```
cgggtcaaggcgacgtgggacccgaggaacgtcttcagccaccgtctcggcatccgcaagccgtagccgtagccacagcca
 R  V  K  A  T  W  D  P  R  N  V  F  S  H  R  L  G  I  R  K  P  –
cagccgtggaaagaagccccgccggcctcgggccggcgggggctttcgcgtgcggtggg gcgtaatcaacttctcctga
gagccgggtgggggcggaattcggggtcgagggccatcagttcgccgatcgtccgagctggttgtcgatgaagtcgaagag
                              –  N  A  S  R  G  L  Q  N  D  I  F  D  F  L
```

Figure 87. Séquence nucléotidique résultant de la délétion de l'ORF *cpkC* chez *Streptomyces coelicolor*. Le gène *scF* présent en aval de *cpkC* est représenté en gris clair et le gène codant la sous-unité CpkB retrouvée en amont est représenté en gris foncé. La séquence du protospacer n'est ici ni représentée ni mutée car elle se trouve au sein du gène délété. La position à laquelle la fusion entre les extrémités 3′ et 5′ du gène *cpkC* a été faite est représentée par un rectangle vide.

La même méthode a été utilisée pour la construction des souches délétées de *cpkA*, *cpkC* ou *cpkF*. Toutes les souches obtenues lors de ces travaux de thèse sont présentées dans le Tableau 20.

Tableau 20. Liste des souches construites au cours de ce travail pour la localisation de la coelimycine synthase.

Souche	Génotype	Description
FM145-MT-L1	*mYPet-cpkA*	Insertion de *mYPet* à l'extrémité 5′ de *cpkA*
FM145-MT-L2	*cpkC-mTurquoise2*	Insertion de *mTurquoise2* à l'extrémité 3′ de *cpkC*
FM145-MT-L3	*mYPet-cpkA, cpkC-mTurquoise2*	Insertion de *mYPet* à l'extrémité 5′ de *cpkA* et insertion de *mTurquoise2* à l'extrémité 3′ de *cpkC*
FM145-MT-L4	*mYPet-cpkA*, *cpkF-mCherry*	Insertion de *mYPet* à l'extrémité 5′ de *cpkA* et insertion de *mCherry* à l'extrémité 3′ de *cpkF*
FM145-MT-L5	*cpkC-mTurquoise2*, *cpkF-mCherry*	Insertion de *mTurquoise2* à l'extrémité 3′ de *cpkC* et insertion de *mCherry* à l'extrémité 3′ de *cpkF*
FM145-MT-L6	Δ*cpkA*	Délétion de *cpkA*
FM145-MT-L7	Δ*cpkC*	Délétion de *cpkC*
FM145-MT-L8	Δ*cpkA*, *cpkC-mTurquoise2*	Délétion de *cpkA* et insertion de *mTurquoise2* à l'extrémité 3′ de *cpkC*
FM145-MT-L9	Δ*cpkF*, *cpkC-mTurquoise2*	Délétion de *cpkF* et insertion de *mTurquoise2* à l'extrémité 3′ de *cpkC*
FM145-MT-L10	Δ*cpkC*, *mYPet-cpkA*	Délétion de *cpkC* et insertion de *mYPet* à l'extrémité 5′ de *cpkA*
FM145-MT-L11	Δ*cpkA*, *cpkF-mCherry*	Délétion de *cpkA* et insertion de *mCherry* à l'extrémité 3′ de *cpkF*
FM145-MT-L12	Δ*cpkC*, *cpkF-mCherry*	Délétion de *cpkC* et insertion de *mCherry* à l'extrémité 3′ de *cpkF*

II.C. Observation de la localisation de Cpk par microscopie confocale

L'observation en microscopie confocale des différentes souches a été faite en collaboration avec l'équipe de Pr. Angel Manteca (Université d'Oviedo, Espagne). Un échange de deux mois au sein de cette équipe a été rendu possible par l'obtention d'une bourse *Doctor, Eplore and Achieve More* (DrEAM) du dispositif Lorraine Université d'Excellence (LUE). Dans un premier temps, les paramètres expérimentaux ont dû être déterminés et optimisés, afin par la suite d'observer les souches d'intérêts.

II.C.1. Ajustement des paramètres pour l'observation des souches de Streptomyces coelicolor *par microscopie confocale*

L'objectif de ce projet est de localiser certaines protéines du complexe enzymatique de la coelimycine synthase fusionnées à des protéines fluorescentes par microscopie confocale. Durant sa thèse, le Dr. Jean-Malo Massicard avait établi que 24 h était le temps de culture optimal pour ces observations microscopiques (Massicard, 2020). L'équipe de Wellington *et coll*. a également étudié l'expression des gènes du cluster *cpk* par la quantification des ARNm correspondants entre 20 et 60 heures de culture, et a établi que ceux-ci sont exprimés à leur maximum aux alentours de 27 h de culture à 30°C (Nieselt et al., 2010). Des échantillons ont été observés à 27 h, 48 h et 72 h de culture.

Les protéines fluorescentes ont été observées aussi bien après 27 h et 48 h de culture, cependant au bout de 72 h aucun signal n'a pu être observé. Les conditions préalablement établies ont donc été suivies, et un temps de culture de 27 h a été choisi pour toutes les souches observées.

Deux méthodes ont également été testées afin de récupérer le mycélium aérien à observer au microscope. La première méthode consiste à placer des lamelles en verre dans la gélose avec un angle de 45°. L'échantillon est ensuite déposé au pied de la lamelle, à la jonction avec la gélose. La seconde consiste à placer une feuille de cellophane sur la gélose et d'y étaler l'échantillon. Ensuite, environ 1cm² de cellophane est découpé et placé entre lame et lamelle avant l'observation au microscope. La première méthode, utilisée par le Dr. Jean-Malo Massicard durant sa thèse pour observer les premières souches du projet, n'a ici pas permis d'obtenir de cliché. En effet, les cellules bactériennes n'ont pas pu être récupérées sur les lamelles présentes dans la gélose. En revanche, la méthode utilisant le cellophane a permis d'observer des colonies bactériennes et d'obtenir des clichés. C'est donc cette seconde approche qui est utilisée par la suite.

Un autre paramètre important a dû être vérifié et optimisé : les longueurs d'onde d'excitation pour l'observation de mYPet et mTurquoise2. Le spectre d'excitation de mTurquoise2 montre deux *maxima*. Le premier est à 435 nm et le second à 453 nm. Cependant, le microscope utilisé ne permettait pas d'exciter les échantillons à ces longueurs d'onde précises. Le microscope utilisé était équipé cependant d'un laser pouvant exciter à 405 nm, et d'un laser pouvant être ajusté afin d'exciter à la longueur d'onde choisie entre 470 et 600 nm. Ainsi, les longueurs d'onde de 405 et 470 nm ont été testées, aucune autre ne pouvant être testée entre ces deux valeurs. Le signal émis par mTurquoise2 a été respectivement collecté de 470 à 525 nm et de 480 à 525 nm selon la longueur d'onde d'excitation utilisée. Les clichés obtenus montrent que la souche FM145 a une auto-fluorescence importante à 405 nm (Figure 88). Cette auto-fluorescence n'est cependant pas observée à 470 nm. C'est donc cette dernière longueur d'onde qui a été sélectionnée pour l'observation de mTurquoise2. De plus, la longueur d'onde permettant l'excitation de mYPet (517 nm) n'induit pas d'auto-fluorescence chez FM145, ce qui permet par la suite d'observer le signal de la protéine fluorescente au sein des souches mutées (Figure 88).

La protéine mYPet a un maximum d'excitation aux alentours de 517 nm, qui a donc été la longueur d'onde d'excitation utilisée. L'émission a dans ce cas été récoltée de 525 à 600 nm. Les résultats obtenus lors de l'observation de la souche sauvage FM145 montrent que celle-ci ne montre pas d'auto-fluorescence dans ces conditions (Figure 88). Ces paramètres sont donc sélectionnés pour la suite.

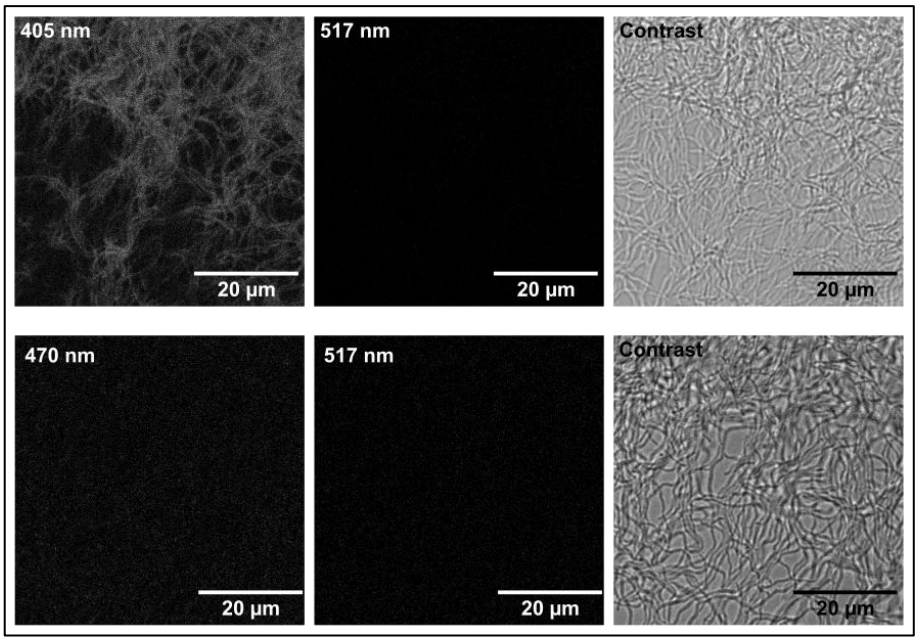


Figure 88. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145 pour l'optimisation des longueurs d'onde d'excitation. Deux longueurs d'onde ont été testées dans le cadre de l'excitation de mTurquoise2 : 405 et 470 nm. Pour ce qui est de mYPet, la longueur d'onde d'excitation est de 517 nm. Afin de valider l'utilisation de ces paramètres, la souche sauvage FM145 a été observée. Lorsque celle-ci est excitée à 405 nm, l'auto-fluorescence est trop importante pour y distinguer le signal d'une protéine fluorescente (*haut*). A 470 nm, l'auto-fluorescence de la souche est très faible, quasi-inexistante (*bas*). Pour ce qui est de l'excitation à 517 nm, aucune auto-fluorescence n'est détectée.

Cependant, la longueur d'onde d'excitation sélectionnée afin d'observer mTurquoise2 (470 nm) permet également l'excitation de mYPet. Afin de valider les résultats obtenus lors de l'observation des deux protéines au sein d'une même souche, il a fallu s'assurer que les signaux observés soient bien propres aux protéines fluorescentes d'intérêt. Pour cela, la souche exprimant la protéine mYPet fusionnée à CpkA a été observée dans les conditions d'observation de la protéine mTurquoise2. La longueur d'onde d'excitation utilisée a dans ce cas était 470 nm, et l'émission a été récoltée de 480 à 525 nm (Figure 89).

Les résultats obtenus permettent de valider les conditions d'observation de mTurquoise2, car celles-ci ne permettent pas l'observation de mYPet. Le signal détecté dans ces conditions au sein de la souche exprimant les deux protéines de fusion mYPet-CpkA et CpkC-mTurquoise2 sera donc bien spécifique de cette seconde.

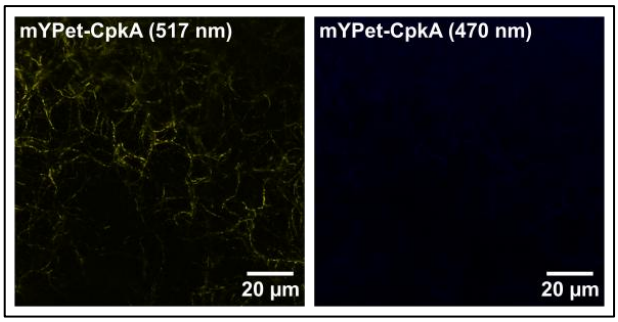


Figure 89. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L1 dans le but de valider les paramètres d'observation. La souche FM145-MT-L1 exprime la première sous-unité de la coelimycine synthase (CpkA) fusionnée en position N-terminale à la protéine fluorescente mYPet. Cette protéine fluorescente présente un maximum d'excitation à 517 nm, mais peut également dans une moindre mesure être excitée à 470 nm, soit la longueur d'onde utilisée pour l'excitation de mTurquoise2. La souche FM145-MT-L1 a donc été observé à la fois en excitant mYPet à une longueur d'onde de 517 nm et en récoltant l'absorbance entre 525 et 600 nm (*gauche*) et en l'excitant à une longueur d'onde de 470 nm, en récoltant cette fois-ci le signal émis entre 480 et 525 nm (*droite*). Les résultats obtenus montrent que la protéine mYPet n'est pas observable dans les conditions utilisées pour l'observation de mTurquoise2.

II.C.2. Localisation des sous-unités de la coelimycine synthase fusionnées à des protéines fluorescentes

Une fois les paramètres d'observation des protéines fluorescentes validés, les souches mutées ont pu être observées. Dans un premier temps, la localisation des protéines mYPet-CpkA et CpkC-mTurquoise2 a été étudiée. Pour cela, les souches FM145-MT-L1 et FM145-MT-L2 ont été utilisées. La première exprime la sous-unité CpkA fusionnée en N-terminal à la protéine fluorescente mYPet, tandis que la seconde exprime la sous-unité CpkC fusionnée en C-terminal à la protéine fluorescente mTurquoise2. Les clichés obtenus avec la souche FM145-MT-L1 ont permis d'observer que mYPet-CpkA n'est pas exprimée dans toute la colonie bactérienne, mais n'est retrouvée que dans quelques îlots d'expression. Dans ces cas, elle s'organise en foci au sein des cellules (Figure 90).

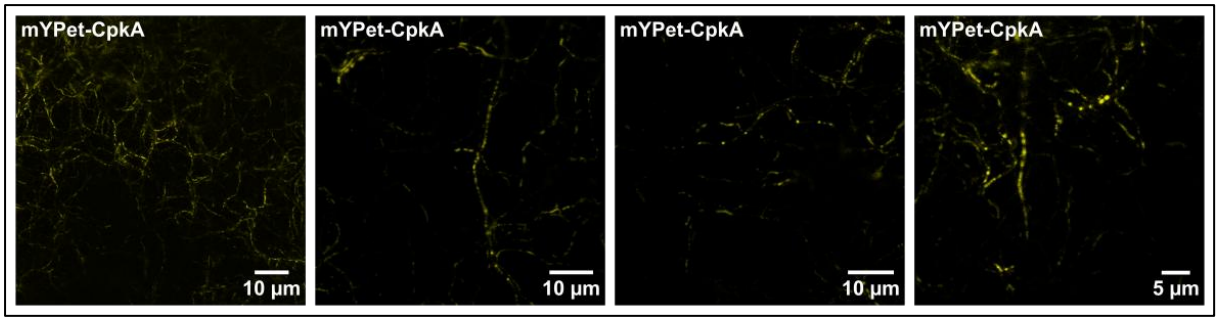


Figure 90. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L1. La longueur d'onde d'excitation est de 517 nm et l'émission est captée de 525 à 600 nm. La souche FM145-MT-L1 exprime la protéine de fusion mYPet-CpkA. Les clichés montrent que la protéine de fusion mYPet-CpkA forme des foci au sein des hyphes bactériens. Ces foci ne sont pas observés au sein de toute la colonie mais seulement au sein de quelques îlots d'expression.

Les mêmes observations ont pu être faites avec la protéine CpkC-mTurquoise2 exprimée dans la souche FM145-MT-L2. Cependant, la longueur d'onde utilisée pour exciter mTurquoise2 n'étant pas la plus optimale, son signal d'émission était faible et difficile à détecter. Un cliché a cependant pu être obtenu montrant des foci au sein des hyphes bactériens.

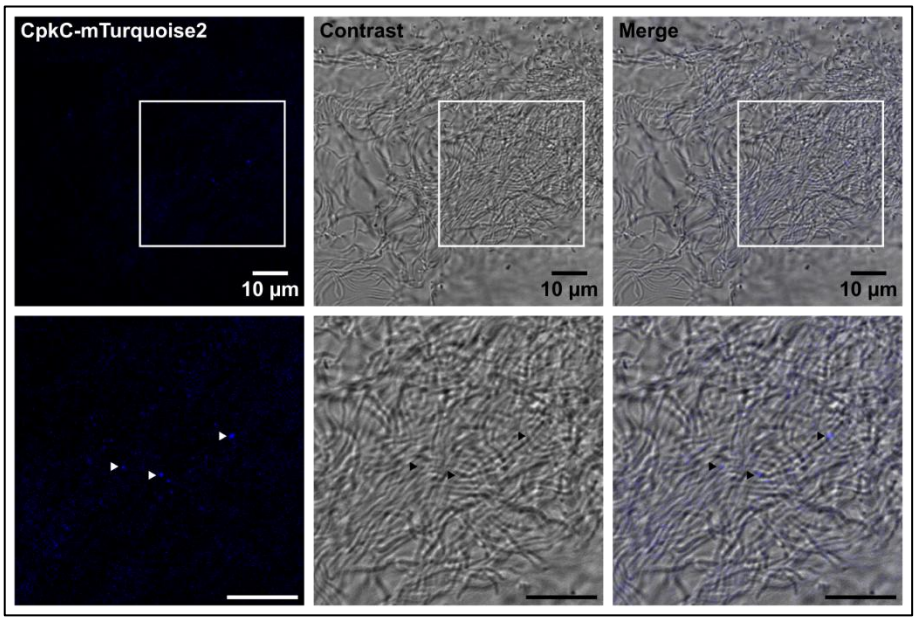


Figure 91. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L2. La longueur d'onde d'excitation est de 470 nm et le signal émis est récolté de 480 à 525 nm. La souche FM145-MT-L2 exprime la protéine de fusion CpkC-mTurquoise2. Les clichés présentés en bas sont des grossissements numériques des images de microscopie (*haut*). Les flèches indiquent les foci observés.

II.C.3. Colocalisation des protéines mYPet-CpkA et CpkC-mTurquoise2

L'un des objectifs de ce projet est d'étudier la colocalisation potentielle des deux sous-unités au sein des colonies de *S. coelicolor*. Pour cela, la souche FM145-MT-L3 a été construite. Celle-ci exprime les protéines de fusion mYPet-CpkA et CpkC-mTurquoise2. L'observation de cette souche révèle que les deux protéines forment également des foci au sein d'îlots d'expression (Figure 92). De plus, ces foci colocalisent parfaitement au sein des hyphes bactériens. Ici, la protéine de fusion CpkC-mTurquoise2 a été observée avec plus de facilité que précédemment. Cela peut s'expliquer par le fait que les protéines mYPet-CpkA et CpkC-mTurquoise2 colocalisent, et ainsi la présence du signal lié à mYPet-CpkA permet d'orienter la recherche du signal lié à mTurquoise2.

