



**HAL**  
open science

# Polycétide synthases : étude de l'influence de l'organisation macromoléculaire sur la fonction par ingénierie génétique

Jean-Malo Massicard

► **To cite this version:**

Jean-Malo Massicard. Polycétide synthases : étude de l'influence de l'organisation macromoléculaire sur la fonction par ingénierie génétique. Sciences du Vivant [q-bio]. Université de Lorraine, 2020. Français. NNT : 2020LORR0278 . tel-03275171

**HAL Id: tel-03275171**

**<https://hal.univ-lorraine.fr/tel-03275171>**

Submitted on 30 Jun 2021

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

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



## AVERTISSEMENT

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.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

## École Doctorale BioSE (Biologie-Santé-Environnement)

### Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

**DOCTEUR DE L'UNIVERSITE DE LORRAINE**

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

par **Jean-Malo MASSICARD**

## **Polycétide synthases : étude de l'influence de l'organisation macromoléculaire sur la fonction par ingénierie génétique**

**Le 18 décembre 2020**

### **Membres du jury :**

<b>Rapporteuses :</b>	<b>Mme Yanyan LI</b>	<b>Chargée de recherche, CNRS, Paris</b>
	<b>Mme Valérie LECLERE</b>	<b>Professeure, Université de Lille, Lille</b>
<b>Examineurs :</b>	<b>Mme Muriel GONDRY</b>	<b>Chercheuse, CEA, Paris</b>
	<b>M. Jean-Luc PERNODET</b>	<b>Directeur de recherche, CNRS, Paris</b>
	<b>M. Bertrand AIGLE</b>	<b>Professeur, Université de Lorraine, Nancy</b>
	<b>Mme Kira WEISSMAN</b>	<b>Professeure, Université de Lorraine, Nancy, directrice de thèse</b>
	<b>M. Christophe JACOB</b>	<b>Maître de Conférences, Université de Lorraine, Nancy, co-directeur de thèse</b>

---

**Laboratoire IMoPA: Ingénierie Moléculaire et Physiopathologie Articulaire  
UMR 7365 CNRS-UL, Biopôle de l'Université de Lorraine, Campus Brabois-Santé,  
9 Avenue de la Forêt de Haye, BP 20199, 54505 Vandœuvre-lès-Nancy**





## École Doctorale BioSE (Biologie-Santé-Environnement)

### Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

**DOCTEUR DE L'UNIVERSITE DE LORRAINE**

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

par **Jean-Malo MASSICARD**

## **Polycétide synthases : étude de l'influence de l'organisation macromoléculaire sur la fonction par ingénierie génétique**

**Le 18 décembre 2020**

### **Membres du jury :**

<b>Rapporteuses :</b>	<b>Mme Yanyan LI</b>	<b>Chargée de recherche, CNRS, Paris</b>
	<b>Mme Valérie LECLERE</b>	<b>Professeure, Université de Lille, Lille</b>
<b>Examineurs :</b>	<b>Mme Muriel GONDRY</b>	<b>Chercheuse, CEA, Paris</b>
	<b>M. Jean-Luc PERNODET</b>	<b>Directeur de recherche, CNRS, Paris</b>
	<b>M. Bertrand AIGLE</b>	<b>Professeur, Université de Lorraine, Nancy</b>
	<b>Mme Kira WEISSMAN</b>	<b>Professeure, Université de Lorraine, Nancy, directrice de thèse</b>
	<b>M. Christophe JACOB</b>	<b>Maître de Conférences, Université de Lorraine, Nancy, co-directeur de thèse</b>

---

**Laboratoire IMoPA: Ingénierie Moléculaire et Physiopathologie Articulaire  
UMR 7365 CNRS-UL, Biopôle de l'Université de Lorraine, Campus Brabois-Santé,  
9 Avenue de la Forêt de Haye, BP 20199, 54505 Vandœuvre-lès-Nancy**



# Remerciements



## Remerciements

*Je voudrais tout d'abord remercier grandement ma directrice de thèse, Madame Kira WEISSMAN, professeure à l'Université de Lorraine, et mon co-directeur de thèse, Monsieur Christophe JACOB, maître de conférences à l'Université de Lorraine, pour toute leur aide. Je suis ravi d'avoir travaillé en leur compagnie car outre leur appui scientifique, ils ont toujours été là pour me soutenir et me conseiller au cours de cette thèse. Qu'ils soient aussi remerciés pour leur gentillesse, leur disponibilité permanente et pour les nombreux encouragements qu'ils m'ont prodigués.*

*J'adresse ensuite tous mes remerciements à Madame Yanyan LI, chargée de recherche au CNRS, ainsi que Madame Valérie LECLERE, professeure à l'Université de Lille, de l'honneur qu'elles m'ont fait en acceptant d'être rapporteuses de cette thèse.*

*Je remercie Monsieur Bertrand AIGLE, professeur à l'Université de Lorraine, ainsi que Monsieur Jean-Luc PERNODET directeur de recherche au CNRS, de l'attention qu'ils m'ont porté en acceptant d'être membre de mon comité de suivi de thèse. Ils ont pris le temps de m'écouter et de discuter avec moi. Leurs remarques m'ont permis de faire mûrir considérablement ma thèse. Je les remercie également avec Madame Muriel GONDRY, chercheuse CEA, d'avoir accepté d'être examinateurs de cette thèse.*

*Je tiens aussi à remercier Monsieur Jean-Yves JOUZEAU, directeur du laboratoire IMoPA, et Monsieur Bruno CHARPENTIER, directeur adjoint du laboratoire IMoPA, qui m'ont accueilli pendant trois ans au sein du laboratoire.*

*Je remercie tous les membres du laboratoire IMoPA, qui ont répondu avec calme et patience aux questions quotidiennes dont je les accablais. J'ai une attention toute particulière pour Madame Sabrina COLLIN, ingénieure d'étude à l'Université de Lorraine, pour son soutien, sa gentillesse et sa bienveillance, qui m'ont accompagnés tout le long de ma thèse.*

*My next thanks go to Mister Gilles P. VAN WEZEL, professor at Leiden University, as well as to Madame Arianne BRIEGEL, professor at Leiden University, for welcoming me to their team at the Institute of Biology of Leiden and for their expert advice. I would also like to thank Madame Paula YAGÜE, Post-doctoral fellow, for her invaluable help in my research.*

*I would like to thank Mister Erik KRISTENSEN, Professor at Southern Denmark University, who welcomed me to his marine ecology laboratory for my first internship and made me want to do research. This first professional experience abroad and this idyllic setting were a trigger in my studies.*

*In this same way, I thank Madame Nina ENTELIS, Research director at CNRS, and Mister Ivan TARASSOV, director of the UMR 7156 Molecular Genetics Genomics Microbiology at Strasbourg, for welcoming me to my M1 internship and which in turn gave me this desire to do*

## *Remerciements*

*research. I would also like to thank Madame Anna SMIRNOVA, Post-doctoral fellow, for advancing me technically and in methodology.*

*Je tiens à remercier particulièrement Monsieur Arnaud GRUEZ, maître de conférences à l'Université de Lorraine, sans qui je n'aurais pas entendu parler de cette thèse. Je souhaite porter un remerciement spécial à Monsieur François TALFOURNIER, sans qui je n'aurais pas pris la biochimie pour affection. Je remercie également Monsieur Mathieu REDERSTORFF, Madame Hortense MAZON, Monsieur Stéphane LABIALLE, Monsieur Sylvain MAENNER, Monsieur Benjamin CHAGOT, tous maître de conférences, et tous les chercheurs et spécialistes, trop nombreux pour tous les citer, qui ont pris le temps de discuter de mon sujet. Chacun de ces échanges m'a aidé à faire avancer mes analyses. Je les remercie aussi tous pour toutes nos discussions et leurs conseils qui m'ont accompagnés tout au long de mon cursus à l'Université de Lorraine.*

*Je remercie toutes les personnes avec qui j'ai partagé mes études et notamment ces années de thèse. J'ai une pensée toute particulière à Monsieur Romain FRANÇOIS, en première année de doctorat, qui sera prendre le relais au sein de l'équipe PKS.*

*Je remercie le Ministère de l'Enseignement Supérieure et de la Recherche et l'Université de Lorraine pour le financement doctoral. Je remercie aussi le dispositif « Doctor, Explore and Achieve More! » (DrEAM) dans le cadre de Lorraine Université d'Excellence (LUE) de m'avoir soutenu financièrement pour évoluer à l'Institut de Biologie de Leiden.*

*Je remercie mes amis Matthieu, Justine, Rémy, Fanny, Thomas, Laeya, Louis, Florent, Guillaume, Géraud, Jean-Christophe, qui ont toujours été là pour moi. Leur soutien inconditionnel et leurs encouragements ont été d'une grande aide. À ceux qui manquent à l'appelle...*

*Mes derniers remerciements vont à ma famille et en particulier à Madame Christine COLSON, ma maman, qui a tout fait pour m'aider, qui m'a soutenu et surtout supporté dans tout ce que j'ai entrepris. À ce même titre, je remercie mes grands-parents, Madame Chantal COLSON et Monsieur Claude COLSON, pour leur soutien précieux tout le long de mon parcours. Je remercie Monsieur Jean-Patrick MASSICARD, mon papa, pour ses conseils avisés. Je remercie Madame Maëlle MASSICARD, ma petite sœur, pour ses conseils concernant mon style d'écriture, qui ont grandement facilité la rédaction de cette thèse. Je la remercie également, avec Madame Servane MASSICARD et Madame Solenne MASSICARD, mes grandes sœurs, pour leur fierté, leur soutien et leur intérêt porté à mes recherches. À Madame Francine MASSICARD, ma grand-mère, partie au début de cette année... À mes nièces, Lyne, Léonor et Lore, qui comprendront à leur tour, un jour peut-être, l'exigence des études longues...*

« La vraie science est une ignorance qui se sait. »

Montaigne, *Essais*, Livre II, Chapitre XII



# Sommaire



## Liste des abréviations

## Liste des figures

## Liste des tableaux

<b>État de l'art .....</b>	<b>37</b>
<b>I. Les métabolites spécialisés .....</b>	<b>39</b>
1. Définition.....	39
2. Les classes de métabolites spécialisés .....	40
3. Intérêts pour les métabolites spécialisés.....	42
4. Découvrir de nouveaux polycétides .....	43
A. La recherche dans la biosphère .....	45
B. La synthèse totale ou l'hémi-synthèse.....	47
C. Vers une troisième approche .....	48
<b>II. Les polycétide synthases .....</b>	<b>50</b>
1. Les PKS de type III.....	50
2. Les PKS de type II.....	51
3. Les PKS de type I.....	53
A. Les PKS de type I itératives .....	54
B. Les PKS de type I modulaires .....	54
a. Les PKS modulaires de type <i>cis</i> -AT .....	55
b. Les PKS modulaires de type <i>trans</i> -AT .....	57
4. La biologie synthétique des PKS modulaires.....	59
A. Les PKS modulaires de type <i>cis</i> -AT comme modèle d'étude.....	59
B. Les cibles de la biologie synthétique des PKS modulaires .....	59

### **III. Du contrôle de la stéréochimie à l'ingénierie des domaines KR..... 62**

#### **1. Intérêt pour la stéréochimie ..... 62**

#### **2. Le contrôle de la stéréochimie lors de la synthèse du polycétide..... 63**

A. Le domaine *acyl carrier protein* (ACP) ..... 64

B. Le domaine acyltransférase (AT) ..... 66

C. Le domaine cétyosynthase (KS) ..... 69

D. Le domaine cétoréductase (KR) ..... 73

E. Les autres réactions de réduction ..... 78

a. Le domaine déshydratase (DH) ..... 78

b. Le domaine ényl réductase (ER) ..... 79

F. Le domaine thioestérase (TE) ..... 79

#### **3. Les échanges de domaine pour modifier la stéréochimie ..... 81**

A. La construction d'un système biosynthétique modèle ..... 81

B. Les échanges de domaines KR ..... 82

### **IV. Du contexte modulaire à l'ingénierie des modules ..... 85**

#### **1. L'architecture d'un module ..... 86**

#### **2. La communication intra- et inter-modulaire..... 88**

A. La communication intra-modulaire ..... 89

a. Les contacts AT-ACP ..... 89

b. Les contacts KS-ACP ..... 90

c. Les contacts KR-ACP ..... 90

d. Les contacts TE-ACP ..... 91

B. La communication inter-modulaire ..... 92

a. Les linkers inter-modulaires ..... 92

b. Les domaines de docking ..... 92

#### **3. Les échanges de module ..... 95**

A. Les échanges du module de chargement ..... 95

B. Les échanges d'un module d'extension ..... 97

a. Les stratégies non-rationnelles ..... 97

b. La recherche de points de fusion naturels ..... 101

c. Les mécanismes naturels de recombinaison pour générer de la diversité .....	107
<b>4. Vers une nouvelle définition d'un module .....</b>	<b>108</b>
A. Le modèle actuel .....	108
B. Le modèle alternatif .....	109
C. Les dilemmes de l'ingénierie des modules .....	110
<b>V. L'organisation macromoléculaire des PKS modulaires .....</b>	<b>113</b>
<b>1. Le mégacomplexe chez <i>Bacillus subtilis</i> .....</b>	<b>113</b>
<b>2. Le modèle actuel du mode d'assemblage .....</b>	<b>118</b>
<b>Objectifs .....</b>	<b>121</b>
<b>Résultats.....</b>	<b>127</b>
<b>I. Étude de l'influence du contexte modulaire sur le contrôle de la stéréochimie .....</b>	<b>129</b>
<b>1. Échange de modules dans un système biosynthétique modèle.....</b>	<b>129</b>
A. Résumé de l'étude .....	129
B. Publication n°1 .....	134
<b>2. Échange de modules dans la DEBS .....</b>	<b>135</b>
<b>II. Étude de la localisation des systèmes PKS au sein des hôtes bactériens .....</b>	<b>137</b>
<b>1. Activation du BGC <i>cpk</i> et visualisation de la coelimycine P2 .....</b>	<b>137</b>
<b>2. Construction des protéines de fusion.....</b>	<b>140</b>
A. Insertion des protéines fluorescentes par la technique CRISPR-Cas9 .....	140
B. Contrôle de l'expression et de la fonction des sous-unités chimériques .....	143
<b>3. Localisation sous-cellulaire des protéines de PKS .....</b>	<b>145</b>
A. CpkA et CpkC localisent en loci dans les hyphes de <i>S. coelicolor</i> .....	145

## Sommaire

B. Colocalisation de CpkA avec CpkC .....	148
C. Colocalisation de CpkC avec CpkF.....	150
<b>III. Ingénierie de la biosynthèse des produits naturels.....</b>	<b>153</b>
1. Contexte du chapitre.....	153
2. Chapitre d'ouvrage n°1 .....	154
<b>Discussion .....</b>	<b>155</b>
<b>Matériels et méthodes .....</b>	<b>165</b>
<b>I. Matériel biologique.....</b>	<b>167</b>
1. Souches de bactéries et de levure .....	167
2. Vecteurs plasmidiques .....	170
3. Cosmide.....	177
<b>II. Milieux de cultures des micro-organismes.....</b>	<b>178</b>
1. Milieux de culture d' <i>E. coli</i> .....	178
2. Milieux de culture de <i>Saccharomyces cerevisiae</i> .....	179
3. Milieux de culture de <i>Sac. erythraea</i> et de <i>S. coelicolor</i> .....	180
<b>III. Solutions tampon .....</b>	<b>183</b>
<b>IV. Construction des plasmides .....</b>	<b>184</b>
1. Clonage de la cassette URA3-ARS-CEN6 dans le plasmide pMU3-DEBS 1-TE .....	184
A. Préparation de la cassette URA3-ARS-CEN6.....	184
B. Préparation du vecteur pMU3-DEBS 1-TE.....	185
C. Assemblage de type Gibson .....	186

<b>2. Clonage des séquences codant les modules de PKS étudiés par recombinaison <i>in vivo</i></b>	<b>187</b>
.....	187
A. Préparation des fragments d'ADN .....	187
B. Préparation du vecteur pMU3-DEBS 1-TE-URA3 .....	190
C. Transformation des cellules de levure .....	190
D. Isolement des plasmides recombinants .....	191
<b>3. Clonage de la cassette <i>tetA-sacB</i> par <i>PCR-targeting</i>.....</b>	<b>192</b>
A. Préparation de la cassette <i>tetA-sacB</i> .....	192
B. Recombinaison entre le plasmide pBF21 et la cassette <i>tetA-sacB</i> par <i>PCR-targeting</i> ..	193
<b>4. Clonage du module 11 de l'AMPHS par <i>PCR-targeting</i> .....</b>	<b>194</b>
A. Préparation du fragment d'ADN .....	194
B. Recombinaison entre le fragment de onzième module de AMPHS et le pBF21- <i>tetA-sacB</i> par <i>PCR-targeting</i> .....	194
C. Isolement des plasmides recombinants.....	195
<b>5. Clonage des matrices d'édition et des séquences espaceurs dans le plasmide pCRISPomyces-2.....</b>	<b>196</b>
A. Assemblage des matrices d'édition par PCR de chevauchement.....	196
B. Préparation du vecteur .....	197
C. Ligation de la matrice d'édition et du plasmide par assemblage de type Gibson .....	198
D. Identification des séquences espaceurs .....	198
E. Insertion des séquences spaceur par assemblage de type Golden Gate .....	199
<b>6. Clonage des matrices d'édition et des séquences espaceurs dans le plasmide pCRISPR-Cas9 .....</b>	<b>202</b>
A. Préparation des matrices d'édition .....	203
B. Linéarisation du plasmide par <i>StuI</i> .....	203
C. Ligation de la matrice d'édition et du plasmide pCRISPR-Cas9 .....	203
D. Identification et préparation des séquences espaceurs .....	203
E. Linéarisation du plasmide par <i>SnaBI</i> .....	204
F. Ligation de la séquence spaceur et du plasmide .....	204
<b>V. Extraction d'ADN génomique, d'ADN plasmidique et de fragments d'ADN.....</b>	<b>205</b>

## Sommaire

<b>1. Extraction de l'ADN génomique de différentes espèces de <i>Streptomyces</i> .....</b>	<b>205</b>
<b>2. Extraction de l'ADN total de cellules de levure.....</b>	<b>205</b>
<b>3. Extraction d'ADN plasmique par midipréparation et minipréparation .....</b>	<b>205</b>
<b>VI. Préparation et transformation des cellules bactériennes.....</b>	<b>207</b>
<b>1. Préparation et transformation de cellules compétentes d'<i>E. coli</i> .....</b>	<b>207</b>
<b>2. Préparation et transformation des cellules électro-compétentes d'<i>E. coli</i>.....</b>	<b>207</b>
<b>3. Préparation des spores et transfert des plasmides d'<i>E. coli</i> à <i>S. coelicolor</i> par conjugaison .....</b>	<b>208</b>
A. Extraction des spores.....	208
B. Préparation, transformation par conjugaison de <i>S. coelicolor</i> et criblage des exconjugants.....	208
a. Transformation par conjugaison de <i>S. coelicolor</i> .....	208
b. Génotypage des exconjugants par PCR sur colonies .....	209
c. Elimination des plasmides génétiquement modifiés .....	212
C. Transformation par conjugaison et culture de <i>Sac. erythraea</i> JC2.....	212
<b>VII. Production des métabolites spécialisés .....</b>	<b>214</b>
<b>1. Production de coelimycine.....</b>	<b>214</b>
A. Production de coelimycine en milieu solide.....	214
B. Production de coelimycine en milieu liquide .....	214
<b>2. Production, extraction, et analyses des lactones.....</b>	<b>214</b>
A. Production de lactones.....	214
B. Production des standards .....	214
C. Extraction des lactones .....	215
D. Analyse des lactones .....	215
<b>VIII. Détection des sous-unités de la CPKS par immuno-empreinte.....</b>	<b>216</b>
<b>1. Préparation des échantillons .....</b>	<b>216</b>

2. Analyse par Dot Blot .....	216
3. Analyse par Western Blot.....	216
<b>IX. Localisation des sous-unités de la CPKS chez <i>S. coelicolor</i> par microscopie.....</b>	<b>218</b>
1. Préparation des échantillons .....	218
2. Acquisition des images par microscopie.....	218
<b>Références bibliographiques .....</b>	<b>219</b>
<b>Annexes.....</b>	<b>251</b>



# **Liste des abréviations**



6-dEB = 6-désoxyérythronolide B

A = adénylation, de l'anglais *adenylation*

ACP = domaine protéique porteur d'acyl, de l'anglais *acyl carrier protein*

ADN = acide désoxyribonucléique

AMPHS = polycétide synthase responsable de la synthèse de l'amphotéricine

ARN = acide ribonucléique

ARNg = ARN guide

ARS = séquence de réplication autonome, de l'anglais *autonomous replicating sequence*

AT = acyltransférase, de l'anglais *acyltransferase*

ATP = adénosine triphosphate

AVES = polycétide synthase responsable de la synthèse de l'ivermectine

BSA : albumine de sérum bovin, de l'anglais *bovine serum albumin*

$\beta$ -ME =  $\beta$ -mercaptoéthanol

C = condensation, de l'anglais *condensation*

Cas9 = endonucléase d'ADN Cas9, de l'anglais *CRISPR associated protein 9*

CoA = coenzyme-A

CPKS = polycétide synthase responsable de la synthèse de la coelimycine

CRISPR = courtes répétitions palindromiques regroupées et régulièrement espacées, de l'anglais *clustered regularly interspaced short palindromic repeats*

Cryo-CLEM = microscopie lumineuse corrélée à la cryo-microscopie électronique de l'anglais *Cryo-correlative light and electron microscopy*

Cryo-EM = cryo-microscopie électronique, de l'anglais *cryo-electronic microscopy*

Da = dalton

DEBS = polycétide synthase responsable de la synthèse de la 6-désoxyérythronolide B

DD = domaine de docking, de l'anglais *docking domain*

DH = déshydratase, de l'anglais *dehydratase*

DMSO = diméthylsulfoxyde

## Liste des abréviations

DO = densité optique

DTT : dithiothréitol

E = épimérase, de l'anglais *epimerase*

EDTA = acide éthylène diamine tétra-acétique

ER = énoyl réductase, de l'anglais *enoyl reductase*

eV = électronvolt

FAS = acide gras synthase, de l'anglais *fatty acid synthase*

GC-MS = chromatographie en phase gazeuse couplée à de la spectrométrie de masse, de l'anglais *gas chromatography-mass spectrometry*

HEPES = acide 4-(2-hydroxyéthyl)-1- pipérazine éthane sulfonique

HR = hautement réductrice

IPTG = isopropyl  $\beta$ -D-1-thiogalactopyranoside

KR = cétoréductase, de l'anglais *ketoreductase*

KS = cétosynthase, de l'anglais *ketosynthase*

LC-MS = chromatographie liquide couplée à de la spectrométrie de masse, de l'anglais *liquid chromatography-mass spectrometry*

MAT = malonyl-acétyl transférase, de l'anglais *malonyl-acetyl transferase*

MCAT = (ACP) S-malonyltransférase, de l'anglais *malonyl-CoA: acyl carrier protein transacylase*

MES : acide 2-(N-morpholino) éthane sulfonique

MDR = déshydrogénase/réductase à courte chaîne, de l'anglais *medium-chain dehydrogenase/reductase*

mFAS = synthase d'acide gras de mammifère, de l'anglais *mammalian fatty acid synthase*

MT = méthyltransférase, de l'anglais *methyltransferase*

$m/z$  = rapport masse sur charge

NADPH = nicotinamide adénine dinucléotide phosphate

NR = non réductrice

NRP = peptide non ribosomique, de l'anglais *non-ribosomal peptide*

*Liste des abréviations*

NRPS = enzymes de synthèse des peptides non ribosomiques, de l'anglais *non-ribosomal peptides synthase*

NYSS = polycétide synthase responsable de la synthèse de la nystatine

OLMS = polycétide synthase responsable de la synthèse de l'oligomycine

ORF = cadre ouvert de lecture, de l'anglais *open reading frame*

pb = paire de bases

PCP = domaine protéique porteur de peptide, de l'anglais *peptidyl carrier protein*

PCR = réaction de polymérisation en chaîne, de l'anglais *polymerase chain reaction*

PDB = banque de donnée des protéines, de l'anglais *protein data bank*

PEG = polyéthylène glycol

PIKS = polycétide synthase responsable de la synthèse de la pikromycine

PKS = polycétide synthase, de l'anglais *polyketide synthase*

PPTase = phosphopantéthéinyl transférase

PR = partiellement réductrice

RAPS = polycétide synthase responsable de la synthèse de la rapamycine

RIFS = polycétide synthase responsable de la synthèse de la rifamycine

RiPP = peptide synthétisé par le ribosome et modifié post-traductionnellement, de l'anglais *ribosomally synthesized and post-translationally modified peptide*

RMN = résonance magnétique nucléaire

SAM = S-adénosylméthionine

SAXS = diffusion des rayons X aux petits angles, de l'anglais *small angle X-ray scattering*

SDR = déshydrogénases/réductases à courte chaîne, de l'anglais *short-chain dehydrogenase/reductase*

SDS-PAGE = électrophorèse en gel de polyacrylamide contenant du dodécylsulfate de sodium, de l'anglais *sodium dodecyl sulfate polyacrylamide gel electrophoresis*

Sfp = surfactine phosphopantéthéinyl-transférase

Svp = *Streptomyces verticillus* PPTase

TE = thioestérase, de l'anglais *thioesterase*

*Liste des abréviations*

TIC = Chromatogramme des ions totaux, de l'anglais *total ion chromatogram*

T<sub>m</sub> = température de fusion

TYLS = polycétide synthase responsable de la synthèse de la tylosine

WT = souche sauvage, de l'anglais *wild type*

# Liste des figures



<b>Figure 1 : Structure des membres représentatifs des principales classes de métabolites spécialisés .</b>	40
<b>Figure 2 : Structure de plusieurs polycétides utilisés en tant que médicaments.</b>	44
<b>Figure 3 : Exemple de modification d'un polycétide par l'approche d'hemi-synthèse.</b>	ERREUR ! SIGNET NON DEFINI.
<b>Figure 4 : Représentation du groupe de gènes biosynthétiques codant la PKS synthétisant le 6-désoxyérythronolide B (DEBS).</b>	49
<b>Figure 5 : Structure des membres représentatifs des métabolites spécialisés biosynthétisés par les PKS de type III.</b>	50
<b>Figure 6 : Représentation schématique du mode d'action des PKS de type III.</b>	51
<b>Figure 7 : Structure des membres représentatifs des métabolites spécialisés biosynthétisés par les PKS de type II.</b>	52
<b>Figure 8 : Représentation schématique du mode d'action général des PKS de type II.</b>	53
<b>Figure 9 : Structure des polycétides biosynthétisés à partir des différentes classes de PKS de type I itératives fongiques.</b>	54
<b>Figure 10 : Représentation de la PKS DEBS synthétisant le 6-desoxyerythronolide B.</b>	56
<b>Figure 11 : Représentation de la PKS responsable de la synthèse de la bacillaene.</b>	58
<b>Figure 12 : Cibles théoriques de modification de la structure d'un polycétide.</b>	61
<b>Figure 13 : Représentation schématique de l'interaction entre l'érythromycine A et la sous-unité 50s du ribosome de <i>deinococcus radiodurans</i>.</b>	63
<b>Figure 14 : Modèle harmonisé du stéréocontrôle au sein des PKS modulaires.</b>	64
<b>Figure 15 : Structure RMN du domaine cœur des ACP provenant du module 2 de la PKS DEBS.</b>	65
<b>Figure 16 : Réaction de transfert du groupement phosphopantéthéine du coenzyme A sur la sérine du domaine ACP catalysée par la PPTase.</b>	65
<b>Figure 17 : Structure cristallographique du domaine AT provenant du module 5 de la PKS DEBS.</b>	66
<b>Figure 18 : Mécanisme catalytique du domaine AT.</b>	68
<b>Figure 19 : Structure cristallographique du domaine KS provenant du module 5 de la PKS DEBS.</b>	70
<b>Figure 20 : Représentation des domaines LINKS provenant d'un domaine KS représentatif chez une PKS de type <i>trans</i>-AT.</b>	71
<b>Figure 21 : Mécanisme catalytique du domaine KS.</b>	72

## Liste des figures

<b>Figure 22 : Structure cristallographique du domaine KR provenant du module 2 de la PKS responsable de la synthèse de l'amphotéricine.</b> .....	73
<b>Figure 23 : Mécanisme catalytique du domaine KR.</b> .....	74
<b>Figure 24 : Modèle du contrôle de la stéréochimie lors de la réduction catalysée par le domaine KR.</b> .....	76
<b>Figure 25 : Mécanisme catalytique du domaine KR.</b> .....	78
<b>Figure 26 : Structure cristallographique du domaine TE de la PKS DEBS.</b> .....	80
<b>Figure 27 : Mécanisme catalytique du domaine TE.</b> .....	81
<b>Figure 28 : Création du système biosynthétique modèle DEBS 1-TE par ingénierie génétique.</b> .....	82
<b>Figure 29 : Stratégie d'échange de la boucle réductrice.</b> .....	83
<b>Figure 30 : Structure cryo-EM de PikAIII en l'absence de substrat.</b> .....	87
<b>Figure 31 : Structure cristallographique du di-domaine KS-AT provenant du module 5 de la PKS DEBS.</b> .....	88
<b>Figure 32 : Représentation schématique des différentes conformations adoptées par PikAIII lors des différentes étapes du cycle d'extension du polycétide.</b> .....	89
<b>Figure 33 : Les domaines de docking chez les PKS modulaires.</b> .....	94
<b>Figure 34 : Ingénierie du module de chargement dans la PKS DEBS.</b> .....	96
<b>Figure 35 : Approche d'ingénierie génétique pour l'échange de modules entier (adapté de Gokhale <i>et al.</i>, 1999).</b> .....	98
<b>Figure 36 : Approche d'ingénierie génétique pour l'échange de modules entier (adapte de Ranganathan <i>et al.</i>, 1999 ; Tang <i>et al.</i>, 2000).</b> .....	99
<b>Figure 37 : Approche d'ingénierie génétique pour l'échange de domaine AT.</b> .....	102
<b>Figure 38 : Evolution des PKS responsables de la synthèse de l'auréothine (AURS) et de la néoauréothine (NORS).</b> .....	105
<b>Figure 39 : Expérience d'inactivation d'un domaine KR dans la PKS responsable de la synthèse de l'auréothine.</b> .....	106
<b>Figure 40 : Analyses des sites de jonction dans un module de la PKS RAPS obtenus par l'approche d'évolution accélérée.</b> .....	108
<b>Figure 41 : Localisation des sous-unités PksR et PksJ de la PKS responsable de la synthèse de la bacillaene et du domaine <i>trans</i>-AT PksE dans <i>B. subtilis</i> 3610 par microscopie confocale.</b> .....	114
<b>Figure 42 : Assemblage d'un mégacomplexe chez <i>B. subtilis</i>.</b> .....	116

<b>Figure 43 : Représentation schématique de l'assemblage vertical de la PKS modulaire de type <i>trans</i>-AT responsable de la synthèse de la bacillaene. ....</b>	<b>119</b>
<b>Figure 44 : Structures des polycétides dont les polycétide synthases (PKS) ont été utilisées comme sources de modules.....</b>	<b>130</b>
<b>Figure 45 : Stratégie d'ingénierie de la PKS modèle DEBS dans le but de modifier la stéréochimie de la 6-désoxyérythronolide B. ....</b>	<b>135</b>
<b>Figure 46 : Représentation de la PKS responsable de la synthèse de la coelimycine. ...</b>	<b>138</b>
<b>Figure 47 : Production du pigment jaune associé à la coelimycine P2 sur DNA.....</b>	<b>139</b>
<b>Figure 48 : Stratégie pour modifier le génome de <i>S. coelicolor</i> FM145 par la technique CRISPR/CAS9.....</b>	<b>141</b>
<b>Figure 49 : Contrôle de l'expression des protéines de fusion.....</b>	<b>143</b>
<b>Figure 50 : Production de la coelimycine P2 sur le milieu DNA modifié par l'ajout de glutamate.....</b>	<b>144</b>
<b>Figure 51 : Localisation de la sous-unité CpkA dans <i>S. coelicolor</i> FM145.....</b>	<b>146</b>
<b>Figure 52 : Localisation de la sous-unité CpkC dans <i>S. coelicolor</i> FM145.....</b>	<b>147</b>
<b>Figure 53 : Colocalisation des sous-unités CpkA et CpkC dans <i>S. coelicolor</i> FM145. ..</b>	<b>149</b>
<b>Figure 54 : Localisation du transporteur CpkF dans <i>s. coelicolor</i> FM145.....</b>	<b>151</b>
<b>Figure 55 : Colocalisation du transporteur CpkF et de la sous-unité CpkC dans <i>S. coelicolor</i> FM145. ....</b>	<b>152</b>
<b>Figure 56 : Carte du plasmide pMU3-DEBS 1-TE. ....</b>	<b>171</b>
<b>Figure 57 : Carte du plasmide pFL38. ....</b>	<b>172</b>
<b>Figure 58 : Carte des plasmides codant la DEBS et les enzymes post-PKS associées....</b>	<b>173</b>
<b>Figure 59 : Carte du plasmide pCRISPomyces-2.....</b>	<b>174</b>
<b>Figure 60 : Carte du plasmide pCRISPR-CAS9. ....</b>	<b>175</b>
<b>Figure 61 : Carte du plasmide pSS172. ....</b>	<b>176</b>
<b>Figure 62 : Carte du plasmide pIJ8630.....</b>	<b>177</b>
<b>Figure 63 : Principe de la méthode Gibson.....</b>	<b>186</b>
<b>Figure 64 : Principe de la méthode Golden Gate. ....</b>	<b>200</b>
<b>Figure 65 : Stratégie pour le criblage de l'évènement de modification génétique souhaité. ....</b>	<b>210</b>
<b>Figure 66 : Représentation schématique du protocole de conjugaison entre des cellules d'<i>E. coli</i> ET12567 (pUZ8002) et de <i>Sac. erythraea</i> JC2, ainsi que l'évènement de modification génétique.....</b>	<b>213</b>
<b>Figure 67 : Exemple de préparation d'un échantillon pour l'analyse par microscopie.</b>	<b>218</b>



# Liste des tableaux



<b>Tableau 1 : Échange des domaines KR au sein de la PKS modèle DEBS 1-TE et analyse des produits obtenus.....</b>	<b>124</b>
<b>Tableau 2 : Dérivés de la souche FM145 de <i>S. coelicolor</i> obtenus ou en cours pour l'étude de la localisation d'un mégacomplexe. ....</b>	<b>142</b>
<b>Tableau 3 : Séquence nucléotidique des désoxyribonucléotides, composition du mélange réactionnel et conditions expérimentales utilisées pour la réaction de PCR de la cassette URA3-ARS-CEN6. ....</b>	<b>185</b>
<b>Tableau 4 : Séquence nucléotidique des oligodésoxyribonucléotides, composition du mélange réactionnel et conditions expérimentales utilisées pour l'amplification des fragments d'ADN codant les modules de PKS. ....</b>	<b>187</b>
<b>Tableau 5 : Séquence ADN des oligodésoxyribonucléotides, composition du mélange réactionnel et conditions expérimentales utilisées pour l'amplification de la cassette <i>tetA-sacB</i>. ....</b>	<b>192</b>
<b>Tableau 6 : Séquence ADN des oligodésoxyribonucléotides pour la réaction de PCR du fragment ADN du onzième module de la AMPHS. ....</b>	<b>194</b>
<b>Tableau 7 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'assemblage des matrices d'édition et pour le séquençage du pCRSPomyces-2. ....</b>	<b>196</b>
<b>Tableau 8 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'amplification des séquences spaceurs. ....</b>	<b>198</b>
<b>Tableau 9 : Mélange réactionnel et programme utilisé pour l'assemblage de la séquence spaceur avec le plasmide pCRISPomyces-2 contenant la matrice d'edition. ....</b>	<b>201</b>
<b>Tableau 10 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'assemblage des matrices d'édition et le séquençage du pCRSPR-CAS9. ....</b>	<b>202</b>
<b>Tableau 11 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'amplification des séquences spaceurs. ....</b>	<b>204</b>
<b>Tableau 12 : Composition du mélange réactionnel et conditions expérimentales utilisées pour les réactions de PCR sur colonies dans le cadre du criblage des évènements de modification dans l'ADN génomique des <i>Streptomyces</i>. ....</b>	<b>211</b>



# État de l'art



# I. Les métabolites spécialisés

## 1. Définition

La biosphère est un véritable réservoir à biomolécules produites par toutes les espèces du vivant. Parmi cette abondance de biomolécules, les produits naturels représentent une large famille de molécules chimiquement différentes et possédant une grande variété d'activités biologiques tels que les antibiotiques, anticancéreux ou encore antiparasitaires. Chez les bactéries (Nappi *et al.*, 2019), les champignons (Bills and Gloer, 2016), les plantes (Zaynab *et al.*, 2018) et les animaux (Torres and Schmidt, 2019), les produits naturels sont issus du métabolisme spécialisé : des molécules qui ne sont pas requises pour la survie de l'organisme au moins dans des conditions de laboratoire, mais qui lui confèrent vraisemblablement un avantage évolutif dans son environnement. Les macromolécules (ADN, ARN et protéines), leurs précurseurs et leurs blocs de construction, ainsi que les intermédiaires du métabolisme primaire, sont donc exclus de la définition de produit naturel (Katz and Baltz, 2016).