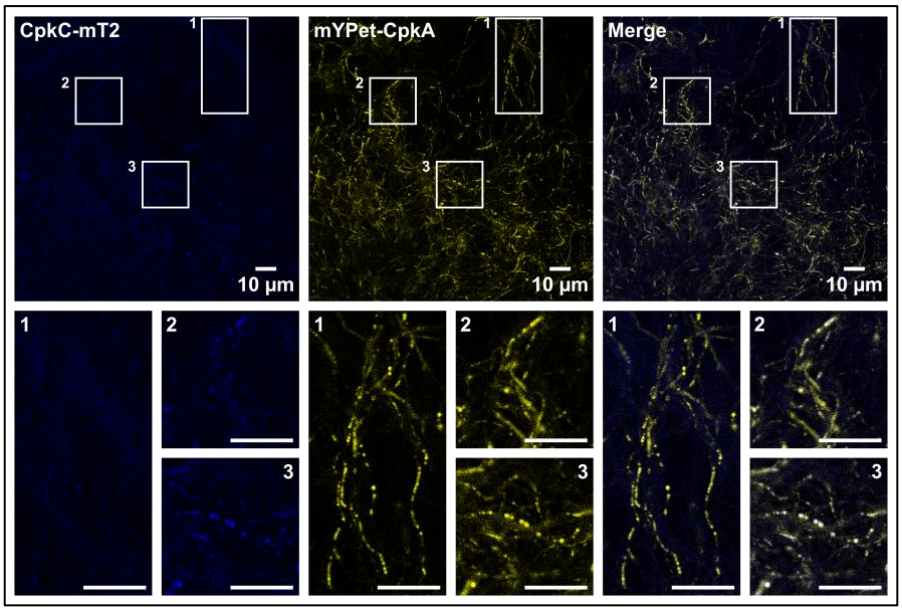


Figure 92. Cliché obtenu par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L3. La souche FM145-MT-L3 permet d'observer la localisation conjointe des protéines de fusion mYPet-CpkA et CpkC-mTurquoise2. La longueur d'onde d'excitation est de 470 nm pour le canal CpkC-mTurquoise2, et l'émission est captée de 480 à 525 nm. Pour le canal mYPet-CpkA, la longueur d'onde d'excitation est de 517 nm et l'émission est captée de 525 à 600 nm. Les images 1, 2 et 3 sont des grossissements numériques des images de microscopie. Les clichés obtenus montrent qu'au sein de cette souche, les protéines mYPet-CpkA et CpkC-mTurquoise2 s'organisent également en foci au sein des hyphes bactériens, et localisés dans des îlots d'expression. De plus, les signaux des deux protéines de fusion colocalisent parfaitement.

II.C.1. Localisation des sous-unités marquées en absence de leur partenaire

L'un des objectifs durant ce projet a été de déterminer l'organisation des sous-unités CpkA et CpkC en l'absence de leur partenaire, dans le but d'établir si la macro-organisation d'une sous-unité est dépendante de la présence des autres ou si les bases moléculaires de celle-ci leur sont intrinsèques. Pour cela, les souches FM145-MT-L8 et FM145-MT-L10 ont été conçues. La première – FM145-MT-L8 – exprime la sous-unité CpkC fusionnée à la protéine fluorescente mTurquoise2, et le gène *cpkA* y est délété. Au regard des résultats obtenus précédemment lors de l'observation de la protéine CpkC-mTurquoise2, l'utilisation de cette souche n'a pas été possible car le signal lié à mTurquoise2 est trop faible. En effet, le cliché présentant l'organisation en foci de CpkC-mTurquoise2 étant le seul obtenu lors des observations faites à partir de la souche FM145-MT-L2, les données récoltées n'auraient pas été suffisante pour analyser d'autres mutants exprimant uniquement mTurquoise2 et non pas mYPet.

La seconde – FM145-MT-L10 – contient à l'inverse la sous-unité CpkA marquée par la fusion à la protéine mYPet, et le gène *cpkC* y est délété. Les clichés obtenus lors des observations de cette souche en microscopie confocale montrent que mYPet-CpkA est toujours exprimée au sein d'îlots de la colonie, et qu'elle semble également s'organiser en foci (Figure 93).

Ainsi, l'organisation et la localisation de mYPet-CpkA ne semble pas dépendre de la présence de la dernière sous-unité du système CpkC. Une analyse informatique est en cours afin de pouvoir déterminer automatiquement la taille des foci observés et ainsi valider si l'absence de CpkC impacte ou non l'organisation de mYPet-CpkA.

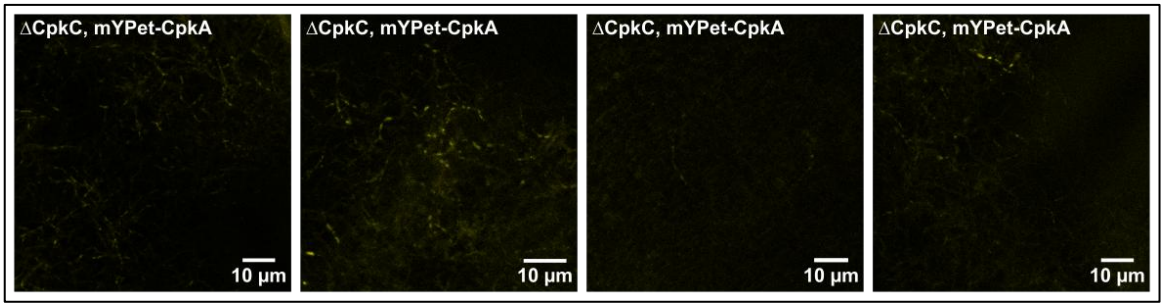


Figure 93. Clichés obtenus par microscopie confocale de la souche *Streptomyces coelicolor* FM145-MT-L10. La longueur d'onde d'excitation est de 517 nm et l'émission est captée de 525 à 600 nm. La souche FM145-MT-L10 permet d'observer la localisation de la protéine de fusion mYPet-CpkA en l'absence de la sous-unité CpkC. Ces observations montrent que mYPet-CpkA forme des foci au sein des hyphes bactériens, et donc que cette organisation est indépendante de la présence de la sous-unité CpkC.

II.C.2. Localisation du transporteur CpkF fusionné à la protéine fluorescente mCherry

La localisation du transporteur de la coelimycine – CpkF – a également été étudiée. Pour cela, la souche FM145-JMM4 a été utilisée. Le transporteur CpkF y est fusionné à la protéine fluorescente mCherry. La protéine mCherry a été excitée à une longueur d'onde de 587 nm et l'émission a été récolée entre 600 et 750 nm. Les clichés obtenus par microscopie ont été traités par déconvolution. Les observations faites en collaboration avec le Dr. Manteca montrent que CpkF se trouve au niveau de la membrane bactérienne, comme attendue, et qu'il y forme également des foci (Figure 94).

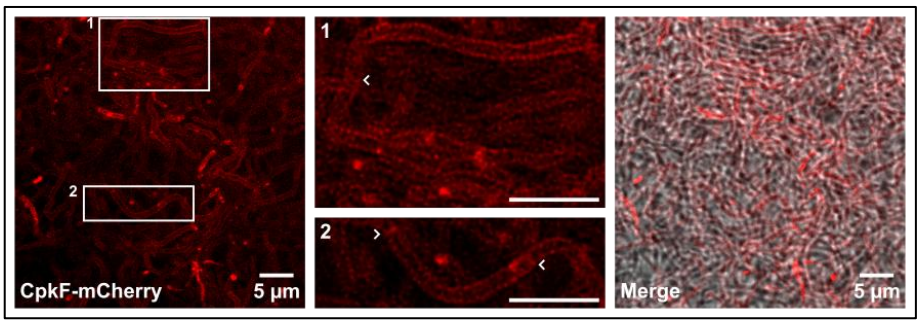


Figure 94. Cliché obtenu par microscopie confocale de la souche *Streptomyces coelicolor* FM145-JMM4. La souche FM145-JMM4 exprime le transporteur CpkF fusionné à la protéine fluorescence mCherry. La longueur d'onde d'excitation utilisée est de 580 nm et l'émission est récoltée entre 600 et 750 nm. Les images au milieu représentent des agrandissements numériques du cliché de gauche, et celui de droite correspond à la superposition avec l'image obtenue des cellules. Les analyses de cette souche ont permis d'observer que le transporteur est exprimé au sein de tous les hyphes des colonies. De plus, il est retrouvé organisé en foci au sein de la membrane bactérienne mais également dans les *septa*, indiqués par des flèches blanches.

A la différence des sous-unités CpkA et CpkC, le transporteur CpkF est retrouvé au niveau de toute la colonie.

L'étude de la localisation et de l'organisation de protéines issues du complexe enzymatique de la coelimycine synthase a permis d'obtenir des informations cruciales pour l'étude des PKS *cis*-AT. Tout d'abord, la localisation des première et dernière sous-unités – CpkA et CpkC – a permis de démontrer que celles-ci s'organisent en foci qui colocalisent au sein d'îlots d'hyphes bactériens. La délétion du partenaire CpkC ne semble pas affecter l'organisation de CpkA, impliquant que la macro-organisation du complexe ne dépendrait pas de la présence de tous les partenaires et que les bases moléculaires pourraient être intrinsèques aux protéines seules. En parallèle, la localisation du transporteur CpkF a permis de déterminer que celui-ci se trouve au niveau de la membrane de tous les hyphes de la colonie.

Les résultats présentés dans cette partie sur la localisation et l'organisation de la coelimycine synthase chez *S. coelicolor* a déjà fait l'objet de la rédaction d'un article scientifique qui sera soumis sous peu une fois les derniers contrôles et les dernières expérimentations effectués. Le brouillon de l'article peut être retrouvé en annexe (page xxxvii)

Discussion

L'organisation des PKS modulaires en fait un objet d'étude privilégié pour l'ingénierie moléculaire dans le but d'obtenir et optimiser de nouvelles molécules d'intérêt. Cependant à ce jour, la majorité des tentatives d'ingénierie des PKS se sont soldées par l'obtention de systèmes enzymatiques aux rendements de production plus faibles que les systèmes natifs (Ranganathan et al., 1999; Su et al., 2022; Wlodek et al., 2017). Une voie émergente et prometteuse afin d'augmenter ces rendements consiste scinder une sous-unité de ces systèmes (Buyachuihan et al., 2023; Liu et al., 2025; Miyazawa et al., 2020). Cela augmente la flexibilité de ces systèmes puisque, en majorité, les domaines d'interactions introduits ne forment pas de liaison covalente.

De plus, l'un des freins à l'ingénierie efficace des PKS est le manque d'information disponible concernant leur macro-organisation *in vivo*, qui permettrait de mieux appréhender l'impact de l'ingénierie sur leur efficacité. A ce propos, des études ont déjà été menées sur l'organisation des PKS modulaires *trans*-AT, et des motifs d'interactions ont pu être identifiés, mais rien de cela n'a été observé chez les PKS *cis*-AT (Gay et al., 2016; Straight et al., 2007; Tittes et al., 2022).

Dans ce cadre, ces travaux de thèse ont été menés autour de deux principaux objectifs. Le premier a été d'investiguer la scission de deux PKS modulaires dans le but de déterminer si cela peut avoir un effet positif sur les rendements de production. Le second a été d'étudier la localisation et macro-organisation d'une PKS modulaire *cis*-AT au sein de son hôte bactérien.

I. Investiguer l'effet de la scission sur les rendements de production

La stratégie de scission de PKS dans le but d'optimiser les rendements de production a été menée sur les systèmes de la rapamycine synthase (RAPS) chez *S. hygroscopicus* et de la stambomycine synthase de *S. ambofaciens*.

I.A. Etude de la rapamycine synthase

La rapamycine synthase est une PKS modulaire *cis*-AT composée de quatorze modules d'élongation répartis au sein de trois sous-unités (RAPSA–C) (Figure 51, page 71) (Aparicio et al., 1996). La seconde sous-unité – RAPSB – a été sélectionnée dans le cadre de la scission car c'est la plus grande, composée de six modules, tandis que les sous-unités RAPSA et RAPSC contiennent respectivement cinq et quatre modules. RAPS contient donc deux paires de domaines de docking (DD) à la jonction des sous-unités RAPSA-RAPSB et RAPSB-RAPSC. L'analyse de ces domaines a permis de classer la première paire au sein des DD de type 1b et la seconde dans la famille des DD de type 1a.

Le point de scission choisi se trouve entre le troisième et le quatrième module de RAPSB, afin d'obtenir deux chaînes polypeptidiques de trois modules chacune. Par la suite, d'autres combinaisons seront envisageables, afin d'obtenir une chaîne de deux modules suivie d'une chaîne de quatre modules ou inversement.

Afin de scinder RAPSB, une paire de DD hétérologue a été sélectionnée, provenant du système de la 6-désoxyérythronolide B synthase (DEBS) qui est une PKS *cis*-AT largement étudiée depuis sa caractérisation en 1990 (Cortes et al., 1990). DEBS est composée également de trois sous-unités (DEBS1–3) et contient donc deux paires de DD de type 1a. L'analyse des résidus clefs impliqués dans les interactions entre chaque paire a permis d'établir que la paire $^{C}DD_{DEBS2}$ / $^{N}DD_{DEBS3}$ avait un profil différent de $^{C}DD_{RAPSB}$ / $^{N}DD_{RAPSC}$ et c'est donc celle-ci qui a été sélectionnée (Figure 67, page 120). Afin de valider cette première analyse *in silico*, des analyses biophysiques ont été effectuées afin de caractériser les interactions entre chaque domaine.

I.A.1. Analyse biophysique des interactions entre les domaines de docking

Pour cela, les domaines utilisés provenant de DEBS – $^{C}DD_{DEBS2}$ et $^{N}DD_{DEBS3}$ – ont été produits et purifiés. Certains domaines provenant de RAPS ont pu également être produits et purifiés ($^{C}DD_{RAPSA}$ et $^{C}DD_{RAPSB}$), ou commandés sous forme de peptide synthétique ($^{N}DD_{RAPSC}$) ou n'ont pas pu être obtenus à ce jour ($^{C}DD_{RAPSB}$). Le domaine $^{N}DD_{RAPSC}$ cependant n'a pas pu être solubilisé et le domaine $^{C}DD_{RAPSA}$ forme des agrégats. Par manque de temps, ces domaines n'ont pas pu être obtenus durant ces travaux de thèse, et donc étudiés. Les seuls domaines disponibles pour l'étude de leurs interactions ont été les domaines $^{C}DD_{DEBS2}$, $^{N}DD_{DEBS3}$ et $^{C}DD_{RAPSB}$.

Afin de valider l'utilisation du titrage calorimétrique isotherme (ITC) comme méthode d'observation des interactions, le couple $^{C}DD_{DEBS2}$ / $^{N}DD_{DEBS3}$ a tout d'abord été étudié comme contrôle. Dans un premier temps, le domaine $^{C}DD_{DEBS2}$ a été utilisé afin de titrer son partenaire $^{N}DD_{DEBS3}$. L'interaction entre ces deux domaines a bien été observée et caractérisée par une constante de dissociation (K_D) d'environ 22,17 ± 15,67 µM et une stœchiométrie de 1,42 ± 0,272 (Figure 76A, page 131). D'autres paires de DD ont également été étudiées par ITC, qui peuvent montrer des valeurs de K_D plus faibles que celle obtenue pour la paire $^{C}DD_{DEBS2}$ / $^{N}DD_{DEBS3}$ – de 3,2 ± 0,3 – 5,7 ± 0,6 µM pour les DD retrouvés au sein de PKS *trans*-AT synthétisant le toblerol (Scat et al., 2024) – ou environ égale : 14,5 ± 0,9 µM dans le cas de l'étude d'une paire de DD provenant de la stambomycine synthase (Su et al., 2022). Cette valeur obtenue ici ne semble donc pas aberrante, cependant la marge d'erreur sur la valeur est élevée et il est donc nécessaire d'affiner les données obtenues lors de nouvelles expérimentations.

De plus, cette paire de DD présente entre DEBS2 et DEBS3 a déjà pu être caractérisée par l'équipe de Ma *et coll*. et la valeur de K_D obtenue ici n'est pas en adéquation avec la leur, qui est de 0,28 ± 0,08 µM (Sun et al., 2022). Cette valeur a été déterminée par interférométrie. Pour cela, le domaine $^{N}DD_{DEBS3}$ a été biotinylé puis immobilisé et l'interaction avec le domaine $^{C}DD_{DEBS2}$ a été caractérisée par plusieurs expérimentations à différentes concentrations de ce second partenaire. Les constantes de dissociation ont été déterminées par la variation de la longueur d'onde appliquée sur l'échantillon. Cette méthode est plus sensible que l'ITC utilisé durant cette thèse, il est donc nécessaire d'approfondir les mesures faites ici dans le but de valider ou non les données obtenues. Bien que les données obtenues au cours de cette thèse démontrent la présence d'une interaction entre les domaines $^{C}DD_{DEBS2}$ et $^{N}DD_{DEBS3}$, les constantes associées à cette interaction peuvent être affinées, notamment en variant les concentrations des partenaires utilisés afin d'obtenir les plateaux nécessaires à une meilleure modélisation des données.

Les analyses par ITC ont également été menées en titrant le domaine $^{C}DD_{DEBS2}$ par le domaine $^{N}DD_{DEBS3}$ (Figure 76B, page 131). Dans ce cas, les données obtenues sont caractéristiques de la présence de deux événements au cours de la titration. Cela signifie que deux sites d'interaction sont observés durant cette expérimentation. Ce phénomène n'est pas observé dans le sens inverse – lorsque le domaine $^{N}DD_{DEBS3}$ est utilisé pour titrer $^{C}DD_{DEBS2}$ – ce qui signifie que cela doit être dû à la concentration à laquelle le domaine $^{N}DD_{DEBS3}$ est utilisé. En effet, dans le cas précédent ayant permis d'observer un unique événement, le domaine $^{N}DD_{DEBS3}$ a été utilisé à une concentration quasi-constante de 80 µM au sein de la cellule du dispositif. Le domaine $^{C}DD_{DEBS2}$, quant à lui, a été placé dans la seringue du dispositif à une concentration de 800 µM. Lorsque la titration se fait dans l'autre sens, la concentration en domaine $^{N}DD_{DEBS3}$ varie dans le temps de 0 à 96,30 µM au sein de la cuve. Il suffit alors que ce domaine médie une interaction – comme sa dimérisation par la formation de *coiled coil* (Figure 33) – dont le K_D est inférieur à 80 µM pour que celui soit observé dans cette seconde condition et non lors de la première configuration. Les données obtenues restent à être affinées également car l'absence de plateaux au début et à la fin de la courbe à la suite du titrage ne permet pas de déterminer les constantes des événements observés.

Il sera également nécessaire d'effectuer les tests d'interaction avec les autres domaines de cette étude, une fois que ceux-ci auront pu être obtenus soit par production et purification, soit sous forme de peptide synthétique. Une fois les interactions entre chaque domaine caractérisées, il sera nécessaire d'évaluer l'impact de la scission sur les rendements de production. Pour cela, les souches sauvage – *S. hygroscopicus* DSM41530 – et surproductrice de la rapamycine – *S. hygroscopicus* ISOM-3410 – seront utilisées. La souche DSM41530 a pu ainsi être cultivée au sein du laboratoire.

Cependant, la souche ISOM-3410 pose plus de soucis puisque l'étape de sporulation de celle-ci n'a pas pu être atteinte dans les conditions étudiées. Il sera donc nécessaire d'optimiser les conditions de culture de cette souche afin de récolter des spores et pouvoir construire les souches mutantes par conjugaison. Les conditions de conjugaison de la souche DSM41650 ont pu être déterminées à partir du vecteur pKCCpf1 vide, et il sera par la suite nécessaire de voir si ces conditions sont applicables à la souche surproductrice de la rapamycine.

I.A.2. Stratégies d'ingénierie de la rapamycine synthase

Une fois les souches mutantes de *S. hygroscopicus* obtenues, des analyses par spectrométrie de masse permettront d'évaluer l'impact de la scission effectuée sur les rendements de production. Si ces rendements sont augmentés par la scission *in vivo* de la PKS, cela permettra de conclure quant au lien positif entre la taille des sous-unités d'un système et sa productivité. Autrement, cela signifiera soit que la taille des sous-unités n'est pas un point d'optimisation pour la production de polycétides et de dérivés, ou que la stratégie de scission adoptée n'est pas optimale en elle-même.

Par la suite, il sera intéressant d'utiliser d'autres interfaces d'interaction afin d'étudier l'impact de différents types de scission. L'étude menée par Grininger *et coll.* a permis de comparer l'utilisation de différentes interfaces d'interaction entre des sous-unités de PKS (Buyachuihan et al., 2023). Dans ce cadre, cette équipe a, entre autres, utilisé la paire de DD présente entre les sous-unités 2 et 3 de DEBS ou la paire de SynZip 3:4 afin de remplacer les domaines natifs retrouvés entre les deux sous-unités de la vénémycine synthase. Les résultats *in vitro* obtenus ont permis de démontrer que l'utilisation de DD hétérologues ou de domaines SynZip impacte la productivité de la PKS ainsi modifiée de la même manière (Figure 57, 3 et 5, page 78). Scinder *in vivo* le système de la rapamycine synthase en utilisant également les domaines SynZip 3:4 apporterait donc de nouvelles informations complémentaires concernant l'efficacité de ces domaines dans le cadre de l'ingénierie des PKS. De plus, il sera intéressant d'utiliser des DD de type 2 ou provenant de PKS *trans*-AT, afin de déterminer leur impact sur la productivité de RAPS. Utiliser de tels domaines au sein de systèmes PKS contenant des DD natifs de type 1 permettrait à la fois de s'affranchir des études d'interactions entre chaque domaine et d'utiliser des interfaces d'interactions natifs de systèmes PKS.

L'étude menée par Wang *et coll*., quant à elle, s'est intéressée à l'étude de la scission *in vivo* de la buténylе-spinosyne synthase (Bus) (Liu et al., 2025). Ce travail a étudié l'impact de la scission de la première sous-unité par l'insertion des DD présents entre les sous-unités 1 et 2 de la salinomycine synthase (Sln). Les systèmes modifiés ont ensuite été exprimés de manière hétérologue chez *S. albus* J1074. Cette scission a montré que la synthèse de sous-unités de taille réduite a un impact positif sur la productivité de la PKS (Figure 59, page 81).

Ainsi, les systèmes scindés de Bus ont une productivité environ 5 à 6 fois supérieure à celle de la PKS sauvage (Figure 59) (Liu et al., 2025). Dans cette étude, le promoteur constitutif fort *kas*Op* a également été intégré en amont de la séquence codant le ^{N}DD intégré. Cela a ainsi diminué l'impact de la présence d'ARN messager (ARNm) tronqués sur l'efficacité de synthèse de la seconde sous-unité, et ainsi permis d'améliorer la productivité de la PKS scindée d'un facteur d'environ 11 à 13 par rapport à la PKS native.

L'un des objectifs futurs dans l'étude de l'impact de la scission de RAPS *in vivo* pourrait consister à insérer un promoteur au sein de l'opéron *rapsA-rapsB*. Dans le but de garder un système PKS modifié qui reste proche de ce qui peut être nativement retrouvé, le promoteur présent en amont de l'opéron contenant les gènes codant les sous-unités RAPSA et RAPSB pourra dans un premier temps être dupliqué en amont du gène *rapsB*. De plus, ce promoteur pourra également être intégré entre les séquences codant les domaines $^{C}DD_{DEBS2}$ et $^{N}DD_{DEBS3}$ au sein de la souche exprimant la rapamycine synthase scindée. Cette stratégie pourra également être appliquée à l'utilisation des domaines SynZip et des domaines de type 2 ou provenant de PKS *trans*-AT. Ces expérimentations permettraient ainsi d'établir des connaissances générales sur l'utilisation de ces différents outils dans le but d'optimiser les rendements de production de PKS modifiées *in vivo*. Cela pourra servir à la constitution d'une véritable boîte à outil de biologie synthétique des PKS modulaires.

I.B. Etude de la stambomycine synthase

La stambomycine synthase est un système bien plus grand que RAPS. Elle comprend vingt-quatre modules d'élongation répartis au sein de neuf sous-unités (Pks1–9) (Figure 48, page 67) (Laureti et al., 2011). Elle a déjà fait l'objet d'ingénierie dans le but de produire des composés de taille réduite : les mini-stambomycines (Figure 49, page 68) (Su et al., 2022). Ce système modifié est délété des sous-unités Pks5 à Pks8 et ne contient donc plus que dix-sept modules d'élongation, répartis au sein de cinq sous-unités : Pks1–4 et Pks9.

Dans le cadre de cette thèse, la stratégie de scission a été appliquée au système natif de la stambomycine synthase et au système modifié produisant les mini-stambomycines. La troisième sous-unité Pks3 a été sélectionnée, car c'est la plus grande. Elle est composée de cinq modules et elle est présente dans les deux systèmes. Ici, la scission a été faite par l'insertion de domaines synthétiques appelés SynZip (Figure 55, page 76) (Thompson et al., 2012) entre le deuxième et le troisième module de la sous-unité Pks3. Cette stratégie a été préférée par rapport à l'insertion de domaines DD provenant d'une autre PKS dans le but d'éviter les interactions non canoniques avec les DD natifs, au nombre de huit paires dans le système natif.

Discussion

Deux paires de SynZip ont été sélectionnées pour ces travaux de thèse. Tout d'abord, la paire de SynZip3:4, dont les partenaires interagissent de manière parallèle pour former un dimère, avec un K_D inférieure à 30 nM (Thompson et al., 2012). Ces domaines ont déjà pu être utilisés dans le cadre de la scission de la vénémycine synthase afin de remplacer les DD natifs (Buyachuihan et al., 2023). L'utilisation des SynZip a montré le même effet sur les rendements de production que l'utilisation de la paire de DD présente entre les sous-unités 2 et 3 de DEBS, soit une diminution des rendements de production d'un facteur d'environ 2 par rapport à la PKS native. La paire de SynZip17:18 a également été choisie. Cette paire est la seule formant un dimère par le biais d'une interaction antiparallèle, avec un K_D inférieur à 10 nM (Thompson et al., 2012). A ce jour cette seconde paire n'a pas été utilisée dans le cadre d'ingénierie de mégacomplexes.

Les domaines SynZip ont une affinité bien plus forte que les domaines de docking natifs des PKS, qui ont un K_D de l'ordre du µM. Ils permettent tout de même d'intégrer de la flexibilité au sein du système en étant intégrés là ou une liaison covalente est nativement retrouvée. Il sera alors intéressant d'évaluer *in vivo* l'impact de la scission de PKS par l'utilisation de SynZip à celui de l'utilisation de domaines plus flexibles que sont les DD, utilisés dans le cas de la scission de RAPS. De plus, les paires de SynZip utilisées ici forment des interactions parallèles (SynZip3:4) et antiparallèle (SynZip17:18), et l'étude des rendements de production de chaque scission permettra d'obtenir des informations quant à l'orientation idéale des domaines.

Si cette stratégie de scission s'avère efficace pour l'optimisation des rendements de production de la PKS miniaturisée produisant les mini-stambomycines (Figure 49, page 68), l'objectif par la suite sera d'optimiser cette scission et l'extraction des composés dans le but d'en obtenir une quantité suffisante à leur caractérisation par résonnance magnétique nucléaire. A ce jour, la stambomycine synthase miniaturisée la plus optimisée ne permet un rendement de production que d'environ 0,1 $mg \cdot L^{-1}$ à 1,3 $mg \cdot L^{-1}$ des mini-stambomycines (**22**) et (**23**), contre un rendement de 30 ± 2 $mg \cdot L^{-1}$ de stambomycines (**21**) dans le cas de la PKS sauvage (Figure 50, page 69).

Les mini-stambomycines obtenus par le Dr. Li Su durant sa thèse sont dépourvues du résidu mycaminose présent au sein des stambomycines. Ce résidu est très certainement clef dans l'activité antibactérienne de ces composés, dont les mini-stambomycines sont donc dépourvue. L'objectif par la suite sera donc d'obtenir une quantité suffisante de composés afin d'effectuer l'ajout du résidu mycaminose par hémisynthèse chimique. Les activités biologiques des mini-stambomycines et de leur dérivés portant le résidu mycaminose seront ensuite étudiées et pourront ainsi être comparées.

II. Etudier l'organisation d'une PKS *cis*-AT au sein de son hôte bactérien

Le second objectif de cette thèse a été d'étudier la localisation et l'organisation d'une PKS *cis*-AT au sein de son hôte bactérien. Pour cela, le système de la coelimycine synthase (Cpk) retrouvé chez *S. coelicolor* a été sélectionné. Ce système est composé d'un module de chargement et cinq modules d'élongation, répartis au sein de trois sous-unités (CpkA–C) (Figure 18, page 26) (Gomez-Escribano et al., 2012). Cette PKS présente l'avantage d'être de petite taille et sa fonctionnalité peut être vérifiée par la production des pigments jaunes coelimycine P1 et P2 en milieu solide (Gomez-Escribano et al., 2012; Gottelt et al., 2010), ce qui en font un sujet d'étude idéal.

Ce projet de thèse constitue la continuité d'un projet du Dr. Jean-Malo Massicard durant sa thèse (Massicard, 2020). Il avait pu créer les souches de *S. coelicolor* FM145 exprimant la première sous-unité du système – CpkA – fusionnée en position N-terminale à la protéine fluorescente mCherry et la dernière sous-unité fusionnée en C-terminal à la protéine fluorescente eGFP (Figure 82, page 139). Les observations de ces souches faites en microscopie confocale de ces fusions, en collaboration avec l'équipe de Pr. Van Wezel à l'Université de Leiden (Pays-Bas), ont permis d'observer que ces deux protéines de fusion s'organisent en foci au sein d'îlots d'hyphes de la colonie bactérienne. Le fait que toute la colonie n'exprime pas le BGC lié à la production de la coelimycine est cohérent avec le modèle de division du travail décrit chez *Streptomyces* (Zhang et al., 2020).

Cependant, ces foci ne colocalisent pas au sein des cellules. Cela peut être dû à une réelle absence de colocalisation, ou à un biais induit par l'oligomérisation de la protéine eGFP pouvant interférer avec l'organisation native de CpkC (Jain et al., 2001). De plus, le signal lié à la présence de la protéine mCherry a été difficilement détectable à travers l'auto-fluorescence de la souche *S. coelicolor* FM145. L'objectif de ce projet de thèse a donc été de construire de nouveaux mutants, exprimant cette fois-ci les sous-unités CpkA et CpkC fusionnées respectivement aux protéines fluorescentes mYPet et mTurquoise2. L'observation en microscopie confocale de ces souches modifiées exprimant mYPet-CpkA et CpkC-mTurquoise2, en collaboration avec l'équipe du Dr. Ángel Manteca à l'Université d'Oviedo (Espagne), a permis également d'identifier la formation de foci au sein d'îlots d'hyphes de la colonie. De plus, ces foci colocalisent parfaitement dans les hyphes bactériens (Figure 92, page 150). Cela implique que la coelimycine synthase s'organise en usines de production au sein des cellules, et que cette observation (également faites au sein des PKS *trans*-AT) peut être étendue à toutes les PKS de type *cis*-AT. Le signal lié à la protéine mTurquoise2 a cependant été difficile à observer. En effet, la longueur d'onde optimale pour l'excitation de mTurquoise2 est de 435 nm.

Or, Cette longueur d'onde n'a pas pu être utilisée lors de l'observation des souches car le laser permettant d'exciter à cette longueur d'onde n'était pas disponible. La longueur d'onde de 405 nm a donc été utilisée, qui permet une excitation relative de 0,5 par rapport à la longueur d'onde de 435 nm. De plus, mTurquoise2 a une intensité d'excitation plus faible que mYPet d'environ 35 % (Goedhart et al., 2012; Nguyen & Daugherty, 2005). Le signal lié à la protéine mTurquoise2 au sein de la souche de *S. coelicolor* FM145 exprimant les protéines de fusion mYPet-CpkA et CpkC-mTurquoise2 a été plus simple à observer, car les foci colocalisant, il suffisait ainsi de trouver le signal lié à l'excitation de mYPet afin de trouver celui de son partenaire.

L'utilisation de protéines fluorescentes pour la visualisation d'éléments intracellulaires est un défi grandissant avec le nombre de protéines utilisées simultanément. Il faut dans ce cas que les protéines soient monomériques et que leur spectre soit assez distinct pour permettre leur observation. Cependant, il est difficile de trouver un couple de protéines remplissant ces critères et dont les intensités de fluorescence sont semblables. Les protéines monomériques vertes et jaunes sont couramment utilisées, couplées à des protéines bleues et rouges. Cependant, les protéines bleues et rouges ont souvent des intensités plus faibles. En 2022, une équipe a publié une nouvelle protéine fluorescente, appelée Electra2, qui a été utilisée dans le cadre d'imagerie cellulaire en 2025 (Hashimura et al., 2025; Papadaki et al., 2022). Cependant, cette protéine a un maximum d'excitation aux alentours de 405 nm, et la souche *S. coelicolor* utilisée montre une auto-fluorescence importante à cette longueur d'onde. Il n'est donc à ce jour pas envisageable ni possible d'utiliser un couple de protéines monomériques plus adéquat, à moins de développer de nouvelles protéines fluorescentes.

L'observation du transporteur CpkF, fusionné à la protéine fluorescente mCherry, montre quant à elle que cette protéine est exprimée au sein de tous les hyphes de la colonie et semble être membranaire. Ces résultats sont cohérents vis-à-vis de la localisation de certains complexes enzymatiques proches des membranes bactériennes et liés à leur système d'excrétion (Hantash & Earhart, 2000; Jain & Cox, 2005; Straight et al., 2007). En effet, cette omniprésence permet très certainement de limiter la cytotoxicité du produit synthétisé ainsi que de libérer les métabolites spécialisés impliqués dans la défense de la colonie, ici la coelimycine A qui est un antibiotique.

Durant sa thèse, le Dr. Jean-Malo Massicard avait également étudié la localisation des protéines de fusion CpkC-eGFP et CpF-mCherry au sein d'une même souche. Les données qu'il a ainsi obtenues n'ont cependant pas permis de conclure quant à un lien potentiel entre la localisation de ces deux protéines.

Au cours de cette thèse, de nouvelles souches ont été conçues dans le but d'étudier la localisation conjointe du transporteur CpkF et des sous-unités CpkA et CpkC. Les souches *S. coelicolor* FM145 exprimant à la fois la protéine de fusion mYPet-CpkA et CpkF-mCherry (FM145-MT-L4), ainsi que la souche exprimant CpkC-mTurquoise2 et CpkF-mCherry (FM145-MT-L5) ont été créées durant ces travaux de thèse. Cependant l'expression des protéines fluorescentes n'a pas pu être étudiée par microscopie confocale pour le moment car la fonctionnalité de Cpk fusionnées aux protéines mYPet et mTurquoise2 n'a pas pu être vérifiée.

L'étude du système de la bacillaene synthase au sein de son hôte bactérien *B. subtilis* a permis d'observer que la première et la dernière sous-unité de ce système se colocalisent au sein des cellules et sont retrouvés proches de la membrane (Figure 37B, page 50) (Straight et al., 2007). Cette proximité a été validée par des clichés obtenus en microscopie électronique à transmission (TEM) (Figure 38, page 50). La PKS responsable de la synthèse du dimycocérosate de phthiocérol (PDIM) retrouvée chez *M. tuberculosis* a également été caractérisée proche de la membrane bactérienne (Jain & Cox, 2005). Cette fois-ci, cette proximité a été observée par l'étude de l'interactome du transporteur MmpL7, responsable notamment de l'excrétion du PDIM. Ainsi, cette équipe a pu identifier une interaction entre la dernière enzyme impliquée dans la production du PDIM – la PKS PpsE – et le transporteur de cette molécule MmpL7. Cette proximité entre synthèse et excrétion des métabolites spécialisés permet à la fois de diminuer l'impact cytotoxique de la molécule pour son hôte et de faciliter son adressage. Il est donc possible d'envisager une proximité entre le complexe enzymatique Cpk et la membrane bactérienne, par le biais du transporteur CpkF. Cette hypothèse pourra être dans un premier temps vérifiée par l'analyse des souches FM145-MT-L4 et FM145-MT-L5.

Ce système marqué de la coelimycine synthase permettra par la suite d'établir les bases moléculaires d'une telle macro-organisation. Les domaines LINKS identifiés au sein des PKS *trans*-AT et impliqués dans leur macro-organisation étant absent des PKS *cis*-AT (Figure 41, page 53), leur complexation dépend d'autres motifs protéiques.

Ce système marqué de la coelimycine synthase pourra également faire l'objet de tentatives d'ingénierie dans le but d'étudier leur impact sur la localisation des enzymes du système. Il serait dans ce cas intéressant de scinder *in vivo* l'une des sous-unités de Cpk, en introduisant des DD provenant d'une autre PKS ou des domaines SynZip, et d'évaluer si cette scission peut avoir un impact sur la colocalisation des sous-unités du système. De plus, il serait intéressant d'étudier l'impact de l'échange de domaines ou de modules sur l'organisation générale du système. En effet, l'une des raisons pouvant expliquer la baisse de productivité des PKS modifiées peut être liée à une perturbation de leur macro-organisation.

En conclusion, ces différents axes de recherche menés au cours de cette thèse convergent en un objectif final unique : améliorer la compréhension des polycétide synthases afin de mieux en appréhender les tentatives d'ingénierie. Les deux axes de cette thèse ont eu pour but de combler un manque dans l'étude de ces méga-complexes enzymatique : leur localisation et leur optimisation au sein de leur hôte natif. Les études *in vivo* menées au cours de ces travaux de thèse et qui seront poursuivies conjointement par les équipes des laboratoires IMoPA et DynAMic contribueront à une ingénierie efficace des polycétide synthases.