Dans la littérature scientifique, la notion de métabolisme spécialisé est souvent assimilée à celle de métabolisme secondaire. Cependant, ces deux notions ne sont pas synonymes. En effet, la notion de métabolisme secondaire, la plus ancienne, est issue d'interprétations de physiologie microbienne restreintes aux conditions de laboratoire. Les métabolites secondaires étaient considérés comme étant produits durant la phase tardive de croissance, les différenciant ainsi des molécules du métabolisme primaire indispensables à la nutrition, à la croissance et au développement d'un organisme. Si les fonctions des métabolites secondaires ont été identifiées dans des conditions de laboratoire, le plus souvent dans le but de découvrir une application pour l'Homme, c'est bien leur utilisation en pharmacologie, en cosmétologie, ou en agronomie qui a été à l'origine de l'intérêt de l'industrie et de la recherche pour leur découverte. Cependant, ces fonctions ne sont pas représentatives des phénomènes ayant lieu dans la Nature (Davies, 2013).

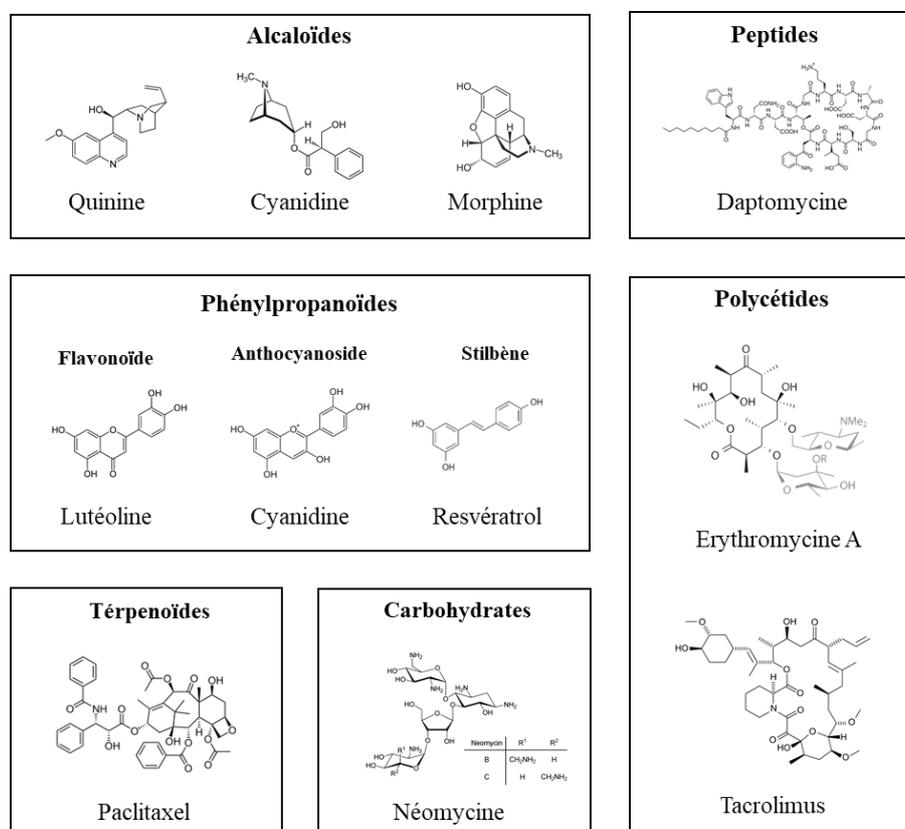
La réelle fonction des métabolites spécialisés est aujourd'hui sujette à débat. Il a été proposé que la concentration de ces molécules dans leurs environnements soit insuffisante pour avoir un effet d'inhibition sur la croissance de microorganismes compétiteurs et sur leur mort (Barka *et al.*, 2016). Cependant, il a été montré que la pression de sélection agit même à des concentrations inférieures à la concentration minimale inhibitrice, ce qui confère tout de même un avantage compétitif pour la bactérie biosynthétisant l'antibiotique (Andersson and

Hughes, 2014 ; Gullberg *et al.*, 2014 ; Westhoff *et al.*, 2017). Néanmoins, le développement de notre compréhension des processus génétiques et biochimiques impliqués dans la production des métabolites secondaires a permis de donner à nouveaux les lettres de noblesse à la fonction de ces biomolécules, comme molécules de signalisations ou de modulations de fonctions physiologiques, et dans les processus de signalisation inter-cellulaire chez les microorganismes (Paul *et al.*, 1981 ; Davies *et al.*, 2006 ; Linares *et al.*, 2006 ; Romero *et al.*, 2011).

Si la notion de métabolites secondaires relève d'un anthropocentrisme en ce qui concerne la fonction de ce type de produit naturel (une activité antibactérienne confère une application comme antibiotique en pharmacologie), il est plus approprié d'employer la notion de métabolites spécialisés pour mettre l'accent sur leur importance en écologie microbienne (Davies, 2013).

## 2. Les classes de métabolites spécialisés

Les métabolites spécialisés sont classifiés selon leur origine biosynthétique, et quelques exemples représentatifs de chaque famille sont représentés dans la **Figure 1**.



**Figure 1 : Structure des membres représentatifs des principales classes de métabolites spécialisés.** Les carbohydrates inclus dans les métabolites spécialisés et condensés à un cœur aglycone sont représentés en gris dans la structure de l'érythromycine A.

Les alcaloïdes sont des composés organiques azotés, basiques, et qui se structurent le plus souvent en hétérocycle (Ziegler and Facchini, 2008). Cette classe est très majoritairement retrouvée chez les plantes. Les premiers découverts comme la quinine (antipyrétique, analgésique et antipaludique) dans certains quinquinas (Mohammadi *et al.*, 2020), l'atropine (anticholinergique) dans certaines *Solanaceae* vireuses (Tran *et al.*, 2018), ou la morphine (antalgique) dans l'opium (Devereaux *et al.*, 2018), sont parmi les premiers métabolites spécialisés révélés pour leur application en pharmacologie.

Chez les plantes, les phénylpropanoïdes sont générés à l'intersection de la synthèse des polycétides et du catabolisme des acides aminés aromatiques (Vogt, 2010). Cette famille de métabolites spécialisés regroupe les flavonoïdes (Imran *et al.*, 2019), les anthocyanines (Wallace and Giusti, 2015), ou encore les stilbènes (Kulkarni and Cantó, 2015), qui peuvent présenter des propriétés bénéfiques pour la santé humaine.

Les terpénoïdes sont une classe de composés organiques essentiellement chez les plantes et les champignons. Ils sont issus de précurseurs contenant cinq carbones (l'isopentényl-pyrophosphate et le diméthylallyl-pyrophosphate) et sont classés selon leur taille (Tholl, 2015). Un exemple notable est le paclitaxel qui est utilisé en chimiothérapie anticancéreuse, et vendu sous le nom « Taxol » (Rowinsky and Donehower, 1995).

Les carbohydrates sont des composés organiques largement représentés dans le métabolisme primaire, et pourtant certains d'entre eux peuvent être classés dans le métabolisme spécialisé. En effet, les saccharides retrouvés dans le métabolisme secondaire inclus ceux pouvant posséder une activité antibiotique (McCranie and Bachmann, 2014), comme la néomycine qui interfère avec la traduction des ARN messagers chez les bactéries (Waksman and Lechevalier, 1949). En outre, certains saccharides peuvent se condenser avec le cœur aglycone d'autres métabolites spécialisés pour former un hétéroside. Ces glycosylations affectent largement leurs propriétés biochimiques, mais également leurs effets biologiques (Schlünzen *et al.*, 2001).

Les peptides sont des polymères d'acides aminés reliés entre eux par des liaisons peptidiques. Lorsque ces molécules désignent une classe de métabolites spécialisés particulièrement présente chez les bactéries, les peptides peuvent être d'origine ribosomique ou non ribosomique. Les RiPP (de l'anglais *ribosomally synthesized and post-translationally modified peptide*), aussi appelés produits naturels ribosomaux, sont des peptides qui sont biosynthétisés par les ribosomes et qui peuvent subir des modifications post-traductionnelles (Arnison *et al.*, 2013 ; Li and Rebuffat, 2019). Les peptides peuvent aussi être d'origine non-

ribosomique, alors appelé NRP (de l'anglais *non-ribosomal peptide*), et sont biosynthétisés par des complexes multi-enzymatiques spécifiques, les NRPS (de l'anglais *non-ribosomal peptides synthase*) (Süssmuth and Mainz, 2017). La plupart des NRP ont des activités biologiques. Par exemple, le NRP daptomycine est un lipopeptide efficace contre les bactéries à Gram positif (Taylor and Palmer, 2016).

Les polycétides sont issus de la condensation itérative d'unités d'acétyles ou de malonyls par un complexe multi-enzymatique spécifique, la PKS (de l'anglais *polyketide synthase*) (Robbins *et al.*, 2016 ; Shimizu *et al.*, 2017 ; Chen *et al.*, 2018 ; Herbst *et al.*, 2018). A l'instar des NRP, les polycétides possèdent une large gamme de bioactivités, comme antibiotique avec les macrolides, dont l'érythromycine, (Arsic *et al.*, 2018), ou encore comme immunosuppresseur avec le tacrolimus (Shrestha, 2017).

### 3. Intérêts pour les métabolites spécialisés

À la fin du XIX<sup>ème</sup> siècle, un médecin français, Ernest Duchesne, publie un premier article précurseur dans sa thèse (Contribution à l'étude de la concurrence vitale chez les micro-organismes : Antagonisme entre les moisissures et les microbes, thèse présentée à la faculté de médecine et de pharmacie de Lyon soutenue publiquement le 17 décembre 1897) sur la thérapie au moyen des microorganismes (*Penicillium glaucum* contre *Escherichia coli*) (Shama, 2016). Postérieurement, par sérendipité, le chercheur écossais Alexander Fleming observa qu'un contaminant fongique affectait la croissance des bactéries voisines. Le champignon s'est avéré être *Penicillium notatum*, et la molécule antibactérienne qu'il a produite a été nommée pénicilline (Fleming, 1929).

Les chercheurs Howard Walter Florey et Sir Ernst Boris Chain ont ensuite réalisé au début des années 1940, l'exploit de purifier et de caractériser la pénicilline (Lobanovska and Pilla, 2017). Cette avancée scientifique récompensa Fleming, Florey et Chain du prix Nobel de médecine en 1945. L'industrialisation à grande échelle de cette molécule a permis de mettre sur le marché des traitements efficaces contre des maladies d'origine bactérienne comme les *Streptococcus* (par exemple, *Streptococcus pneumoniae* qui est à l'origine des pneumonies et de certaines méningites). Par la suite, les chercheurs Selman Waksman (qui a introduit la notion d'antibiotique pour la première fois en 1941) ainsi que ses étudiants Elisabeth Bugie et Albert Schatz, découvrirent la streptomycine (Schatz *et al.*, 1944), le premier antibiotique efficace contre la tuberculose, permettant à Selman Waksman d'obtenir le prix Nobel de médecine en 1952 (Davies, 2013). Lors de cette même année, le premier macrolide

découvert, l'érythromycine, purifié à partir de la bactérie *Saccharopolyspora erythraea* par la société Eli Lilly, a été commercialisée sous le nom de Ilosone (dérivé de Iloilo, le nom de la région des Philippines d'où provenait l'échantillon de sol contenant la bactérie) (Wohlleben *et al.*, 2016).

La découverte de ces molécules thérapeutiques est à l'origine de l'essor de la découverte des produits naturels et, couplée à l'identification et la caractérisation des pathogènes microbiens majeurs de l'époque, a conduit à une période extraordinaire : l'ère des antibiotiques. Entre les années 1940 et aujourd'hui, plus de 23 000 produits naturels, issus largement des Actinobactéries, ont été isolés et testés pour leur activité antibactérienne, antifongique et antiparasitaire sur les hommes et les animaux (Bérdy, 2012). Parmi les produits naturels mis sur le marché, les NRP et les polycétides sont les métabolites spécialisés les plus représentés du fait de leur vaste gamme de bioactivités permettant de traiter de nombreuses pathologies (Butler *et al.*, 2014, 2017). En effet, depuis plus de 30 ans, la moitié des médicaments approuvée par la US Food and Drug Administration, incluant 74 % des anticancéreux et 73 % des antibiotiques, est issue des produits naturels ou leurs dérivés synthétiques (Giddings and Newman, 2013 ; Cragg *et al.*, 2014 ; Demain, 2014 ; Newman and Cragg, 2016).

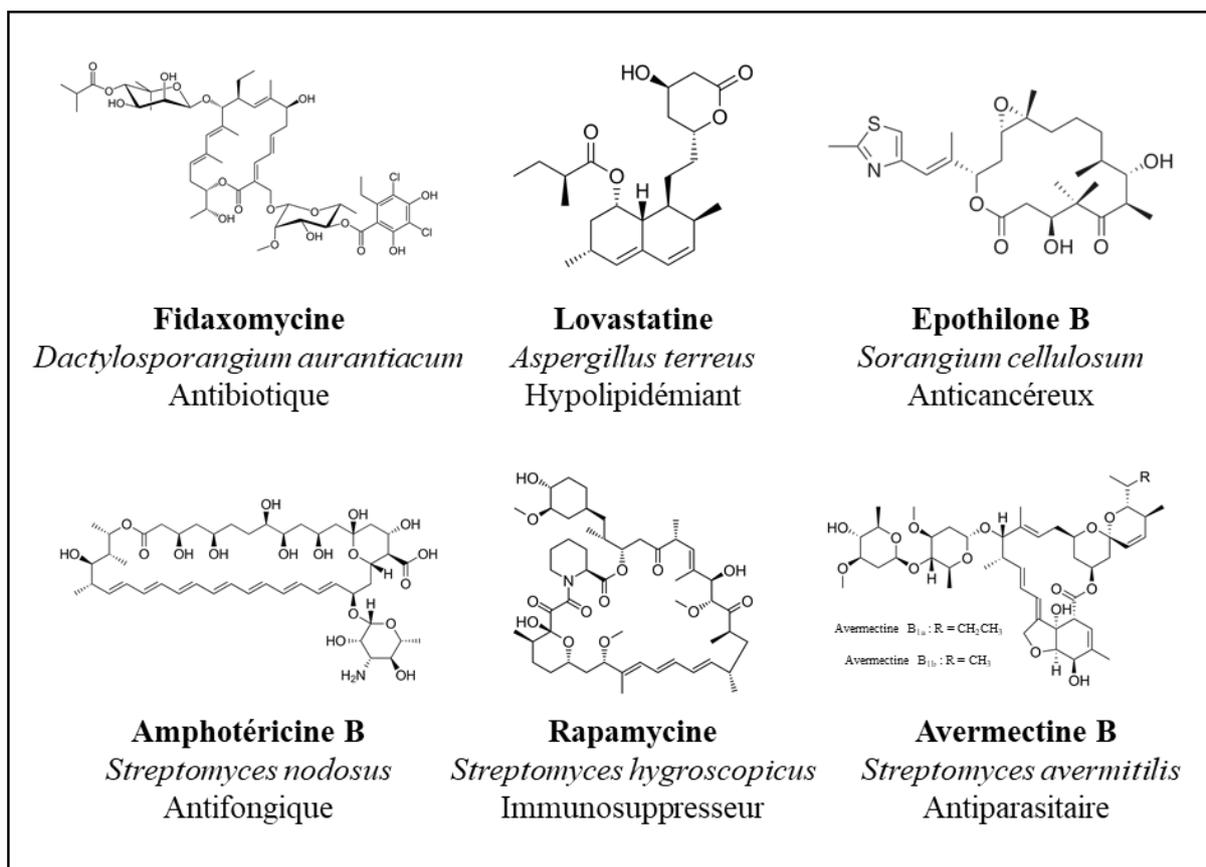
Au vu et au su de la grande diversité de bioactivité des produits naturels, la découverte de nouvelles structures permet potentiellement de découvrir de nouvelles cibles thérapeutiques. Si les métabolites spécialisés confèrent un avantage évolutif certain pour l'organisme qui les produits, ces molécules n'ont pas évolué pour être utilisées par l'Homme de manière optimale. C'est pourquoi il est possible d'optimiser ces structures pour pouvoir améliorer les propriétés pharmacocinétique et pharmacodynamique du composé, de réduire des effets secondaires, ou encore de contourner un mécanisme de résistance (Fischbach and Walsh, 2009 ; Olesen *et al.*, 2018). D'un point de vue plus fondamental, étudier de telles molécules permet de comprendre quels sont les processus génétiques et biochimiques employés par la Nature pour générer des biomolécules d'une telle complexité (Weissman, 2016).

#### **4. Découvrir de nouveaux polycétides**

L'utilité des NRP et des polycétides en thérapie suscite un grand intérêt pour découvrir de nouvelles molécules appartenant à ces familles de métabolites spécialisés et, parmi elles, les polycétides constituent une classe de molécules privilégiées. En effet, les polycétides présentent une grande richesse de groupements et sont majoritairement cycliques, ce qui conduit alors à une rigidification de la structure et favorise donc des interactions spécifiques avec leurs cibles.

## État de l'art

Cette grande spécificité permet d'utiliser les polycétides comme antibiotiques (érythromycine A, fidaxomicine), hypolipémiants (lovastatine), anticancéreux (épothilone), antifongiques (amphotéricine), immunosuppresseurs (rapamycine), ou antiparasitaires (ivermectine). Des exemples de structures de ces différentes biomolécules sont montrées dans la **Figure 2**.



**Figure 2 : Structure de plusieurs polycétides utilisés en tant que médicaments.**

Il existe deux approches efficaces pour découvrir de nouvelles structures de polycétides. Les produits naturels peuvent être directement recherchés dans la biosphère et peuvent être directement utilisés sous leur forme naturelle. Cette première approche est souvent complétée par des méthodes de synthèse chimique, comme l'hémi-synthèse qui est un procédé de synthèse chimique d'une molécule réalisée à partir de composés naturels qui possèdent déjà une partie de la molécule visée, qui ont été obtenues par fermentation. Une troisième approche est étudiée depuis près de trente ans, complémentaire aux deux approches déjà existantes, et est encore à l'étude pour pouvoir en exploiter son plein potentiel : celle de l'ingénierie des enzymes qui biosynthétisent les polycétides.

## A. La recherche dans la biosphère

Au cours de ces huit dernières décennies, la découverte de produits naturels a grandement évolué. Les stratégies étaient relativement simples de 1940 à 1970 reposant sur l'acquisition des microorganismes par l'échantillonnage de l'environnement, la fermentation des microorganismes et l'isolement des métabolites secondaires. Puis, l'accélération des progrès de la science et de la technologie entre 1970 et 2000 ont diversifié et complexifié la recherche de produits naturels. Cependant les efforts globaux de découverte des produits naturels ont ralenti avec le déclin de l'intérêt des industries pharmaceutiques pour les produits naturels au cours des vingt dernières années (Davies, 2013).

L'histoire des antibiotiques est donc intimement liée à la découverte de nouveaux microorganismes, et ces sont les Actinobactéries qui produit la plupart des antibiotiques utilisés en thérapie. Les Actinobactéries sont des bactéries à Gram positif qui comprennent de nombreux genres connus pour leur production de métabolites secondaires, comme les genres *Actinomyces*, *Actinoplanes*, *Amycolatopsis*, *Kitasatospora*, *Micromonospora*, *Nocardia*, *Nocardiopsis*, *Pseudonocardia*, *Saccharopolyspora* ou *Streptomyces* (van der Meij *et al.*, 2017). Parmi eux, ce sont les *Streptomyces* qui constituent le genre bactérien à la source d'environ deux tiers de tous les antibiotiques naturels connus (Lucas *et al.*, 2013). Depuis plus d'un demi-siècle, l'espèce *Streptomyces coelicolor* a été le modèle majeur pour l'étude de la production des antibiotiques et le contrôle de leur synthèse. Parmi les antibiotiques produits par *Streptomyces coelicolor*, les polycétides sont largement représentés, incluant l'actinorhodin (Rudd and Hopwood, 1979), l'undécylprodigiosine (Feitelson *et al.*, 1985), le Cda (de l'anglais *calcium-dependent antibiotic*) (Hopwood and Wright, 1983) et la méthylénomycine (Wright and Hopwood, 1976). Les gènes codant les protéines biosynthétisant ces polycétides sont généralement regroupés ensemble dans un même groupe de gènes biosynthétiques (BGC, de l'anglais *biosynthetic gene cluster*). Le BGC regroupe aussi le ou les gène(s) de résistance et les régulateurs transcriptionnels qui contrôlent la biosynthèse des gènes codant les protéines contenues dans le BGC.

À partir du XXI<sup>ème</sup> siècles, le séquençage des génomes microbiens a ouvert des stratégies totalement nouvelles pour la découverte de médicaments à partir des métabolites spécialisés. Le séquençage de la bactérie *Streptomyces coelicolor* a été un événement historique qui a permis de révéler une source inattendue de produits naturels encore inconnus (Bentley *et al.*, 2002 ; Challis and Hopwood, 2003 ; van Keulen and Dyson, 2014). Parmi les 20 nouveaux

BGC découverts, l'un d'eux s'est révélé être associé à la synthèse d'un polycétide (Pawlik *et al.*, 2007) qui ne peut être produit que dans certaines conditions de laboratoire (Gottelt *et al.*, 2010 ; Pawlik *et al.*, 2010 ; Bednarz *et al.*, 2019), appelé coelimycin P1 (Gomez-Escribano *et al.*, 2012).

L'exploitation des banques de données disponibles de génomes des microorganismes a permis de révéler le potentiel des microorganismes pour la production de nouveaux produits naturels (Cimermanic *et al.*, 2014). En effet, c'est grâce à la diminution du prix du séquençage à haut débit ces dernières années que les banques de données en génomes microbiens ont pu s'enrichir de manière exponentielle (Kalkreuter *et al.*, 2020). Les études bio-informatiques ont montré qu'un génome bactérien contient en moyenne une trentaine de BGC associés à la synthèse de produits naturels, et que la majorité génère des produits naturels encore inconnus (Mukherjee *et al.*, 2017). Le séquençage de bactéries des Actinobactéries a démontré que de nombreux BGC étaient présents dans leurs génomes, et certaines espèces peuvent contenir jusqu'à 50 BGC différents, représentant jusqu'à 10 % de leur génome (Udwary *et al.*, 2007).

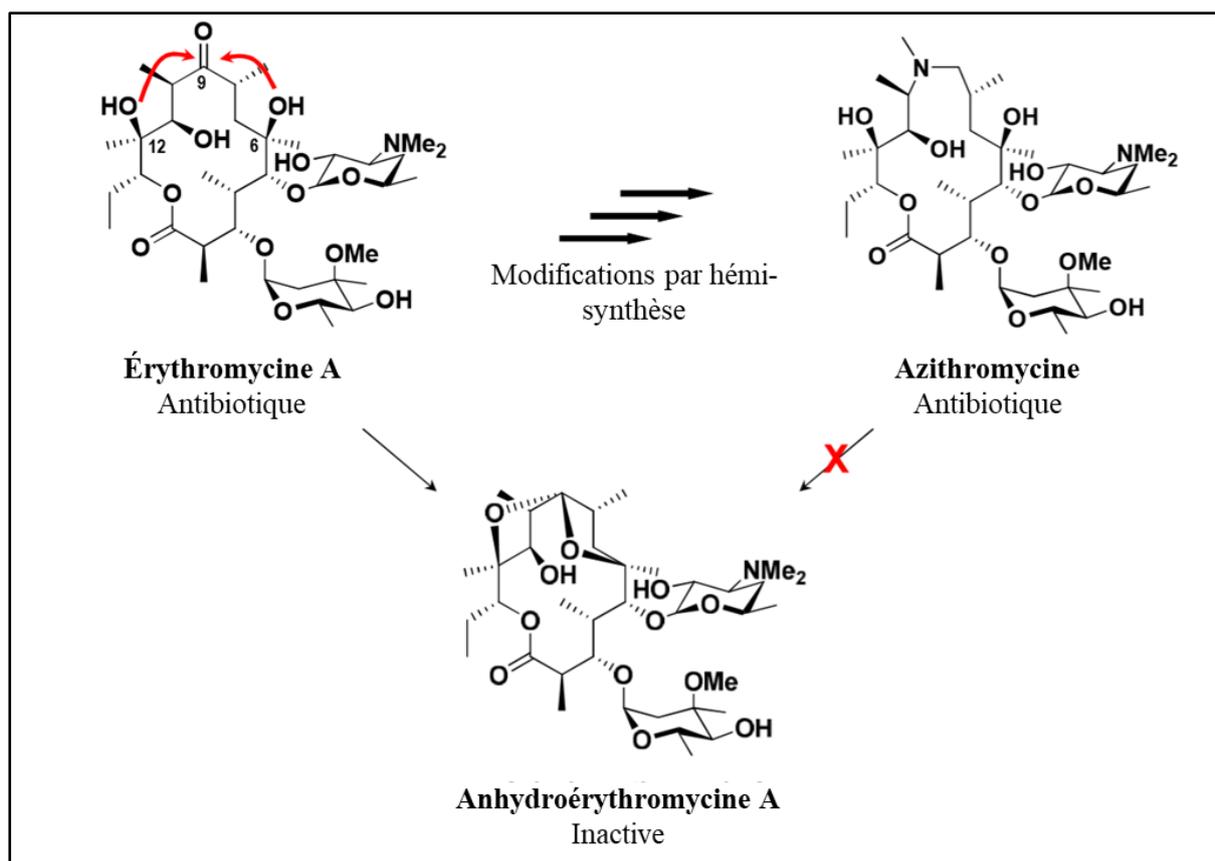
Dans ce contexte, les Actinobactéries sont des bactéries privilégiées pour la synthèse des produits naturels. Par exemple, la bactérie *Streptomyces avermitilis* contient 40 BGC, dont 23 sont silencieux dans les conditions normales de laboratoire (Ikeda *et al.*, 2014). Si cette dernière bactérie fait partie des plus étudiées des Actinobactéries, il est facile d'imaginer le plein potentiel des milliers de génomes disponibles dans les bases de données. En effet, il est estimé que moins de 10 % des BGC associés aux métabolites spécialisés sont exprimés en à un niveau suffisant pour être détectés et caractérisés dans des conditions normales de laboratoire, et que les autres BGC nécessitent d'être activés.

L'approche de l'exploitation du génome, aussi appelée par néologisme *genome mining*, implique l'identification des BGC jusque-là non caractérisés dans les génomes d'organismes séquencés, l'analyse de séquence des protéines codées par ces BGC et l'identification expérimentale de leurs produits (Helfrich *et al.*, 2014). En effet, il est aujourd'hui possible de cribler les génomes bactériens pour identifier des BGCs grâce à des outils bioinformatiques comme antiSMASH (Medema *et al.*, 2011), CLUSEAN (Weber *et al.*, 2009) ou PRISM (Röttig *et al.*, 2011), et de prédire la structure du polycétide. Puis, les BGC peuvent être activés en suivant plusieurs stratégies qui peuvent être combinées comme l'optimisation des conditions de culture (Gottelt *et al.*, 2010 ; Pawlik *et al.*, 2010), l'utilisation des molécules de signalisation (Bednarz *et al.*, 2019), l'exploitation d'un système hétérologue (Gomez-Escribano and Bibb, 2011), l'ingénierie du ribosome (Ochi, 2007 ; Gomez-Escribano *et al.*, 2012), ou encore la manipulation de la régulation transcriptionnelle du BGC (Laureti *et*

*al.*, 2011 ; Wang *et al.*, 2019). Enfin, le polycétide produit peut être détecté, identifié et caractérisé au moyen de différentes méthodes analytiques, comme la spectrométrie de masse complétée avec la résonance magnétique nucléaire (Helfrich *et al.*, 2014). L'un des premiers produits naturels découvert par l'approche de *genome mining* est le polycétide ECO-02301, qui possède une activité antifongique (McAlpine *et al.*, 2005). L'ensemble des approches précédentes de *genome mining* nécessite la mise en culture de la bactérie. Cependant, moins de 2 % des bactéries sont cultivables sous des conditions normales de laboratoire (Kalkreuter *et al.*, 2020), ce qui signifie que la quantité de BGC encore inconnue est bien plus importante que le montrent les études qui se basent sur les banques de données disponibles. Une approche alternative de *genome mining* tient son l'avantage de l'organisation des gènes produisant les polycétides en BGC, et est qualifié de métagénomique (Handelsman *et al.*, 1998). Il est alors possible d'extraire l'ADN des bactéries directement à partir d'échantillons environnementaux, de l'identifier grâce aux outils de bio-informatique et de le caractériser (Seow *et al.*, 1997 ; Wawrik *et al.*, 2005, 2007). Le BGC ainsi capturé peut être ensuite exprimé de manière hétérologue dans le but de produire de potentiels nouveaux polycétides (Feng *et al.*, 2011 ; Helfrich *et al.*, 2014).