Annexe

The coelimycin *cis*-AT polyketide synthase forms macro-complexes in its host *Streptomyces coelicolor*

Magali Tirado[1,2], Jean-Malo Massicard[1], Paula Yagüe[3,4], Gemma Fernández-García[3], Ariane Briegel[4], Gilles P. van Wezel[4], Ángel Manteca[3], Bertrand Aigle[2]*, Christophe Jacob[1]*, Kira J. Weissman[1]*

[1]Université de Lorraine, CNRS, IMoPA, F-54000 Nancy, France

[2]Université de Lorraine, INRAE, DynAMic, F-54000 Nancy, France

[3]Department of Functional Biology, Microbiology Area, IUOPA and ISPA, Faculty of Medicine, University of Oviedo, Oviedo, Spain

[4]Institute of Biology, Leiden University, Leiden, Netherlands

*Corresponding authors: bertrand.aigle@univ-lorraine.fr; christophe.jacob@univ-lorraine.fr; kira.weissman@univ-lorraine.fr

INTRODUCTION

Polyketides are the basis of several drugs of high therapeutic and economic value, such as antibiotics, immunosuppressants and anti-cancer agents (Newman & Cragg, 2016). It is thus of great interest to find new molecules to overcome issues such as the evolution of new diseases or the raising of antibiotic resistance (Johnston et al., 2024). This can be achieved by either find new natural products (Laureti et al., 2011; Pawlik et al., 2007; Wang et al., 2024), chemically modify existing structures (Peltier, 1992) or engineer the enzymes producing polyketides: the polyketide synthases (PKSs) (Weissman, 2016). To date, this last promising approach has been widely studied but attempts in modifying PKSs have led to reduced yields compared to native enzymatic systems (Miyazawa et al., 2021; Su et al., 2022; Wlodek et al., 2017). Modular PKSs are mega-enzymes composed of modules ranging from 150 to 200 kDa. A whole enzymatic complex can easily weigh several MDa (Fallon et al., 2024). This feature may explain the low rate of success of heterologous production or PKSs engineering. Understanding the relationship between the structure and the function of such megacomplexes is a key to improving production and engineering of new compounds.

To date, the main lines of the enzymatic mechanisms of PKSs have been established. They are true “assembly lines” as each step of polyketide synthesis is carried out by a specific catalytic domain (Weissman, 2009). Three essential domains are required for each elongation cycle: an acyltransferase (AT) domain recruiting a specific acyl-CoA building block; a central acyl carrier protein (ACP) to which the polyketide in elongation will be covalently yet temporary linked; and a ketosynthase (KS) domain catalyzing the binding of the new building block to the polyketide by Claisen condensation. Additional domains might be found, modifying the ketone function, such as ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER). Domains are grouped into functional units called “modules”, each catalyzing the addition of a new acyl-CoA unit and its potential modifications.

This feature gave the name of "modular PKS" to these complex systems. Modules are distributed among several polypeptide subunits communicating with each other *via* short protein sequences at the C- and N-termini called docking domains (DDs) (Broadhurst et al., 2003; Smith et al., 2021).

Thanks to their architecture being both multi-modular and multi-protein, PKSs are among the largest enzymatic complexes, in size and complexity, in the field of living organisms. Each module weighs from 150 to 200 kDa, forming mega-complexes of several MDa. A major remaining question about PKSs is the role of the macromolecular context in their functioning. This aspect has been largely overlooked during previous engineering attempts, which may explain in part the low success rate (Dodge et al., 2018; Weissman, 2017). So far, only one study reports information on the macromolecular organization of these multi-enzymatic systems in bacterial cells (Straight et al., 2007). Kolter *et coll*. used the *trans*-AT-type PKS producing bacillaene in *Bacillus subtilis*, an inhibitor of protein synthesis in prokaryotes (Patel et al., 1995), as a model. This PKS forms megacomplexes of the size comparable to an organelle, with an estimated molecular weight of 10 to 100 MDa. In addition, observations using electron microscopy indicate that this mega-complex does not appear to be associated with the membrane. This organelle-sized complex is believed to be a biosynthetic factory promoting coupled synthesis and export of the metabolite, as observed for other PKSs (Jain & Cox, 2005; Simunovic et al., 2003).

Streptomyces have been widely studied for they produce a third of all known polyketides (Chassagne et al., 2019). *S. coelicolor* is a key model used for decades as it produces four metabolites of high interest: undecylprodigiosin, actinorhodin, a calcium-dependent antibiotic and coelimycin. The organization of the gene cluster responsible to produce coelimycin A (CA) – called Cpk – and its regulation have been extensively studied (Figure 1A) (Pawlik et al., 2007; Bednarz et al., 2019; Kotowska et al., 2024), making it an interesting model to go further and study the macro-organization of the PKS itself. This system is composed of three subunits (CpkA–C) producing a precursor, further modified by post-PKS enzymes to produce CA (Figure 1B). When supplemented with N-acetylcysteine or L-glutamic acid yellow pigments are obtained, respectively coelimycin P1 (CP1) and P2 (CP2) (Gottelt et al., 2010; Pawlik et al., 2010).

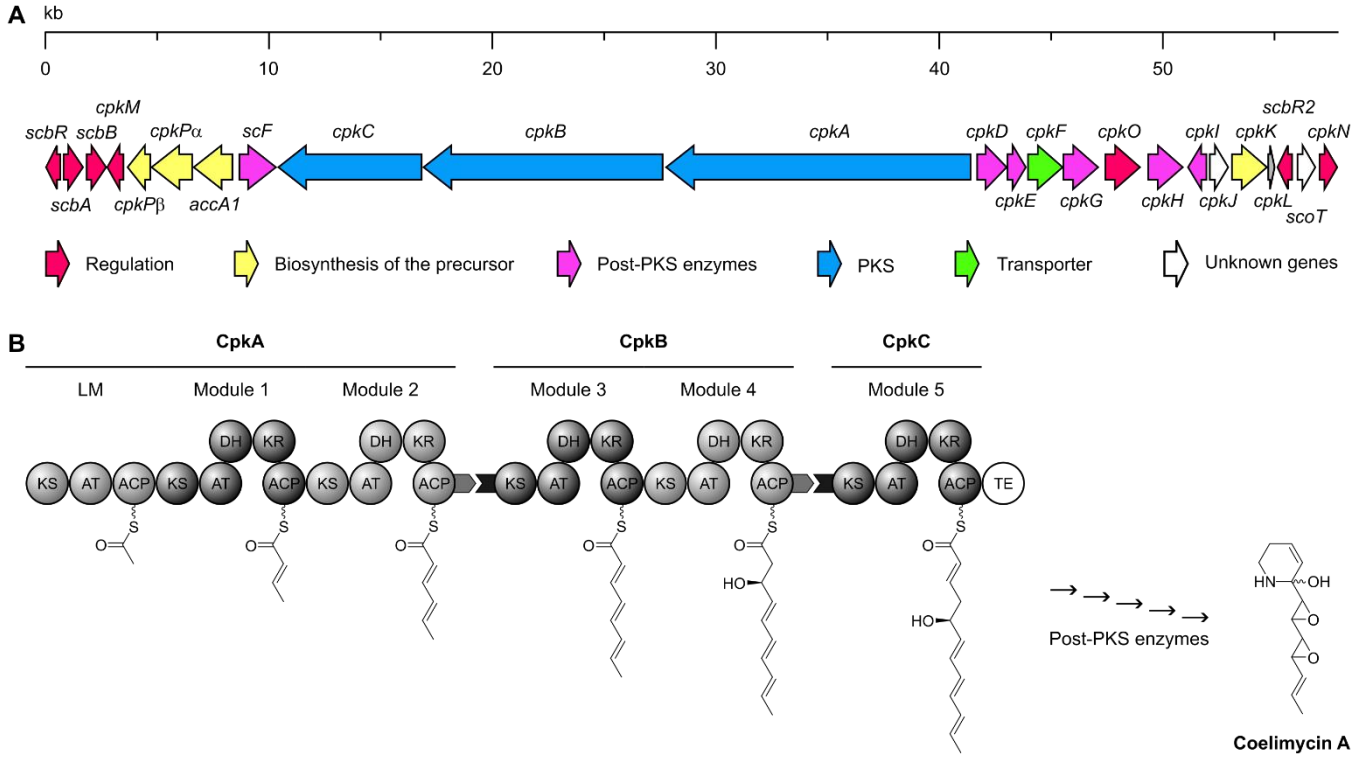


Figure 1. Organization of the gene cluster and the PKS producing coelimycin in *Streptomyces coelicolor* (adapted from (Gomez-Escribano et al., 2012). (A) Organization of the gene cluster associated with Cpk in *S. coelicolor*. Among the twenty-four genes composing the cluster, three encode PKS subunits (*cpkA–C*). Seven are dedicated to either regulation or biosynthesis of the precursor and six encode post-PKS enzymes. The gene *cpkF* is encoding the transporter associated with this system and two genes have functions yet to be discovered. **(B)** Organization in modules of the Cpk PKS in *S. coelicolor*. The system is composed of a loading module (LM) and five elongation modules, split into three subunits (CpkA–CpkC). The molecule produced is then modified by post-PKS enzymes to obtain the antibiotic coelimycin A.

In this study, we investigated the organization of the PKS synthesizing coelimycin in *S. coelicolor* (Gomez-Escribano et al., 2012). We focused on the subunits composing the core PKS (CpkA–C) and the transporter associated (CpkF). We studied the localization in the cells of CpkA, CpkC and CpkF by fusing them with fluorescent proteins.

RESULTS

Engineered PKSs complexes are functionals

Among the wide repertoire of *S. coelicolor* bioactive molecules, coelimycin A is an antibiotic produced by a modular type I *cis*-AT PKS. The gene cluster associated with this system contains, amongst others, three genes encoding PKSs subunits (*cpkA–cpkC*) and a transporter (*cpkF*). In this study, we modified genetically this cluster to express enzymes fused to fluorescent proteins and verified that the modified PKSs were functionals.

The CPK cluster in *S. coelicolor* is known to be expressed on DNAgar media (Gottelt et al., 2010), and to produce an antibiotic called coelimycin A. When supplemented with L-glutamic acid, coelimycin A is modified into a yellow pigment called coelimycin P2 (CP2). We identified that 50 mM of L-glutamic acid was enough for CP2 production, the compound being visible after 24 hours of culture. Production of CP2 was further used to verify the PKS functionality.

mCherry-CpkA and CpkC-eGFP are organized in foci

To determine the localization of CpkA and CpkC, we first constructed strains expressing the fused proteins mCherry-CpkA and / or CpkC-eGFP. We observed these strains with the collaboration of Dr. Paula Yagüe (Leiden University, Netherlands). Strains were cultured on DNAgar media for 26 hours to express coelimycin as previously described (Nieselt et al., 2010) before observation using confocal microscopy.

Results show that both subunits are forming foci in the cells, meaning that they are macro-complexing in the host (Figure 2). These foci were not found in all the cells nor in all samples prepared and we also observed diffuse signals throughout the hyphae. We then compared the localization of CpkA and CpkC by observing the strain expressing both fused proteins. As previously, about 10 percent of the samples expressed one of the proteins of interest. When the cells expressed both, we found that the signals did not overlap, meaning that the subunits are not colocalizing. These observations may be due to the fluorescent protein disturbing the assembly line of the PKS, or the interference between the organization of the subunits and the potential oligomerization of eGFP.

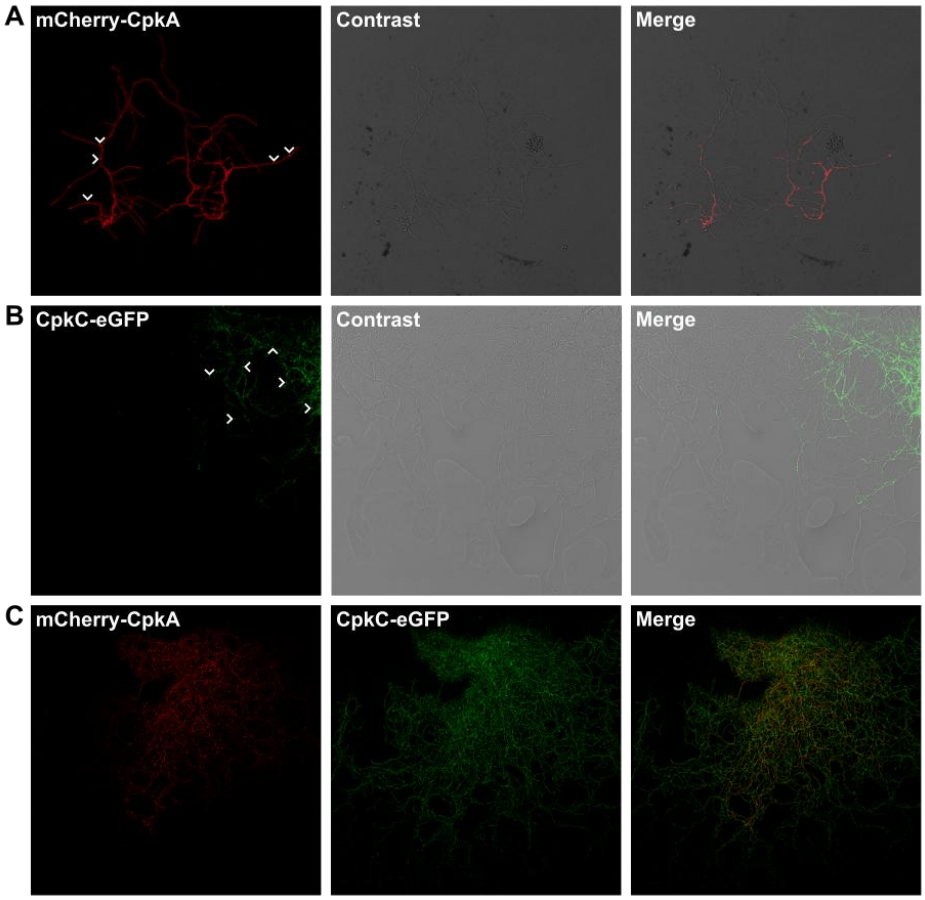


Figure 2. Localization of mCherry-CpkA and CpkC-eGFP by confocal microscopy in *Streptomyces coelicolor*. (A) Observation of the strain expressing mCherry-CpkA. CpkA is organized in foci inside the cells pointed by arrows. Signal of the foci is sometime hard to distinguish throughout diffuse signal. **(B)** Observation of the strain expressing CpkC-eGFP. CpkC is organized in foci as well in the cells, pointed by arrows. **(C)** Observation of the strain expressing mCherry-CpkA and CpkC-eGFP.

mYPet-CpkA and CpkC-mTurquoise2 are organized in foci colocalizing

To ensure that the absence of colocalization was not due to the fluorescent proteins interfering with the PKS assembly line, we created new fused proteins mYPet-CpkA and CpkC-mTurquoise2. We observed the strains obtained with the collaboration of Pr. Angel Manteca (Universidad de Oviedo, Spain). In this case, CpkA and CpkC also appeared as forming foci in the cells. In the mutant expressing both fusions, we observed that CpkA and CpkC are perfectly colocalizing (Figure 3). As for the previous construction, we could only observe the proteins in part of the colony and not in all hyphae. Foci length is between 0.7 and 1 µm, which make it impossible to distinguish if CpkA and CpkC are associated to the membrane or not with such resolution.

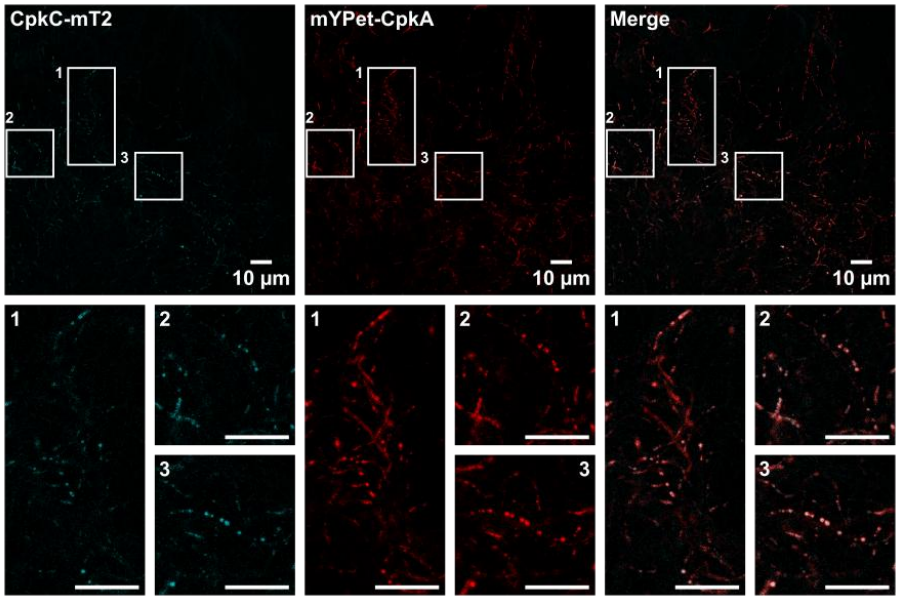


Figure 3. Colocalization of the subunits CpkA and CpkC in *S. coelicolor*. CpkA and CpkC are organized in foci colocalizing in the double mutant strain mYPet-CpkA, CpkC-mTurquoise2. The fusion CpkC-mTurquoise2 (mT2) is not expressed in all the cells and is organized in foci (*Left*). We observed the same situation with the fusion mYPet-cpkA (*Middle*), however mYPet-CpkA is found expressed in more hyphae than CpkC-mT2. When CpkC-mT2 is observed, it always colocalizes with mYPet-CpkA (*Right*).

These results show that mCherry is not a suitable fluorescent protein in *Streptomyces* spp. as the detection of the signal throughout autofluorescence depends on the level of expression of the protein of interest. We can also conclude that using eGFP may not be suitable to observe localization of proteins forming complexes as it can oligomerize and create aggregates.

CpkF is a transmembrane transporter expressed in all the cells

We also studied the localization and organization of the transporter CpkF. Using the online tool TMHMM 2.0 (Hallgren et al., 2022), we found CpkF to be a protein composed of 14 transmembrane helixes. We aimed to analyze its localization in the cells, fusing it to the fluorescent protein mCherry at the C-terminal position. The fluorescent protein should therefore be intracellular.

We observed using confocal microscopy that CpkF is expressed in all the cells and forms foci localized in the membrane, as expected (Figure 4). The transporter is also found in the septa of the bacteria, as shown by the arrows on figure 4 middle.

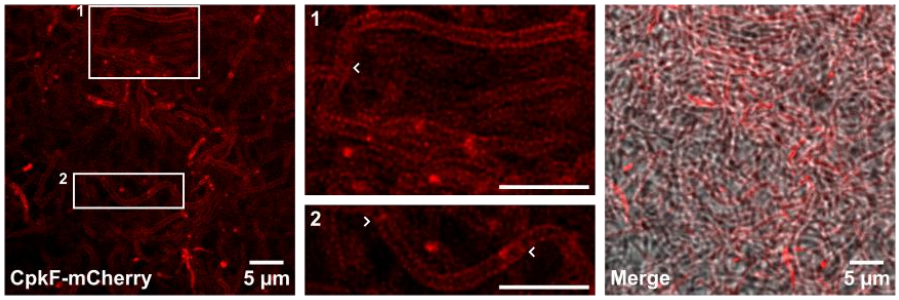


Figure 4. Localization of the transporter CpkF in *S. coelicolor*. The CpkF-mCherry fusion forms dots in the membrane of the bacteria (*Left* and *Middle*). Arrows are pointing to septa in which CpkF is found (*Middle*). (*Right*) is the merge of the contrast showing hyphae and the signal of mCherry (*Left*).

DISCUSSION

Studies led on *trans*-AT PKSs unraveled the way they organize themselves, in bacterial cells and *in vitro* (Gay et al., 2016; Straight et al., 2007; Tittes et al., 2022). They form mega-complexes concentrating the many proteins composing them, including subunits, *trans*-acting enzymes, post-synthesis modifying enzymes, as well as precursors and extension units. The mega-complexes create an environment that might favor communication between PKSs subunits, as their interaction are weak *in vitro* (Broadhurst et al., 2003) and protect the reaction intermediates. Colocalization of CpkA and CpkC subunits supports these previous observations and theories.

During the observation of the mutated strains, CpkA and CpkC were not observed in all the cells nor were all samples. This means that all the cells are not expressing the complex without it being critical for the colony. Such observations are relevant regarding the division of labor previously described in *Streptomyces* colonies (Zhang et al., 2020). Moreover, we mono-cultured the strains in a sterile media, so it is not necessary for the colony to over-produce antibiotics such as coelimycin.

A hypothesis explaining this organization would be an association *via* the membrane proteins, serving as a platform for the association of PKSs (Straight et al., 2007). Such association have already been observed between a PKS subunit and the transporter of the produced lipid, acting as a virulence factor in *Mycobacterium tuberculosis* (Jain & Cox, 2005). Other PKSs and similar biosynthetic systems have been found to interact with the membranes of bacteria in many ways (Hantash & Earhart, 2000; Imperi & Visca, 2013; Kiesau et al., 1997; Porter et al., 2013; Xu et al., 2008). All these analyses support the hypothesis of the formation of a biosynthetic factory within the bacterial hosts, located near the membrane. This proximity might be a strategy conserved in bacteria, aiming to couple biosynthesis with secretion, thus avoiding the accumulation in the cytoplasm of potentially toxic molecules (Imperi & Visca, 2013; Simunovic et al., 2003; Straight et al., 2007).

All these observations support the localization of the trans-membrane transporter CpkF in all the cells, to reduce auto-toxicity. Detailed analyses of the localization of the Cpk subunits could therefore prove the potential association of the complex to the membrane.

This macro-organization may also be induced by interaction between homo-modules of different complexes. Inter-modules interactions were identified in the PKS producing azalomycin F (Zhai et al., 2023). The first module here catalyzes the addition of two acyl-CoA units iteratively. The ER domain found in this first module is believed to act with the following module as a "*trans*" domain. It may be that macro-complexation of this PKS allows the ER domain of one PKS to interact with module 3 of a second system. Such interaction may either facilitate or be due to macro-organization. Interactions in the PKS producing 6-deoxyeyrhonolid B (DEBS) have also been observed and characterized. A 4′-phosphopantetheine (P-pant) arm, prosthetic group tethered to ACP domains, was found to interact with the KS domain of another subunit (Cogan et al., 2024). It might also be interesting to investigate the role of docking domains in such macro-organization as they are already involved in the organization of subunits. All these elements highlight the importance of understanding accurately the organization of such mega-complexes to improve their engineering.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Streptomyces coelicolor FM145 was used to fuse fluorescent proteins to Cpk component, for its reduced auto-fluorescence in red and green (Willemse & van Wezel, 2009). This strain was grown on *Soya Flour Mannitol* (SFM) medium (20 g soya flour, 20 g mannitol, 20 g agar for 1000 mL of media, pH 7.2) for sporulation and conjugation, and on *Difco Nutrient Agar* (DNAgar) media for *cpk* production and observation by confocal microscopy. *Escherichia coli* ET12567(pUZ8002) was used for transformation of *S. coelicolor* FM145 by conjugation and *E. coli* DH5α for common DNA manipulation. *E. coli* strains were grown on *Lysogeny Broth* (LB) media (10 g tryptone, 5 g yeast extract, 5 g sodium chloride for 1000 mL of media, pH 7.5), supplemented when needed with 17 g of agar per liter for solid media and adequate antibiotics.

Genetic engineering using CRISPR-Cas9

The vector pCRISPomyces-2 expressing CRISPR-Cas9 editing system was used to edit *S. coelicolor* genome (Cobb et al., 2015). The crRNA was selected to hybridize near the target region, to a sequence containing NGG motif on its 3'end (N being any nucleotide and NGG being the protospacer adjacent motif (PAM)). The sequences were designed using CRISPy-web online tool (Blin et al., 2016). The crRNA sequence was created by annealing of primers, then inserted into pCRISPomyces-2 by golden gate using *Bbs*I restriction site (Cobb et al., 2015). Editing matrices composed of homologous regions (HR1 and HR2), and if needed DNA fragments for fluorescent proteins insertion, were then cloned into pCRISPomyces-2-crRNA vectors by Gibson assembly using *Xba*I restriction site.

In addition, the sequence recognized by the crRNA was mutated into the corresponding HR to avoid cleavage of the plasmid. Transformation of *S. coelicolor* FM145 was performed as previously described (Cobb et al., 2015).

Assess the functionality of modified strains and PKSs

Expression of fluorescent protein in mutated strains was measured using dot plot method. Strains were cultured into DNB media for 24h. Proteins were then precipitated using acetone and 10 µL of the extract were poured on a nitrocellulose membrane.

Verifying the functionality of modified PKSs was performed by production of the yellow pigment coelimycin P2 (CP2). CP2 is produced by the addition of L-glutamic acid to coelimycin A. As previously described, DNAgar media supplemented with L-glutamic acid was used for CP2 production (Gottelt et al., 2010).

Observation of the strains using confocal microscopy

Strains were grown on *Difco Nutrient Agar* (DNA) media, supplemented with 10 mM glutamate. Stock solutions of spores concentrated at 10^8 spores per mL were prepared, of which 100 µL were plated on DNA media covered with cellophane. After 26h hours of incubation at 30°C, pieces of cellophane were cut and observed using a Leica TCS SP8 confocal microscope with hybrid detectors. To observe mCherry, 522 nm laser was used for excitation of mCherry-CpkA and emission was collected between 586–789 nm. For CpkF-mCherry, the fluorescent protein was observed at 587 nm excitation and emission was collected between 600 and 750 nm. The protein eGFP was excited at 488 nm and emission collected between 500 and 548 nm. Signals related to mYPet were observed using 517 nm wavelength for excitation, and emission was collected between 525 and 600 nm. When two fluorescent proteins were observed in one strain, sequential mode was used. Images were analyzed with Fiji is just ImageJ (Fiji) software (Schindelin et al., 2012). When indicated, images were deconvolved with Huygens Essential version 22.10 (Scientific Volume Imaging, The Netherlands, http://svi.nl).

BIBLIOGRAPHY

Bednarz, B., Kotowska, M., & Pawlik, K. J. (2019). Multi-level regulation of coelimycin synthesis in Streptomyces coelicolor A3(2). *Applied Microbiology and Biotechnology*, *103*(16), 6423–6434. https://doi.org/10.1007/s00253-019-09975-w

Blin, K., Pedersen, L. E., Weber, T., & Lee, S. Y. (2016). CRISPy-web: An online resource to design sgRNAs for CRISPR applications. *Synthetic and Systems Biotechnology*, *1*(2), 118–121. https://doi.org/10.1016/j.synbio.2016.01.003

Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., & Weissman, K. J. (2003). The Structure of Docking Domains in Modular Polyketide Synthases. *Chemistry & Biology*, *10*(8), 723–731. https://doi.org/10.1016/S1074-5521(03)00156-X

Chassagne, F., Cabanac, G., Hubert, G., David, B., & Marti, G. (2019). The landscape of natural product diversity and their pharmacological relevance from a focus on the Dictionary of Natural Products®. *Phytochemistry Reviews*, *18*(3), 601–622. https://doi.org/10.1007/s11101-019-09606-2

Cobb, R. E., Wang, Y., & Zhao, H. (2015). High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system. *ACS Synthetic Biology*, *4*(6), 723–728. https://doi.org/10.1021/sb500351f

Cogan, D. P., Soohoo, A. M., Chen, M., Liu, Y., Brodsky, K. L., & Khosla, C. (2024). Structural basis for intermodular communication in assembly-line polyketide biosynthesis. *Nature Chemical Biology*. https://doi.org/10.1038/s41589-024-01709-y

Dodge, G. J., Maloney, F. P., & Smith, J. L. (2018). Protein–protein interactions in " *cis* -AT" polyketide synthases. *Natural Product Reports*, *35*(10), 1082–1096. https://doi.org/10.1039/C8NP00058A

Fallon, T. R., Shende, V. V., Wierzbicki, I. H., Pendleton, A. L., Watervoot, N. F., Auber, R. P., Gonzalez, D. J., Wisecaver, J. H., & Moore, B. S. (2024). Giant polyketide synthase enzymes in the biosynthesis of giant marine polyether toxins. *Science*, *385*(6709), 671–678. https://doi.org/10.1126/science.ado3290

Gay, D. C., Wagner, D. T., Meinke, J. L., Zogzas, C. E., Gay, G. R., & Keatinge-Clay, A. T. (2016). The LINKS motif zippers trans-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *Journal of Structural Biology*, *193*(3), 196–205. https://doi.org/10.1016/j.jsb.2015.12.011

Gomez-Escribano, J. P., Song, L., Fox, D. J., Yeo, V., Bibb, M. J., & Challis, G. L. (2012). Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of Streptomyces coelicolor M145. *Chemical Science*, *3*(9), 2716–2720. https://doi.org/10.1039/C2SC20410J

Gottelt, M., Kol, S., Gomez-Escribano, J. P., Bibb, M., & Takano, E. (2010). Deletion of a regulatory gene within the cpk gene cluster reveals novel antibacterial activity in Streptomyces coelicolor A3(2). *Microbiology (Reading, England)*, *156*(Pt 8), 2343–2353. https://doi.org/10.1099/mic.0.038281-0

Hallgren, J., Tsirigos, K. D., Pedersen, M. D., Almagro Armenteros, J. J., Marcatili, P., Nielsen, H., Krogh, A., & Winther, O. (2022). *DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks*. https://doi.org/10.1101/2022.04.08.487609

Hantash, F. M., & Earhart, C. F. (2000). Membrane Association of the Escherichia coli Enterobactin Synthase Proteins EntB/G, EntE, and EntF. *Journal of Bacteriology*, *182*(6), 1768–1773.

Imperi, F., & Visca, P. (2013). Subcellular localization of the pyoverdine biogenesis machinery of Pseudomonas aeruginosa: A membrane-associated "siderosome". *FEBS Letters*, *587*(21), 3387–3391. https://doi.org/10.1016/j.febslet.2013.08.039

Jain, M., & Cox, J. S. (2005). Interaction between polyketide synthase and transporter suggests coupled synthesis and export of virulence lipid in M. tuberculosis. *PLoS Pathogens*, *1*(1), e2. https://doi.org/10.1371/journal.ppat.0010002

Johnston, H. E., Batt, S. M., Brown, A. K., Savva, C. G., Besra, G. S., & Fütterer, K. (2024). Cryo-electron microscopy structure of the di-domain core of Mycobacterium tuberculosis polyketide synthase 13, essential for mycobacterial mycolic acid synthesis. *Microbiology (Reading, England)*, *170*(10), 001505. https://doi.org/10.1099/mic.0.001505

Kiesau, P., Eikmanns, U., Gutowski-Eckel, Z., Weber, S., Hammelmann, M., & Entian, K. D. (1997). Evidence for a multimeric subtilin synthetase complex. *Journal of Bacteriology*, *179*(5), 1475–1481. https://doi.org/10.1128/jb.179.5.1475-1481.1997

Kotowska, M., Wenecki, M., Bednarz, B., Ciekot, J., Pasławski, W., Buhl, T., & Pawlik, K. J. (2024). Coelimycin inside out—Negative feedback regulation by its intracellular precursors. *Applied Microbiology and Biotechnology*, *108*(1), 531. https://doi.org/10.1007/s00253-024-13366-1

Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G. L., & Aigle, B. (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6258–6263. https://doi.org/10.1073/pnas.1019077108

Miyazawa, T., Fitzgerald, B. J., & Keatinge-Clay, A. T. (2021). Preparative production of an enantiomeric pair by engineered polyketide synthases. *Chemical Communications*, *57*(70), 8762–8765. https://doi.org/10.1039/D1CC03073F

Newman, D. J., & Cragg, G. M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products*, *79*(3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055

Nieselt, K., Battke, F., Herbig, A., Bruheim, P., Wentzel, A., Jakobsen, Ø. M., Sletta, H., Alam, M. T., Merlo, M. E., Moore, J., Omara, W. A., Morrissey, E. R., Juarez-Hermosillo, M. A., Rodríguez-García, A., Nentwich, M., Thomas, L., Iqbal, M., Legaie, R., Gaze, W. H., … Wellington, E. M. (2010). The dynamic architecture of the metabolic switch in Streptomyces coelicolor. *BMC Genomics*, *11*(1), 10. https://doi.org/10.1186/1471-2164-11-10

Patel, P. S., Huang, S., Fisher, S., Pirnik, D., Aklonis, C., Dean, L., Meyers, E., Fernandes, P., & Mayerl, F. (1995). Bacillaene, a novel inhibitor of procaryotic protein synthesis produced by Bacillus subtilis: Production, taxonomy, isolation, physico-chemical characterization and biological activity. *The Journal of Antibiotics*, *48*(9), 997–1003. https://doi.org/10.7164/antibiotics.48.997

Pawlik, K., Kotowska, M., Chater, K. F., Kuczek, K., & Takano, E. (2007). A cryptic type I polyketide synthase (cpk) gene cluster in Streptomyces coelicolor A3(2). *Archives of Microbiology*, *187*(2), 87–99. https://doi.org/10.1007/s00203-006-0176-7

Pawlik, K., Kotowska, M., & Kolesiński, P. (2010). Streptomyces coelicolor A3(2) produces a new yellow pigment associated with the polyketide synthase Cpk. *Journal of Molecular Microbiology and Biotechnology*, *19*(3), 147–151. https://doi.org/10.1159/000321501

Peltier, K. (1992). The new macrolactones: Clarithromycin (Biaxin) and azithromycin (Zithromax). *Connecticut Medicine*, *56*(7), 371–374.

Porter, J. L., Tobias, N. J., Pidot, S. J., Falgner, S., Tuck, K. L., Vettiger, A., Hong, H., Leadlay, P. F., & Stinear, T. P. (2013). The cell wall-associated mycolactone polyketide synthases are necessary but not sufficient for mycolactone biosynthesis. *PloS One*, *8*(7), e70520. https://doi.org/10.1371/journal.pone.0070520

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Simunovic, V., Gherardini, F. C., & Shimkets, L. J. (2003). Membrane localization of motility, signaling, and polyketide synthetase proteins in Myxococcus xanthus. *Journal of Bacteriology*, *185*(17), 5066–5075. https://doi.org/10.1128/jb.185.17.5066-5075.2003

Smith, H. G., Beech, M. J., Lewandowski, J. R., Challis, G. L., & Jenner, M. (2021). Docking domain-mediated subunit interactions in natural product megasynth(et)ases. *Journal of Industrial Microbiology and Biotechnology*, *48*(3–4), kuab018. https://doi.org/10.1093/jimb/kuab018

Straight, P. D., Fischbach, M. A., Walsh, C. T., Rudner, D. Z., & Kolter, R. (2007). A singular enzymatic megacomplex from Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 305–310. https://doi.org/10.1073/pnas.0609073103

Su, L., Hôtel, L., Paris, C., Chepkirui, C., Brachmann, A. O., Piel, J., Jacob, C., Aigle, B., & Weissman, K. J. (2022). Engineering the stambomycin modular polyketide synthase yields 37-membered mini-stambomycins. *Nature Communications*, *13*(1), 515. https://doi.org/10.1038/s41467-022-27955-z

Tittes, Y. U., Herbst, D. A., Martin, S. F. X., Munoz-Hernandez, H., Jakob, R. P., & Maier, T. (2022). The structure of a polyketide synthase bimodule core. *Science Advances*, *8*(38), eabo6918. https://doi.org/10.1126/sciadv.abo6918

Wang, Y., Duan, Y., Zhang, M., Liang, C., Li, W., Liu, C., & Ye, Y. (2024). Genome Sequencing and Metabolic Potential Analysis of Irpex lacteus. *Journal of Fungi*, *10*(12), 846. https://doi.org/10.3390/jof10120846

Weissman, K. J. (2009). Introduction to polyketide biosynthesis. *Methods in Enzymology*, *459*, 3–16. https://doi.org/10.1016/S0076-6879(09)04601-1

Weissman, K. J. (2016). Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Natural Product Reports*, *33*(2), 203–230. https://doi.org/10.1039/c5np00109a

Weissman, K. J. (2017). Correction: Uncovering the structures of modular polyketide synthases. *Natural Product Reports*, *34*(8), 1035–1035. https://doi.org/10.1039/C7NP90027A

Willemse, J., & van Wezel, G. P. (2009). Imaging of Streptomyces coelicolor A3(2) with reduced autofluorescence reveals a novel stage of FtsZ localization. *PloS One*, *4*(1), e4242. https://doi.org/10.1371/journal.pone.0004242

Wlodek, A., Kendrew, S. G., Coates, N. J., Hold, A., Pogwizd, J., Rudder, S., Sheehan, L. S., Higginbotham, S. J., Stanley-Smith, A. E., Warneck, T., Nur-E-Alam, M., Radzom, M., Martin, C. J., Overvoorde, L., Samborskyy, M., Alt, S., Heine, D., Carter, G. T., Graziani, E. I., … Gregory, M. A. (2017). Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nature Communications*, *8*(1), 1206. https://doi.org/10.1038/s41467-017-01344-3

Xu, X.-P., Wang, Z.-J., Fan, K.-Q., Wang, S.-L., Jia, C.-J., Han, H., Ramalingam, E., & Yang, K.-Q. (2008). Localization of the ActIII actinorhodin polyketide ketoreductase to the cell wall. *FEMS Microbiology Letters*, *287*(1), 15–21. https://doi.org/10.1111/j.1574-6968.2008.01289.x

Zhai, G., Zhu, Y., Sun, G., Zhou, F., Sun, Y., Hong, Z., Dong, C., Leadlay, P. F., Hong, K., Deng, Z., Zhou, F., & Sun, Y. (2023). Insights into azalomycin F assembly-line contribute to evolution-guided polyketide synthase engineering and identification of intermodular recognition. *Nature Communications*, *14*(1), 612. https://doi.org/10.1038/s41467-023-36213-9