## B. La synthèse totale ou l'hémi-synthèse

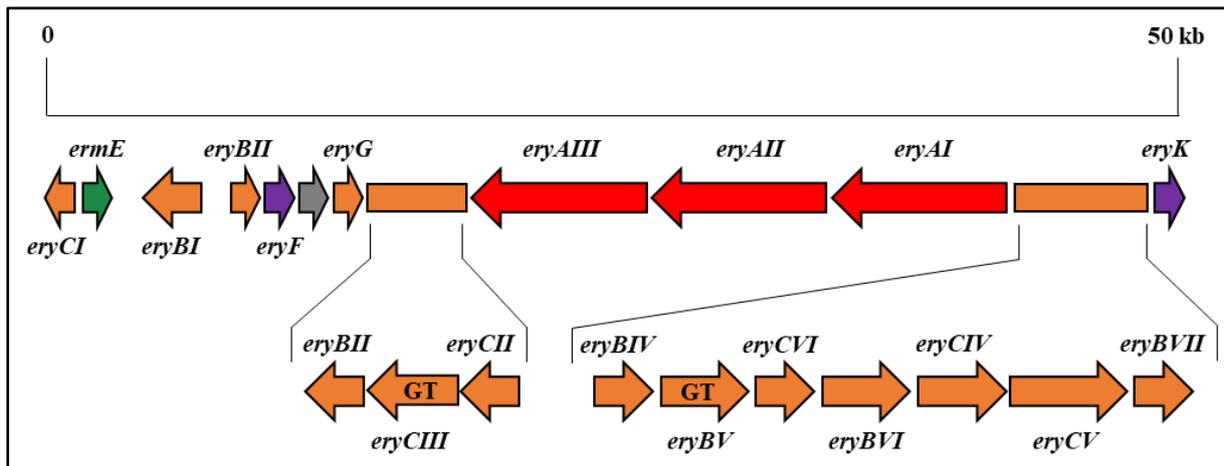
La deuxième stratégie repose sur la modification de structures déjà existantes par l'hémi-synthèse chimique ou l'obtention complète de la structure par synthèse chimique. Par exemple, l'azithromycine et la clarithromycine sont des composés hémi-synthétiques dérivés de l'érythromycine A qui ont des propriétés pharmacodynamiques et pharmacocinétiques améliorées, qui sont efficaces contre des bactéries résistantes à l'érythromycine A, et qui ont aussi permis de réduire des effets secondaires (Peltier, 1992). En effet, l'érythromycine peut s'inactiver par l'attaque nucléophile des groupements hydroxyles portés par les carbones C6 et C12 au niveau du centre électrophile de la fonction cétone porté par le carbone C9, générant alors l'anhydroérythromycine A (**Figure 3**). Pour contourner ce problème de stabilité et augmenter significativement la durée de vie du médicament, la fonction cétone portée par le carbone C9 de l'érythromycine a été modifiée en fonction azane par l'addition d'un atome d'azote dans le cycle lactone de l'érythromycine A (**Figure 3**), générant alors l'azithromycine qui est l'antibiotique le plus vendu depuis trente ans et qui est toujours prescrit lors d'infection bactérienne (Jin *et al.*, 2019).



**Figure 3 : Exemple de modification d'un polycétide par l'approche d'hémi-synthèse (D'après Peltier, 1992).** La fonction cétone au niveau du carbone C9 de l'érythromycine a été modifiée en fonction azane pour générer l'azithromycine. Cette modification permet de ne pas produire de forme inactive de l'érythromycine A par l'attaque nucléophile (indiquée par les flèches rouges) des groupements hydroxyles au niveau des carbones C12 et C6 et au niveau de la fonction cétone portée par le carbone C9.

### C. Vers une troisième approche

En complément de ces stratégies, une troisième approche a été suggérée pour la première fois dans les années 1990, avec la découverte du BGC à l'origine de la biosynthèse de l'érythromycine A (**Figure 4**). En effet, l'analyse des gènes associés à la synthèse de l'érythromycine A chez *Saccharopolyspora erythraea* a révélé que ce polycétide était biosynthétisé par un complexe multi-enzymatique spécifique appelé PKS (Cortes *et al.*, 1990 ; Donadio *et al.*, 1991), où chaque bloc de construction du polycétide est assemblé au moyen de plusieurs activités enzymatiques. Alors, l'ingénierie de ces PKS pourrait permettre de produire des analogues de structures prédictibles avec un rendement comparable au produit initial. Cette approche est appelée biologie synthétique des PKS (Weissman, 2016), et nécessite la compréhension en détail du fonctionnement et de l'organisation de ces systèmes multi-enzymatiques hautement complexes.



**Figure 4 : Représentation du groupe de gènes biosynthétiques codant la PKS synthétisant le 6-désoxyérythronolide B (DEBS).** La PKS est composée de trois sous-unités multi-enzymatiques codées au sein du génome par les trois gènes *eryAI*, *eryAII* et *eryAIII* (en rouge). Ces gènes sont présents au sein du groupe comprenant le gène de résistance à l'érythromycine *ermE* (en vert), et les gènes codant les enzymes post-PKS du processus de maturation du 6-désoxyérythronolide B (en orange et en violet).

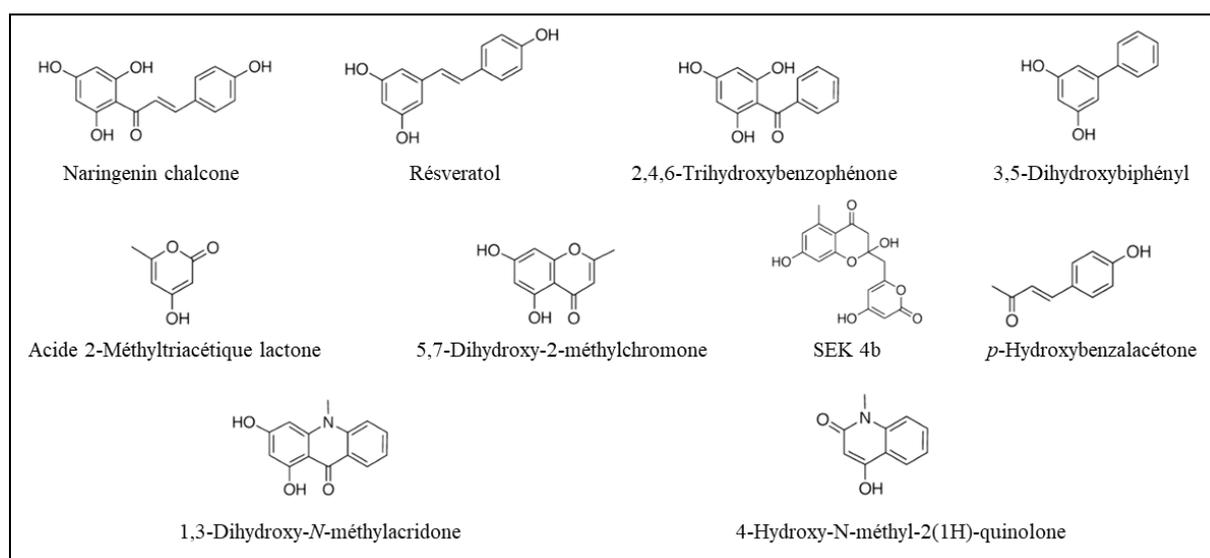
Légende : GT = glycosyltransférase.

## II. Les polycétide synthases

Les polycétide synthases (PKS, de l'anglais *polyketide synthase*) sont des enzymes qui réalisent les différentes étapes d'assemblage des polycétides *via* la condensation de simples blocs de construction que sont les unités acyl-CoA. Les PKS sont classées en trois types distincts : les types I et II qui présentent des caractéristiques similaires avec les enzymes impliquées lors de la biosynthèse des acides gras, et le type III. Cette classification prend aussi en compte la nature des produits obtenus, ainsi les PKS de type II et III synthétisent des polycétides aromatiques, tandis que les PKS de type I permettent l'assemblage de polycétides au moins partiellement réduits. De surcroît le mode opératoire, itératif ou modulaire, permet aussi de classer les PKS.

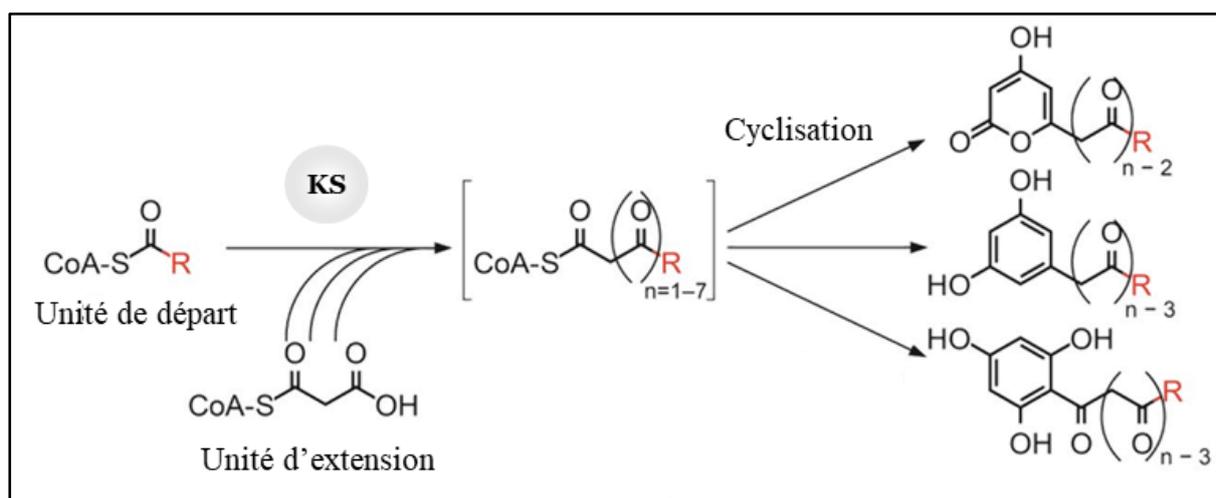
### 1. Les PKS de type III

Les enzymes de la famille des PKS de type III, aussi appelées chalcones ou stilbènes synthases, sont principalement retrouvées chez les plantes et les bactéries où elles interviennent lors de la première étape de la biosynthèse des flavonoïdes (Shimizu *et al.*, 2017). Ce type de PKS génère une grande diversité de produits (**Figure 5**), due au nombre variable d'étapes d'extension (2 à 8), et au mécanisme de cyclisation (Morita *et al.*, 2019).



**Figure 5 : Structure des membres représentatifs des métabolites spécialisés biosynthétisés par les PKS de type III.**

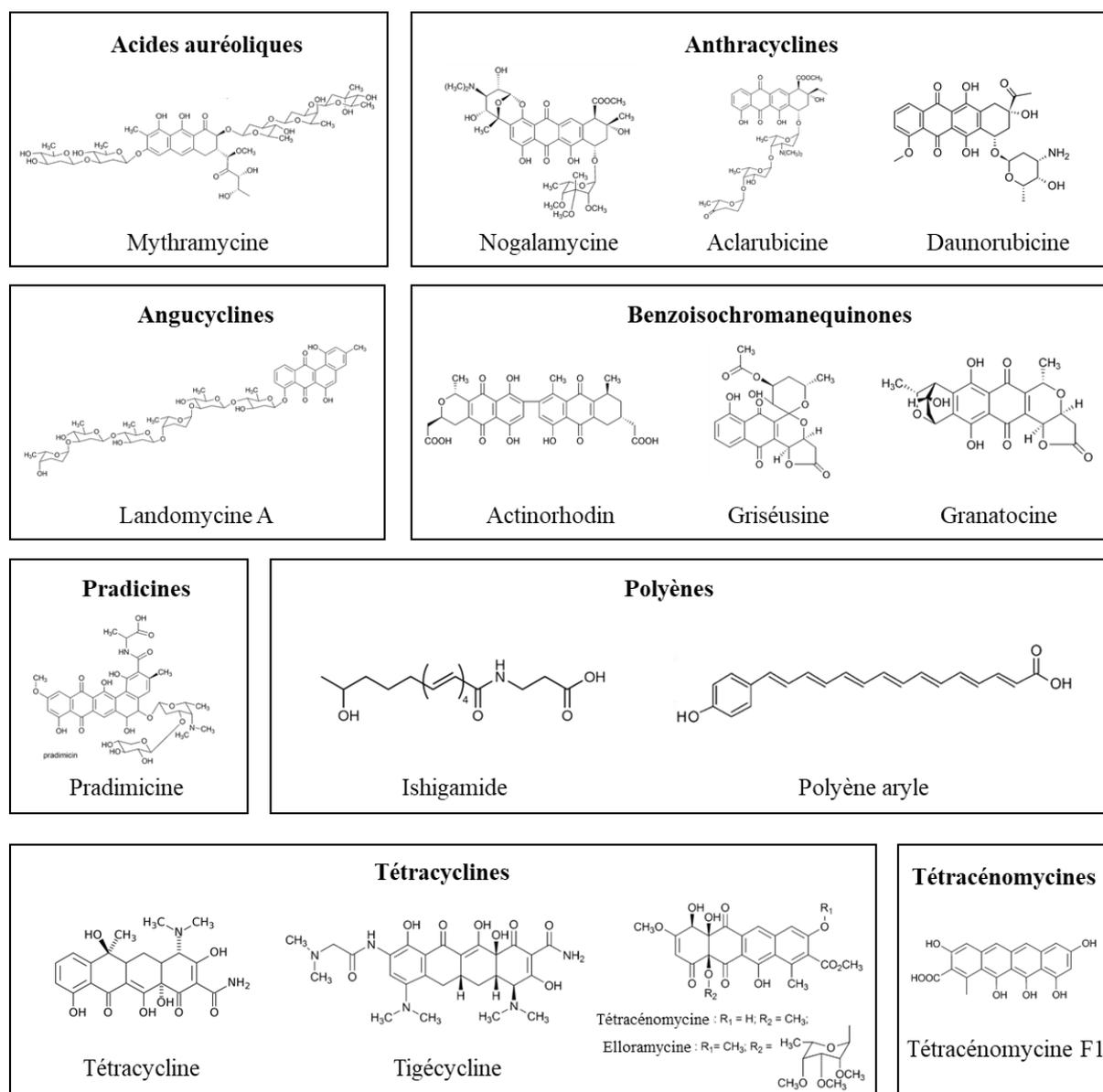
Les PKS de type III montrent l'architecture la plus simple parmi les différents types de PKS. En effet, ces enzymes sont composées d'un domaine cétoynthase (KS, de l'anglais *ketosynthase*) multi-fonctionnel (**Figure 6**), contrairement aux PKS de type I et de type II qui présentent une organisation multi-enzymatique. Ce type d'enzyme a la particularité d'être itératif, c'est-à-dire que seul le site actif du domaine KS est utilisé pour réaliser la condensation répétée de plusieurs acyl-CoA dans le but de former le polycétide. Les unités d'extension utilisées par les PKS de type III restent chargées sur des molécules de coenzyme A. Les condensations itératives sont suivies d'une étape de cyclisation qui opère suivant différents mécanismes de réactions intramoléculaires (Shimizu *et al.*, 2017). Des étapes supplémentaires de modification sont possibles pour terminer la formation des produits *via* des enzymes dites post-PKS.



**Figure 6 : Représentation schématique du mode d'action des PKS de type III (Adapté de Shimizu *et al.*, 2017).** Les PKS de type III catalysent la formation du polycétide au moyen d'un seul site actif et par l'utilisation des substrats activés par le coenzyme A.

## 2. Les PKS de type II

Les produits naturels synthétisés par les PKS de type II forment un groupe de métabolites spécialisés important retrouvé chez les bactéries et les champignons. Ces composés sont majoritairement aromatiques, polyinsaturés, et diversifiés dans leurs structures. La **Figure 7** montre des membres représentatifs des 8 classes de polycétides issus de PKS de type II.



**Figure 7 : Structure des membres représentatifs des métabolites spécialisés biosynthésés par les PKS de type II.**

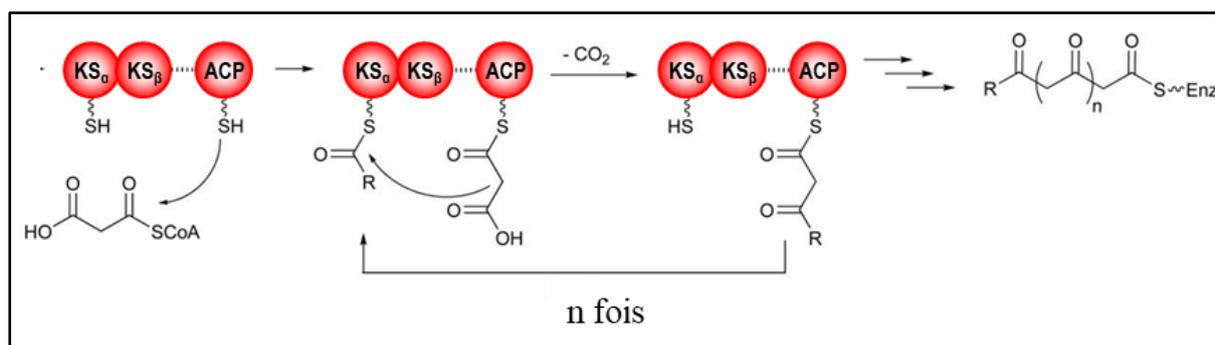
Les PKS de type II présentent une homologie avec les acides gras synthases bactériennes, en termes de mécanisme d'extension de la chaîne du polycétide et de composition en domaines au sein de la PKS (Chen *et al.*, 2018). À l'instar des PKS de type III, les enzymes de type II sont itératives. Cependant, elles présentent une organisation plus complexe, qui est multi-enzymatique (Hertweck *et al.*, 2007).

Deux domaines cétoynthases (KS<sub>α</sub> et KS<sub>β</sub>) et un domaine ACP (de l'anglais *acyl carrier protein*) vont former une PKS minimale (**Figure 8**). Des sous-unités supplémentaires peuvent s'ajouter à cette PKS minimale, telles que des cétoréductases, des cyclases ou encore des aromatasés, qui réalisent la maturation de la chaîne en cours de croissance. Les unités

d'extension sont attachées au niveau du domaine ACP, pour pouvoir être condensées au sein du complexe multi-enzymatique, contrairement aux PKS de type III.

La synthèse débute par la condensation itérative d'unités malonyl-CoA *via* les deux KS. Pour ce faire, le domaine  $KS_{\beta}$  fournit l'unité de départ, l'acétyl-ACP, qui est obtenue par la décarboxylation du malonyl-ACP. Le domaine  $KS_{\alpha}$  catalyse les réactions de condensation entre un malonate greffé à l'ACP et la chaîne en extension attachée à cette KS.

La sélection des unités d'extension se fait par l'action de la protéine MCAT (de l'anglais *malonyl-CoA: acyl carrier protein transacylase*) dont le gène est absent du groupe de gènes codant l'ensemble des PKS de type II à l'instar des FAS (Fernández-Moreno *et al.*, 1992 ; Hitchman *et al.*, 1998), ou bien les domaines ACP sont capables d'auto-malonylation (Arthur *et al.*, 2005). Des enzymes post-PKS vont terminer la formation du polycétide que ce soient par exemple des oxygénases, ou des glycosyl- et méthyl-transférases.



**Figure 8 : Représentation schématique du mode d'action général des PKS de type II (d'après Chen *et al.*, 2018).** Les PKS de type II sont dits itératives et donc constituées d'un seul module composé de plusieurs enzymes individuelles, qui catalysent l'intégralité de l'extension de la chaîne par itération de leurs activités.

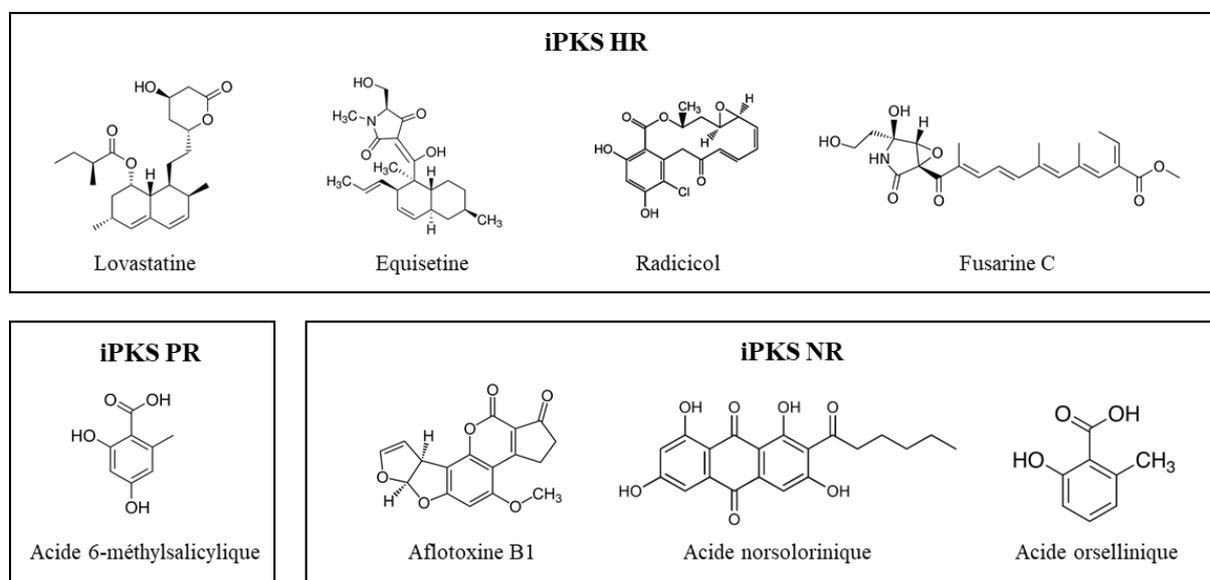
### 3. Les PKS de type I

C'est le type de PKS le plus complexe et le plus variable. Il est caractérisé par la présence de plusieurs sites actifs au sein de la même chaîne polypeptidique. Les PKS de type I possèdent un mode opératoire de type itératif (A) ou de type modulaire (B). Comme les PKS de type II, ce type de PKS présente une homologie avec les acides gras synthases (Weissman, 2009 ; Herbst *et al.*, 2018).

## A. Les PKS de type I itératives

Ce type de PKS réalise l'extension de la chaîne du polycétide au sein du même module. Un seul module multi-enzymatique est ainsi utilisé à répétition. Les PKS de type I itératives synthétisent des acides gras poly-insaturés à longue chaîne, ou des polycétides de type ènediynes chez les bactéries (Ahlert *et al.*, 2002). En outre, il existe dans le règne des champignons trois catégories de PKS de type I itératives. Elles sont désignées sous les noms de non-réductrices, partiellement réductrices et hautement réductrices, en fonction de la composition des complexes multi-enzymatiques en domaines méthylase, cétoréductase et déshydratase (Weissman, 2009).

Les polycétides générés par ces PKS fongiques sont très nombreux, dont beaucoup sont des toxines (**Figure 9**). Cette diversité provient du fait que les BGC associés à la synthèse des métabolites spécialisés fongiques contiennent une à deux PKS, 5 à 20 enzymes de modification et peuvent former des systèmes hybrides avec les NRPS (Herbst *et al.*, 2018). De surcroît, les voies de biosynthèse se recoupent entre PKS fongiques apportant alors des étapes modifications supplémentaires, conférant un plus grand degré de diversité en polycétides de ce type.



**Figure 9 : Structure des polycétides biosynthétisés à partir des différentes classes de PKS de type I itératives fongiques.** iPKS HR = PKS itérative hautement réductrice, iPKS PR = PKS itérative partiellement réductrice, iPKS NR = PKS itérative non-réductrice.

## B. Les PKS de type I modulaires

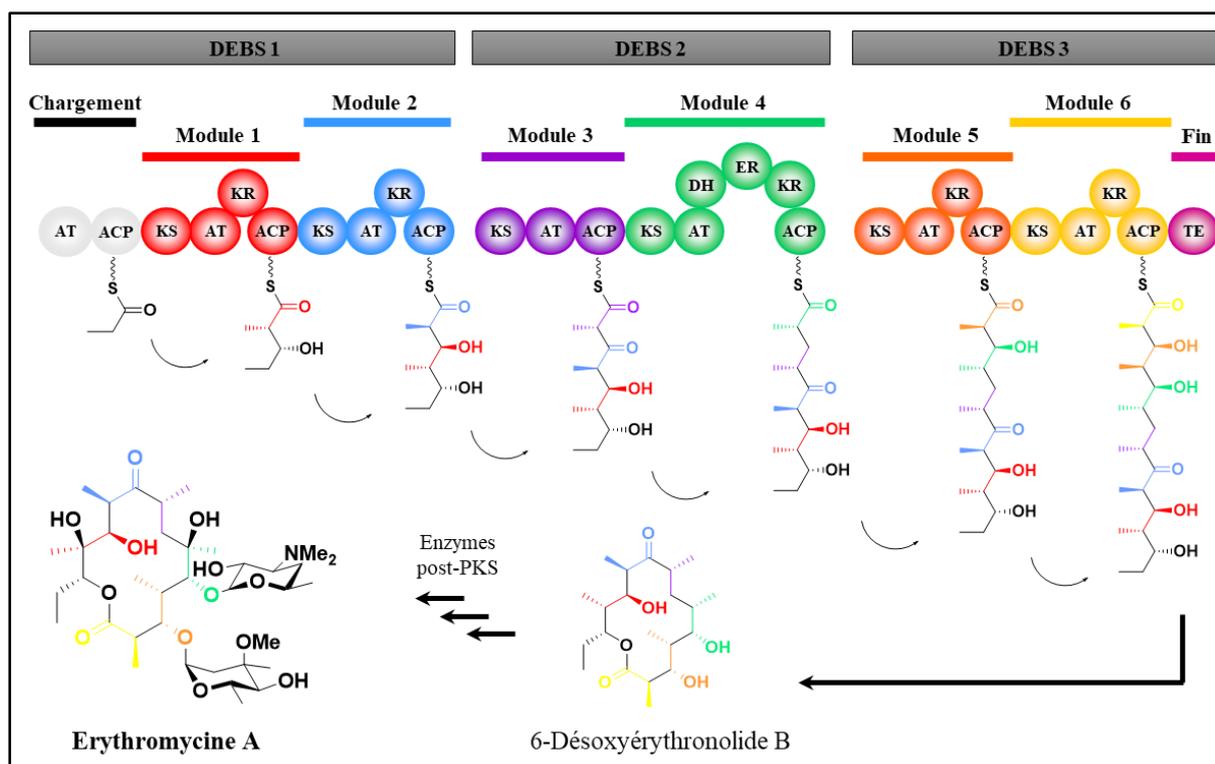
Au contraire des PKS de type I itératives, les PKS de type I modulaires ajoutent des unités d'extension spécifiques au niveau des différents modules qui se succèdent. Les PKS de type I

modulaires sont également caractérisées par leur immense taille, pouvant posséder plusieurs modules qui sont répartis sur différentes chaînes polypeptidiques. Par exemple, la néomédiomycine synthase possède 30 modules, et le BGC *nmd* le codant est le plus grand connu actuellement (Zhang *et al.*, 2017).

#### a. Les PKS modulaires de type *cis*-AT

Chez les PKS modulaires de type *cis*-AT, un module, au sein de la chaîne polypeptidique, catalyse la condensation et la modification d'un seul acyl-CoA (unité d'extension) sur l'intermédiaire en extension. Les modules sont constitués au minimum par les domaines KS, acyltransférase (AT, de l'anglais *acyltransferase*) et ACP. Des domaines optionnels cétoréductase (KR, de l'anglais *ketoreductase*), déshydratase (DH, de l'anglais *dehydratase*) et énoyl réductase (ER, de l'anglais *enoyl reductase*) peuvent être présents pour permettre la modification de l'intermédiaire nouvellement formé.

Le système modèle pour l'étude des PKS *cis*-AT est la désoxyérythronolide B synthase (DEBS), qui est impliqué dans la synthèse de l'érythromycine A (Cortes *et al.*, 1990 ; Donadio *et al.*, 1991). La DEBS a servi de modèle pour la compréhension des mécanismes catalytiques ainsi que des relations/fonctions des PKS modulaires de type *cis*-AT (Khosla *et al.*, 2007 ; Weissman, 2015 ; Dodge *et al.*, 2018), et certains aspects seront détaillés dans les sections **III** et **IV** de cet état de l'art. Cette PKS est constituée de trois chaînes polypeptidiques : DEBS 1 constituée du module de chargement (nécessaire pour recruter l'unité de départ et la transférer sur le premier module) ainsi que des deux premiers modules ; DEBS 2 formée par les modules 3 et 4 ; et DEBS 3 composée des module 5 et 6, de même que du domaine TE nécessaire à la libération du polycétide (**Figure 10**).



**Figure 10 : Représentation de la PKS DEBS synthétisant le 6-désoxyérythronolide B (d'après KJ Weissman et Cortes *et al.*, 1990).** La synthèse de l'érythromycine A est réalisée *via* la condensation d'une unité de propionate et de six unités de méthylmalonate grâce à l'activité successive des six modules d'extension composant la PKS. La chaîne est libérée et cyclisée par l'activité du domaine TE, afin de générer le produit 6-désoxyérythronolide B. La maturation de ce dernier par les enzymes post-PKS permet l'obtention de l'érythromycine A. Légende : AT = acyltransférase, ACP = *acyl carrier protein*, KS = cétyosynthase, KR = cétyoréductase, DH = déshydratase, ER = ényol réductase, TE = thioestérase. Bien que la PKS, montrée ci-dessus, opère en homodimère, seul un ensemble de sites actifs est montré pour plus de clarté. Les domaines liés de manière covalente par des « linkers » se touchent entre eux sur le schéma.

La synthèse de l'érythromycine débute par le recrutement de l'unité de départ, un propionate, *via* le domaine AT du module de chargement. Le propionate est ensuite transféré au groupement phosphopantéthéine attaché de manière covalente au domaine ACP. L'extension du polycétide s'effectue par la translocation du propionate du domaine ACP au domaine KS du premier module, qui est responsable de la première extension. Le domaine AT du premier module sélectionne la première unité d'extension, un méthylmalonate, et le transfert au niveau de l'ACP. Puis, le domaine KS catalyse une condensation décarboxylative, dite de Claisen, permettant la formation d'une nouvelle liaison carbone-carbone entre l'unité de départ et l'unité d'extension par la décarboxylation du méthylmalonate. L'unité de départ a alors été allongé de deux carbones, formant donc un dicétide. Puis, suivant la composition du module en domaines optionnels, le groupement méthyl en C2 peut être épimérisé et le groupement cétone

en C3 est modifié par réduction. À la fin des six condensations réalisées par les six modules successifs d'extension, le domaine TE libère le polycétide non-mature en le cyclisant pour générer le 6-désoxyérythronolide B (6-dEB). Enfin, le 6-dEB est modifié par cinq enzymes post-PKS (deux hydroxylases, deux glycosyltransférases dont une, EryCII, qui agit en complexe avec un partenaire homologue au cytochrome P450 (Moncrieffe *et al.*, 2012), et une méthyltransférase) pour générer l'érythromycine A, qui est le produit final bioactif (Khosla *et al.*, 2007).