Zhang, Z., Du, C., de Barsy, F., Liem, M., Liakopoulos, A., van Wezel, G. P., Choi, Y. H., Claessen, D., & Rozen, D. E. (2020). Antibiotic production in Streptomyces is organized by a division of labor through terminal genomic differentiation. *Science Advances*, *6*(3), eaay5781. https://doi.org/10.1126/sciadv.aay5781

Notes bibliographiques

Alekseyev, V. Y., Liu, C. W., Cane, D. E., Puglisi, J. D., & Khosla, C. (2007). Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Science: A Publication of the Protein Society*, *16*(10), 2093–2107. https://doi.org/10.1110/ps.073011407

Aminov, R. (2017). History of antimicrobial drug discovery: Major classes and health impact. *Biochemical Pharmacology*, *133*, 4–19. https://doi.org/10.1016/j.bcp.2016.10.001

Annaval, T., Paris, C., Leadlay, P. F., Jacob, C., & Weissman, K. J. (2015). Evaluating Ketoreductase Exchanges as a Means of Rationally Altering Polyketide Stereochemistry. *ChemBioChem*, *16*(9), 1357–1364. https://doi.org/10.1002/cbic.201500113

Aparicio, J. F., Molnár, I., Schwecke, T., König, A., Haydock, S. F., Ee Khaw, L., Staunton, J., & Leadlay, P. F. (1996). Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus: Analysis of the enzymatic domains in the modular polyketide synthase. *Gene*, *169*(1), 9–16. https://doi.org/10.1016/0378-1119(95)00800-4

Austin, M. B., & Noel, J. P. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports*, *20*(1), 79–110. https://doi.org/10.1039/b100917f

Barajas, J. F., McAndrew, R. P., Thompson, M. G., Backman, T. W. H., Pang, B., De Rond, T., Pereira, J. H., Benites, V. T., Martín, H. G., Baidoo, E. E. K., Hillson, N. J., Adams, P. D., & Keasling, J. D. (2019). Structural insights into dehydratase substrate selection for the borrelidin and fluvirucin polyketide synthases. *Journal of Industrial Microbiology and Biotechnology*, *46*(8), 1225–1235. https://doi.org/10.1007/s10295-019-02189-z

Baranasic, D., Gacesa, R., Starcevic, A., Zucko, J., Blažič, M., Horvat, M., Gjuračić, K., Fujs, Š., Hranueli, D., Kosec, G., Cullum, J., & Petković, H. (2013). Draft Genome Sequence of Streptomyces rapamycinicus Strain NRRL 5491, the Producer of the Immunosuppressant Rapamycin. *Genome Announcements*, *1*(4), e00581-13. https://doi.org/10.1128/genomeA.00581-13

Bednarz, B., Millan-Oropeza, A., Kotowska, M., Świat, M., Quispe Haro, J. J., Henry, C., & Pawlik, K. (2021). Coelimycin Synthesis Activatory Proteins Are Key Regulators of Specialized Metabolism and Precursor Flux in Streptomyces coelicolor A3(2). *Frontiers in Microbiology*, *12*, 616050. https://doi.org/10.3389/fmicb.2021.616050

Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., … Hopwood, D. A. (2002). Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature*, *417*(6885), 141–147. https://doi.org/10.1038/417141a

Blin, K., Pedersen, L. E., Weber, T., & Lee, S. Y. (2016). CRISPy-web: An online resource to design sgRNAs for CRISPR applications. *Synthetic and Systems Biotechnology*, *1*(2), 118–121. https://doi.org/10.1016/j.synbio.2016.01.003

Bozhüyük, K. A. J., Fleischhacker, F., Linck, A., Wesche, F., Tietze, A., Niesert, C.-P., & Bode, H. B. (2018). De novo design and engineering of non-ribosomal peptide synthetases. *Nature Chemistry*, *10*(3), 275–281. https://doi.org/10.1038/nchem.2890

Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., & Weissman, K. J. (2003). The Structure of Docking Domains in Modular Polyketide Synthases. *Chemistry & Biology*, *10*(8), 723–731. https://doi.org/10.1016/S1074-5521(03)00156-X

Buchholz, T. J., Geders, T. W., Bartley, F. E., Reynolds, K. A., Smith, J. L., & Sherman, D. H. (2009). Structural Basis for Binding Specificity between Subclasses of Modular Polyketide Synthase Docking Domains. *ACS Chemical Biology*, *4*(1), 41–52. https://doi.org/10.1021/cb8002607

Bury-Moné, S., Thibessard, A., Lioy, V. S., & Leblond, P. (2023). Dynamics of the Streptomyces chromosome: Chance and necessity. *Trends in Genetics: TIG*, *39*(11), 873–887. https://doi.org/10.1016/j.tig.2023.07.008

Butcher, R. A., Schroeder, F. C., Fischbach, M. A., Straight, P. D., Kolter, R., Walsh, C. T., & Clardy, J. (2007). The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences*, *104*(5), 1506–1509. https://doi.org/10.1073/pnas.0610503104

Buyachuihan, L., Zhao, Y., Schelhas, C., & Grininger, M. (2023). Docking Domain Engineering in a Modular Polyketide Synthase and Its Impact on Structure and Function. *ACS Chemical Biology*, acschembio.3c00074. https://doi.org/10.1021/acschembio.3c00074

Caffrey, P. (2003). Conserved Amino Acid Residues Correlating With Ketoreductase Stereospecificity in Modular Polyketide Synthases. *ChemBioChem*, *4*(7), 654–657. https://doi.org/10.1002/cbic.200300581

Chain, E., Florey, H. W., Gardner, A. D., Heatley, N. G., Jennings, M. A., Orr-Ewing, J., Sanders, A. G., & Peltier, L. F. (1940). Penicillin as a Chemotherapeutic Agent. *The Lancet*, *236*(6104), 226–228.

Chassagne, F., Cabanac, G., Hubert, G., David, B., & Marti, G. (2019). The landscape of natural product diversity and their pharmacological relevance from a focus on the Dictionary of Natural Products®. *Phytochemistry Reviews*, *18*(3), 601–622. https://doi.org/10.1007/s11101-019-09606-2

Chen, A., Re, R. N., & Burkart, M. D. (2018). Type II fatty acid and polyketide synthases: Deciphering protein-protein and protein-substrate interactions. *Natural Product Reports*, *35*(10), 1029–1045. https://doi.org/10.1039/c8np00040a

Chen, R., Zhang, Q., Tan, B., Zheng, L., Li, H., Zhu, Y., & Zhang, C. (2017). Genome Mining and Activation of a Silent PKS/NRPS Gene Cluster Direct the Production of Totopotensamides. *Organic Letters*, *19*(20), 5697–5700. https://doi.org/10.1021/acs.orglett.7b02878

Chen, X.-H., Vater, J., Piel, J., Franke, P., Scholz, R., Schneider, K., Koumoutsi, A., Hitzeroth, G., Grammel, N., Strittmatter, A. W., Gottschalk, G., Süssmuth, R. D., & Borriss, R. (2006). Structural and Functional Characterization of Three Polyketide Synthase Gene Clusters in *Bacillus amyloliquefaciens* FZB 42. *Journal of Bacteriology*, *188*(11), 4024–4036. https://doi.org/10.1128/JB.00052-06

Claxton, H. B., Akey, D. L., Silver, M. K., Admiraal, S. J., & Smith, J. L. (2009). Structure and functional analysis of RifR, the type II thioesterase from the rifamycin biosynthetic pathway. *The Journal of Biological Chemistry*, *284*(8), 5021–5029. https://doi.org/10.1074/jbc.M808604200

Cobb, R. E., Wang, Y., & Zhao, H. (2015). High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system. *ACS Synthetic Biology*, *4*(6), 723–728. https://doi.org/10.1021/sb500351f

Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., & Leadlay, P. F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature*, *348*(6297), 176–178. https://doi.org/10.1038/348176a0

Cortes, J., Wiesmann, K. E. H., Roberts, G. A., Brown, M. J. B., Staunton, J., & Leadlay, P. F. (1995). Repositioning of a Domain in a Modular Polyketide Synthase to Promote Specific Chain Cleavage. *Science*, *268*(5216), 1487–1489. https://doi.org/10.1126/science.7770773

Curran, S. C., Hagen, A., Poust, S., Chan, L. J. G., Garabedian, B. M., De Rond, T., Baluyot, M.-J., Vu, J. T., Lau, A. K., Yuzawa, S., Petzold, C. J., Katz, L., & Keasling, J. D. (2018). Probing the Flexibility of an Iterative Modular Polyketide Synthase with Non-Native Substrates *in Vitro*. *ACS Chemical Biology*, *13*(8), 2261–2268. https://doi.org/10.1021/acschembio.8b00422

D'Ambrosio, H., Keeler, A., & Derbyshire, E. (2023). Examination of Secondary Metabolite Biosynthesis in Apicomplexa. *ChemBioChem*, e202300263. https://doi.org/10.1002/cbic.202300263

Dasí-Delgado, P., Andreu, C., & Del Olmo, M. (2025). Strategies Used for the Discovery of New Microbial Metabolites with Antibiotic Activity. *Molecules*, *30*(13), 2868. https://doi.org/10.3390/molecules30132868

Derewacz, D. K., Covington, B. C., McLean, J. A., & Bachmann, B. O. (2015). Mapping Microbial Response Metabolomes for Induced Natural Product Discovery. *ACS Chemical Biology*, *10*(9), 1998–2006. https://doi.org/10.1021/acschembio.5b00001

Dietl, A., & Barends, T. R. M. (2022). Dynamics in an unusual acyl carrier protein from a ladderane lipid-synthesizing organism. *Proteins: Structure, Function, and Bioinformatics*, *90*(1), 73–82. https://doi.org/10.1002/prot.26187

Dorival, J., Annaval, T., Risser, F., Collin, S., Roblin, P., Jacob, C., Gruez, A., Chagot, B., & Weissman, K. J. (2016). Characterization of Intersubunit Communication in the Virginiamycin *trans* -Acyl Transferase Polyketide Synthase. *Journal of the American Chemical Society*, *138*(12), 4155–4167. https://doi.org/10.1021/jacs.5b13372

Drufva, E. E., Spengler, N. R., Hix, E. G., & Bailey, C. B. (2021). Site-Directed Mutagenesis of Modular Polyketide Synthase Ketoreductase Domains for Altered Stereochemical Control. *Chembiochem: A European Journal of Chemical Biology*, *22*(7), 1122–1150. https://doi.org/10.1002/cbic.202000613

Duchesne, E. (1897). *Antagonisme entre les moisissures et les microbes*. Faculté de médecine et de pharmacie de Lyon.

Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PloS One*, *3*(11), e3647. https://doi.org/10.1371/journal.pone.0003647

Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. *The British Journal of Experimental Pathology*, *10*, 226–236.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch. (2005). Protein Identification and Analysis Tools on the Expasy Server. In *The Proteomics Protocols Handbook* (John M. Walker, pp. 571–607). Humana Press.

Gay, D. C., Gay, G., Axelrod, A. J., Jenner, M., Kohlhaas, C., Kampa, A., Oldham, N. J., Piel, J., & Keatinge-Clay, A. T. (2014). A Close Look at a Ketosynthase from a Trans-Acyltransferase Modular Polyketide Synthase. *Structure*, *22*(3), 444–451. https://doi.org/10.1016/j.str.2013.12.016

Gay, D. C., Wagner, D. T., Meinke, J. L., Zogzas, C. E., Gay, G. R., & Keatinge-Clay, A. T. (2016). The LINKS motif zippers trans-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *Journal of Structural Biology*, *193*(3), 196–205. https://doi.org/10.1016/j.jsb.2015.12.011

Gay, D. C., You, Y.-O., Keatinge-Clay, A. T., & Cane, D. E. (2013). Structure and Stereospecificity of the Dehydratase Domain from the Terminal Module of the Rifamycin Polyketide Synthase. *Biochemistry*, *52*(49), 8916–8928. https://doi.org/10.1021/bi400988t

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimousin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. J., & Royant, A. (2012). Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nature Communications*, *3*, 751. https://doi.org/10.1038/ncomms1738

Gomez-Escribano, J. P., Song, L., Fox, D. J., Yeo, V., Bibb, M. J., & Challis, G. L. (2012). Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of Streptomyces coelicolor M145. *Chemical Science*, *3*(9), 2716. https://doi.org/10.1039/c2sc20410j

Gottelt, M., Kol, S., Gomez-Escribano, J. P., Bibb, M., & Takano, E. (2010). Deletion of a regulatory gene within the cpk gene cluster reveals novel antibacterial activity in Streptomyces coelicolor A3(2). *Microbiology (Reading, England)*, *156*(Pt 8), 2343–2353. https://doi.org/10.1099/mic.0.038281-0

Grant, S. G., Jessee, J., Bloom, F. R., & Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(12), 4645–4649. https://doi.org/10.1073/pnas.87.12.4645

Guzman, K. M., Yuet, K. P., Lynch, S. R., Liu, C. W., & Khosla, C. (2021). Properties of a 'Split-and-Stuttering' Module of an Assembly Line Polyketide Synthase. *The Journal of Organic Chemistry*, *86*(16), 11100–11106. https://doi.org/10.1021/acs.joc.1c00120

Hackl, S., & Bechthold, A. (2015). The Gene *bldA* , a Regulator of Morphological Differentiation and Antibiotic Production in *Streptomyces*. *Archiv Der Pharmazie*, *348*(7), 455–462. https://doi.org/10.1002/ardp.201500073

Hagan, R. M., Björnsson, R., McMahon, S. A., Schomburg, B., Braithwaite, V., Bühl, M., Naismith, J. H., & Schwarz-Linek, U. (2010). NMR Spectroscopic and Theoretical Analysis of a Spontaneously Formed Lys–Asp Isopeptide Bond. *Angewandte Chemie International Edition*, *49*(45), 8421–8425. https://doi.org/10.1002/anie.201004340

Hahn, D. R., Gustafson, G., Waldron, C., Bullard, B., Jackson, J. D., & Mitchell, J. (2006). Butenyl-spinosyns, a natural example of genetic engineering of antibiotic biosynthetic genes. *Journal of Industrial Microbiology & Biotechnology*, *33*(2), 94–104. https://doi.org/10.1007/s10295-005-0016-9

Hahn, D. R., Jackson, J. D., Bullard, B. S., Gustafson, G. D., Waldron, C., & Mitchell, J. C. (2002). *Biosynthetic genes for butenyl-spinosyn insecticide production* (Patent EP 1 414 969 B1).

Hantash, F. M., & Earhart, C. F. (2000). Membrane Association of the *Escherichia coli* Enterobactin Synthase Proteins EntB/G, EntE, and EntF. *Journal of Bacteriology*, *182*(6), 1768–1773. https://doi.org/10.1128/JB.182.6.1768-1773.2000

Hashimura, H., Nakagawa, H., & Sawai, S. (2025). Use of blue fluorescent protein Electra2 for live-cell imaging in Dictyostelium discoideum. *microPublication Biology*, *2025*. https://doi.org/10.17912/micropub.biology.001774

Heidary, M., Khosravi, A. D., Khoshnood, S., Nasiri, M. J., Soleimani, S., & Goudarzi, M. (2018). Daptomycin. *Journal of Antimicrobial Chemotherapy*, *73*(1), 1–11. https://doi.org/10.1093/jac/dkx349

Helfrich, E. J. N., & Piel, J. (2016). Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports*, *33*(2), 231–316. https://doi.org/10.1039/c5np00125k

Herbst, D. A., Townsend, C. A., & Maier, T. (2018). The architectures of iterative type I PKS and FAS. *Natural Product Reports*, *35*(10), 1046–1069. https://doi.org/10.1039/c8np00039e

Hong, H., Sun, Y., Zhou, Y., Stephens, E., Samborskyy, M., & Leadlay, P. F. (2016). Evidence for an iterative module in chain elongation on the azalomycin polyketide synthase. *Beilstein Journal of Organic Chemistry*, *12*, 2164–2172. https://doi.org/10.3762/bjoc.12.206

Hopwood, D. A. (2019). Highlights of Streptomyces genetics. *Heredity*, *123*(1), 23–32. https://doi.org/10.1038/s41437-019-0196-0

Horsman, M. E., Hari, T. P. A., & Boddy, C. N. (2016). Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: Logic gate or a victim of fate? *Natural Product Reports*, *33*(2), 183–202. https://doi.org/10.1039/c4np00148f

Imran, M., Rauf, A., Abu-Izneid, T., Nadeem, M., Shariati, M. A., Khan, I. A., Imran, A., Orhan, I. E., Rizwan, M., Atif, M., Gondal, T. A., & Mubarak, M. S. (2019). Luteolin, a flavonoid, as an anticancer agent: A review. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, *112*, 108612. https://doi.org/10.1016/j.biopha.2019.108612

Institut Pasteur. (2018, October 24). *Antibiotiques: Quand les bactéries font de la résistance*. Pasteur. https://www.pasteur.fr/fr/journal-recherche/dossiers/antibiotiques-quand-bacteries-font-resistance

Jaenicke, L. (2007). Centenary of the Award of a Nobel Prize to Eduard Buchner, the Father of Biochemistry in a Test Tube and Thus of Experimental Molecular Bioscience. *Angewandte Chemie International Edition*, *46*(36), 6776–6782. https://doi.org/10.1002/anie.200700390

Jain, M., & Cox, J. S. (2005). Interaction between Polyketide Synthase and Transporter Suggests Coupled Synthesis and Export of Virulence Lipid in M. tuberculosis. *PLoS Pathogens*, *1*(1), e2. https://doi.org/10.1371/journal.ppat.0010002

Jain, R. K., Joyce, P. B. M., Molinete, M., Halban, P. A., & Gorr, S.-U. (2001). Oligomerization of green fluorescent protein in the secretory pathway of endocrine cells. *Biochemical Journal*, *360*(3), 645–649. https://doi.org/10.1042/bj3600645

Jenner, M., Kosol, S., Griffiths, D., Prasongpholchai, P., Manzi, L., Barrow, A. S., Moses, J. E., Oldham, N. J., Lewandowski, J. R., & Challis, G. L. (2018). Mechanism of intersubunit ketosynthase-dehydratase interaction in polyketide synthases. *Nature Chemical Biology*, *14*(3), 270–275. https://doi.org/10.1038/nchembio.2549

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Kaeberlein, T., Lewis, K., & Epstein, S. S. (2002). Isolating 'Uncultivable' Microorganisms in Pure Culture in a Simulated Natural Environment. *Science*, *296*(5570), 1127–1129. https://doi.org/10.1126/science.1070633

Kampa, A., Gagunashvili, A. N., Gulder, T. A. M., Morinaka, B. I., Daolio, C., Godejohann, M., Miao, V. P. W., Piel, J., & Andrésson, Ó. S. (2013). Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses. *Proceedings of the National Academy of Sciences*, *110*(33). https://doi.org/10.1073/pnas.1305867110

Katsuyama, Y., Kita, T., Funa, N., & Horinouchi, S. (2009). Curcuminoid Biosynthesis by Two Type III Polyketide Synthases in the Herb Curcuma longa. *Journal of Biological Chemistry*, *284*(17), 11160–11170. https://doi.org/10.1074/jbc.M900070200

Keatinge-Clay, A. T. (2007). A Tylosin Ketoreductase Reveals How Chirality Is Determined in Polyketides. *Chemistry & Biology*, *14*(8), 898–908. https://doi.org/10.1016/j.chembiol.2007.07.009

Keatinge-Clay, A. T. (2017). Polyketide Synthase Modules Redefined. *Angewandte Chemie (International Ed. in English)*, *56*(17), 4658–4660. https://doi.org/10.1002/anie.201701281

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., & Hopwood, D. A. (Eds.). (2000). *Practical Streptomyces genetics*. Innes.