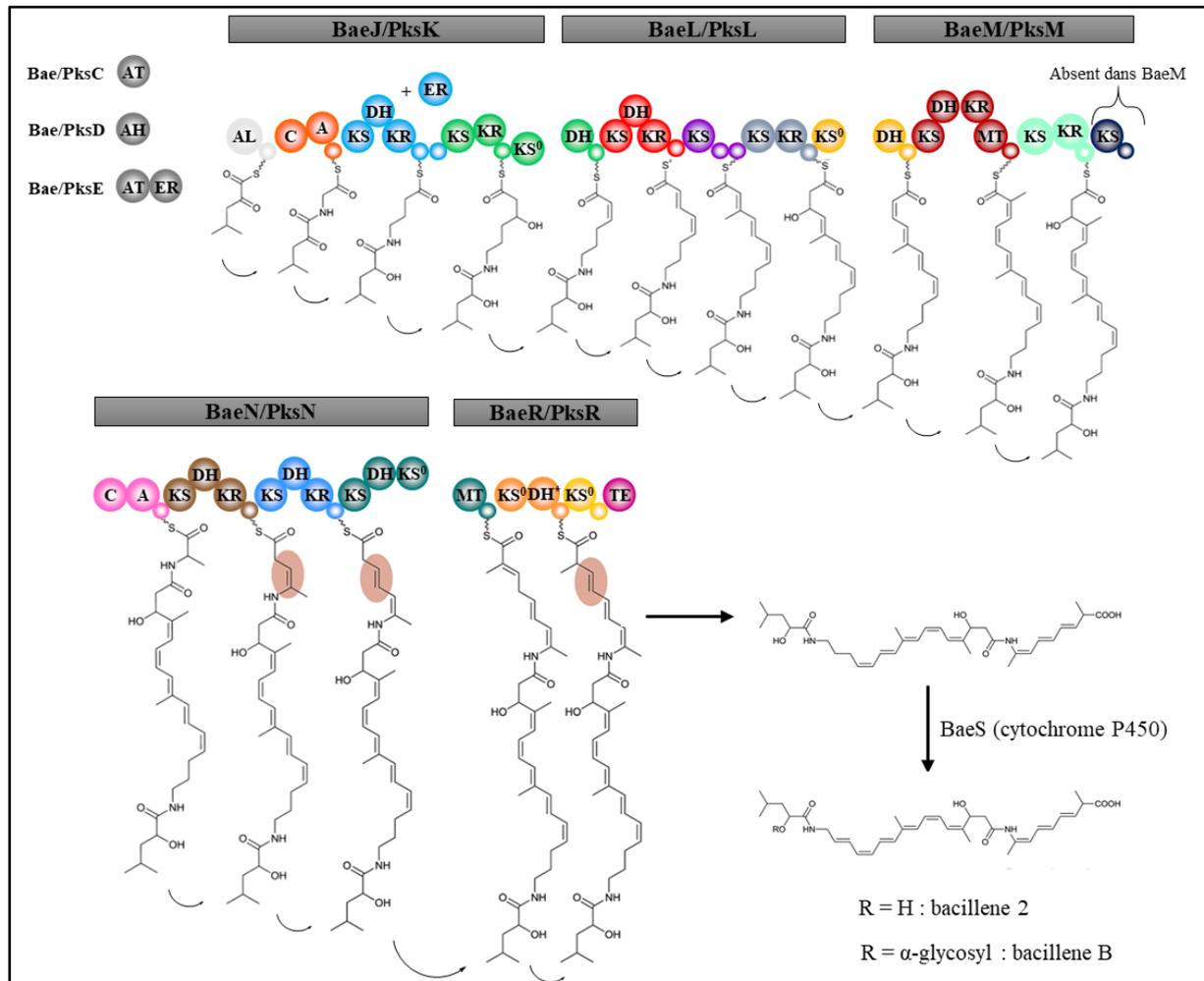
### b. Les PKS modulaires de type *trans*-AT

Initialement, l'étude des PKS modulaires de type *trans*-AT (Piel, 2002 ; Cheng *et al.*, 2003 ; Helfrich and Piel, 2016) a été négligée du fait de la rareté de ces systèmes chez les Actinobactéries. Grâce à l'émergence des séquences génomiques bactériennes, l'étude des gènes codant les PKS a montré que près de 38 % des PKS modulaires sont de type *trans*-AT, suggérant que les polycétides produits par ces systèmes constituent une classe de produit naturel à part (O'Brien *et al.*, 2014). Les PKS modulaires de type *trans*-AT s'organisent aussi en système multi-modulaire, mais montrent des différences significatives par rapport aux PKS modulaires de type *cis*-AT.

La première différence est l'absence de domaine AT au sein de la chaîne polypeptidique des PKS modulaires de type *trans*-AT, qui agit alors en tant qu'enzyme discrète par l'activité de 1 à 3 domaines AT par BGC. En outre, ces systèmes peuvent posséder des modules contenant des domaines de nature différente que les PKS modulaires de types *cis*-AT. Les modules de PKS modulaires de type *trans*-AT peuvent aussi posséder des modules inactifs pour l'activité d'élongation, des modules tronqués, et de nombreuses enzymes discrètes qui permettent d'ajouter des modifications supplémentaires aux polycétides (Helfrich and Piel, 2016).

La **Figure 11** représente un système modèle de PKS modulaire de type *trans*-AT, celui responsable de la synthèse de la bacillaene chez *Bacillus subtilis* (Patel *et al.*, 1995) et *Bacillus amyloliquefaciens* (Chen *et al.*, 2006, 2007), un inhibiteur de la synthèse des protéines chez les procaryotes. Cette PKS est composée de 5 sous-unités comportant des modules de nature différente, eux-mêmes pouvant inclure des modules caractéristiques des NRPS, des modules tronqués et aussi des modules caractéristiques des PKS modulaires de type *cis*-AT. De surcroît, de nombreux domaines agissent en *trans*, incluant 2 domaines *trans*-AT, un domaine acylhydrolase (AH) et un domaine ER. Dans le but de comprendre un niveau d'organisation

macromoléculaire supérieur, l'organisation de cette PKS dans la bactérie sera détaillée dans la section V de cet état de l'art.



**Figure 11 : Représentation de la PKS responsable de la synthèse de la bacillaene (d'après Helfrich and Piel, 2016).** Les PKS modulaires de type *trans*-AT possèdent une grande richesse en nature de modules. La PKS Bae chez *Bacillus amyloliquefaciens* et la PKS Pks chez *B. subtilis* sont représentées. Des différences dans la composition en domaines peuvent exister pour un même système PKS. Par exemple, le module 10 de la sous-unité PksM est absent dans la sous-unité homologue BaeM. Les fonctions chimiques surlignées en ocre sont des doubles liaisons β,γ.

Légende : AL = acyl-ligase, A = domaine d'adénylation, C = domaine de condensation, KS = céto-synthase, KS<sup>0</sup> = céto-synthase qui effectue la réaction de transacylation mais pas la réaction de condensation décarboxylative, AT = acyltransférase, AH = acylhydrolase, KR = céto-réductase, DH = déshydratase, ER = énoyl réductase, MT = méthyltransférase, TE = thioestérase.

## 4. La biologie synthétique des PKS modulaires

### A. Les PKS modulaires de type *cis*-AT comme modèle d'étude

Que ce soit pour la PKS DEBS ou d'autres PKS modulaires de type *cis*-AT, une corrélation est observée entre la composition en domaines actifs des modules de la PKS et la structure du polycétide biosynthétisé. Ce phénomène est désigné sous le terme de règle de colinéarité. Cette observation met en avant le caractère modulaire et la logique de biosynthèse de ce type de PKS qui produit le métabolite sous le principe de la chaîne d'assemblage. De cette manière, le produit formé par une PKS de ce type peut être prédit, grâce à la composition en modules de la chaîne d'assemblage, constituant ainsi un avantage considérable pour la synthèse de nouvelles structures *via* la manipulation des gènes codant les PKS. À ce jour, des centaines de nouvelles structures ont pu être générées par la modification de domaines, de modules et de sous-unités (Weissman, 2016b). Il est à noter que ce lien entre la composition en domaines des polypeptides et les produits est moins clair pour les PKS modulaires de type *trans*-AT (Nguyen *et al.*, 2008). Cette différence entre les PKS modulaires de types *cis*-AT et *trans*-AT sera discutée dans la section IV de cet état de l'art.

En outre des remarquables mécanismes catalytiques des PKS, leur architecture multi-modulaire permet de mettre à l'épreuve leur potentiel pour des expériences de biologie synthétique. D'un point de vue fondamental, il y a de nombreuses relations entre la structure et la fonction des PKS, et l'étude de ces systèmes parmi les plus complexes du vivant est un véritable défi. L'étude de l'évolution, de la diversité et de l'organisation structurale et macromoléculaire des PKS permettra à terme d'améliorer notre compréhension de ces complexes multi-enzymatiques. Ces connaissances pourront être exploitées et permettront d'augmenter considérablement notre arsenal thérapeutique. À cela s'ajoute la découverte de nouveaux systèmes PKS, et donc de nouvelles molécules bioactives et potentiellement utilisables en thérapie. Les approches respectant le mode d'évolution (Peng *et al.*, 2019a) ou exploitant les mécanismes d'évolution (Chemler *et al.*, 2015 ; Wlodek *et al.*, 2017) des PKS ont déjà commencé à faire leur preuve pour générer des PKS chimères fonctionnelles et donc de nouvelles molécules bioactives.

### B. Les cibles de la biologie synthétique des PKS modulaires

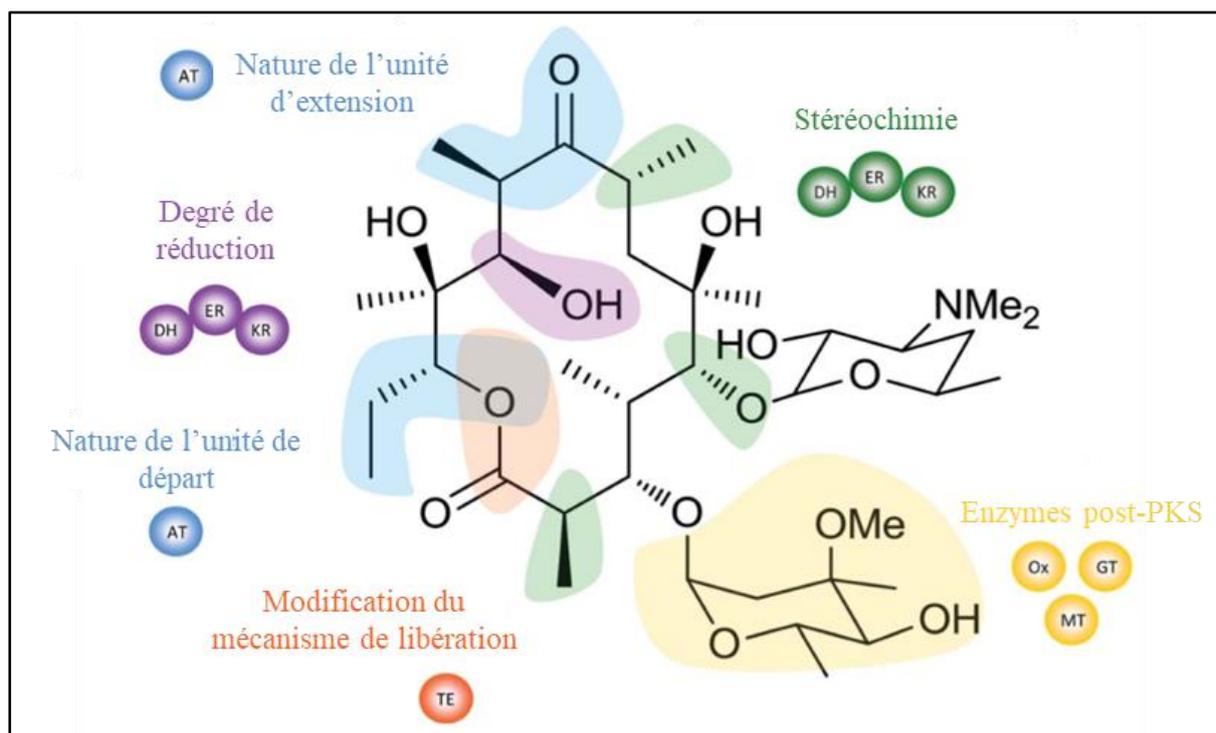
La grande diversité des PKS et des polycétides qu'elles produisent mettent en lumière le potentiel de la biologie synthétique des PKS modulaires. La première approche d'ingénierie de

ces systèmes est de manipuler les domaines (**Figure 12**). En effet, il est possible de modifier un domaine AT pour changer la nature de l'unité de départ ou d'extension, soit en modifiant les résidus du site actif (Reeves *et al.*, 2001 ; Del Vecchio *et al.*, 2003) , soit en échangeant ce domaine avec un autre (Oliynyk *et al.*, 1996). Il est aussi possible de changer le type de modification apporté par un domaine KR, en l'inactivant (Donadio *et al.*, 1991, 1993) ou en modifiant le type de réduction par échange de domaine (Kellenberger *et al.*, 2008), pouvant apporter une épimérisation supplémentaire du groupe méthyle en C2 (Kellenberger *et al.*, 2008 ; Annavaal *et al.*, 2015 ; Eng *et al.*, 2016). La modification du degré de réduction peut aussi être accomplie en modifiant ou en échangeant la boucle de réduction (Hagen *et al.*, 2016). Enfin, la relocalisation du domaine TE conduit à une libération prématurée du polycétide en cours d'extension, permettant ainsi de contrôler la longueur du polycétide (Cortes *et al.*, 1995). La nature du domaine TE peut aussi être modifiée pour changer le mécanisme de libération du polycétide de la PKS, soit par macro-cyclisation, soit par hydrolyse (Weissman, 2016).

Il est également possible de modifier des modules entiers ou même des sous-unités entières : les supprimer pour réduire ou les ajouter pour augmenter la taille de la chaîne (Peng *et al.*, 2019a), ou les échanger pour modifier une ou plusieurs caractéristiques structurales du polycétide à la fois (Gokhale *et al.*, 1999 ; Ranganathan *et al.*, 1999). Compte tenu de leur contribution essentielle à la bioactivité du polycétide mature, les enzymes post-PKS sont également des cibles attractives pour la biologie synthétique, par la suppression ou par l'ajout d'étapes de modifications spécifiques, ou par des changements dans la régiospécificité ou la stéréospécificité des réactions (Dhakal *et al.*, 2019).

Si tous les changements ci-dessus pourraient en principe être combinés pour introduire simultanément plusieurs types de modifications, les trois dernières décennies d'ingénierie des PKS modulaires ont montré que toutes ces modifications ne peuvent pas être obtenues (Klaus and Grninger, 2018 ; Weissman, 2016b). Les PKS modulaires sont des systèmes multi-enzymatiques hautement spécialisés, et il existe de nombreuses lacunes dans nos connaissances à tous les niveaux moléculaires, de l'activité catalytique des domaines individuels (Weissman, 2017), en passant par la communication entre les domaines d'un même module et entre deux modules de la PKS lors de la biosynthèse du polycétide (Dodge *et al.*, 2018), jusqu'à sa localisation et son organisation dans la cellule (Straight *et al.*, 2007). Au vu et au su de la complexité de ces systèmes, il est donc préférable d'identifier des stratégies d'ingénierie pour constituer une véritable boîte à outil de biologie synthétique des PKS modulaires. Actuellement, la modification de la stéréochimie du polycétide est un véritable défi, car elle peut combiner à la fois une réaction de réduction et d'épimérisation, toutes-deux

catalysées par un même domaine KR qui doit agir dans un contexte modulaire précis. Dans la section **III** de cet état de l'art, nous nous intéresserons au contrôle de la stéréochimie lors de la biosynthèse du polycétide et aux stratégies d'ingénierie des PKS visant à modifier la stéréochimie d'un polycétide en manipulant le domaine KR qui ont conduit à la nécessité de définir ce qu'est un contexte modulaire.



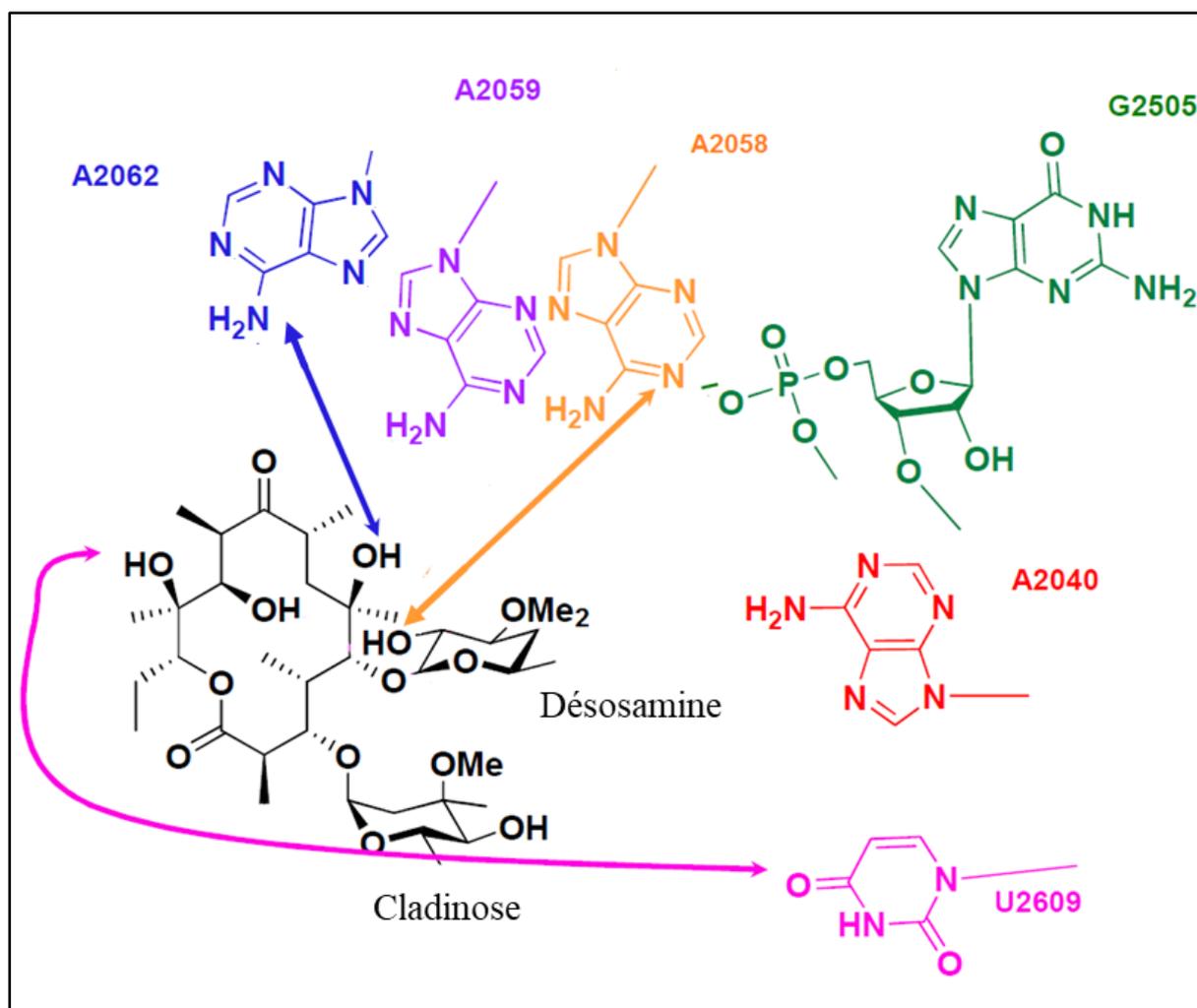
**Figure 12 : Cibles théoriques de modification de la structure d'un polycétide (Adapté de Weissman, 2016).** Les types de modifications apportées sont surlignés en couleur et en fonction des domaines ou des enzymes post-PKS impliqués. Toutes les modifications indiquées peuvent être combinées.

### III. Du contrôle de la stéréochimie à l'ingénierie des domaines KR

#### 1. Intérêt pour la stéréochimie

La richesse en bioactivités des polycétides provient de leurs structures complexes, qui intègrent à la fois une densité de groupes fonctionnels élevée et une stéréochimie riche. Ces caractéristiques, associées au fait que la majorité des polycétides réduits sont macrocycliques, confèrent aux polycétides une affinité pour leur cible biologique élevée ( $10^{-7}$  à  $10^{-9}$  M) (Clardy and Walsh, 2004).

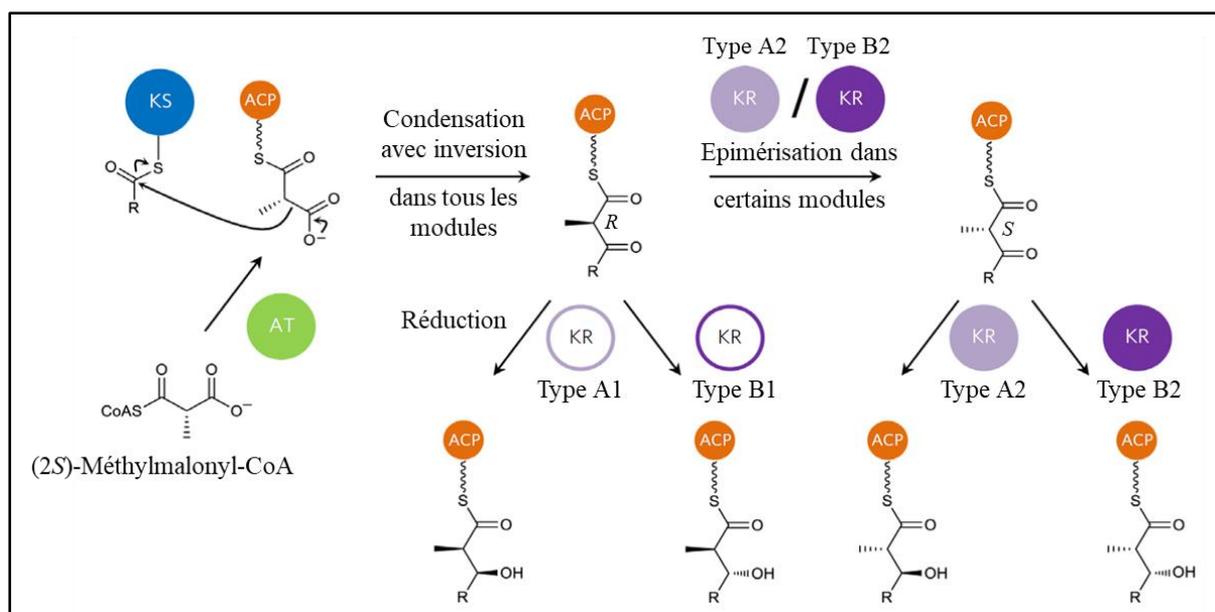
Contrairement à la structure du polycétide, il n'y a pas de corrélation évidente entre la composition en domaines et en modules de la PKS et la stéréochimie du polycétide synthétisé. Par exemple, l'érythromycine A possède 10 centres asymétriques dans sa structure, soit 1024 ( $2^{10}$ ) stéréoisomères possibles. Pourtant, un seul stéréoisomère est produit par la PKS, démontrant le contrôle strict de la stéréochimie des groupements du polycétide. À chaque étape d'extension, la stéréochimie est contrôlée au niveau du groupement hydroxyle et aussi au niveau du groupement méthyle adjacent. L'importance de la stéréochimie des polycétides est illustrée par la structure de l'ARNr 23S de la sous-unité 50S du ribosome de *Deinococcus radiodurans* co-cristallisé en présence d'érythromycine A (**Figure 13**) (Schlünzen *et al.*, 2001). La configuration des carbones est un paramètre critique afin que les groupements hydroxyles (au niveau des carbones C6, C11 et C12) et le désosamine (attaché de manière covalente au groupement hydroxyle au niveau du carbone C5) de l'érythromycine A puissent interagir avec les bases nucléotidiques du ribosome. De cette manière, la translocation des peptides et la formation de polypeptides sont bloquées : ainsi la biosynthèse des protéines est entravée, empêchant alors la croissance des bactéries. Ce lien intime entre la stéréochimie du polycétide et l'activité biologique fait du contrôle de la stéréochimie une cible attractive pour générer de nouvelles structures de polycétides par l'approche de biologie synthétique (Weissman, 2016).



**Figure 13 :** Représentation schématique de l'interaction entre l'érythromycine A et la sous-unité 50S du ribosome de *Deinococcus radiodurans* (Adapté par KJ Weissman d'après Schlünzen *et al.*, 2001). L'érythromycine A agit en tant qu'antibiotique en se fixant au niveau de la sous-unité 50S du ribosome et plus particulièrement à l'ARNr 23S. Cette fixation se fait *via* trois nucléotides de l'ARNr 23S et trois groupements fonctionnels de l'érythromycine A. Ces groupements fonctionnels possèdent une stéréochimie bien définie qui est contrôlée par la PKS. Ce contrôle de la stéréochimie est essentiel pour permettre l'interaction avec l'ARNr 23S et est donc indispensable à l'activité biologique du polycétide.

## 2. Le contrôle de la stéréochimie lors de la synthèse du polycétide

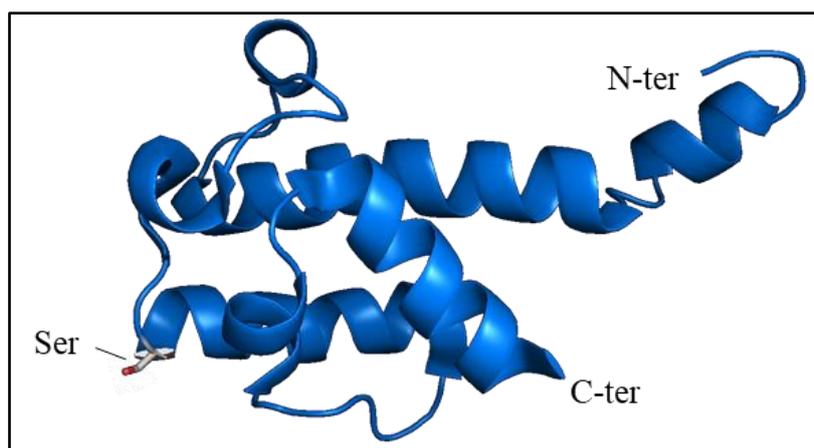
A partir de la structure de l'érythromycine A (**Figure 13**), il est clair que les polycétides sont riches en centres stéréogènes, et que les PKS peuvent générer les deux configurations des centres méthyle et hydroxyle. Les trois domaines qui sont impliqués dans la synthèse de ces parties de la molécule (KS, AT et KR, lorsque cette dernière est seule dans la boucle réductrice) peuvent donc être impliqués dans le contrôle de la stéréochimie (**Figure 14**).



**Figure 14 : Modèle harmonisé du stéréocontrôle au sein des PKS modulaires (adapté de AnnaVal *et al.*, 2015).** Les domaines AT (en vert) sélectionnent un seul isomère (2*S*) de l'unité d'extension méthylmalonate pour ensuite le transférer vers le domaine ACP (en orange) modifié post-traductionnellement par ajout d'un groupement phosphopantéthéine. Le domaine KS (en bleu) catalyse la condensation de type Claisen entre l'unité d'extension et la chaîne du polycétide en élongation, en respectant une inversion de la stéréochimie au niveau du centre méthyle et génère uniquement la stéréochimie *R*. Le groupement cétone au niveau du carbone C3 peut être réduit dans les deux directions par le domaine KR (en nuance de violet, plein ou vide suivant le type de domaine KR considéré) afin de générer un produit de configuration *S* ou *R*, pour les domaines KR de type A ou B, respectivement. Le centre méthyle peut également être épimérisé par les domaines KR du type A2 ou B2 avant la réduction, produisant la configuration *S*. Il est à noter que les configurations *R* et *S* du carbone C3 dépendent du radical (*R*).

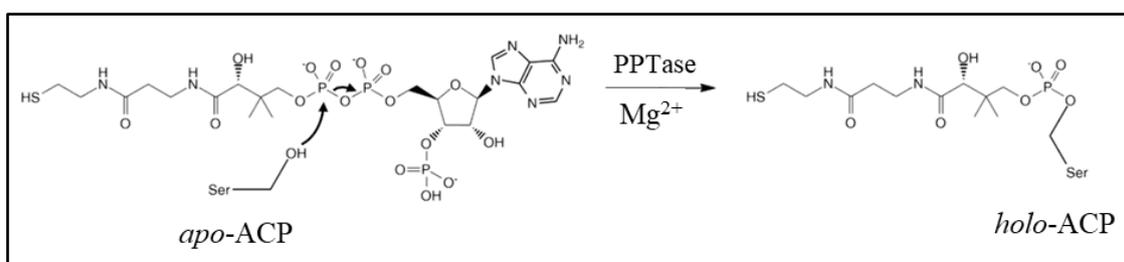
### A. Le domaine *acyl carrier protein* (ACP)

Le domaine ACP est responsable du support de la chaîne du polycétide au cours des différentes étapes catalytiques au sein de la PKS. Ce domaine non-catalytique est composé d'environ 80 acides aminés et adopte un repliement en 4 hélices  $\alpha$  organisées en fagot (**Figure 15**, Alekseyev *et al.*, 2007), ainsi que d'une 5<sup>ème</sup> hélice à l'extrémité N-terminale (Moretto *et al.*, 2019). Le domaine ACP des FAS (de l'anglais *fatty acid synthase*) de type II, des PKS de type II et des NRPS présentent le même repliement (Keatinge-Clay, 2012).



**Figure 15 :** Structure RMN du domaine cœur des ACP provenant du module 2 de la PKS DEBS (PDB : 2JU2 ; Alekseyev *et al.*, 2007). La sérine qui sert de point d'attachement au bras phosphopantéthéine est indiquée sur la structure.

Le domaine ACP a la particularité d'être activé par une modification post-traductionnelle avec un groupement prosthétique. Comme montré en **Figure 16**, une enzyme appelée phosphopantéthéinyl transférase (PPTase) catalyse le transfert du groupement phosphopantéthéine au niveau de la sérine du motif DSL conservé au sein des domaines ACP (Lambalot *et al.*, 1996). Lorsque l'*apo*-ACP est modifié par ce groupement prosthétique, il est appelé *holo*-ACP, et lorsque le polycétide est attaché, le domaine est alors appelé acyl-ACP. Les conséquences de ces modifications sont : (i) de diminuer le pK<sub>a</sub> du substrat pour favoriser son activation pour la réaction de transfert du substrat, par le remplacement de le groupement hydroxyle de la sérine du motif DSL du domaine ACP par le groupement thiol du bras phosphopantéthéine, (ii) de permettre à la chaîne du polycétide d'accéder aux différents sites actifs d'un module et (iii) d'augmenter la sélectivité pour le substrat qui devient la chaîne du polycétide et l'*holo*-ACP (Beld *et al.*, 2015).

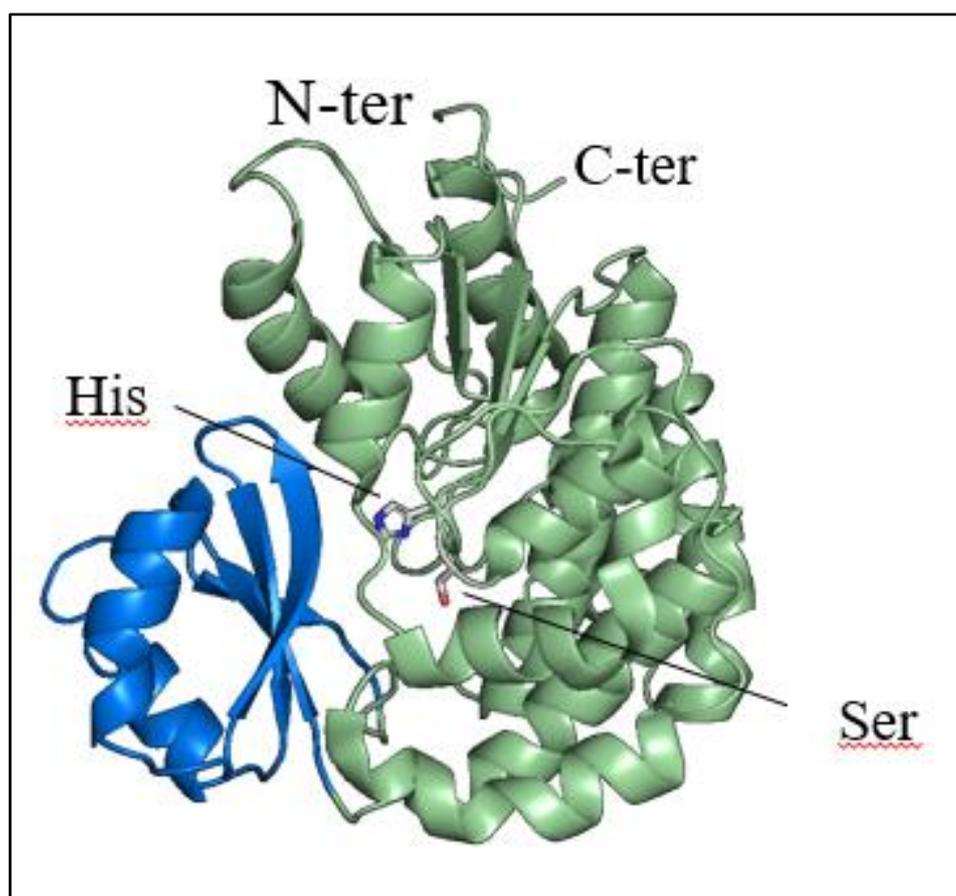


**Figure 16 :** Réaction de transfert du groupement phosphopantéthéine du coenzyme A sur la sérine du domaine ACP catalysée par la PPTase (adapté de Lambalot *et al.*, 1996). Le bras phosphopantéthéine est attaché de manière covalente par une liaison thioester à la sérine du domaine ACP pour générer l'*holo*-ACP.

## B. Le domaine acyltransférase (AT)

Le domaine AT est responsable du recrutement de l'unité de départ et des unités d'extension, ainsi que de leur transfert au niveau du bras phosphopantéthéine du domaine ACP. Ce domaine est présent au sein de la chaîne polypeptidique chez les PKS modulaires de type *cis*-AT et est présent en tant que protéine discrète chez les PKS modulaires de type *trans*-AT. Ce domaine enzymatique est monomérique et est composé d'environ 300 acides aminés.

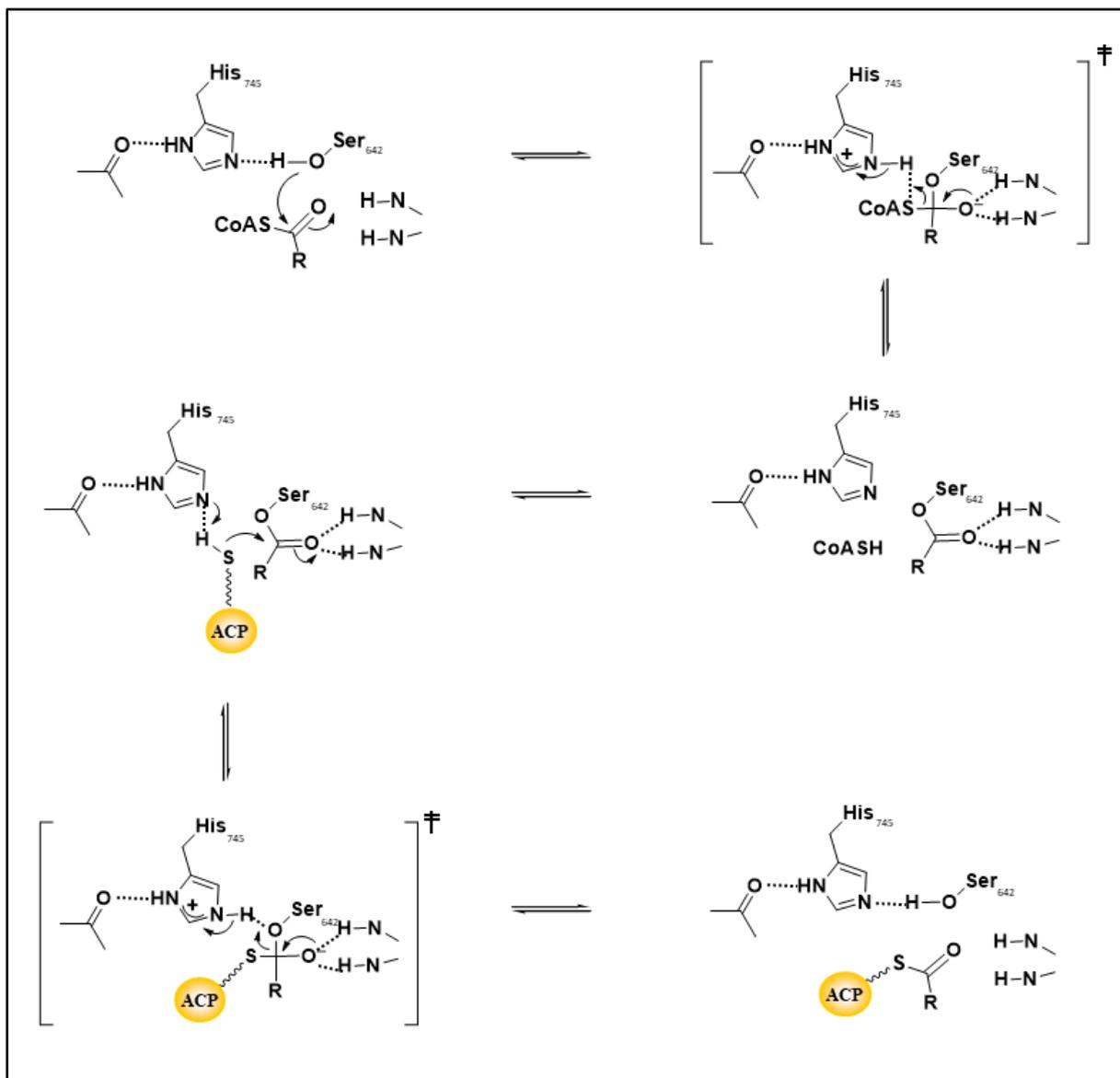
Ce domaine est divisé en deux sous-domaines, un sous-domaine catalytique d'environ 240 acides aminés structuré avec un repliement de type  $\alpha/\beta$  hydrolase, et un sous-domaine non-catalytique d'environ 60 acides aminés structuré en repliement de type ferrédoxine (**Figure 17**). Les structures d'un domaine AT provenant d'un système de PKS modulaire de type *cis*-AT (Tang et al., 2006, 2007) ou *trans*-AT présentent les mêmes repliements (Wong *et al.*, 2011).



**Figure 17 : Structure cristallographique du domaine AT provenant du module 5 de la PKS DEBS (PDB : 2HG4 ; Tang *et al.*, 2006).** Le sous-domaine non-catalytique est représenté en bleu et le sous-domaine catalytique est représenté en vert pâle. Les résidus histidine et sérine impliqués dans la catalyse sont indiqués sur la structure.

Le mécanisme catalytique du recrutement des unités de départ ou d'extension par le domaine AT est de type ping-pong et est semblable à celui des protéases à sérine (**Figure 18**). L'histidine active la sérine catalytique par une catalyse basique au niveau du groupement hydroxyle. La sérine activée peut effectuer une attaque nucléophile au niveau du carbone de la fonction thioester de l'unité d'extension liée de manière covalente au CoA. La charge négative portée sur l'oxygène de l'intermédiaire réactionnel est stabilisée par le trou de l'oxyanion ce qui permet le départ du CoA, générant ainsi le domaine AT lié de manière covalente à son substrat (Dunn *et al.*, 2013).

Le transfert du substrat au niveau du groupement phosphopantéthéine du domaine ACP est présenté dans la **Figure 18**. Ce mécanisme catalytique est semblable à celui du recrutement du substrat par le domaine AT. Pour cela, le groupement thiol du bras phosphopantéthéine de l'*holo*-ACP réalise une attaque nucléophile au niveau du carbone de la fonction ester au niveau de la sérine catalytique du domaine AT lié à son substrat. L'intermédiaire réactionnel est stabilisé par le trou de l'oxyanion porté par un groupement amine de la chaîne polypeptidique ce qui provoque le départ du domaine AT, conduisant au groupement phosphopantéthéine lié de manière covalente à son substrat (Dunn, *et al.*, 2013).



**Figure 18 : Mécanisme catalytique du domaine AT (adapté de Tsai and Ames, 2009).** Réaction de recrutement de l'unité d'extension par le domaine AT et réaction du transfert du polycétide du domaine AT vers le domaine ACP. Le NH du trou de l'oxyanion est fourni par la chaîne polypeptidique et permet de stabiliser les charges sur les intermédiaires réactionnels. R = chaîne du polycétide.

Le domaine AT recrute le substrat de manière stéréospécifique (**Figure 14**). Dans le cas de la PKS modèle DEBS, les domaines AT recrutent le (2*S*)-méthylmalonyl-CoA, à l'exception de l'AT du module de chargement qui recrute un propionyl-CoA (Marsden *et al.*, 1994). Les domaines AT sont homologues parmi les PKS modulaires de type *cis*-AT, et possèdent très probablement la même stéréospécificité (Dunn *et al.*, 2013). Si les domaines AT présentent la même stéréospécificité pour leur substrat, ils ne possèdent pas pour autant la même sélectivité. En effet, cette sélectivité est médiée par l'existence de motifs consensus dans le site actif du domaine AT. Par exemple, les domaines AT sélectionnant le méthylmalonyl-CoA possèdent

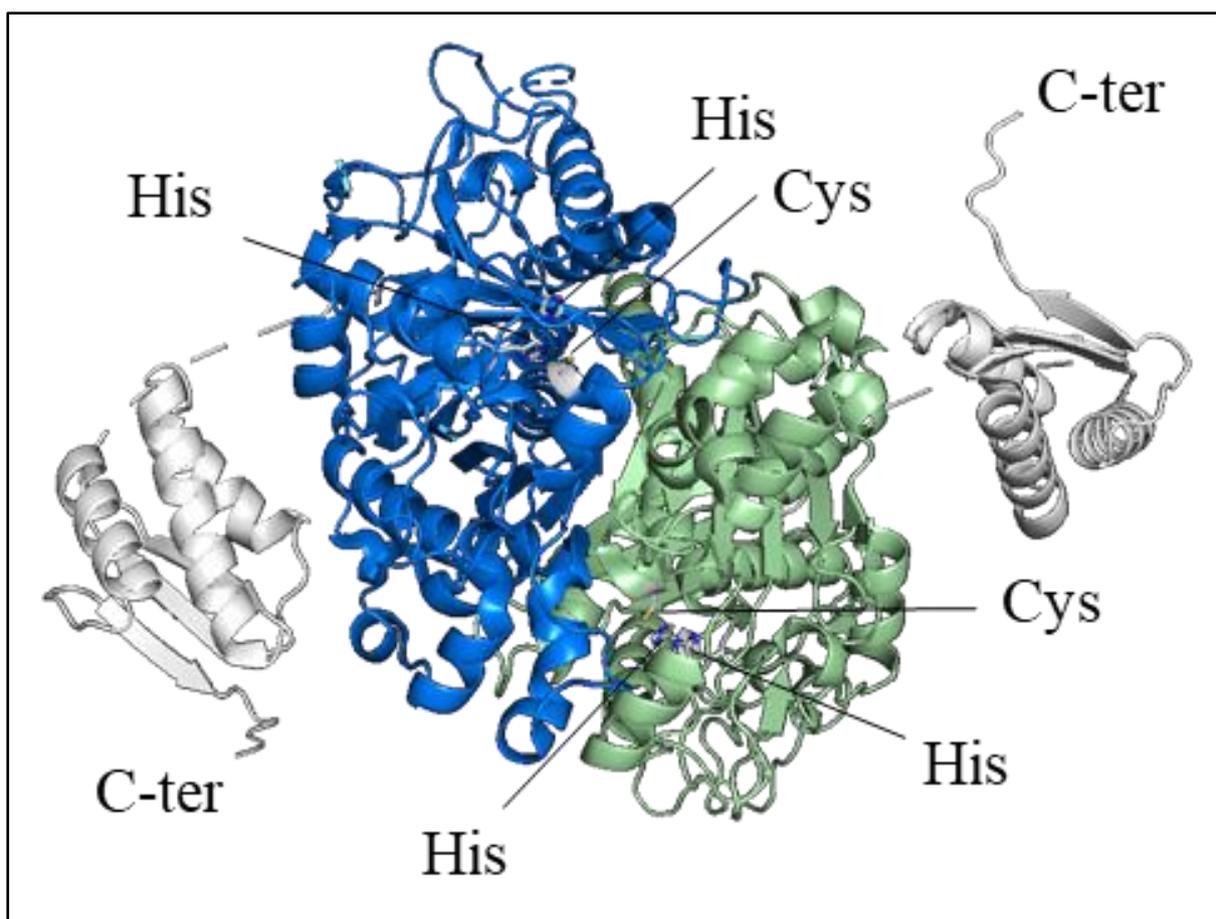
un motif **YASH** alors que ceux sélectionnant le malonyl-CoA possèdent un motif **HAFH** (Tang *et al.*, 2006), où l'histidine en gras est impliquée dans la catalyse. Si la PKS DEBS utilise le propionyl-CoA comme unité de départ et le méthylmalonyl-CoA comme unité d'extension, il existe cependant d'autres unités d'extension pouvant être utilisées par les PKS modulaires (Ray and Moore, 2016).

Des expériences de biologie synthétique ont eu pour objectif de convertir des domaines AT spécifiques du (2S)-méthylmalonyl-CoA en domaines AT spécifiques du malonyl-CoA en échangeant les motifs caractéristiques. Cependant, les domaines AT générés ont montré une promiscuité de substrat en étant capable de sélectionner les deux types d'unités d'extension. Ce résultat montre que l'échange de motif seul ne suffit pas pour modifier la sélectivité du domaine, et que d'autres éléments au sein du domaine AT sont impliqués dans la sélectivité du substrat (Reeves *et al.*, 2001 ; Del Vecchio *et al.*, 2003 ; Sundermann *et al.*, 2013). Plus récemment, le domaine AT spécifique du (2S)-méthylmalonyl-CoA provenant du sixième module de DEBS a été converti en domaine AT spécifique d'un substrat non-naturel, le propargylmalonyl-CoA (Koryakina *et al.*, 2017). Ce domaine a été obtenu par la mutation d'un seul acide aminé (Tyr186Arg) qui n'avait pas été encore identifié comme étant important pour la sélectivité des domaines AT. Ce premier exemple d'inversion de la spécificité d'un domaine AT montre notamment que d'autres acides aminés que les motifs caractéristiques peuvent aussi être impliqués dans le recrutement du substrat.

### C. Le domaine cétoynthase (KS)

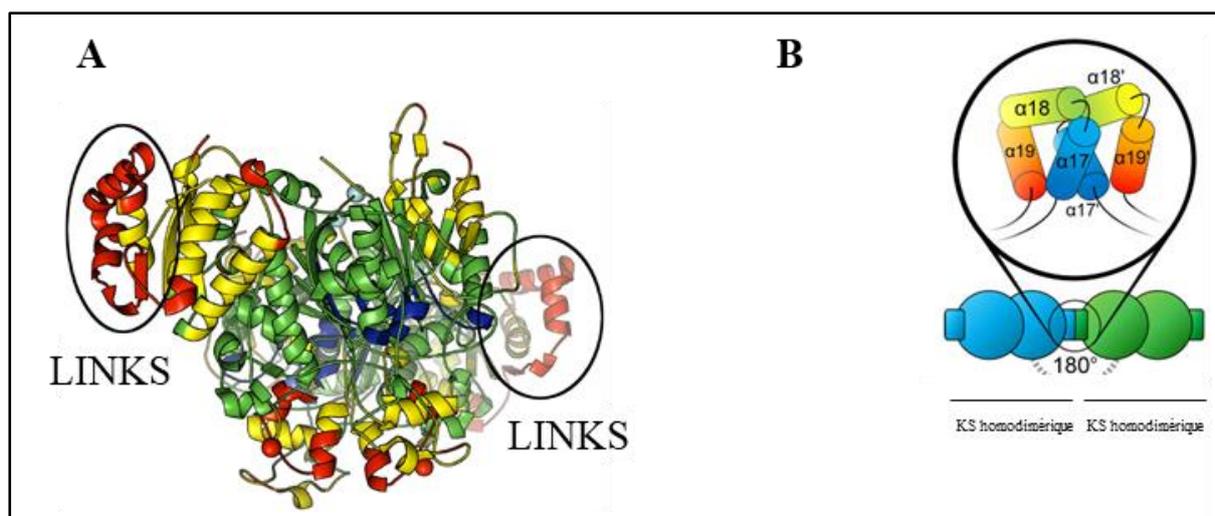
Le domaine KS est responsable de la condensation entre une unité d'extension et la chaîne du polycétide en cours d'extension. Ce domaine enzymatique est homodimérique et chaque monomère est composé d'environ 430 acides aminés. Ce dernier se structure avec un repliement de type thiolase  $\alpha/\beta/\alpha/\beta$  et la dimérisation s'opère à l'interface des deux feuillets  $\beta$  (Tang *et al.*, 2006, 2007 ; Bretschneider *et al.*, 2013).

La KS est lié de manière covalente au domaine AT *via* une région « linker » KS-AT d'environ 100 acides aminés qui possède un repliement de type ferrédoxine (**Figure 19**). Les structures du fragment de module KS-AT des modules 3 et 5 de la PKS DEBS (Tang *et al.*, 2006, 2007) ont révélé que cette région est similaire aux FAS (Maier *et al.*, 2008) et aux PKS modulaires de type *trans*-AT. Ces derniers n'intègrent pourtant pas de domaine AT au sein de la chaîne polypeptidique, mais possèdent néanmoins ce motif structural (Lohman *et al.*, 2015) connu pour servir de région favorisant l'interaction avec les domaines *trans*-AT.



**Figure 19 : Structure cristallographique du domaine KS provenant du module 5 de la PKS DEBS (PDB : 2HG4 ; Tang *et al.*, 2006).** Les monomères du domaine KS sont représentés en bleu et en vert pâle. Le linker KS-AT est représenté en blanc. Les résidus impliqués dans la catalyse sont indiqués sur la structure.

Récemment, Keatinge-Clay *et coll.* (Gay *et al.*, 2016) ont révélé l'existence d'un motif qui serait impliqué dans des interactions de type ionique et hydrophobe entre les domaines KS de plusieurs PKS modulaires de type *trans*-AT. Ces motifs appelés LINKS (*Laterally-Interacting Ketosynthase Sequence*) ont été trouvés au sein du *linker* KS-AT et seraient à l'origine d'interactions latérales entre les sous-unités d'un même système PKS modulaire de type *trans*-AT pour former un mégacomplexe (**Figure 20A et 20B**). Les conséquences de ce type d'interactions sur l'organisation macromoléculaire des PKS modulaires de type *trans*-AT seront détaillées dans la section V de cet état de l'art.

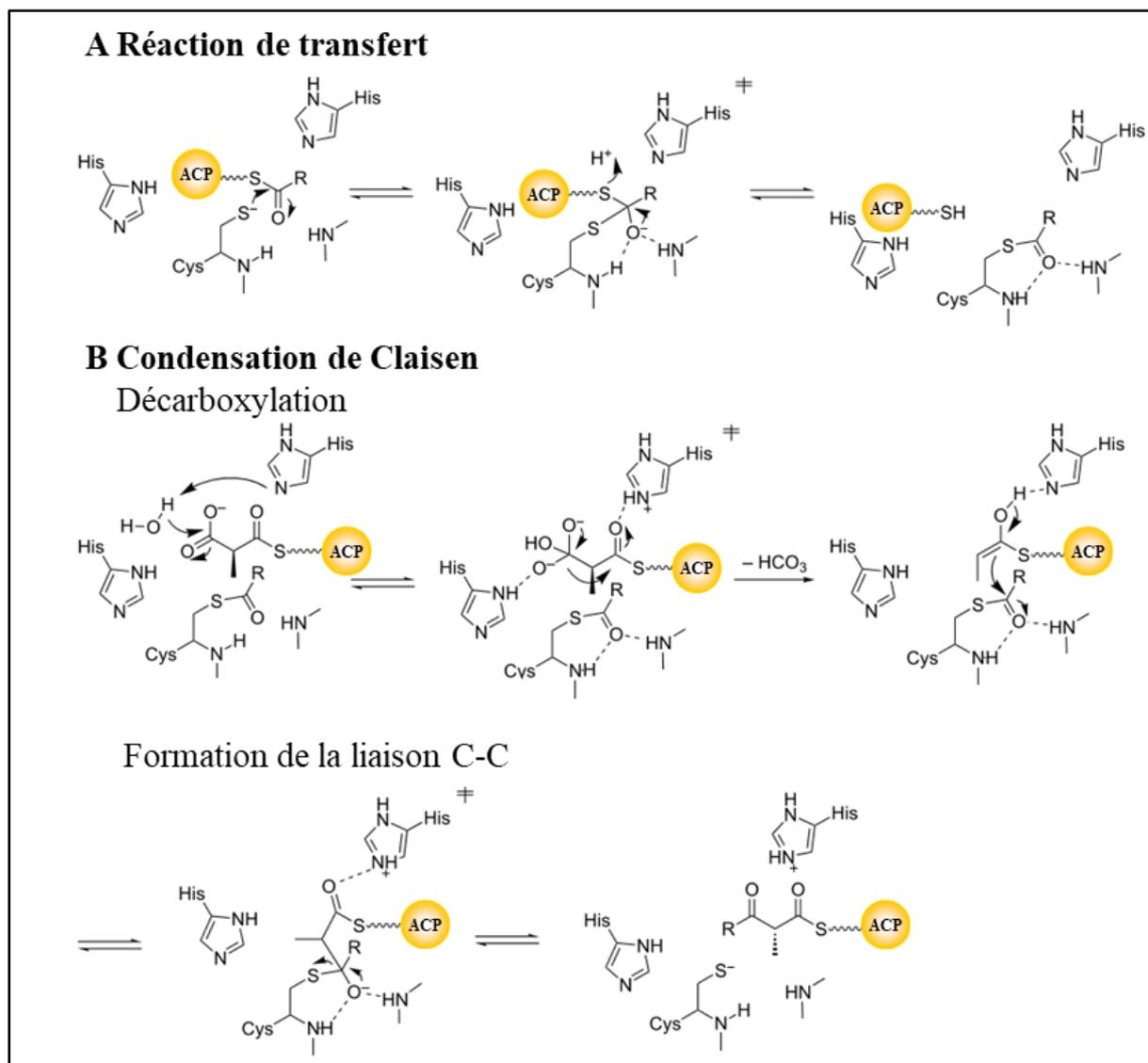


**Figure 20 : Représentation des domaines LINKS provenant d'un domaine KS représentatif chez une PKS de type *trans*-AT (Gay *et al.*, 2016). (A) Modèle d'un domaine KS possédant le motif LINKS. La couleur bleue représente les séquences conservées, tandis que les couleurs vers le rouge sont les séquences les moins conservées. (B) Représentation schématique du mode d'interaction des motifs LINKS. Les KS homodimériques effectueraient ainsi des interactions latérales, permettant un assemblage vertical des PKS.**

Les réactions de translocation et d'extension peuvent être vu comme deux demi-réactions comprenant un mécanisme de type ping-pong (Robbins *et al.*, 2016). Le mécanisme de translocation de la chaîne du polycétide du domaine ACP du module précédent au niveau du domaine KS est illustré dans la **Figure 21A**. Cette étape implique l'attaque du groupement thiolate de la cystéine catalytique sur le carbone de la fonction thioester liant la chaîne polycétidique au groupement phosphopantéthéine du domaine ACP. Cette étape nécessite que la cystéine catalytique soit préalablement activée, ce qui est rendu possible par la présence du macrodipôle d'une hélice  $\alpha$  à proximité qui permet de diminuer le  $pK_a$  local. La charge négative portée par l'intermédiaire réactionnel est stabilisée par le trou de l'oxyanion formé par la chaîne polypeptidique ce qui provoque le départ du groupement phosphopantéthéine du domaine ACP, générant l'intermédiaire de polycétide lié de manière covalente au domaine KS.

Le mécanisme de condensation de la chaîne du polycétide et de l'unité d'extension se déroule en deux étapes et est illustré dans la **Figure 21B**. Ce mécanisme est extrapolé à partir du domaine KS présent chez la FAS animale (Witkowski *et al.*, 2002). La formation de l'état de transition génère un hydroxyle à partir d'une molécule d'eau, ce qui favorise l'attaque nucléophile au niveau du groupement carboxylate de l'unité d'extension portée par l'*holo*-ACP. L'intermédiaire réactionnel est ensuite stabilisé par formation d'un énol, ce qui implique une deuxième histidine qui agit comme catalyseur acide. La décarboxylation de l'unité d'extension

permet la création d'un carbone nucléophile qui réagit alors avec le carbone électrophile de la fonction thioester de la chaîne du polycétide en élévation liée à la cystéine du domaine KS, permettant la régénération de cette dernière. Le polycétide est ainsi allongé d'une unité acétyle, soit deux carbones (Witkowski *et al.*, 2002).



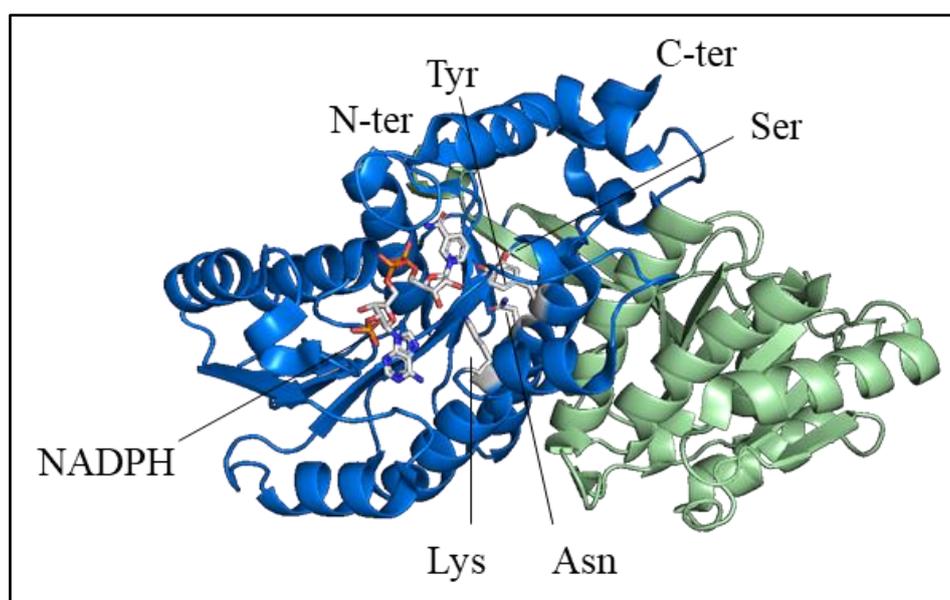
**Figure 21 : Mécanisme catalytique du domaine KS (adapté de Weissman, 2017).** (A). Réaction du transfert du polycétide du domaine ACP vers le domaine KS. (B). Réaction de condensation décarboxylative, aussi appelée réaction de Claisen, suivie de la formation de la nouvelle liaison carbone-carbone (C-C) entre l'intermédiaire polycétidique en cours de biosynthèse et l'unité d'extension.

L'étape de condensation est accompagnée d'une inversion de la configuration du groupe méthyle en position C2 (**Figure 14**). En effet, la configuration de la fonction énolate responsable de l'attaque nucléophile est plane. La stéréochimie finale de la liaison dépend donc

du côté de l'énolate qui est utilisé pour former la nouvelle liaison carbone-carbone. Dans ce cas, la direction de l'attaque nucléophile au sein d'un domaine KS entraîne une inversion de la stéréochimie du groupe méthyle en position C2 (Wiesmann *et al.*, 1995 ; Weissman *et al.*, 1997 ; Weissman, 2017).

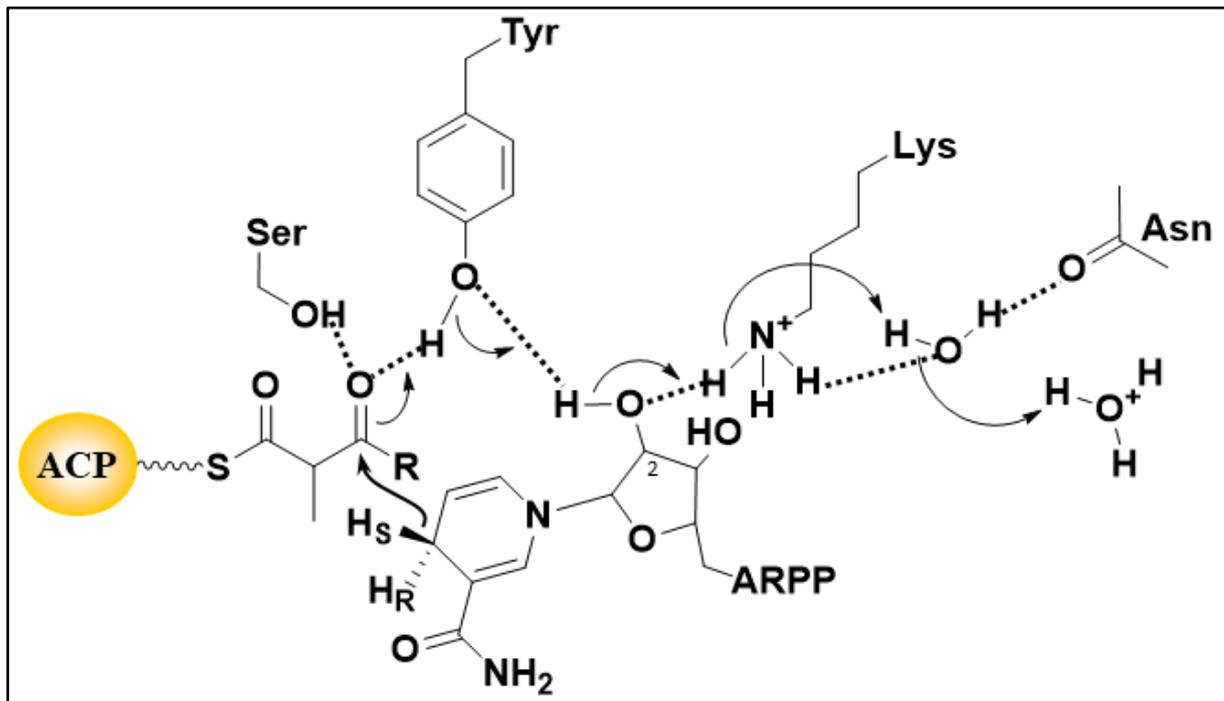
#### D. Le domaine cétoréductase (KR)

Le domaine KR est responsable de la réduction de la fonction cétone en position C3 en alcool secondaire (McPherson *et al.*, 1998) et peut être responsable de l'épimérisation du groupement méthyle en position C2 (Keatinge-Clay, 2007). Ce domaine catalytique optionnel est monomérique et est composé d'environ 450 acides aminés (Zheng *et al.*, 2010). Le domaine KR est divisé en deux sous-domaines, un sous-domaine catalytique d'environ 240 acides aminés, et un sous-domaine non-catalytique d'environ 200 acides aminés, tous deux structurés avec le même repliement de type Rossmann (**Figure 22**) (Zheng *et al.*, 2010). Plus précisément, le sous-domaine catalytique contient une tétrade catalytique caractéristique de la famille des déshydrogénases/réductases à courte chaîne (de l'anglais *Short-chain dehydrogenase/reductase* (SDR)), et est capable de fixer le cofacteur NADPH (Reid *et al.*, 2003 ; Zheng *et al.*, 2010).