Kotowska, M., & Pawlik, K. (2014). Roles of type II thioesterases and their application for secondary metabolite yield improvement. *Applied Microbiology and Biotechnology*, *98*(18), 7735–7746. https://doi.org/10.1007/s00253-014-5952-8

Kreuzaler, F., & Hahlbrock, K. (1975). Enzymic synthesis of an aromatic ring from acetate units. Partial purification and some properties of flavanone synthase from cell-suspension cultures of Petroselinum hortense. *European Journal of Biochemistry*, *56*(1), 205–213. https://doi.org/10.1111/j.1432-1033.1975.tb02223.x

Kumar, Y., & Goodfellow, M. (2008). Five new members of the Streptomyces violaceusniger 16S rRNA gene clade: Streptomyces castelarensis sp. nov., comb. nov., Streptomyces himastatinicus sp. nov., Streptomyces mordarskii sp. nov., Streptomyces rapamycinicus sp. nov. and Streptomyces ruanii sp. nov. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, *58*(6), 1369–1378. https://doi.org/10.1099/ijs.0.65408-0

Kuscer, E., Coates, N., Challis, I., Gregory, M., Wilkinson, B., Sheridan, R., & Petković, H. (2007). Roles of rapH and rapG in positive regulation of rapamycin biosynthesis in Streptomyces hygroscopicus. *Journal of Bacteriology*, *189*(13), 4756–4763. https://doi.org/10.1128/JB.00129-07

Kusebauch, B., Busch, B., Scherlach, K., Roth, M., & Hertweck, C. (2010). Functionally Distinct Modules Operate Two Consecutive α,β→β,γ Double-Bond Shifts in the Rhizoxin Polyketide Assembly Line. *Angewandte Chemie International Edition*, *49*(8), 1460–1464. https://doi.org/10.1002/anie.200905467

Kwan, D. H., Sun, Y., Schulz, F., Hong, H., Popovic, B., Sim-Stark, J. C. C., Haydock, S. F., & Leadlay, P. F. (2008). Prediction and Manipulation of the Stereochemistry of Enoylreduction in Modular Polyketide Synthases. *Chemistry & Biology*, *15*(11), 1231–1240. https://doi.org/10.1016/j.chembiol.2008.09.012

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, *47*(W1), W171–W174. https://doi.org/10.1093/nar/gkz365

Larcombe, D. E., Braes, R. E., Croxford, J. T., Wilson, J. W., Figurski, D. H., & Hoskisson, P. A. (2024). Sequence and origin of the Streptomyces intergenetic-conjugation helper plasmid pUZ8002. *Access Microbiology*, *6*(6). https://doi.org/10.1099/acmi.0.000808.v3

Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G. L., & Aigle, B. (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6258–6263. https://doi.org/10.1073/pnas.1019077108

Li, L., Fierer, J. O., Rapoport, T. A., & Howarth, M. (2014). Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. *Journal of Molecular Biology*, *426*(2), 309–317. https://doi.org/10.1016/j.jmb.2013.10.021

Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W., & Lu, Y. (2018). CRISPR-Cpf1-Assisted Multiplex Genome Editing and Transcriptional Repression in Streptomyces. *Applied and Environmental Microbiology*, *84*(18). https://doi.org/10.1128/AEM.00827-18

Li, W., Wu, J., Tao, W., Zhao, C., Wang, Y., He, X., Chandra, G., Zhou, X., Deng, Z., Chater, K. F., & Tao, M. (2007). A genetic and bioinformatic analysis of Streptomyces coelicolor genes containing TTA codons, possible targets for regulation by a developmentally significant tRNA. *FEMS Microbiology Letters*, *266*(1), 20–28. https://doi.org/10.1111/j.1574-6968.2006.00494.x

Li, Y., Zhang, W., Zhang, H., Tian, W., Wu, L., Wang, S., Zheng, M., Zhang, J., Sun, C., Deng, Z., Sun, Y., Qu, X., & Zhou, J. (2018). Structural Basis of a Broadly Selective Acyltransferase from the Polyketide Synthase of Splenocin. *Angewandte Chemie International Edition*, *57*(20), 5823–5827. https://doi.org/10.1002/anie.201802805

Liang, Z., Liang, H., Guo, Y., & Yang, D. (2021). Cyanidin 3-O-galactoside: A Natural Compound with Multiple Health Benefits. *International Journal of Molecular Sciences*, *22*(5), 2261. https://doi.org/10.3390/ijms22052261

Lioy, V. S., Lorenzi, J.-N., Najah, S., Poinsignon, T., Leh, H., Saulnier, C., Aigle, B., Lautru, S., Thibessard, A., Lespinet, O., Leblond, P., Jaszczyszyn, Y., Gorrichon, K., Varoquaux, N., Junier, I., Boccard, F., Pernodet, J.-L., & Bury-Moné, S. (2021). Dynamics of the compartmentalized Streptomyces chromosome during metabolic differentiation. *Nature Communications*, *12*(1), 5221. https://doi.org/10.1038/s41467-021-25462-1

Liu, C., Yuan, M., Xu, X., Wang, L., Keatinge-Clay, A. T., Deng, Z., Lin, S., & Zheng, J. (2018). Substrate-bound structures of a ketoreductase from amphotericin modular polyketide synthase. *Journal of Structural Biology*, *203*(2), 135–141. https://doi.org/10.1016/j.jsb.2018.04.001

Liu, G., Chater, K. F., Chandra, G., Niu, G., & Tan, H. (2013). Molecular regulation of antibiotic biosynthesis in streptomyces. *Microbiology and Molecular Biology Reviews: MMBR*, *77*(1), 112–143. https://doi.org/10.1128/MMBR.00054-12

Liu, H., Xue, R., Wang, Y., Stirling, E., Ye, S., Xu, J., & Ma, B. (2021). FACS-iChip: A high-efficiency iChip system for microbial 'dark matter' mining. *Marine Life Science & Technology*, *3*(2), 162–168. https://doi.org/10.1007/s42995-020-00067-7

Liu, L., Thamchaipenet, A., Fu, H., Betlach, M., & Ashley, G. (1997). Biosynthesis of 2-Nor-6-deoxyerythronolide B by Rationally Designed Domain Substitution. *Journal of the American Chemical Society*, *119*(43), 10553–10554. https://doi.org/10.1021/ja972451y

Liu, Y., Song, C., Cui, Q., Sun, H., Jiang, C., Guo, R., He, R., Li, Z., Luan, J., & Wang, H. (2025). Improving polyketide biosynthesis by rescuing the translation of truncated mRNAs into functional polyketide synthase subunits. *Nature Communications*, *16*(1), 774. https://doi.org/10.1038/s41467-025-55973-0

López-Muñoz, F., González, E., Serrano, M. D., Antequera, R., & Alamo, C. (2011). Una visión histórica de las drogas de abuso desde la perspectiva criminológica. *Cuadernos de Medicina Forense*, *17*(1), 21–33.

Luo, Z., Li, P., Zhang, D., Zhu, J., Wang, W., Zhao, W., Li, P., & Yuan, G. (2024). A Novel Antimicrobial Mechanism of Azalomycin F Acting on Lipoteichoic Acid Synthase and Cell Envelope. *Molecules (Basel, Switzerland)*, *29*(4), 856. https://doi.org/10.3390/molecules29040856

MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H., & MacNeil, T. (1992). Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene*, *111*(1), 61–68. https://doi.org/10.1016/0378-1119(92)90603-M

Marsden, A. F. A., Wilkinson, B., Cortés, J., Dunster, N. J., Staunton, J., & Leadlay, P. F. (1998). Engineering Broader Specificity into an Antibiotic-Producing Polyketide Synthase. *Science*, *279*(5348), 199–202. https://doi.org/10.1126/science.279.5348.199

Massicard, J.-M. (2020). *Polycétides syntases: Étude de l'influence de l'organisation macromoléculaire sur la fonction par ingénerie génétique*. Université de Lorraine.

Massicard, J.-M., Soligot, C., Weissman, K. J., & Jacob, C. (2020). Manipulating polyketide stereochemistry by exchange of polyketide synthase modules. *Chemical Communications (Cambridge, England)*, *56*(84), 12749–12752. https://doi.org/10.1039/d0cc05068g

McCullough, T. M., Dhar, A., Akey, D. L., Konwerski, J. R., Sherman, D. H., & Smith, J. L. (2023). Structure of a modular polyketide synthase reducing region. *Structure (London, England: 1993)*, *31*(9), 1109-1120.e3. https://doi.org/10.1016/j.str.2023.05.019

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A., & Khosla, C. (1993). Engineered Biosynthesis of Novel Polyketides. *Science*, *262*(5139), 1546–1550. https://doi.org/10.1126/science.8248802

McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., & Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel 'unnatural' natural products. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(5), 1846–1851. https://doi.org/10.1073/pnas.96.5.1846

McLean, T. C., Wilkinson, B., Hutchings, M. I., & Devine, R. (2019). Dissolution of the Disparate: Co-ordinate Regulation in Antibiotic Biosynthesis. *Antibiotics*, *8*(2), 83. https://doi.org/10.3390/antibiotics8020083

Miles, A. J., Janes, R. W., & Wallace, B. A. (2021). Tools and methods for circular dichroism spectroscopy of proteins: A tutorial review. *Chemical Society Reviews*, *50*(15), 8400–8413. https://doi.org/10.1039/d0cs00558d

Miyazawa, T., Fitzgerald, B. J., & Keatinge-Clay, A. T. (2021). Preparative production of an enantiomeric pair by engineered polyketide synthases. *Chemical Communications*, *57*(70), 8762–8765. https://doi.org/10.1039/D1CC03073F

Miyazawa, T., Hirsch, M., Zhang, Z., & Keatinge-Clay, A. T. (2020). An in vitro platform for engineering and harnessing modular polyketide synthases. *Nature Communications*, *11*(1), 80. https://doi.org/10.1038/s41467-019-13811-0

Mumm, R., Posthumus, M. A., & Dicke, M. (2008). Significance of terpenoids in induced indirect plant defence against herbivorous arthropods. *Plant, Cell & Environment*, *31*(4), 575–585. https://doi.org/10.1111/j.1365-3040.2008.01783.x

Mutak, S. (2007). Azalides from azithromycin to new azalide derivatives. *The Journal of Antibiotics*, *60*(2), 85–122. https://doi.org/10.1038/ja.2007.10

Nair, A. V., Robson, A., Ackrill, T. D., Till, M., Byrne, M. J., Back, C. R., Tiwari, K., Davies, J. A., Willis, C. L., & Race, P. R. (2020). Structure and mechanism of a dehydratase/decarboxylase enzyme couple involved in polyketide β-methyl branch incorporation. *Scientific Reports*, *10*(1), 15323. https://doi.org/10.1038/s41598-020-71850-w

National Institute for Health and Care Excellence. (2025). *A new model for evaluating and purchasing antimicrobials in the UK*. National Institute for Health and Care Excellence (NICE). https://www.nice.org.uk/what-nice-does/life-sciences-how-to-get-your-product-to-market/a-new-model-for-evaluating-and-purchasing-antimicrobials-in-the-uk

Neelam, Khatkar, A., & Sharma, K. K. (2020). Phenylpropanoids and its derivatives: Biological activities and its role in food, pharmaceutical and cosmetic industries. *Critical Reviews in Food Science and Nutrition*, *60*(16), 2655–2675. https://doi.org/10.1080/10408398.2019.1653822

Nguyen, A. W., & Daugherty, P. S. (2005). Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nature Biotechnology*, *23*(3), 355–360. https://doi.org/10.1038/nbt1066

Nichols, D., Cahoon, N., Trakhtenberg, E. M., Pham, L., Mehta, A., Belanger, A., Kanigan, T., Lewis, K., & Epstein, S. S. (2010). Use of Ichip for High-Throughput *In Situ* Cultivation of "Uncultivable" Microbial Species. *Applied and Environmental Microbiology*, *76*(8), 2445–2450. https://doi.org/10.1128/AEM.01754-09

Nicolaou, K. C., Pan, S., Shelke, Y., Rigol, S., Bao, R., Das, D., & Ye, Q. (2022). A unified strategy for the total syntheses of eribulin and a macrolactam analogue of halichondrin B. *Proceedings of the National Academy of Sciences*, *119*(32), e2208938119. https://doi.org/10.1073/pnas.2208938119

Nieselt, K., Battke, F., Herbig, A., Bruheim, P., Wentzel, A., Jakobsen, Ø. M., Sletta, H., Alam, M. T., Merlo, M. E., Moore, J., Omara, W. A., Morrissey, E. R., Juarez-Hermosillo, M. A., Rodríguez-García, A., Nentwich, M., Thomas, L., Iqbal, M., Legaie, R., Gaze, W. H., … Wellington, E. M. (2010). The dynamic architecture of the metabolic switch in Streptomyces coelicolor. *BMC Genomics*, *11*(1), 10. https://doi.org/10.1186/1471-2164-11-10

Oliynyk, M., Brown, M. J. B., Cortés, J., Staunton, J., & Leadlay, P. F. (1996). A hybrid modular polyketide synthase obtained by domain swapping. *Chemistry & Biology*, *3*(10), 833–839. https://doi.org/10.1016/S1074-5521(96)90069-1

Papacostas, G., & Gaté, J. (1928). *Les associations microbiennes: Leurs applications thérapeutiques*. Doin.

Papadaki, S., Wang, X., Wang, Y., Zhang, H., Jia, S., Liu, S., Yang, M., Zhang, D., Jia, J.-M., Köster, R. W., Namikawa, K., & Piatkevich, K. D. (2022). Dual-expression system for blue fluorescent protein optimization. *Scientific Reports*, *12*(1), 10190. https://doi.org/10.1038/s41598-022-13214-0

Payen, A., & Persoz, J.-F. (1833). Mémoire sur la Diastase, les principaux Produits de ses Réactions, et leurs applications aux arts industriels. *Annales de Chimie et de Physique*, *53*, 73–92.

Peltier, K. (1992). The new macrolactones: Clarithromycin (Biaxin) and azithromycin (Zithromax). *Connecticut Medicine*, *56*(7), 371–374.

Petkovic, H., Lill, R. E., Sheridan, R. M., Wilkinson, B., McCormick, E. L., McArthur, H. A. I., Staunton, J., Leadlay, P. F., & Kendrew, S. G. (2003). A novel erythromycin, 6-desmethyl erythromycin D, made by substituting an acyltransferase domain of the erythromycin polyketide synthase. *The Journal of Antibiotics*, *56*(6), 543–551. https://doi.org/10.7164/antibiotics.56.543

Pinnert-Sindico, S. (1954). [A new species of Streptomyces producing antibiotics Streptomyces ambofaciens n. Sp., cultural characteristics]. *Annales De l'Institut Pasteur*, *87*(6), 702–707.

Plackett, B. (2020). Why big pharma has abandoned antibiotics. *Nature*, *586*(7830), S50–S52. https://doi.org/10.1038/d41586-020-02884-3

Pulsawat, N., Kitani, S., & Nihira, T. (2007). Characterization of biosynthetic gene cluster for the production of virginiamycin M, a streptogramin type A antibiotic, in Streptomyces virginiae. *Gene*, *393*(1–2), 31–42. https://doi.org/10.1016/j.gene.2006.12.035

Ranganathan, A., Timoney, M., Bycroft, M., Cortés, J., Thomas, I. P., Wilkinson, B., Kellenberger, L., Hanefeld, U., Galloway, I. S., Staunton, J., & Leadlay, P. F. (1999). Knowledge-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: A route to simple statin analogues. *Chemistry & Biology*, *6*(10), 731–741. https://doi.org/10.1016/s1074-5521(00)80020-4

Rauf, A., Imran, M., Butt, M. S., Nadeem, M., Peters, D. G., & Mubarak, M. S. (2018). Resveratrol as an anti-cancer agent: A review. *Critical Reviews in Food Science and Nutrition*, *58*(9), 1428–1447. https://doi.org/10.1080/10408398.2016.1263597

Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C. R., & McDaniel, R. (2003). A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry*, *42*(1), 72–79. https://doi.org/10.1021/bi0268706

Reinke, A. W., Grant, R. A., & Keating, A. E. (2010). A Synthetic Coiled-Coil Interactome Provides Heterospecific Modules for Molecular Engineering. *Journal of the American Chemical Society*, *132*(17), 6025–6031. https://doi.org/10.1021/ja907617a

Résistance aux antimicrobiens: Accélérer les ripostes nationales et mondiales (A77/5; Assemblée mondiale de la santé). (2024). Organisation mondiale de la santé. https://apps.who.int/gb/ebwha/pdf_files/WHA77/A77_5-fr.pdf

Ridley, C. P., Lee, H. Y., & Khosla, C. (2008). Evolution of polyketide synthases in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(12), 4595–4600. https://doi.org/10.1073/pnas.0710107105

Risser, F., Collin, S., Dos Santos-Morais, R., Gruez, A., Chagot, B., & Weissman, K. J. (2020). Towards improved understanding of intersubunit interactions in modular polyketide biosynthesis: Docking in the enacyloxin IIa polyketide synthase. *Journal of Structural Biology*, *212*(1), 107581. https://doi.org/10.1016/j.jsb.2020.107581

Ruan, X., Pereda, A., Stassi, D. L., Zeidner, D., Summers, R. G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M. J., Donadio, S., & Katz, L. (1997). Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. *Journal of Bacteriology*, *179*(20), 6416–6425. https://doi.org/10.1128/jb.179.20.6416-6425.1997

Savitzky, Abraham., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Scat, S., Weissman, K. J., & Chagot, B. (2024). Insights into docking in megasynthases from the investigation of the toblerol trans-AT polyketide synthase: Many α-helical means to an end. *RSC Chemical Biology*, *5*(7), 669–683. https://doi.org/10.1039/d4cb00075g

Schatz, A., Bugle, E., & Waksman, S. A. (1944). Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Experimental Biology and Medicine*, *55*(1). https://doi.org/10.3181/00379727-55-14461

Schlimpert, S., Wasserstrom, S., Chandra, G., Bibb, M. J., Findlay, K. C., Flärdh, K., & Buttner, M. J. (2017). Two dynamin-like proteins stabilize FtsZ rings during *Streptomyces* sporulation. *Proceedings of the National Academy of Sciences*, *114*(30). https://doi.org/10.1073/pnas.1704612114

Schwecke, T., Aparicio, J. F., Molnar, I., Konig, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortes, J., & Lester, J. B. (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proceedings of the National Academy of Sciences*, *92*(17), 7839–7843. https://doi.org/10.1073/pnas.92.17.7839

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. Red fluorescent protein. *Nature Biotechnology*, *22*(12), 1567–1572. https://doi.org/10.1038/nbt1037

Shi, C., Miller, B. R., Alexander, E. M., Gulick, A. M., & Aldrich, C. C. (2020). Design, Synthesis, and Biophysical Evaluation of Mechanism-Based Probes for Condensation Domains of Nonribosomal Peptide Synthetases. *ACS Chemical Biology*, *15*(7), 1813–1819. https://doi.org/10.1021/acschembio.0c00411

Shimizu, Y., Ogata, H., & Goto, S. (2017). Type III Polyketide Synthases: Functional Classification and Phylogenomics. *ChemBioChem*, *18*(1), 50–65. https://doi.org/10.1002/cbic.201600522

Smith, H. G., Beech, M. J., Lewandowski, J. R., Challis, G. L., & Jenner, M. (2021). Docking domain-mediated subunit interactions in natural product megasynth(et)ases. *Journal of Industrial Microbiology and Biotechnology*, *48*(3–4), kuab018. https://doi.org/10.1093/jimb/kuab018

Staunton, J., & Weissman, K. J. (2001). Polyketide biosynthesis: A millennium review. *Natural Product Reports*, *18*(4), 380–416. https://doi.org/10.1039/a909079g

Straight, P. D., Fischbach, M. A., Walsh, C. T., Rudner, D. Z., & Kolter, R. (2007). A singular enzymatic megacomplex from Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 305–310. https://doi.org/10.1073/pnas.0609073103

Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, *189*(1), 113–130. https://doi.org/10.1016/0022-2836(86)90385-2

Su, L. (2021). *Synthèse d'analogues de stambomycine, une famille de polycétides présentant des activités antitumorales, par ingénierie génétique et approches associées*. Université de Lorraine.