**Figure 22 : Structure cristallographique du domaine KR provenant du module 2 de la PKS responsable de la synthèse de l'amphotéricine (PDB : 3MJC ; Zheng *et al.*, 2010). Le sous-domaine catalytique est représenté en bleu. Il est responsable de la fixation du cofacteur NADPH/NADP<sup>+</sup> qui est représenté en bâton. Le sous-domaine structural est représenté en vert pâle. Les quatre résidus impliqués dans la catalyse sont indiqués.**

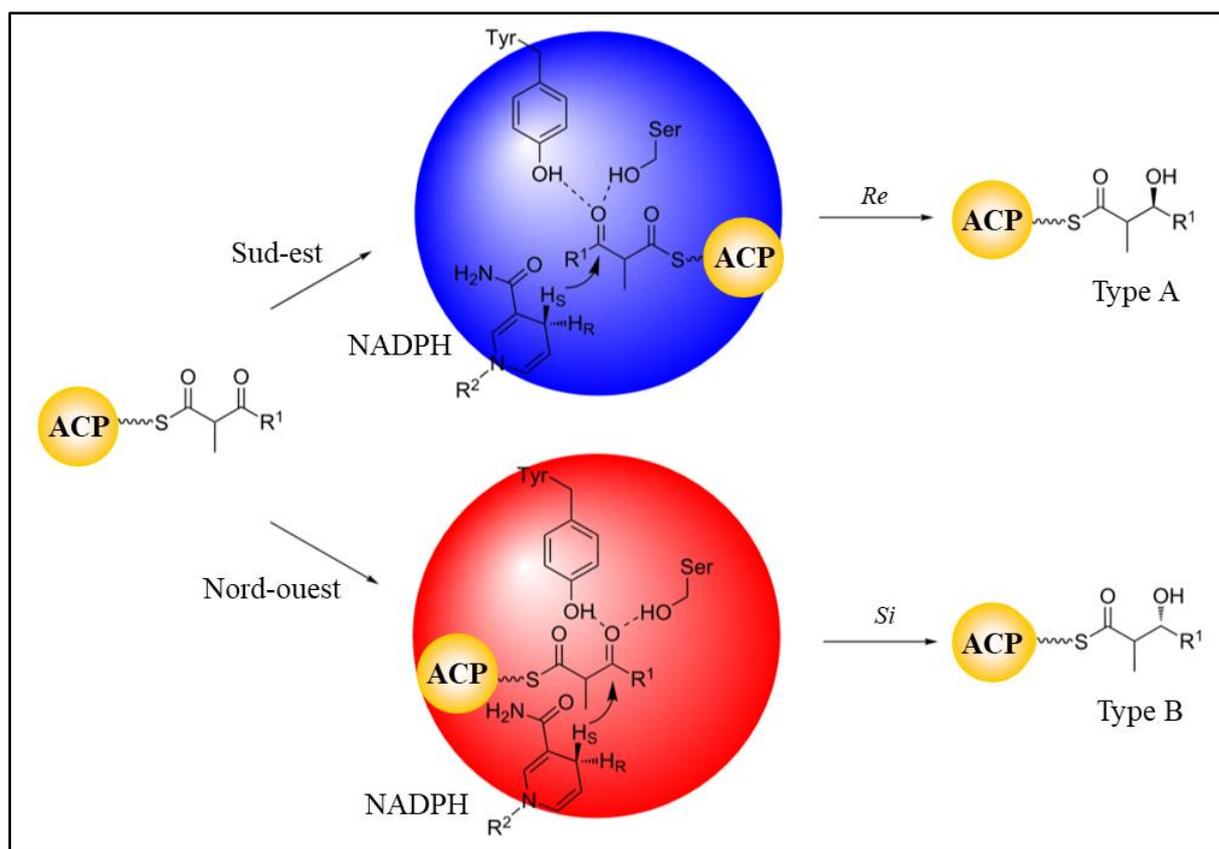
Le mécanisme catalytique illustré dans la **Figure 23** implique le transfert d'un hydrure 4-pro-*S* du cofacteur NADPH au niveau du carbone C3 portant la fonction cétone (McPherson *et al.*, 1998). La réaction enzymatique est favorisée par l'augmentation du caractère électrophile de la fonction cétone car l'état de transition dans lequel il est attaqué est stabilisé par deux stratégies : la formation d'une liaison hydrogène par la sérine 144 et la catalyse acide générale par la tyrosine 157, qui réduisent toutes deux la charge négative sur l'oxygène. Ce dernier acide aminé effectue une catalyse acide au niveau de la fonction cétone pour stabiliser l'oxyanion formé, générant alors un groupement hydroxyle. La Tyr catalytique est finalement stabilisée grâce à un relais de proton qui implique une lysine, une glutamine, l'hydroxyle du ribose et quatre molécules d'eau (Reid *et al.*, 2003).



**Figure 23 : Mécanisme catalytique du domaine KR (d'après Tsai and Ames, 2009).** La réaction de réduction de la fonction cétone au niveau du carbone C3 du polycétide en cours de biosynthèse implique une tétrade catalytique et le cofacteur NADPH. La tyrosine catalytique joue le rôle de catalyseur acide et cède un proton pour stabiliser l'oxyanion généré par la réduction. Cette tyrosine sera ensuite régénérée par un relai de proton impliquant une lysine, une glutamine, l'hydroxyle du ribose, ainsi que quatre molécules d'eau.

Les domaines KR catalysent une réduction stéréospécifique du groupement cétone au niveau du carbone C3, avec pour conséquence la génération d'un groupement hydroxyle pouvant adopter deux configurations possibles (**Figure 14**) (Siskos *et al.*, 2005). Le mécanisme

catalytique implique le cofacteur NADPH qui est toujours fixé dans la même orientation dans le site actif quelle que soit la direction finale de réduction. Par conséquent, la direction de réduction découle des deux modes d'entrée du substrat lié au groupement phosphopantéthéine du domaine ACP dans le site actif du domaine KR (**Figure 24**). Plusieurs motifs au sein des domaines KR ont été identifiés par alignement de séquences (Caffrey, 2003), et ont permis de les classer en deux grands types. En premier lieu, il existe le type A conduisant à la formation d'un hydroxyle de configuration *S* et en second lieu, le type B amenant à la configuration *R* lorsque les substituants en C2 sont prioritaires par rapport au C4. Il est à noter que la nomenclature de la stéréochimie est inversée lorsque les substituants en C4 sont prioritaires par rapport au C2. De surcroît, un troisième type de domaine KR existe, le type C qui ne présente pas d'activité réductase. La résolution de nombreuses structures de domaines KR a révélé que les motifs clés sont localisés à proximité du site actif (Keatinge-Clay, 2007 ; Zheng and Keatinge-Clay, 2011 ; Zheng *et al.*, 2010, 2012, 2013 ; Bonnett *et al.*, 2013 ; Piasecki *et al.*, 2014 ; Liu *et al.*, 2018). En effet, le motif LDD est caractéristique des domaines KR de type B et est trouvé au sein d'une boucle à proximité du site actif. Quant aux domaines KR de type A, ils ne possèdent pas le motif LDD, mais un acide aminé tryptophane conservé. Ces acides aminés se situent du côté opposé à tyrosine catalytique.



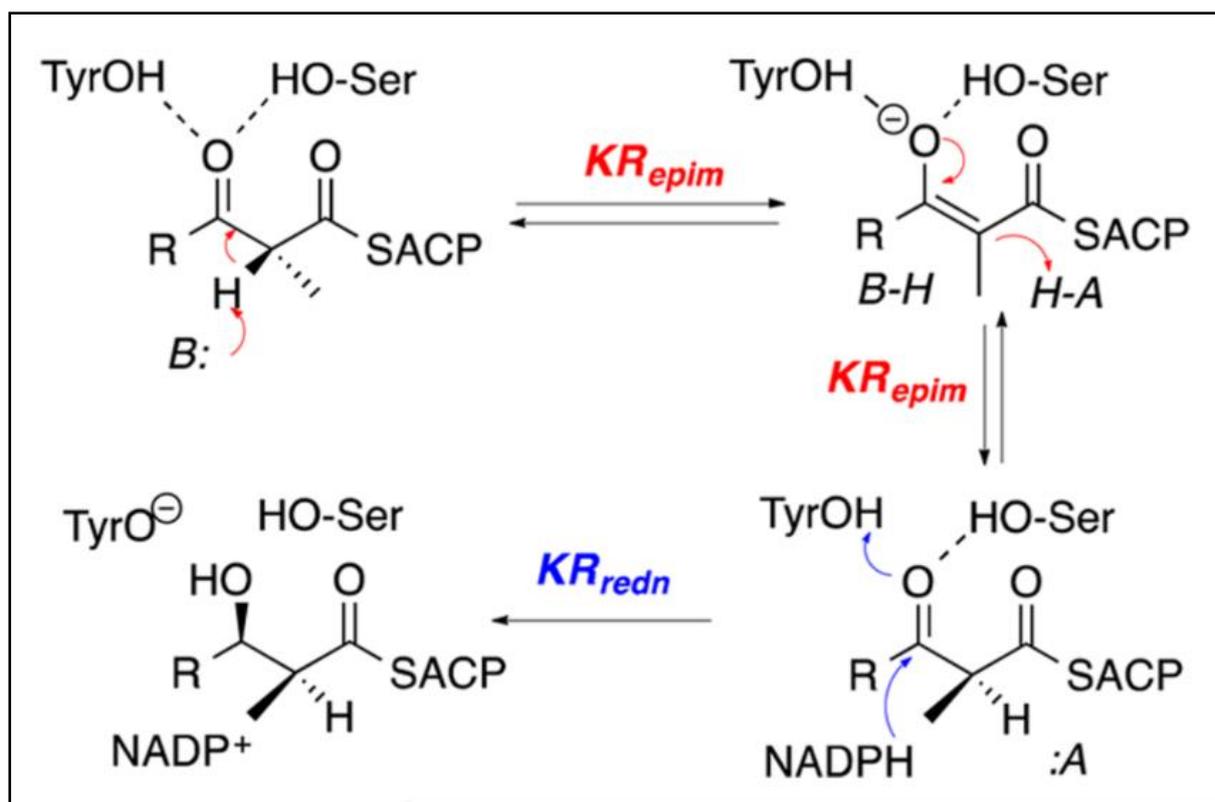
**Figure 24 : Modèle du contrôle de la stéréochimie lors de la réduction catalysée par le domaine KR (Adapté de Weissman, 2017).** Le positionnement de la tétrade catalytique et du cofacteur NADPH sont conservés au sein de tous les domaines KR. C'est donc la direction de l'entrée du substrat dans le site actif du domaine KR (cercles pleins en rouge et en bleu) qui détermine la direction de la réaction de réduction. L'entrée par la gauche génère une stéréochimie de type A, tandis que l'entrée par la droite génère une stéréochimie de type B. Dans les deux cas, le NADPH délivre son hydrure par le dessous du substrat.

Le rôle du tryptophane caractéristique des domaines KR de type A a été mis en lumière grâce à la structure à haute résolution du domaine KR du second module de la PKS responsable de la synthèse de l'amphotéricine co-cristallisé avec le cofacteur et le substrat (Liu *et al.*, 2018). L'entrée du substrat est favorisée par l'interaction entre le tryptophane conservé et le bras phosphopantéthéine. En effet, il avait été montré précédemment que la mutation de cet acide aminé entraîne une modification de la stéréospécificité (Baerga-Ortiz *et al.*, 2006 ; Zheng *et al.*, 2010). Cependant, cette interaction ne serait pas universelle pour tous les domaines KR de type A, puisque cet acide aminé est plus enfoui dans certains domaines KR de type A comme le domaine KR<sub>11</sub> de la PKS responsable de la synthèse de l'amphotéricine, ce qui indique qu'il n'y aurait pas d'interaction directe entre le bras phosphopantéthéine et le tryptophane. Par conséquent, d'autres bases moléculaires seraient impliquées dans la réduction (Bonnett *et al.*, 2013 ; Zheng *et al.*, 2013). Concernant l'implication des motifs caractéristiques des

domaines KR de type B, et se basant sur des études récentes (Dutta *et al.*, 2014 ; Kim *et al.*, 2014), Keatinge-Clay *et coll.* (Liu *et al.*, 2018) propose qu'un aspartate équivalent au troisième résidu du motif LDD caractéristique des domaines KR de type B aiderait à orienter le bras phosphopantéthéine en formant une liaison hydrogène avec le groupe carbonyle en C4.

Un modèle alternatif propose que la direction de la réduction soit contrôlée par un degré de structuration du site actif différent entre les domaines KR de type A et B (Bonnett *et al.*, 2013 ; Weissman, 2017). Pour les domaines KR de type A, la présence du cofacteur NADPH permet de maintenir un site actif catalytiquement compétent, dans lequel un acide aminé clé bloque l'entrée dans le site actif par le nord-ouest. Cette configuration contraindrait le substrat à entrer dans le site actif uniquement par le sud-est. Le tryptophane caractéristique des domaines KR de type A pourrait alors aider à positionner le bras phosphopantéthéine par l'intermédiaire d'une liaison hydrogène. En ce qui concerne les domaines KR de type B, le substrat est supposé entrer par les deux entrées du site actif. Cependant, seule l'entrée du substrat du côté nord-ouest génère un site actif catalytiquement compétent. Dans ce modèle, le motif LDD n'interagirait pas directement avec le substrat, mais favoriserait la configuration catalytiquement compétente du site actif.

Certains domaines KR possèdent en plus de leur activité réductase, une activité épimérase du groupement méthyl au niveau du carbone C2 en inversant sa configuration (Keatinge-Clay, 2007 ; Garg *et al.*, 2013, 2014). Ainsi, il faut différencier les domaines KR de type A1 et B1 sans activité épimérase et celles de type A2 et B2 avec activité épimérase (**Figure 14**). Le type C se divise également en deux sous-types, C1 et C2, dont le dernier est capable de catalyser une épimérisation (Keatinge-Clay, 2007). Comme montré **Figure 25**, le mécanisme catalytique impliquerait la tyrosine et la sérine catalysant la réduction (Xie *et al.*, 2016), mais le modèle proposé est encore incomplet car les catalyseurs acide et basique potentiellement impliqués ne sont pas encore identifiés. En effet, les études bio-informatiques n'ont pas permis d'identifier des motifs caractéristiques des domaines KR épimérisants (Keatinge-Clay, 2007 ; Kitsche and Kalesse, 2013), et donc d'autres bases moléculaires sont potentiellement impliquées. Un mécanisme hypothétique propose que le mode de fixation du substrat pour la réaction d'épimérisation ne soit pas le même que pour la réaction de réduction (Weissman, 2017).



**Figure 25 : Mécanisme catalytique du domaine KR (d'après Xie *et al.*, 2016).** La réaction d'épimérisation de la fonction cétone au niveau du carbone C2 du polycétide en cours de biosynthèse implique les deux acides aminés tyrosine et sérine de la tétrade catalytique de la réaction de réduction.

Néanmoins, même si les mécanismes catalytiques impliquant la double activité des domaines KR ne sont pas parfaitement élucidés, des expériences de biologie synthétique ayant pour but de modifier la stéréochimie en manipulant les domaines KR devraient permettre de modifier la stéréochimie d'un polycétide. Ces expériences ont confirmé que d'autres bases moléculaires sont impliquées dans l'activité des domaines KR (Kellenberger *et al.*, 2008 ; Annaval *et al.*, 2015 ; Eng *et al.*, 2016) et ces résultats ainsi que les conséquences sur la biologie synthétique des PKS modulaires seront détaillés dans la partie suivante.

## E. Les autres réactions de réduction

### a. Le domaine déshydratase (DH)

Le domaine DH est responsable de la déshydratation de l'alcool secondaire généré par le domaine KR en fonction alcène (Xie and Cane, 2018). Ce domaine catalytique optionnel est homodimérique chez les PKS de type *cis*-AT, et chaque monomère est composé d'environ 280 acides aminés avec un repliement de type double hot-dog  $\beta\beta\alpha\alpha\beta\beta$  (Keatinge-

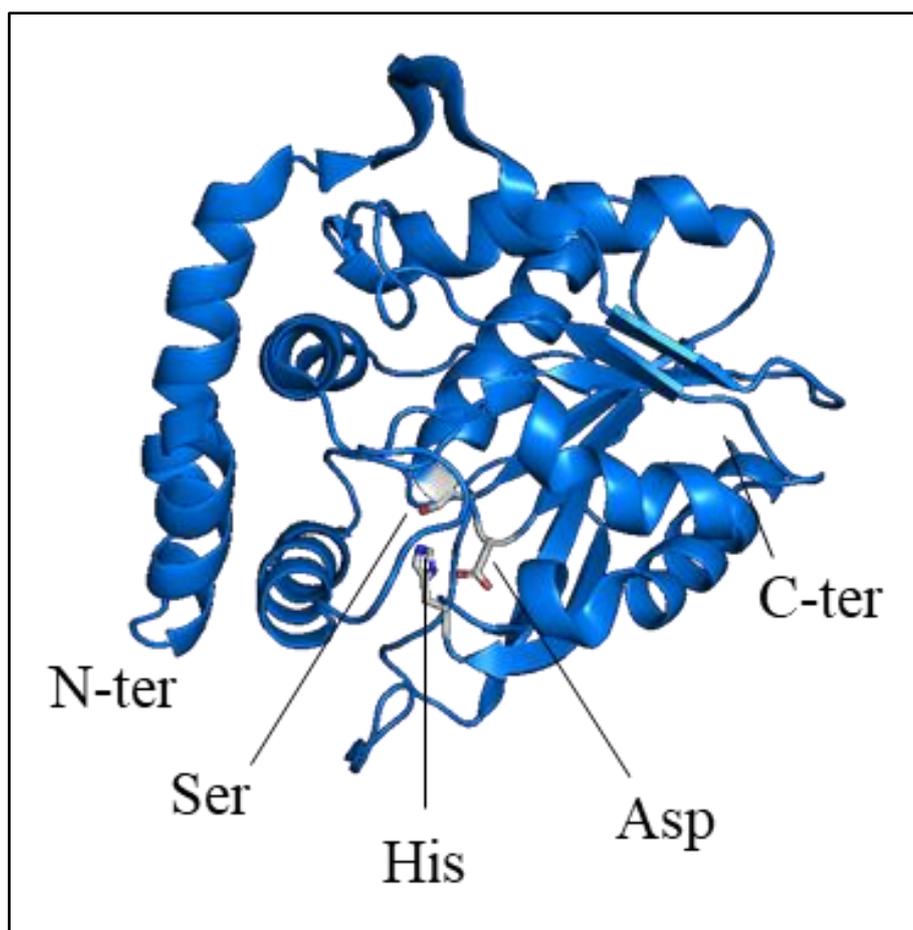
Clay, 2008 ; Barajas *et al.*, 2019). La déshydratation effectuée par le domaine DH permet d'engendrer deux stéréochimies possibles de la double liaison formée. La stéréochimie de cette double liaison serait déterminée par la direction de réduction au niveau du domaine KR, et donc par la configuration du groupement hydroxyle dans le site actif du domaine DH (Weissman, 2017).

### **b. Le domaine énoyl réductase (ER)**

Le domaine ER est responsable de la réduction de la fonction alcène générée par le domaine DH en fonction méthylène (Kwan *et al.*, 2008 ; Kwan and Leadlay, 2010). Ce domaine catalytique optionnel est monomérique chez les PKS de type *cis*-AT et composé d'environ 310 acides aminés (Zheng *et al.*, 2012). Il appartient à la famille des déshydrogénase/réductase à courte chaîne (MDR, de l'anglais *medium-chain dehydrogenase/reductase*), qui sont des enzymes à cofacteur NADP(H). Le domaine ER est divisé en deux sous-domaines : un responsable de la fixation du cofacteur NADPH composé d'environ 140 acides aminés et structuré avec un repliement de type Rossmann, et un autre responsable de la fixation du substrat composé d'environ 170 acides aminés (Ames *et al.*, 2012 ; Zheng *et al.*, 2012). Le domaine catalytique ER effectue une catalyse stéréospécifique. En effet, la stéréochimie de la réaction va être contrôlée par la nature du résidu impliqué dans la catalyse acide (Kwan *et al.*, 2008 ; Kwan and Leadlay, 2010 ; Weissman, 2017). La conséquence est la même que pour les domaines KR : le sens du transfert de l'hydrure du NADPH conditionne la stéréochimie finale, générant ainsi deux configurations possibles (Zhang *et al.*, 2018). À ce stade, la fonction cétone initiale ciblée par la boucle de réduction des PKS modulaires est entièrement réduite en fonction méthylène.

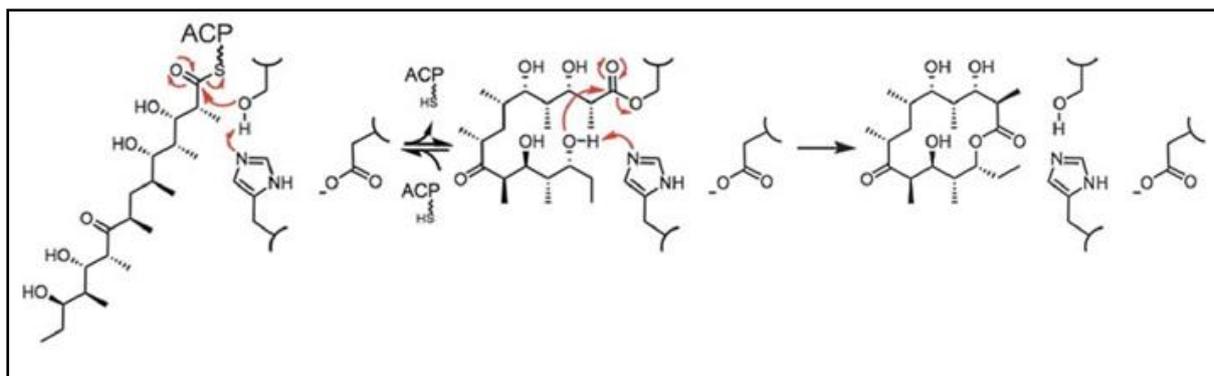
### **F. Le domaine thioestérase (TE)**

Le domaine TE est responsable de la libération de la chaîne du polycétide de la PKS en fin de synthèse. Ce domaine catalytique est homodimérique. Comme présenté en **Figure 26**, un monomère est composé d'environ 260 acides aminés avec un repliement de type  $\alpha/\beta$  hydrolase composé de huit feuillets  $\beta$  et de six hélices  $\alpha$ . La dimérisation s'effectue *via* les deux hélices  $\alpha$  en N-terminal (Tsai *et al.*, 2001, 2002 ; Scaglione *et al.*, 2010).



**Figure 26 : Structure cristallographique du domaine TE de la PKS DEBS (PDB : 1KEZ ; Tsai *et al.*, 2001).** Le domaine TE se dimérise par l'intermédiaire de deux hélices  $\alpha$  en N-terminal. Les résidus impliqués dans la catalyse sont indiqués.

Le mécanisme catalytique des domaines TE est présenté en **Figure 27**. La réaction de libération du polycétide de la PKS est similaire au mécanisme catalytique des domaines AT (Keatinge-Clay, 2012). D'abord, la sérine catalytique du domaine TE, activée préalablement au moyen d'une catalyse basique générale effectuée par une histidine, réalise une attaque nucléophile au niveau de l'acyl-ACP, résultant alors en un transfert de la chaîne du polycétide vers le site actif du domaine TE. Puis, le domaine TE catalyse la libération de la chaîne du polycétide grâce à un nucléophile activé à nouveau par une catalyse basique générale réalisée par l'histidine, qui peut être soit une molécule d'eau, soit un nucléophile interne à la chaîne du polycétide. Le produit ainsi généré peut être soit une macrolactone ou un macrolactame, soit un polycétide linéaire qui sera cyclisé par une enzyme de maturation. Il est à noter que les polycétides ne subissent pas nécessairement une étape de cyclisation et peuvent donc rester sous une forme linéaire.

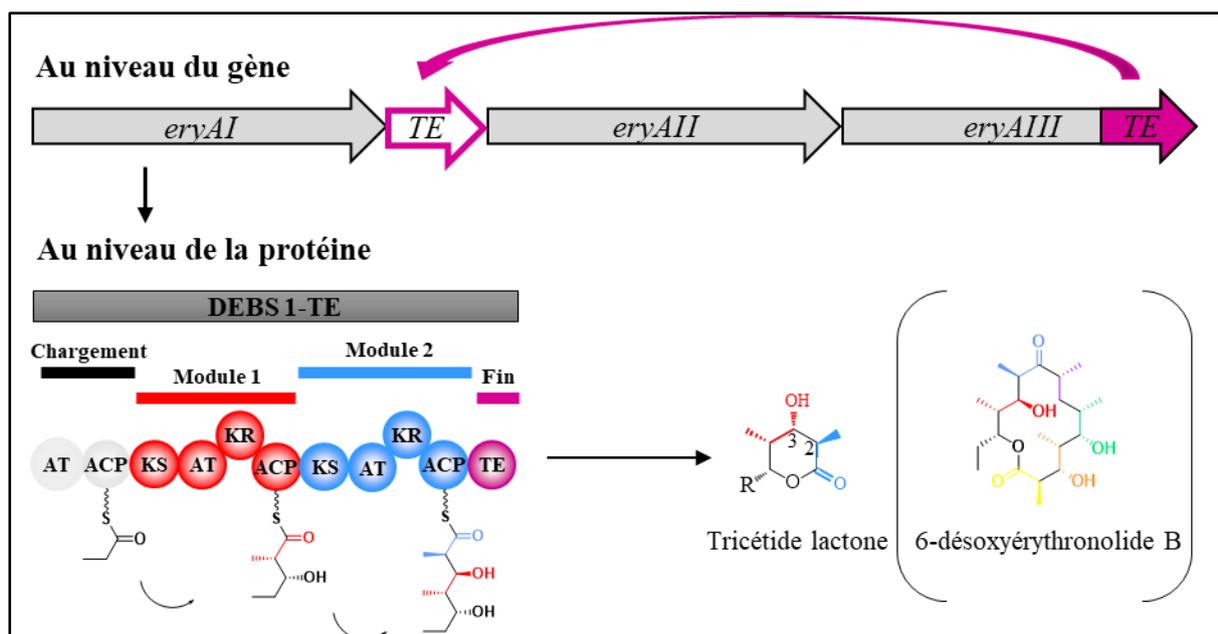


**Figure 27 : Mécanisme catalytique du domaine TE (Tsai *et al.*, 2002 ; Scaglione *et al.*, 2010).** La sérine catalytique, préalablement activée par l'histidine au moyen d'une catalyse basique générale, réalise une attaque nucléophile au niveau du groupement thioester formé par le polycétide lié au domaine ACP. Puis, l'histidine va à nouveau effectuer une catalyse basique générale pour activer un nucléophile qui peut être soit une molécule d'eau, soit un nucléophile interne à la chaîne du polycétide. Le nucléophile activé va alors réaliser une attaque nucléophile au niveau de l'intermédiaire acyl-enzyme, permettant ainsi la macrocyclisation du polycétide.

### 3. Les échanges de domaine pour modifier la stéréochimie

#### A. La construction d'un système biosynthétique modèle

La relocalisation du domaine TE à l'extrémité carboxy-terminale des domaines ACP est une stratégie permettant de libérer la chaîne du polycétide en cours de biosynthèse de manière prématurée. La relocalisation du domaine TE de la PKS DEBS (normalement présent à l'extrémité carboxy-terminale de la troisième sous-unité de la PKS DEBS, DEBS 3) à la fin de la première sous-unité, a permis de biosynthétiser et de libérer une tricétide lactone (**Figure 28**) (Kao *et al.*, 1994 ; Cortes *et al.*, 1995). Cette molécule incorpore les deux stéréochimies au niveau des groupes hydroxyle en C3 et méthyle en C2, au lieu de la 6-désoxyérythronolide B qui est une molécule plus complexe (Weissman, 2016b). Le rendement de production du tricétide lactone est de 10 à 15 mg.L<sup>-1</sup>, ce qui est un peu moins que le rendement de production de l'érythromycine A produit par la souche native (20 mg.L<sup>-1</sup>). Le DEBS 1-TE est donc un système modèle pour l'étude du stéréocontrôle *in vivo* et a été notamment exploité dans des expériences d'échange de domaine KR dans le but d'échanger la stéréochimie du polycétide (Kellenberger *et al.*, 2008 ; Annaval *et al.*, 2015).

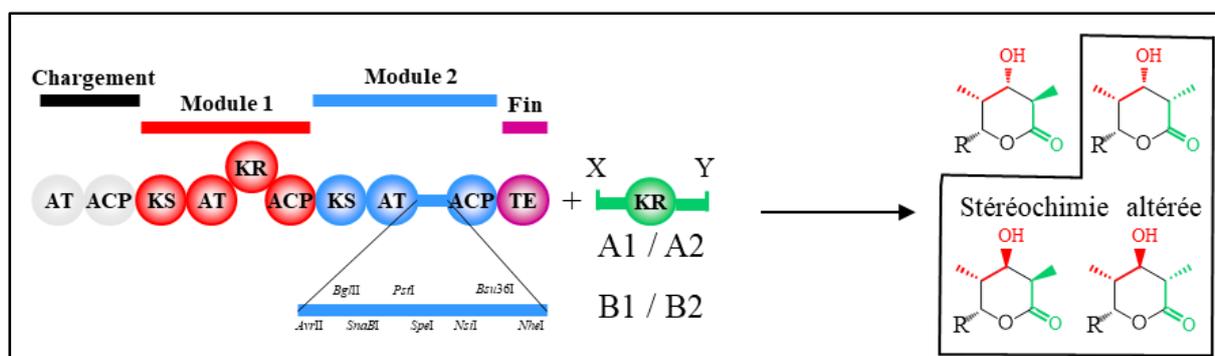


**Figure 28 : Création du système biosynthétique modèle DEBS 1-TE par ingénierie génétique (Adapté par KJ Weissman de Cortes *et al.*, 1995).** La région du gène *eryAIII* codant la thioestérase (en rose) a été relocalisée à la fin du gène *eryAI* codant le module de chargement et les deux premiers modules d'extension de la PKS de l'érythromycine A. La protéine résultante, DEBS 1-TE, produit une tricétide lactone au lieu de l'heptacétide 6-désoxyérythronolide B. La stéréochimie contrôlée par le domaine KR du module 1 et par le domaine KR du module 2 est suivie au niveau des carbones C4 et C5 et des carbones C2 et C3 de la lactone produite, respectivement.

## B. Les échanges de domaines KR

Une approche fructueuse pour manipuler la réaction catalysée par les domaines KR consiste à échanger ce domaine pour un autre de stéréospecificité opposée. Cette stratégie a apporté des résultats convaincants dès les premières tentatives de l'ingénierie des PKS modulaires (Bedford *et al.*, 1996 ; Kao *et al.*, 1998). Cette approche a ensuite été appelée « stratégie d'échange de la boucle réductrice », en faisant référence à une stratégie d'échange de domaine KR au sein du système biosynthétique DEBS 1-TE (Kellenberger *et al.*, 2008). Pour ce faire, la séquence codant le domaine KR<sub>2</sub> de type A1 (le produit est réduit sans être épimérisé) du deuxième module de DEBS 1-TE a été remplacée par une séquence polylinker contenant plusieurs sites de restriction, permettant une variation des sites de clonage des domaines KR introduits (**Figure 29**). Le choix des sites de fusion a été basé sur l'analyse des séquences des domaines KR et domaines adjacents, en identifiant des régions qui ne présentent pas d'homologie de séquence, et donc qui sont compatibles avec une région linker. Par la suite, les structures à haute résolution des domaines KR ont confirmé que ces sites se situent à l'extérieur des domaines conservés dans les régions connectant le domaine AT précédent et le

domaine ACP succédant le domaine KR dans le module. L'introduction de séquences codant une large gamme de domaines KR, provenant soit de la PKS DEBS soit d'autres systèmes PKS dans la région polylinker, s'est avérée efficace pour modifier de manière prédictible la stéréochimie de la cétoréduction sur une cible donnée (Kellenberger *et al.*, 2008).



**Figure 29 : Stratégie d'échange de la boucle réductrice (Adapté de Kellenberger *et al.*, 2008 ; Annaval *et al.*, 2015).**

Cette dernière approche a été étendue à un plus grand nombre de domaines KR épimérisants en utilisant les sites de clonage de la région polylinker qui ont permis d'obtenir les meilleurs rendements (Annaval *et al.*, 2015). Sur un total de huit domaines KR de type A2 testés (introduits dans le contexte modulaire contenant un domaine KR de type A1), seuls deux domaines (Amph KR<sub>1</sub> et KR<sub>11</sub>) ont permis d'obtenir la lactone attendue de manière quantifiable, mais aussi un mélange de lactone non-épimérisée et de cétolactone (lactone non-réduite) sous forme de traces. Ces premiers résultats ont montré que le contexte modulaire pourrait avoir une influence sur les activités du domaine KR. En outre, l'introduction des domaines KR de type B2, devant conduire à une cétoréduction dans le sens non-natif et à un changement de la stéréochimie du groupement méthyle, a abouti à des traces de produits attendus, voire aucun produit. Les domaines KR de type B2 ont donc montré une portabilité globalement faible dans un contexte modulaire de type A1. Une étude similaire mais exploitant un autre système biosynthétique (Eng *et al.*, 2016) a conduit aux mêmes conclusions.

L'ensemble de ces observations confirme que le mécanisme d'épimérisation ne dépend pas seulement du domaine KR, mais aussi du contexte modulaire dans lequel le domaine KR se trouve. En effet, il a été montré que le module 2 de la PKS DEBS est plus résistant à la protéolyse que le module 1 qui le précède qui est capable d'épimériser, ce qui implique que le module 2 de la DEBS adopte une architecture plus compacte (Hong *et al.*, 2005). Ainsi, l'accès

## *État de l'art*

d'une molécule d'eau agissant en qualité de catalyseur basique potentiel dans la réaction d'épimérisation peut être compromis. Ces résultats suggèrent que pour permettre une cétoréduction efficace couplée à un changement de la stéréochimie du groupement méthyle, les modules complets devraient être échangés entre les systèmes PKS au lieu des domaines KR. En effet, l'efficacité de la réaction de réduction et d'épimérisation dépendrait du contexte modulaire dans lequel il agit (Annaval *et al.*, 2015), et particulièrement des contacts spécifiques entre domaines d'un même module, pour acheminer le substrat dans les différents sites actifs de manière optimal. Il est alors possible que l'échange de domaine KR provoque des contacts KR/ACP non-favorables, et donc que cette interface protéique conduise à un mauvais positionnement du substrat dans le site actif du domaine KR (Dodge *et al.*, 2018 ; Moretto *et al.*, 2019).

## IV. Du contexte modulaire à l'ingénierie des modules

Le module est une unité fonctionnelle qui a pour rôle de recruter l'unité de départ, d'allonger le polycétide, ou de libérer le polycétide de la PKS. Ces fonctions sont déterminées par les domaines catalytiques qui composent le module. L'unité fonctionnelle regroupe alors plusieurs activités enzymatiques, qui doivent agir de manière ordonnée et spécifique, car une PKS modulaire de type *cis*-AT biosynthétise toujours le(s) même(s) produit(s). C'est le domaine ACP qui joue un rôle central dans un module, car il achemine le polycétide au sein des différents sites actifs des domaines catalytiques, permettant la communication intra-modulaire, et il achemine le polycétide d'un module à l'autre, donnant lieu à la communication inter-modulaire. Ces deux modes de communication vont définir le contexte modulaire, et cette notion sera précisée tout au long de cette section.

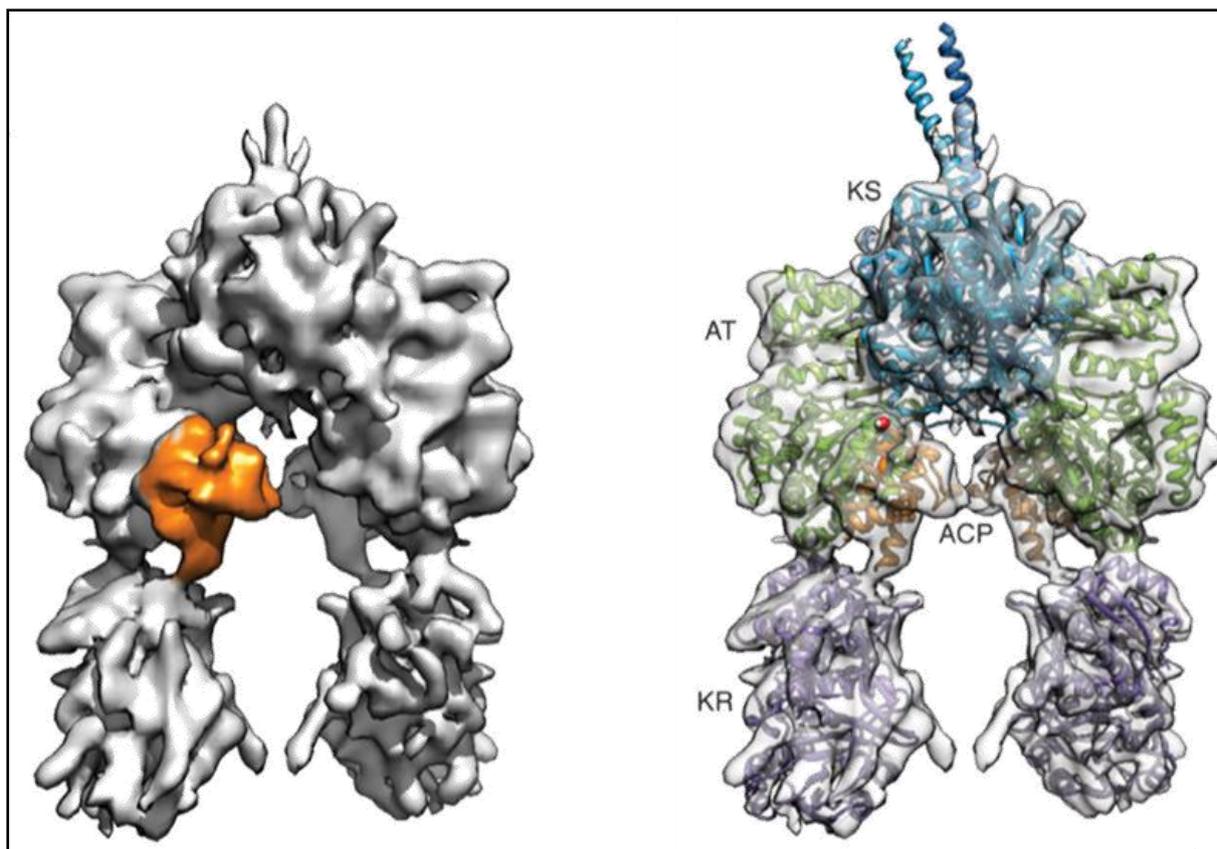
Les deux types de communication ont d'abord été étudiés avec des formes tronquées d'un module. Par exemple, il a été montré *in vitro* que l'hélice II du domaine ACP serait impliquée dans l'interaction avec la PPTase (Weissman *et al.*, 2006), la boucle I et l'hélice II interviendraient dans le cadre de l'extension du polycétide (Kapur *et al.*, 2010, 2012), l'hélice I serait impliquée dans le transfert inter-modulaire du polycétide (Kapur *et al.*, 2010, 2012) et les hélices II et III entreraient en contact avec le domaine AT (Miyanaga *et al.*, 2018). Cependant, ces interactions ne sont pas représentatives du contexte modulaire, car elles ne prennent pas en compte les contraintes physiques imposées par les liaisons covalentes entre les différents domaines d'un module. De surcroît, ces études ont été menées pour un nombre limité de domaines catalytiques, et ne sont donc pas généralisables à toutes les PKS modulaires de type *cis*-AT.

Les informations partielles et incomplètes des contacts protéiques entre les domaines d'une PKS modulaire montrent que la caractérisation structurale et fonctionnelle d'un module entier est nécessaire. Aujourd'hui, il n'existe pas de structure à haute résolution d'un module, mais des données à plus basse résolution sont disponibles pour la PKS responsable de la synthèse de la pikromycine (PIKS). En effet, des structures du cinquième module de la sous-unité PikAIII (PikAIII M<sub>5</sub>) ont été obtenues à une résolution de 7–11 Å par la technique de cryo-EM, et ce à plusieurs étapes du cycle catalytique (Dutta *et al.*, 2014 ; Whicher *et al.*, 2014). Le PikAIII M<sub>5</sub> est responsable du cinquième cycle d'extension de la pikromycine et agit au sein de la PKS PIKS avec un contexte modulaire précis. En ce qui concerne la communication intra-modulaire, le PikAIII M<sub>5</sub> est composé des domaines minimums KS, AT et ACP, et possède une boucle de réduction constituée du domaine optionnel KR. Pour la communication inter-

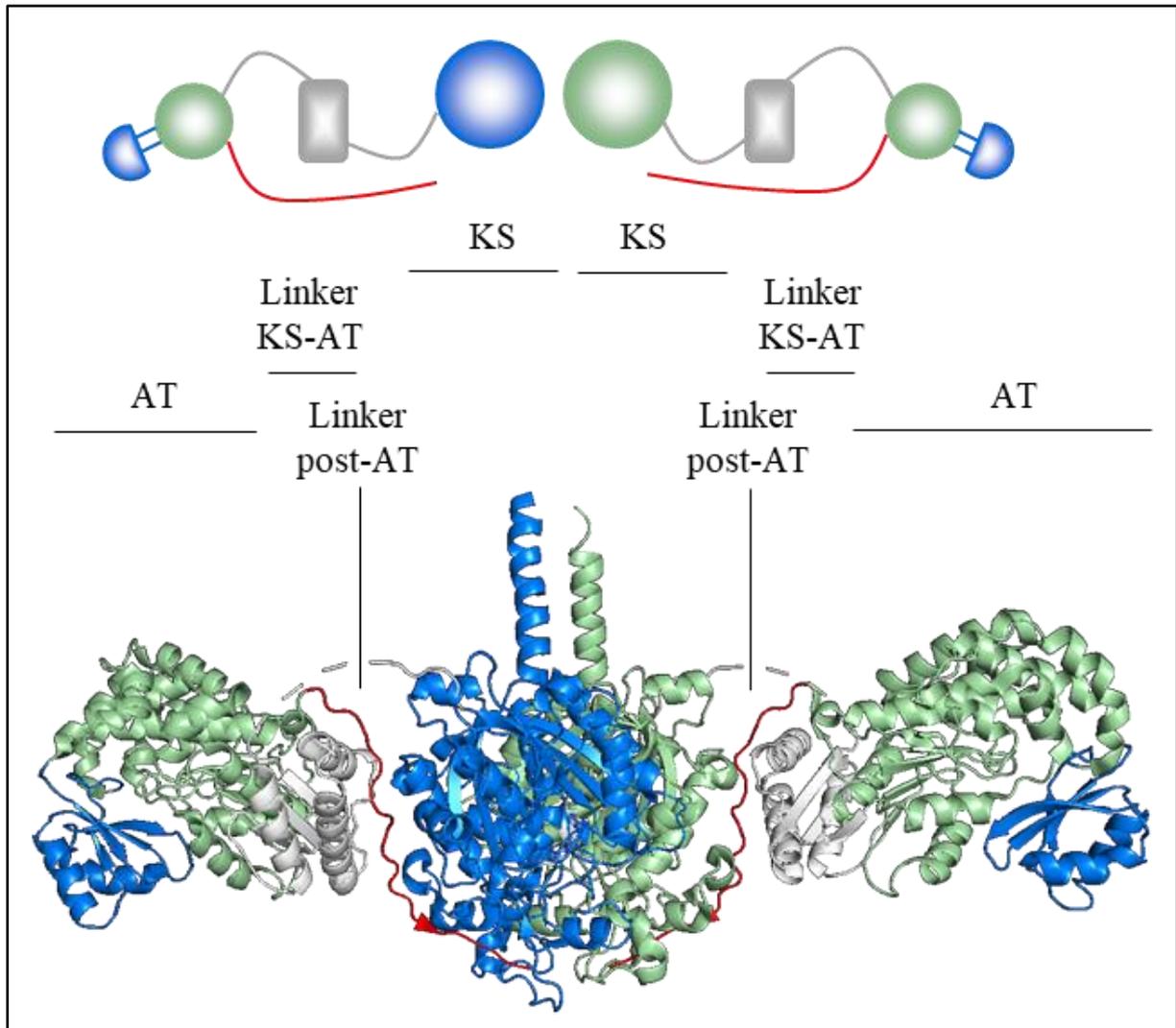
modulaire, le PikAIII M<sub>5</sub> est impliqué dans la translocation de la chaîne du polycétide provenant de la seconde sous-unité de PIKS (PikAII) *via* un domaine de docking, et aussi vers la quatrième sous-unité de PIKS (PikAIV) par l'intermédiaire d'un domaine de docking qui contient un élément de dimérisation post-ACP et une hélice  $\alpha$  de docking.

## **1. L'architecture d'un module**

À la résolution des modèles cryo-EM de PikAIII, les auteurs ont été en mesure d'introduire les structures à haute résolution des domaines individuels et du di-domaine KS-AT disponible de la PKS DEBS (**Figure 30**). En l'absence de substrat, le module s'organise en formant une arche contenant une chambre réactionnelle. La clé de voûte est formée par le domaine homodimérique KS qui forme, avec les domaines AT, les voussoirs de l'arc. Ensuite, les domaines KR forment les jambes de l'arche, tandis que le domaine ACP lié au groupement phosphopantéthéine adopte deux positions équiprobables dans le modèle : à proximité des domaines KR ou AT. Une différence fondamentale de ce modèle avec les structures à haute résolution introduites dans les modèles, est la configuration du di-domaine KS-AT (Tang, *et al.*, 2006 ; Tang, *et al.*, 2007). En effet les di-domaines des modules 3 et 5 de la PKS DEBS montrent une configuration plus étendue (**Figure 31**). Ces deux modèles sont en désaccord, mais il est à noter que les informations structurales apportées par les structures cristallographiques des di-domaines provenant de DEBS proviennent d'un fragment de module et ne représentent pas l'architecture d'un module entier. Néanmoins, la configuration en forme d'arche et la configuration étendue (Li *et al.*, 2018) sont toutes les deux catalytiquement compétentes, ce qui suggère que les modules de PKS de type *cis*-AT pourraient alors subir d'importants changements conformationnels, comme des interconversions entre une conformation en arche et étendue. Des données supplémentaires sont alors nécessaires pour déterminer lequel des deux modèles est correct, notamment par la caractérisation d'un module contenant des enzymes. Récemment, des études cinétique du module 1 de la DEBS ont montré un changement conformationnel du module lors de l'étape de cétoréduction (Cogan *et al.*, 2020). Pour cela, l'anticorps monoclonal anti-KR1 a été employé comme inhibiteur compétitif pour le site de fixation du NADPH. Le module 1 de la DEBS passerait alors d'une configuration en arche vers une configuration étendue lors de l'étape de cétoréduction. Cette observation est une preuve supplémentaire des changements conformationnels d'un module lors de l'étape d'extension, et cela implique une grande précision de la communication intra- et inter-modulaire.



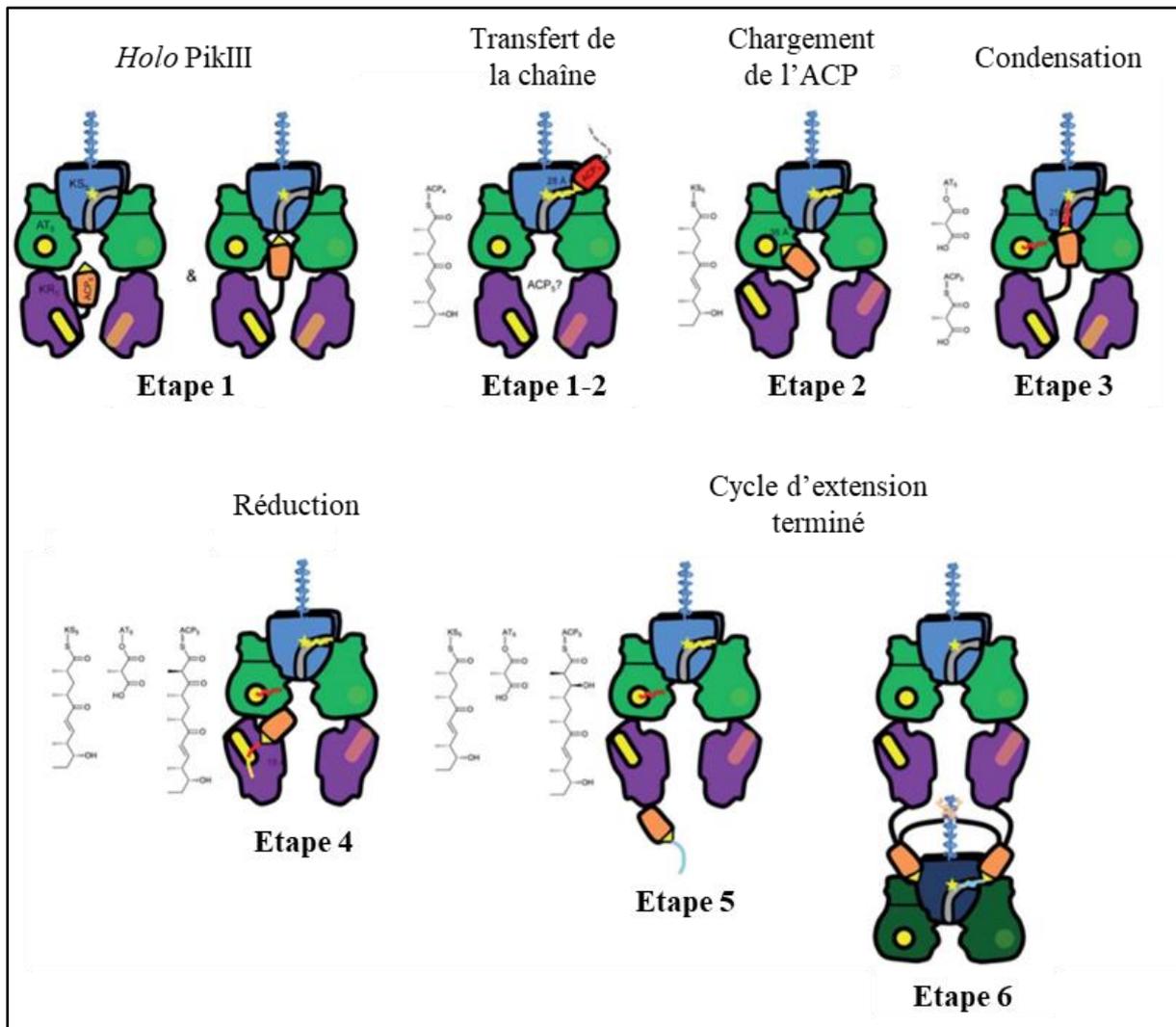
**Figure 30 : Structure cryo-EM de PikAIII en l'absence de substrat (adapté de Dutta *et al.*, 2014).** À gauche, la structure cryo-EM de PikAIII est représentée en solide. À droite, les structures à haute résolutions des domaines KS (en bleu), AT (vert), KR (violet), et ACP (en orange) de la PKS DEBS ont été positionnées au sein de la carte EM. Dans le conformère présenté, le domaine ACP se trouve en contact avec le domaine AT.



**Figure 31 : Structure cristallographique du di-domaine KS-AT provenant du module 5 de la PKS DEBS (PDB : 2HG4 ; Tang *et al.*, 2006).** Les domaines KS et AT présentés dans les **Figures 19** et **17** respectivement sont indiqués. Le linker post-AT est indiqué en rouge.

## 2. La communication intra- et inter-modulaire

Pour caractériser les différentes étapes du cycle catalytique au sein du PikAIII M<sub>5</sub> (**Figure 32**), les auteurs ont modifié le ou les domaines impliqués et analysé les changements conformationnels du module. Ces données renseignent sur les communications intra- et inter-modulaire au sein du PikAIII M<sub>5</sub>. En outre, mises en relation avec les données biophysiques disponibles pour des formes tronquées de module, ces données permettent de comprendre la trajectoire suivie par le domaine ACP lors de l'étape d'extension du polycétide, et de mettre en évidence les contacts essentiels avec les différents domaines catalytiques dans un contexte modulaire précis.



**Figure 32 : Représentation schématique des différentes conformations adoptées par PikAIII lors des différentes étapes du cycle d'extension du polycétide (Adapté de Whicher *et al.*, 2014).** Les domaines ACP<sub>4</sub> (en rouge), KS<sub>5</sub> (en bleu clair), AT<sub>5</sub> (en vert clair), KR<sub>5</sub> (en violet), ACP<sub>5</sub> (en orange), KS<sub>6</sub> (en bleu foncé), et AT<sub>6</sub> (en vert foncé) sont représentés. Sous forme *holo*, le domaine ACP<sub>5</sub> se situe au centre de l'arche formé par le module, soit à proximité du domaine KR<sub>5</sub> (à gauche), soit à proximité du domaine AT<sub>5</sub> (à droite) (**Etape 1**). L'acyl-ACP<sub>4</sub> entre en contact avec le domaine KS<sub>5</sub> pour le transfert inter-modulaire de la chaîne du polycétide (**Etape 1-2**). Il est à noter l'étape 1-2 a été obtenue en l'absence de l'ACP<sub>5</sub>. Lors de l'acylation de l'ACP<sub>5</sub>, ce dernier entre en contact avec le domaine AT<sub>5</sub> (**Etape 2**) puis à l'entrée du site actif du domaine KS<sub>5</sub>, une fois acylé avec une molécule de méthylmalonate (**Etape 3**). Après la formation de l'hexacétide, le domaine ACP<sub>5</sub> se place à l'entrée du site actif de la KR<sub>5</sub> (**Etape 4**). Une fois la réaction de réduction effectuée, le domaine ACP<sub>5</sub> se positionne hors de la chambre réactionnelle de manière à pouvoir transférer le polykétide au domaine KS<sub>6</sub> du module suivant (**Etapes 5 et 6**).

## A. La communication intra-modulaire

### a. Les contacts AT-ACP

L'étape de sélection de l'unité d'acyl-CoA par le domaine AT est suivie du transfert par acylation de l'unité vers le domaine *holo*-ACP, ce qui implique une interface engageant des

contacts spécifiques entre ces deux domaines (Dutta *et al.*, 2014 ; Whicher *et al.*, 2014). Il a été montré que pour les modules 3 et 6 de la PKS DEBS, les domaines AT<sub>3</sub> et AT<sub>6</sub> ont une affinité pour la réaction d'acylation dix fois supérieure pour le domaine ACP appartenant au même module (Wong *et al.*, 2010). Ces données indiquent que les domaines AT et ACP appartenant à un même module effectuent des contacts privilégiés (Dunn *et al.*, 2013). Dans le contexte de l'ingénierie des PKS, ces contacts devraient être soit maintenus, soit optimisés dans le cas d'une interface AT-ACP non-native.

### **b. Les contacts KS-ACP**

L'étape de condensation de la chaîne du polycétide avec l'unité d'extension implique que le domaine ACP chargé avec l'unité d'extension, vienne à proximité du domaine KS pour la réaction de condensation. Le modèle cryo-EM de PikAIII M<sub>5</sub> montre que le domaine ACP<sub>5</sub> subit un important changement de position dans la chambre réactionnelle, qui vient se placer à proximité du site actif de la KS<sub>5</sub>. L'entrée dans le site actif du domaine KS<sub>5</sub> de l'unité d'extension portée par le domaine ACP<sub>5</sub> montre ainsi un mode d'entrée différent pour le substrat porté par le domaine ACP<sub>4</sub> (Whicher *et al.*, 2014). Cela implique un mode d'interaction différent selon que le domaine KS interagit avec le domaine ACP du même module (ACP<sub>5</sub>) ou avec le module en amont (ACP<sub>4</sub>). Les contacts ACP<sub>n</sub>/KS<sub>n+1</sub> et leurs conséquences seront détaillés dans les parties suivants.

### **c. Les contacts KR-ACP**

L'étape de modification par le domaine optionnel KR implique que la chaîne du polycétide venant d'être étendue d'une unité, peut subir une réaction de modification au niveau de la fonction cétone en C3 et de la fonction méthyle en C2. Le modèle cryo-EM de PikAIII M<sub>5</sub> montre que le domaine ACP<sub>5</sub> entre en contact avec les éléments de structure secondaire portant les motifs clés caractéristiques des domaines KR de type A, positionnant la sérine catalytique à 19 Å du site actif (Dutta *et al.*, 2014). Des expériences mesurant l'efficacité catalytique des domaines KR de la PKS DEBS ont montré que la majorité de ces domaines KR ont une efficacité en moyenne dix fois supérieure pour le domaine ACP appartenant au même module (Ostrowski *et al.*, 2016). Mais dans un seul cas, le domaine KR du module 6 de la DEBS a montré une efficacité plus forte pour un domaine ACP appartenant à un autre module, celui du module 6 de la DEBS.

Une étude ultérieure a déterminé l'affinité entre deux paires de domaines KR et ACP provenant de la PKS responsable de la synthèse de la mycolactone (Moretto *et al.*, 2017). L'affinité pour le domaine ACP est de l'ordre de 3  $\mu\text{M}$  lorsque celui-ci est sous sa forme *holo*, et de 0,2  $\mu\text{M}$  sous sa forme acylée. Cette étude a montré que les domaines KR testés présentaient une affinité élevée pour le domaine ACP provenant du même module ou non, ce qui entre en contradiction avec l'étude de Cane et *coll* (Ostrowski *et al.*, 2016). La préférence du domaine KR pour la forme acyl-ACP suggère que la présence du substrat du domaine KR est essentielle pour l'établissement de l'interface KR/ACP. Cette différence d'affinité est cohérente avec les données structurales, car dans la structure du PikAIII M<sub>5</sub>, le domaine ACP entre en contact avec le domaine KR uniquement lorsque l'étape d'extension a eu lieu (Whicher *et al.*, 2014 ; Dodge *et al.*, 2018).

Nonobstant les incohérences de ces études, le domaine KR seul peut être échangé, mais il pourrait être plus efficace de conserver le domaine ACP provenant du même module. En outre, en relation avec la partie **III.3 : Les échanges de domaine pour modifier la stéréochimie** de la section précédente, ces études renforcent l'hypothèse selon laquelle les domaines KR et ACP d'un même module devrait être échangé ensemble pour modifier de manière efficace la stéréochimie d'un polycétide. En effet, s'il y a plusieurs modes d'interaction entre un domaine ACP et un certain type de domaine KR, alors l'établissement de contacts spécifiques lors de l'entrée du substrat dans le site actif du domaine KR serait essentiel pour la réaction d'épimérisation (Weissman, 2017). Par exemple, dans le cas de l'échange d'un domaine KR épimérisant pour un non-épimérisant, l'interface KR/ACP ne serait plus native, ce qui impliquerait alors un mode d'interaction différent, et donc le positionnement du substrat par un domaine ACP spécifique pour un domaine KR non-épimérisant, ne permettrait pas la réaction d'épimérisation.

#### **d. Les contacts TE-ACP**

Des études biophysiques ont montré une absence de contact entre le domaine ACP et le domaine TE. En effet, chez la PKS DEBS, le transfert de la chaîne du polycétide pour la réaction de libération de la PKS a lieu en l'absence de contacts entre le domaine ACP<sub>6</sub> et le domaine TE (Tran *et al.*, 2008, 2010). Ce domaine TE est actif même en présence d'un substrat non-naturel ou d'un domaine ACP hybride, et est souvent utilisé dans les expériences de biologie synthétique pour modifier la taille d'un polycétide (Cortes *et al.*, 1995 ; Menzella *et al.*, 2005 ; Pieper *et al.*, 1995 ; Chen *et al.*, 2000). La plupart de ces modules chimériques ont

été générés pour biosynthétiser des intermédiaires à un stade précoce de la biosynthèse. Cependant, la situation peut différer pour les intermédiaires d'un stade avancé de la biosynthèse, où le TE peut présenter une forte sélectivité pour son substrat naturel (Horsman *et al.*, 2016). Ce phénomène a été montré pour le domaine TE de la PKS PIKS, et ce domaine a été proposé comme un point de contrôle qualité de la biosynthèse de la pikromycine (Hansen *et al.*, 2017 ; Sherman *et al.*, 2020).

## **B. La communication inter-modulaire**

### **a. Les linkers inter-modulaires**

Une fois l'étape d'extension et les possibles étapes de modifications accomplies, le polycétide en extension est transféré du domaine ACP au domaine KS du module suivant. Les modules au sein d'une même sous-unité sont liés de manière covalente par une courte séquence appelée linker inter-modulaire, qui possède une taille d'environ vingt acides aminés chez les PKS de type *cis*-AT et qui est très peu conservée (Gokhale *et al.*, 1999). De surcroît, les linkers inter-modulaires confèrent de la flexibilité entre deux modules. Cette caractéristique a été observée dans la structure d'un bimodule entier obtenue à basse résolution (Edwards *et al.*, 2014). En effet, Edwards *et coll.* ont enregistré les données SAXS de la 3<sup>ème</sup> sous-unité de la PKS DEBS contenant les modules 5 et 6 et le domaine TE (DEBS 3). Le modèle obtenu montre que les modules 5 et 6 de DEBS 3 ont été acquis avec un angle de 70°, et donc que le linker inter-modulaire confère à DEBS 3 une grande flexibilité. Néanmoins, cette observation a besoin d'être confirmée par une structure à haute résolution d'un module, notamment car le linker KS-AT n'a pas pu être visualisé. De plus, cette observation reste à généraliser à d'autres systèmes multi-modulaires.

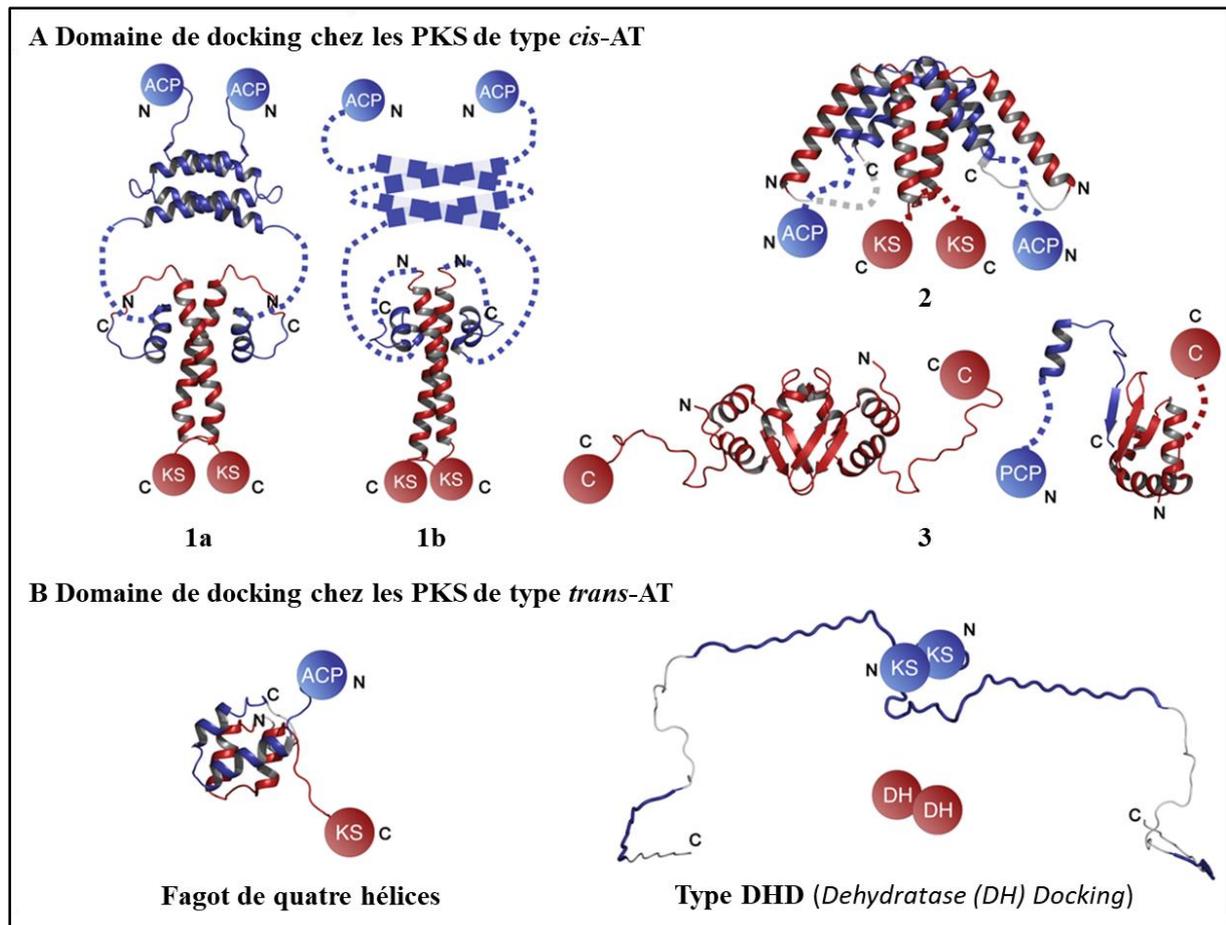
### **b. Les domaines de docking**

Lorsque les modules ne sont pas liés de manière covalente, ils s'associent par l'intermédiaire des domaines de docking (DD) (Broadhurst *et al.*, 2003). Ces derniers sont des courtes séquences protéiques pouvant posséder leur propre repliement structural, et se situant à l'extrémité carboxy-terminale (<sup>C</sup>DD) et amino-terminale (<sup>N</sup>DD) des sous-unités de PKS. Dans le cadre de la biologie synthétique, la nature de l'interaction entre deux modules par une liaison covalente ou non-covalente, est un facteur critique pour l'obtention d'une PKS chimère fonctionnelle. Cependant, l'un des problèmes fondamentaux lié à l'échange de modules qui doit

être encore résolu est le mécanisme par lequel l'intermédiaire de polycétide est transféré d'un module à l'autre.

Les structures à haute résolution des DD ont permis d'identifier quatre classes chez les PKS de type *cis*-AT (**Figure 33A**). Les DD respectent alors une orthogonalité qui est définie par la famille de DD. Pour les classes 1a/1b (Broadhurst *et al.*, 2003 ; Buchholz *et al.*, 2009) et 2 (Whicher *et al.*, 2014), l'interaction entre les domaines de docking s'effectue *via* des hélices  $\alpha$ . Pour la classe 3 (Richter *et al.*, 2008), l'interaction entre les deux domaines de docking s'effectue par le <sup>C</sup>DD possédant un motif linéaire court (SLiM pour *short linear motif*) qui vient s'insérer dans le brin  $\beta$  du <sup>N</sup>DD, formant alors un autre feuillet  $\beta$  comme une épingle à cheveux (Hacker *et al.*, 2018). Du fait de ce mode particulier d'interaction, cette dernière classe de DD a été renommée *SLiM- $\beta$ -hairpin docking* ( $\beta$ HD) *domain* (Kosol *et al.*, 2019). L'état d'oligomérisation entre les DD est différent, les <sup>N</sup>DD étant largement homodimériques, et les <sup>C</sup>DD étant monomériques. Il est à noter que les DD obéissent aussi à un code électrostatique, conférant alors un second degré de sélectivité pour les DD d'une même famille (Broadhurst *et al.*, 2003).