Su, L., Hôtel, L., Paris, C., Chepkirui, C., Brachmann, A. O., Piel, J., Jacob, C., Aigle, B., & Weissman, K. J. (2022). Engineering the stambomycin modular polyketide synthase yields 37-membered mini-stambomycins. *Nature Communications*, *13*(1), 515. https://doi.org/10.1038/s41467-022-27955-z

Su, L., Souaibou, Y., Hôtel, L., Jacob, C., Grün, P., Shi, Y.-N., Chateau, A., Pinel, S., Bode, H. B., Aigle, B., & Weissman, K. J. (2025). Exploiting the inherent promiscuity of the acyl transferase of the stambomycin polyketide synthase for the mutasynthesis of analogues. *Chemical Science*, *16*(12), 5076–5088. https://doi.org/10.1039/d4sc06976e

Suhren, O. (1951). [First clinical experience with pikromycin, a new antibiotic from actinomyces in therapy of pyoderma]. *Medizinische Klinik*, *46*(25–26), 722–723.

Sun, C., Yang, Z., Zhang, C., Liu, Z., He, J., Liu, Q., Zhang, T., Ju, J., & Ma, J. (2019). Genome Mining of *Streptomyces atratus* SCSIO ZH16: Discovery of Atratumycin and Identification of Its Biosynthetic Gene Cluster. *Organic Letters*, *21*(5), 1453–1457. https://doi.org/10.1021/acs.orglett.9b00208

Sun, X., Yuan, Y., Chen, Q., Nie, S., Guo, J., Ou, Z., Huang, M., Deng, Z., Liu, T., & Ma, T. (2022). Metabolic pathway assembly using docking domains from type I cis-AT polyketide synthases. *Nature Communications*, *13*(1), 5541. https://doi.org/10.1038/s41467-022-33272-2

Sung, A., Gromek, S., & Balunas, M. (2017). Upregulation and Identification of Antibiotic Activity of a Marine-Derived Streptomyces sp. Via Co-Cultures with Human Pathogens. *Marine Drugs*, *15*(8), 250. https://doi.org/10.3390/md15080250

Takano, E., Kinoshita, H., Mersinias, V., Bucca, G., Hotchkiss, G., Nihira, T., Smith, C. P., Bibb, M., Wohlleben, W., & Chater, K. (2005). A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of Streptomyces coelicolor. *Molecular Microbiology*, *56*(2), 465–479. https://doi.org/10.1111/j.1365-2958.2005.04543.x

Tang, Y., Kim, C.-Y., Mathews, I. I., Cane, D. E., & Khosla, C. (2006). The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proceedings of the National Academy of Sciences*, *103*(30), 11124–11129. https://doi.org/10.1073/pnas.0601924103

Thanapipatsiri, A., Gomez-Escribano, J. P., Song, L., Bibb, M. J., Al-Bassam, M., Chandra, G., Thamchaipenet, A., Challis, G. L., & Bibb, M. J. (2016). Discovery of Unusual Biaryl Polyketides by Activation of a Silent Streptomyces venezuelae Biosynthetic Gene Cluster. *Chembiochem: A European Journal of Chemical Biology*, *17*(22), 2189–2198. https://doi.org/10.1002/cbic.201600396

Thompson, K. E., Bashor, C. J., Lim, W. A., & Keating, A. E. (2012). SYNZIP protein interaction toolbox: In vitro and in vivo specifications of heterospecific coiled-coil interaction domains. *ACS Synthetic Biology*, *1*(4), 118–129. https://doi.org/10.1021/sb200015u

Tittes, Y. U., Herbst, D. A., Martin, S. F. X., Munoz-Hernandez, H., Jakob, R. P., & Maier, T. (2022). The structure of a polyketide synthase bimodule core. *Science Advances*, *8*(38), eabo6918. https://doi.org/10.1126/sciadv.abo6918

Torres, J. P., & Schmidt, E. W. (2019). The biosynthetic diversity of the animal world. *Journal of Biological Chemistry*, *294*(46), 17684–17692. https://doi.org/10.1074/jbc.REV119.006130

Tsai, S.-C., Miercke, L. J. W., Krucinski, J., Gokhale, R., Chen, J. C.-H., Foster, P. G., Cane, D. E., Khosla, C., & Stroud, R. M. (2001). Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: Versatility from a unique substrate channel. *Proceedings of the National Academy of Sciences*, *98*(26), 14808–14813. https://doi.org/10.1073/pnas.011399198

Tsai, S.-C. S., & Ames, B. D. (2009). Structural enzymology of polyketide synthases. *Methods in Enzymology*, *459*, 17–47. https://doi.org/10.1016/S0076-6879(09)04602-3

Valenzano, C. R., You, Y.-O., Garg, A., Keatinge-Clay, A., Khosla, C., & Cane, D. E. (2010). Stereospecificity of the Dehydratase Domain of the Erythromycin Polyketide Synthase. *Journal of the American Chemical Society*, *132*(42), 14697–14699. https://doi.org/10.1021/ja107344h

Vogt, T. (2010). Phenylpropanoid Biosynthesis. *Molecular Plant*, *3*(1), 2–20. https://doi.org/10.1093/mp/ssp106

Wainwright, M. (1989). Nobel Infallibility. *Nature*, *342*, 336.

Waksman, S. A., & Lechevalier, H. A. (1949). Neomycin, a New Antibiotic Active against Streptomycin-Resistant Bacteria, including Tuberculosis Organisms. *Science*, *109*(2830), 305–307. https://doi.org/10.1126/science.109.2830.305

Wang, Z.-C., Stegall, H., Miyazawa, T., & Keatinge-Clay, A. T. (2025). A CRISPR-Cas9 system for knock-out and knock-in of high molecular weight DNA enables module-swapping of the pikromycin synthase in its native host. *Microbial Cell Factories*, *24*(1), 125. https://doi.org/10.1186/s12934-025-02741-w

Weissman, K. J. (2009). Introduction to polyketide biosynthesis. *Methods in Enzymology*, *459*, 3–16. https://doi.org/10.1016/S0076-6879(09)04601-1

Weissman, K. J. (2016). Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Natural Product Reports*, *33*(2), 203–230. https://doi.org/10.1039/c5np00109a

Weissman, K. J. (2017). Polyketide stereocontrol: A study in chemical biology. *Beilstein Journal of Organic Chemistry*, *13*, 348–371. https://doi.org/10.3762/bjoc.13.39

Whicher, J. R., Smaga, S. S., Hansen, D. A., Brown, W. C., Gerwick, W. H., Sherman, D. H., & Smith, J. L. (2013). Cyanobacterial Polyketide Synthase Docking Domains: A Tool for Engineering Natural Product Biosynthesis. *Chemistry & Biology*, *20*(11), 1340–1351. https://doi.org/10.1016/j.chembiol.2013.09.015

Willemse, J., & van Wezel, G. P. (2009). Imaging of Streptomyces coelicolor A3(2) with reduced autofluorescence reveals a novel stage of FtsZ localization. *PloS One*, *4*(1), e4242. https://doi.org/10.1371/journal.pone.0004242

Wilson, D. J., Xue, Y., Reynolds, K. A., & Sherman, D. H. (2001). Characterization and analysis of the PikD regulatory factor in the pikromycin biosynthetic pathway of Streptomyces venezuelae. *Journal of Bacteriology*, *183*(11), 3468–3475. https://doi.org/10.1128/JB.183.11.3468-3475.2001

Wlodek, A., Kendrew, S. G., Coates, N. J., Hold, A., Pogwizd, J., Rudder, S., Sheehan, L. S., Higginbotham, S. J., Stanley-Smith, A. E., Warneck, T., Nur-E-Alam, M., Radzom, M., Martin, C. J., Overvoorde, L., Samborskyy, M., Alt, S., Heine, D., Carter, G. T., Graziani, E. I., … Gregory, M. A. (2017). Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nature Communications*, *8*(1), 1206. https://doi.org/10.1038/s41467-017-01344-3

World Health Organization. (2024). *Global research agenda for antimicrobial resistance in human health* (World Health Organization).

Xie, X., & Cane, D. E. (2018). pH-Rate profiles establish that polyketide synthase dehydratase domains utilize a single-base mechanism. *Organic & Biomolecular Chemistry*, *16*(47), 9165–9170. https://doi.org/10.1039/c8ob02637h

Xue, Y., Zhao, L., Liu, H., & Sherman, D. H. (1998). A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity. *Proceedings of the National Academy of Sciences*, *95*(21), 12111–12116. https://doi.org/10.1073/pnas.95.21.12111

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., Fedorova, I., Nakane, T., Makarova, K. S., Koonin, E. V., Ishitani, R., Zhang, F., & Nureki, O. (2016). Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell*, *165*(4), 949–962. https://doi.org/10.1016/j.cell.2016.04.003

Yoon, Y. J., Kim, E.-S., Hwang, Y.-S., & Choi, C.-Y. (2004). Avermectin: Biochemical and molecular basis of its biosynthesis and regulation. *Applied Microbiology and Biotechnology*, *63*(6), 626–634. https://doi.org/10.1007/s00253-003-1491-4

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., & Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(12), E690-697. https://doi.org/10.1073/pnas.1115485109

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, *163*(3), 759–771. https://doi.org/10.1016/j.cell.2015.09.038

Zhai, G., Wang, W., Xu, W., Sun, G., Hu, C., Wu, X., Cong, Z., Deng, L., Shi, Y., Leadlay, P. F., Song, H., Hong, K., Deng, Z., & Sun, Y. (2020). Cross-Module Enoylreduction in the Azalomycin F Polyketide Synthase. *Angewandte Chemie International Edition*, *59*(50), 22738–22742. https://doi.org/10.1002/anie.202011357

Zhai, G., Zhu, Y., Sun, G., Zhou, F., Sun, Y., Hong, Z., Dong, C., Leadlay, P. F., Hong, K., Deng, Z., Zhou, F., & Sun, Y. (2023). Insights into azalomycin F assembly-line contribute to evolution-guided polyketide synthase engineering and identification of intermodular recognition. *Nature Communications*, *14*(1), 612. https://doi.org/10.1038/s41467-023-36213-9

Zhang, Z., Du, C., de Barsy, F., Liem, M., Liakopoulos, A., van Wezel, G. P., Choi, Y. H., Claessen, D., & Rozen, D. E. (2020). Antibiotic production in Streptomyces is organized by a division of labor through terminal genomic differentiation. *Science Advances*, *6*(3), eaay5781. https://doi.org/10.1126/sciadv.aay5781

Zhao, L. Y., Shi, J., Xu, Z. Y., Sun, J. L., Yan, Z. Y., Tong, Z. W., Tan, R. X., Jiao, R. H., & Ge, H. M. (2024). Hybrid Type I and II Polyketide Synthases Yield Distinct Aromatic Polyketides. *Journal of the American Chemical Society*, *146*(43), 29462–29468. https://doi.org/10.1021/jacs.4c08803

Zheng, J., Gay, D. C., Demeler, B., White, M. A., & Keatinge-Clay, A. T. (2012). Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nature Chemical Biology*, *8*(7), 615–621. https://doi.org/10.1038/nchembio.964

Résumés en français et anglais

Résumé en français

Face à la propagation de l'antibiorésistance et autres mécanismes de résistance, et au développement de nouvelles pathologies, la recherche de biomolécules innovantes et leur optimisation est un enjeu de santé majeur. Les polycétides sont une source inestimable de composés d'intérêts en médecine et en agriculture, car ils possèdent un large éventail de structures et de propriétés. Ils sont produits par des animaux, des plantes et des champignons, mais les plus grands producteurs de polycétides sont les bactéries du genre *Streptomyces*. Une stratégie prometteuse pour obtenir de nouvelles biomolécules est de modifier les enzymes responsables de leur production : les polycétide synthases (PKS).

L'organisation particulière des PKS de type I modulaires en font des sujets d'études intéressants. Elles sont organisées en complexes de plusieurs chaînes polypeptidiques appelées sous-unités, composées de modules comprenant différents domaines. Ces sous-unités interagissent *via* des domaines de docking (DD). Le nombre, la composition et l'organisation des modules permet de déterminer la structure du squelette polycétidique. Cette règle de colinéarité a inspiré des décennies de tentatives d'ingénierie pour obtenir et optimiser de nouvelles biomolécules. Cependant, ces tentatives ont très souvent conduit à l'obtention de systèmes modifiés dont les rendements de production de dérivés de polycétides étaient bien plus faibles que les rendements obtenus avec le complexe enzymatique parental, voire inexistant.

La majorité des sous-unités caractérisées sont composées d'un nombre restreint de modules. Cette observation nous a conduit à l'hypothèse que cet arrangement pourrait être un avantage évolutif, par la flexibilité que cela leur confère au sein des cellules. Le premier objectif de cette thèse a donc été d'investiguer la scission de grandes sous-unités de PKS en deux comme manière d'augmenter les rendements de production. Pour cela, deux systèmes PKS ont été étudiés : 1) la rapamycine synthase retrouvée chez *S. hygroscopicus* dont la sous-unité B, composée de 6 modules, a été scindée par l'introduction de DD exogènes entre les modules 3 et 4. Dans ce cadre, les vecteurs d'édition génétique ont été construits et les interactions entre les DD natifs et insérés ont en partie pu être caractérisées, afin d'assurer l'absence d'interaction non canonique ; 2) la stambomycine synthase de *S. ambofaciens* dont la sous-unité 3, composée de 5 modules, a été scindée en insérant une interface d'interaction synthétique formée de *Synthetic Zippers* (SynZips) entre les modules 2 et 3. Cette modification génique a été appliquée à la souche exprimant la PKS sauvage, ainsi qu'à une souche mutante codant une PKS miniaturisée synthétisant des mini-stambomycines. Ici, les souches nécessaires à l'étude de la productivité de PKS scindées ont été obtenues, et leurs analyses sont en cours.

Un autre frein aux tentatives d'ingénierie des PKS modulaires est leur macro-complexation. En effet, la sous-famille des PKS *trans*-AT a été caractérisée comme formant des macro-complexes proches de la membrane bactérienne. Les motifs identifiés comme pouvant permettre cette macro-organisation étant absents de la sous-famille des PKS *cis*-AT, et aucune étude n'ayant été menée sur ces systèmes, leur localisation et organisation au sein de leur hôte bactérien reste à ce jour inconnue. Le second objectif de cette thèse a donc été de localiser et observer l'organisation de la coelimycine synthase, retrouvée chez *S. coelicolor*. Le marquage des enzymes de ce complexe par des protéines fluorescentes, couplé à de la microscopie confocale, a permis d'établir que les première et dernière sous-unités forment des foci qui colocalisent au sein d'îlots d'hyphes bactériens dans la colonie. Le transporteur associé est quant à lui retrouvé au niveau de la membrane des hyphes de toute la colonie. Ces résultats permettront par la suite d'investiguer les bases moléculaires d'une telle organisation.

Abstract in English

Antibiotic resistance and the discovery of new resistance profiles and new diseases are a major threat to human health, and the discovery and optimization of new therapeutic compounds is an urgent need. Polyketides are an invaluable source of such compounds, with high structural and proprieties variabilities. They are produced by animals, plants and fungi, but the main producers of polyketides remain bacteria of the genus *Streptomyces*. A promising strategy to obtain and optimize new molecules of human's interest is to engineer the enzymes producing polyketides: polyketide synthases (PKSs).

Modular type I PKS are very interesting to study as their organization induce the possibility to guess the structure of the produced compound from their peptide sequences. Indeed, they are composed of different domains, organized into modules distributed amongst peptide chains called subunits and interacting *via* docking domains (DDs). The composition of these modules allows to determine the structure of the produced polyketide. This collinearity rule has inspired decades of engineering attempts to produce new molecules. However, these attempts often lead to reduced or even non-existent yields compared to the native systems.

Most of the subunits characterized to date are composed of a small number of modules. This might be an evolutionary advantage. The first objective of this thesis project was to study the splitting of PKSs to improve yields. The rapamycin synthase was studied in this aim, using DDs to split the larger B subunit, composed of 6 modules. The chosen DDs were placed between the 3rd and 4th module of the subunit. Plasmids containing the sequences needed to engineer *S. hygroscopicus* genome were obtained and the interactions between native and heterologous DDs were partially characterized to avoid non-canonical interactions. The stambomycins synthase, found in *S. ambofaciens*, was also studied, using synthetic interaction domains called Synthetic Zippers (SynZips) introduced between the 2nd and 3rd module of the 5-modules containing subunit 3. This splitting strategy was also used on an engineered strain producing a reduced PKS producing mini-stambomycins, to try improving its yield. Strains expressing the split PKS were obtained and are currently undergoing further analyses to assess the impact on their productivity.

Another bottleneck to engineered modular PKSs lies in their large weight and macro-organization, influencing on the efficiency of heterologous expression. The *trans*-AT PKSs sub-family has been characterized as forming foci colocalized near the bacterial membrane. However, the motif identified as mediating such macro-organization in *trans*-AT PKSs has not been found in *cis*-AT PKSs, and no study has been led on their own organization. The second objective of this thesis project was to study the localization and organization of the coelimycin synthase in its native host *S. coelicolor*. Labelling the enzymes of this complex with fluorescent proteins, coupled with confocal microscopy, has established that the first and last subunits of this system were found forming foci colocalized in hyphal islands of the colony. The PKS-associated transporter was also studied and observed in the membrane of all hyphae in the colony, forming foci as well. These results pave the way for a better understanding of the molecular basis underlying such an organization.