En plus des DD précédemment décrits, les PKS modulaires de type *trans*-AT possèdent des familles de DD qui ne sont pas présentes chez les PKS modulaires de type *cis*-AT (**Figure 33B**). La première classe de DD chez les PKS de type *trans*-AT comprend des DD sont deux monomères de <sup>C</sup>DD et de <sup>N</sup>DD, formés chacun de deux hélices  $\alpha$ , qui adopte une topologie de fagot de quatre hélices  $\alpha$  (Dorival *et al.*, 2016 ; Zeng *et al.*, 2016). L'autre classe de DD mise en évidence chez les PKS modulaires de type *trans*-AT est la classe DHD (de l'anglais *DeHydratase Docking*). Cette classe de DD est située à l'extrémité carboxy-terminale des domaines KS, et interagit avec l'extrémité amino-terminale d'un domaine DH (Jenner *et al.*, 2018). Cette dernière classe de DD est le premier exemple de l'interaction entre sous-unité de PKS *via* un DD et un domaine catalytique.



**Figure 33 : Les domaines de docking chez les PKS modulaires (Risser *et al.*, 2020).** (A) Les quatre classes de domaine de docking (DD) chez les PKS modulaires de type *cis*-AT sont représentées avec leur domaine adjacent. Les DD présents à l'extrémité carboxy-terminale des domaines ACP sont représentés en bleu, et les DD présents à l'extrémité amino-terminale des domaines KS sont représentés en rouge. Les DD de type 3 sont présents dans les systèmes hybrides PKS/NRPS sous forme dimérique (à gauche), et dans les systèmes purement NRPS sous forme monomérique (à droite). (B) Les classes de domaine de docking (DD) identifiées chez les PKS modulaires de type *trans*-AT.  
Légende : PCP = *peptidyl carrier protein* (chez les NRPS), C = domaine de condensation (chez les NRPS).

L'efficacité de la communication entre les modules de PKS est une question essentielle pour l'obtention d'une PKS fonctionnelle. Les DD et les modules interagissent entre eux par des interactions faibles, avec une constante de dissociation  $K_D$  dont les valeurs sont comprises entre 2 et 24  $\mu\text{M}$  (Whicher *et al.*, 2014 ; Buchholz *et al.*, 2009 ; Dorival *et al.*, 2016). Ces faibles affinités indiquent que les sous-unités de PKS interagissent de manière transitoire (Nooren and Thornton, 2003 ; Acuner Ozbabacan *et al.*, 2011).

Les DD se présentent comme des cibles de choix pour manipuler des modules ou des sous-unités de PKS, pour les faire interagir par des interactions non-covalentes. Cependant, les efforts d'ingénierie des PKS modulaires en exploitant les DD ont montré des résultats limités

(Klaus *et al.*, 2016 ; Weissman, 2016). En effet, d'autres facteurs sont parfois nécessaires, comme des interactions supplémentaires aux DD (Dodge *et al.*, 2018), pour une communication inter-modulaire efficace et seront présentés dans les parties suivantes.

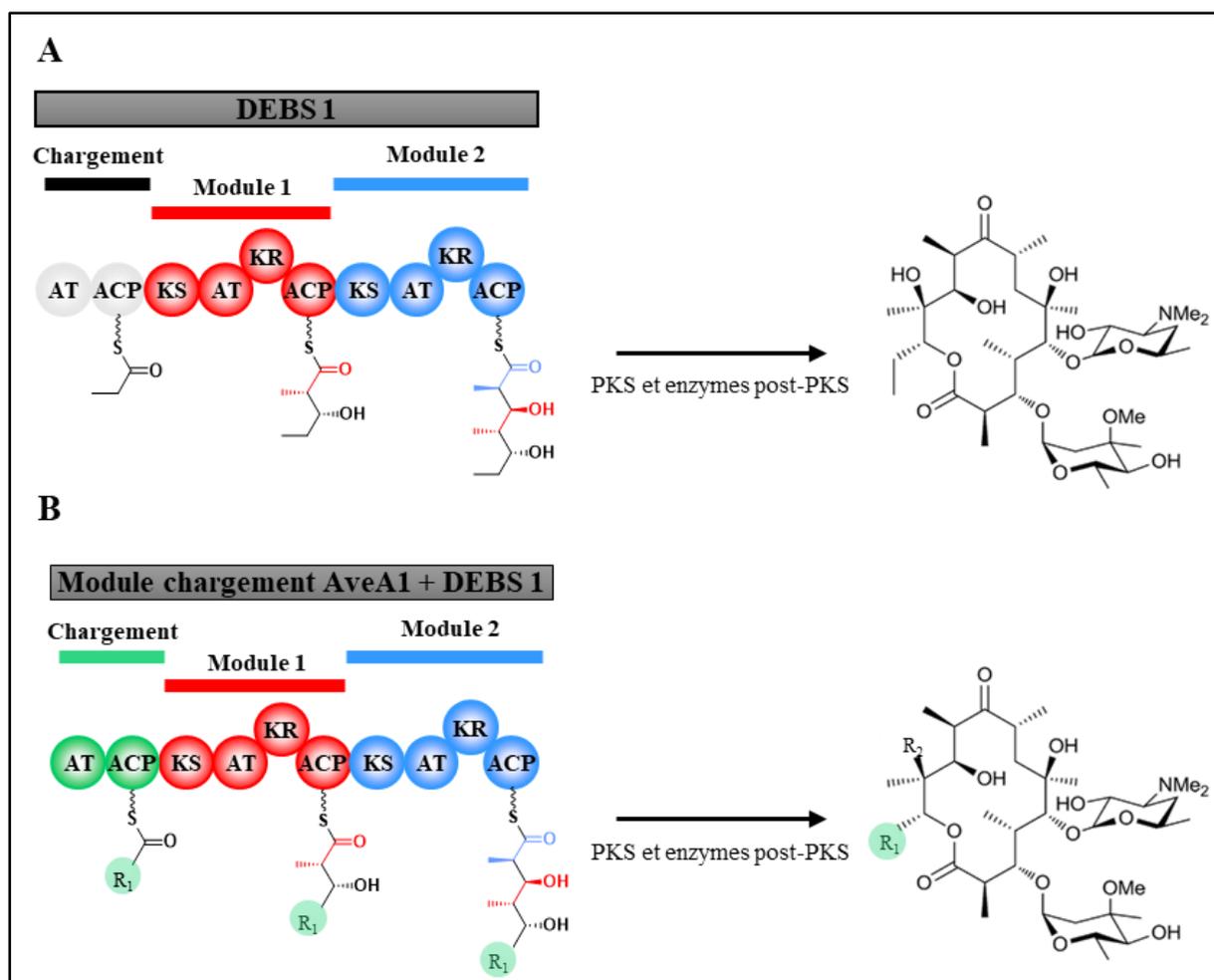
### **3. Les échanges de module**

Des expériences significatives d'échange de modules exploitant notamment la PKS DEBS comme modèle d'étude ont été conduites avec succès. Ces études, mises en relation avec les connaissances apportées par les parties précédentes, montrent l'importance de maintenir certains contacts protéine-protéine dans contexte modulaire natif dans des expériences d'échange de modules.

#### **A. Les échanges du module de chargement**

Une étude pionnière a consisté à échanger le module de chargement de la PKS DEBS avec le module de chargement de la PKS responsable de la synthèse de l'avermectine (AVES) provenant de *Streptomyces avermitilis* (**Figure 34A et 34B**) (Marsden *et al.*, 1998). Cette modification a permis de générer plusieurs dérivés de l'érythromycine, ayant incorporés comme unité de départ soit de l'isobutyrate, soit du 2-méthyl-butyrate, tous deux caractéristiques de l'avermectine.

Dans une autre expérience, des analogues des précurseurs de l'avermectine ont été utilisés pour générer des analogues de ce macrolide (Dutton *et al.*, 1991). Cette expérience a montré que le module de chargement de l'AVES peut accepter d'autres précurseurs que l'isobutyrate et le 2-méthyl-butyrate. La promiscuité de substrat du module de chargement de l'AVES a été exploitée par la suite pour incorporer une librairie d'acide gras à courtes chaînes chez la souche *Saccharopolyspora erythraea*, permettant de générer douze analogues de l'érythromycine, et présentant une activité antimicrobienne (Pacey *et al.*, 1998).



**Figure 34 : Ingénierie du module de chargement dans la PKS DEBS (Adapté de Marsden *et al.*, 1998).** (A) Le module de chargement de la PKS DEBS est spécifique au propionyl-CoA. (B) Le remplacement du module de chargement de la PKS DEBS par celui de la PKS responsable de la synthèse de l'avermectine (AVES) a entraîné une modification de la spécificité de l'initiation de la chaîne. Des analogues de l'érythromycine ont été générés, incorporant les unités de départ isobutyrate et 2-méthyle butyrate caractéristiques de l'avermectine. Ces modifications n'ont pas empêché certaines modifications par les enzymes post-PKS. R<sub>1</sub> = *i*-propyle ou *sec*-butyle, R<sub>2</sub> = H ou OH.

La promiscuité de substrat du module de chargement de l'AVES a aussi été exploitée dans le cadre de la mutasynthèse dans le but d'abolir la production de l'unité de départ utilisé pour la synthèse de l'avermectine par la souche *S. avermitilis* (Goudie *et al.*, 1993). Puis, différentes unités de départ ont été ajoutées dans le milieu de culture de la souche modifiée afin d'être sélectionnées par le module de chargement de l'AVES. Parmi les analogues de l'avermectine ainsi créés, l'un d'eux, la doramectine, ayant incorporé un acide carboxylique cyclohexane, a montré des propriétés antihelminthiques. Ce dérivé, aussi appelé Dectomax, est encore aujourd'hui utilisé comme traitement des infestations par des parasites pour les bovins, les ovins et les porcins.

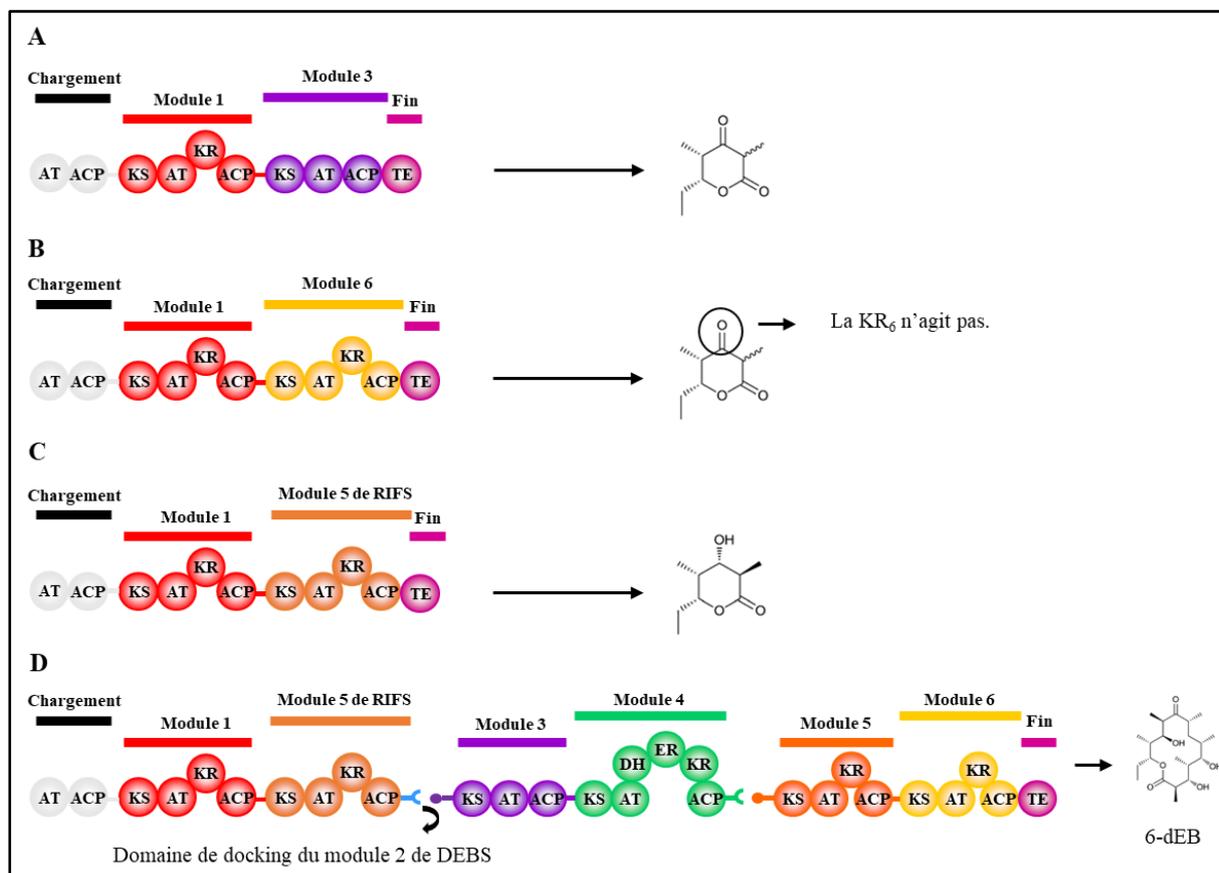
L'AVES a été ensuite modifiée dans le but d'échanger le module de chargement natif avec celui de la PKS responsable de la synthèse de la phospholactomycine qui recrute naturellement l'acide carboxylique cyclohexane comme unité de départ (Palaniappan *et al.*, 2003). La souche *S. avermitilis* ne produisant pas cette molécule de manière naturelle, la voie de biosynthèse a aussi été introduite (Goudie *et al.*, 1993). La souche obtenue produit la doramectine avec un rendement cinq fois supérieur à la souche sauvage supplémentée en acide carboxylique cyclohexane (53 mg.L<sup>-1</sup> contre 9 mg.L<sup>-1</sup>). Cependant, ce rendement est encore insatisfaisant au regard des attentes de la biologie synthétique des PKS modulaires (l'ivermectine est produit avec un rendement de 500 mg.L<sup>-1</sup>), et la PKS modifiée n'est pas entièrement efficace car le produit natif ivermectine avait aussi été obtenu. Toutefois, ces expériences de biologie synthétique démontrent qu'il est possible de générer *in vivo* des analogues de structure de polycétide avec un potentiel thérapeutique. En outre, l'utilisation du vivant au profit des méthodes chimiques permettront de diminuer drastiquement l'impact environnementale de la production des médicaments (Weissman, 2016).

## B. Les échanges d'un module d'extension

### a. Les stratégies non-rationnelles

A l'instar du module de chargement, de nombreuses expériences de biologie synthétique des PKS modulaires ont eu pour but d'échanger un module d'extension. Par exemple, différentes constructions bimodulaires ont été obtenues comportant le module 1 (y compris le module de chargement) et le module 3 ou 6 de la PKS DEBS (**Figure 35A et 35B**) ou le module 5 de la PKS synthétisant la rifamycine (RIFS) (**Figure 35C**). Dans ces constructions, le linker inter-modulaire naturellement présent entre les modules 1 et 2 de la PKS DEBS a été conservé (Gokhale *et al.*, 1999). La construction bimodulaire comportant le module 1 de la PKS DEBS et le module 5 de la PKS RIFS (KR de type A1) a permis d'obtenir la tricétide lactone prédite. Cependant, la construction bimodulaire comportant le module 1 de la PKS DEBS et le module 6 de la PKS DEBS (KR de type A1) a produit une lactone dont la fonction cétone du C3 n'était pas réduite. Cette absence de réduction reste inexplicée, car la même expérience réalisée *in vitro* a généré une lactone réduite (Gokhale *et al.*, 1999). Néanmoins, cette étude a montré pour la première fois que le maintien du linker inter-modulaire présent entre les modules 1 et 2 au sein du DEBS 1-TE pouvait être un paramètre important pour l'obtention du produit attendu. De surcroît, cette étude a été poursuivie au niveau de la PKS entière, en échangeant le module 2 de la PKS DEBS avec le module 5 de la PKS RIFS, et

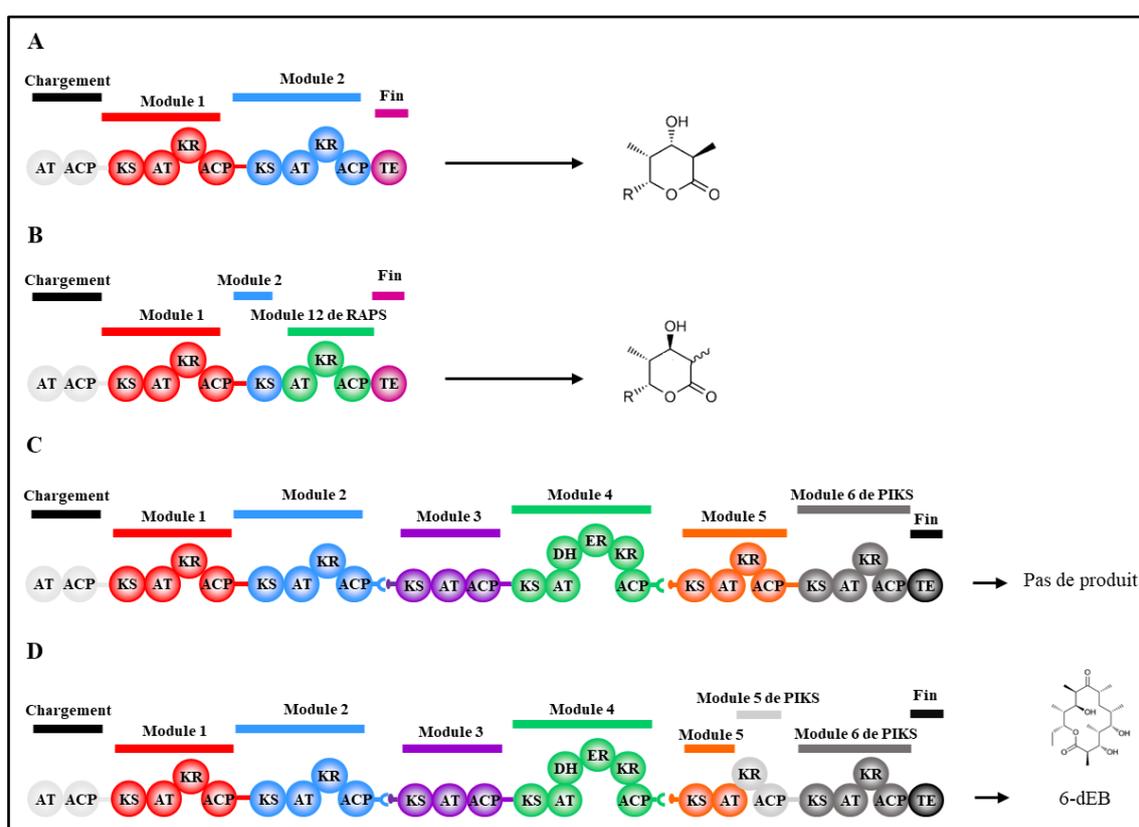
a permis d'obtenir la 6-dEB (**Figure 35D**). Bien que les DD n'étaient pas encore connus à l'époque de cette étude, la paire de DD native permettant l'interaction entre la première et la deuxième sous-unité de la DEBS avaient été conservés.



**Figure 35 : Approche d'ingénierie génétique pour l'échange de modules entier (Adapté de Gokhale *et al.*, 1999).** (A) et (B) Le « linker » intermodulaire natif présent entre les modules 1 et 2 de la PKS DEBS est conservé entre le module 1 (en rouge) de la PKS DEBS et les modules 3 (violet) et 6 (jaune) du même système. Le domaine KR du module 6 n'est pas capable de catalyser la réduction du groupement cétone en groupement hydroxyle. Le module de chargement est indiqué en gris. (C) et (D) La même stratégie a été appliquée pour le module 1 (en rouge) de la PKS DEBS et le module 5 (orange) de la PKS synthétisant la rifamycine, et a permis de produire la macrolactone prédite avec succès.

Une autre étude a montré que le remplacement du module 2 du DEBS 1-TE (**Figure 36A**) par le module 12 de la PKS synthétisant la rapamycine (RAPS) (**Figure 36B**) peut produire de simples analogues de statines, qui sont des molécules possédant des propriétés anticholestérolémiques (Ranganathan *et al.*, 1999). Plusieurs points de fusion ont été utilisés entre la PKS parentale DEBS et les modules de la RAPS, pour délimiter les bornes des modules à échanger. Cette étude a montré l'importance du maintien de l'interface ACP<sub>n</sub>/KS<sub>n+1</sub>, qu'il

appartienne à la même chaîne polypeptidique ou non. Des travaux similaires ont consisté à remplacer le module 6 de la PKS DEBS avec le module 6 de la PKS synthétisant la pikromycine, dans le contexte d'une chaîne d'assemblage entière de PKS (Tang *et al.*, 2000). La jonction inter-modulaire native entre le module 5 et le module 6 de la PKS DEBS a été conservée, comme suggéré par les travaux de Gokhale *et coll.* (Gokhale *et al.*, 1999). Cependant, aucun produit n'a été observé (**Figure 36C**). Lorsque le module 6 et les domaines KR et ACP du module 5 de la PKS synthétisant la pikromycine ont été échangés (**Figure 36D**), le produit désiré a ainsi été obtenu (avec un rendement d'environ 3 mg. L<sup>-1</sup>). Cette étude conforte l'hypothèse selon laquelle le contact entre l'interface ACP<sub>n</sub>/KS<sub>n+1</sub>, devrait être conservé.



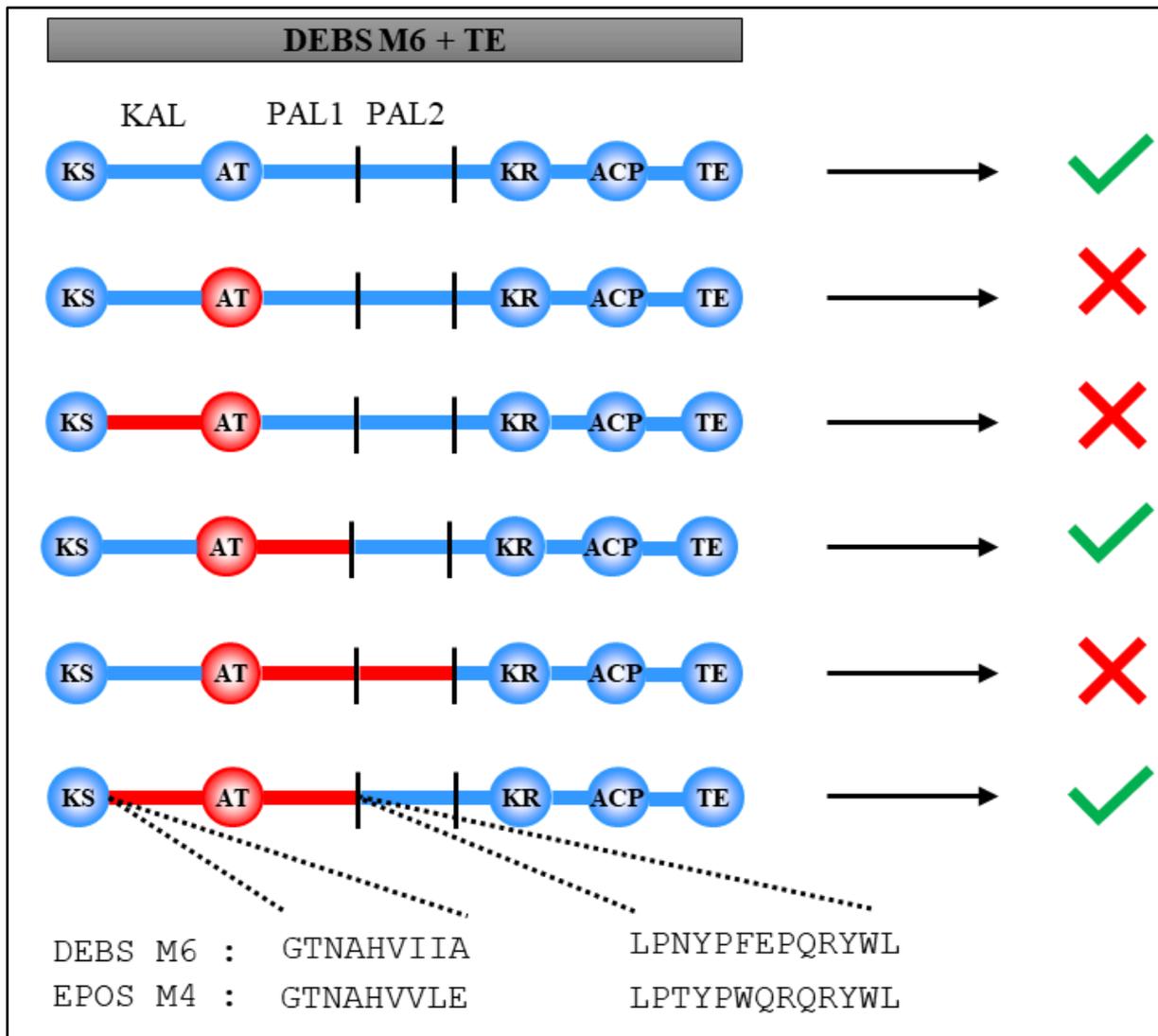
**Figure 36 : Approche d'ingénierie génétique pour l'échange de modules entier (Adapté de Ranganathan *et al.*, 1999 ; Tang *et al.*, 2000).** (A) Le système biosynthétique modèle DEBS 1-TE se compose du module de chargement (en gris), du module 1 (en rouge), du module 2 (en bleu), ainsi que de la thioestérase de la PKS DEBS. (B) Le système biosynthétique modèle DEBS 1-TE en (A) a été modifié en échangeant les domaines AT, KR et ACP du module 2 de la PKS DEBS par les domaines AT, KR et ACP du module 12 de la PKS synthétisant la rapamycine (vert). La lactone prédite a été obtenue (la stéréochimie du groupement méthyle au niveau du carbone C2 n'a pas été caractérisée). (C) La PKS DEBS a été modifiée en échangeant le module 6 par le module 6 de la PIKS. Cependant, le produit attendu n'a pas été obtenu. (D) La PKS DEBS a été modifiée en échangeant les domaines KR et ACP du module 5 et le module 6 par leur équivalent de la PKS PIKS. Le produit attendu a été cette fois-ci obtenu.

La possibilité de modifier la longueur de la chaîne de polycétides a aussi été testée en introduisant un module au sein d'une sous-unité de la PKS DEBS (Rowe *et al.*, 2001). Les modules 2 et 5 de la PKS RAPS ont été introduits entre les modules 1 et 2 du DEBS 1-TE, et ont permis d'obtenir le tétracétide attendu. Cependant, que ce soit dans l'introduction du module 2 ou du module 5 de la PKS RAPS, le produit majoritaire était le tricétide lactone, obtenu par la translocation directe du polycétide du module 1 au module 2 de la PKS DEBS. Les sites d'insertion des modules a ensuite fait l'objet de modifications, en préservant les interfaces  $ACP_n/KS_{n+1}$  natives, mais n'a pas permis d'améliorer la synthèse du tétracétide attendu (Rowe *et al.*, 2001). Cette modification a ensuite été effectuée au niveau de la PKS entière. Le macrolide attendu possédant l'extension de deux carbones a été obtenu, mais comme pour les études avec une PKS trimodulaire, le produit majoritaire a été, comme attendu, le macrolide natif (3 mg.L<sup>-1</sup> contre 10 mg.L<sup>-1</sup>). Ce résultat indique que le cycle d'extension réalisé par le module de la RAPS n'était pas réalisé de manière systématique et que le polycétide a été potentiellement transloqué du domaine  $ACP_n$  de la PKS DEBS au domaine  $ACP_{n+1}$  de la PKS RAPS. Par la suite, ce mécanisme de translocation direct a été élucidé en exploitant le système modèle DEBS 1-TE, montrant qu'au lieu d'être transloqué du domaine  $ACP_n$  de la PKS DEBS au domaine  $KS_{n+1}$  de la PKS RAPS, le polycétide est transloqué directement sur le domaine  $ACP_{n+1}$  de la PKS RAPS, ne permettant pas le cycle d'extension du polycétide par le module de la RAPS introduit (Thomas *et al.*, 2002).

Ces études pionnières ont montré qu'il est possible d'échanger un module en utilisant les bornes classiques d'un module  $KS-AT-(KR-DH-ER)-ACP$  pour obtenir le produit escompté (Gokhale *et al.*, 1999). Cependant, cette stratégie n'est pas universelle et peut conduire à un module non fonctionnel. L'autre stratégie qui consiste à échanger le module  $AT_n-(KR-DH-ER)_n-ACP_n-KS_{n+1}$  permet de maintenir le contact  $ACP_n/KS_{n+1}$  pour le transfert inter-modulaire du polycétide, et dans certain cas d'obtenir le produit attendu (Ranganathan *et al.*, 1999). En effet, plus récemment, ces bornes de module ont été exploitées pour générer des PKS chimères entre la PKS PIKS et la PKS responsable de la synthèse de vénèmycine, et ont permis d'obtenir les produits attendus, avec néanmoins de très faibles rendements par comparaison avec le composé parental (Miyazawa *et al.*, 2020). Grâce aux connaissances grandissantes des relations structures / fonctions des PKS modulaires, les stratégies d'échange de modules qui ont montrées des résultats significatifs ont pu être rationalisées avec les structures à haute résolution des domaines, et ont permis, dans certains cas, d'identifier des points de fusion naturels.

## b. La recherche de points de fusion naturels

L'échange de module permettant de maintenir les contacts  $ACP_n/KS_{n+1}$  nécessite l'identification des points de fusion entre le domaine KS et le domaine AT, et dans ce but, l'échange de domaine AT est une cible attractive. Sur la base des analyses de séquences ADN de différents domaines AT, Keasling *et coll.* (Yuzawa *et al.*, 2017) ont identifié des points de fusion pouvant être exploités dans des expériences d'échange de domaine. Ces points de fusions déterminés *in silico* ont été exploités dans des expériences d'échange de domaine AT en utilisant le sixième module de DEBS fusionné au domaine TE (**Figure 37**) et la première sous-unité de la PKS responsable de la synthèse de la lipomycine (LIPS). En accord avec les données structurales (Tang *et al.*, 2006, 2007), le point de fusion en amont du domaine AT est localisé dans le linker KS-AT, au niveau d'un motif GTNAH conservé qui avait déjà été identifié (Ranganathan *et al.*, 1999). Une étude ultérieure avait exploité le linker KS-AT au niveau de ce motif conservé pour supprimer deux modules de la PKS responsable de la synthèse de l'amphotéricine. La PKS chimère obtenue s'est montrée fonctionnelle, et produisant des rendements élevés du produit attendu (Carmody *et al.*, 2004). Le point de fusion en aval du domaine AT se situe dans la partie amino-terminale du linker post-AT, permettant alors de maintenir la partie conservée de ce linker, et ce point de fusion avait aussi été exploité dans des expériences d'échange de domaine AT (Oliynyk *et al.*, 1996 ; Petković *et al.*, 2008 ; Stassi *et al.*, 1998).



**Figure 37 : Approche d'ingénierie génétique pour l'échange de domaine AT (Adapté de (Yuzawa *et al.*, 2017)).** En bleu, le sixième module de la PKS DEBS. En rouge, le quatrième module de la PKS responsable de la synthèse de l'épothilone (EPOS). Les séquences conservées correspondant aux points de fusion qui ont permis de générer un système fonctionnel sont indiquées. KAL = linker KS-AT, PAL1 = partie amino-terminale non conservée du linker post-AT, PAL2 = partie amino-terminale conservée du linker post-AT.

Des études exploitant le point de fusion dans le linker KS-AT ont montré qu'il est possible de simuler l'évolution d'une PKS, et ainsi d'identifier les points de fusion des insertions ou délétions des gènes codant les modules. Pour cela, le groupe de Hertweck utilise comme système modèle deux systèmes PKS similaires capables de synthétiser des polypropionates. Les PKS responsables de la synthèse de l'auréothine (AURS) et de la noréothine (NORS) possèdent un premier module capable de catalyser plusieurs cycles d'extension (de 2 à 4) (**Figure 38A et 38B**). Une première approche a été de déléter plusieurs gènes codant des modules de la NORS dans le but de créer une AURS dans un système hôte hétérologue (Traitcheva *et al.*, 2007). La première PKS obtenue possède un module chimère 2 et 4 (permettant l'éviction des cycles

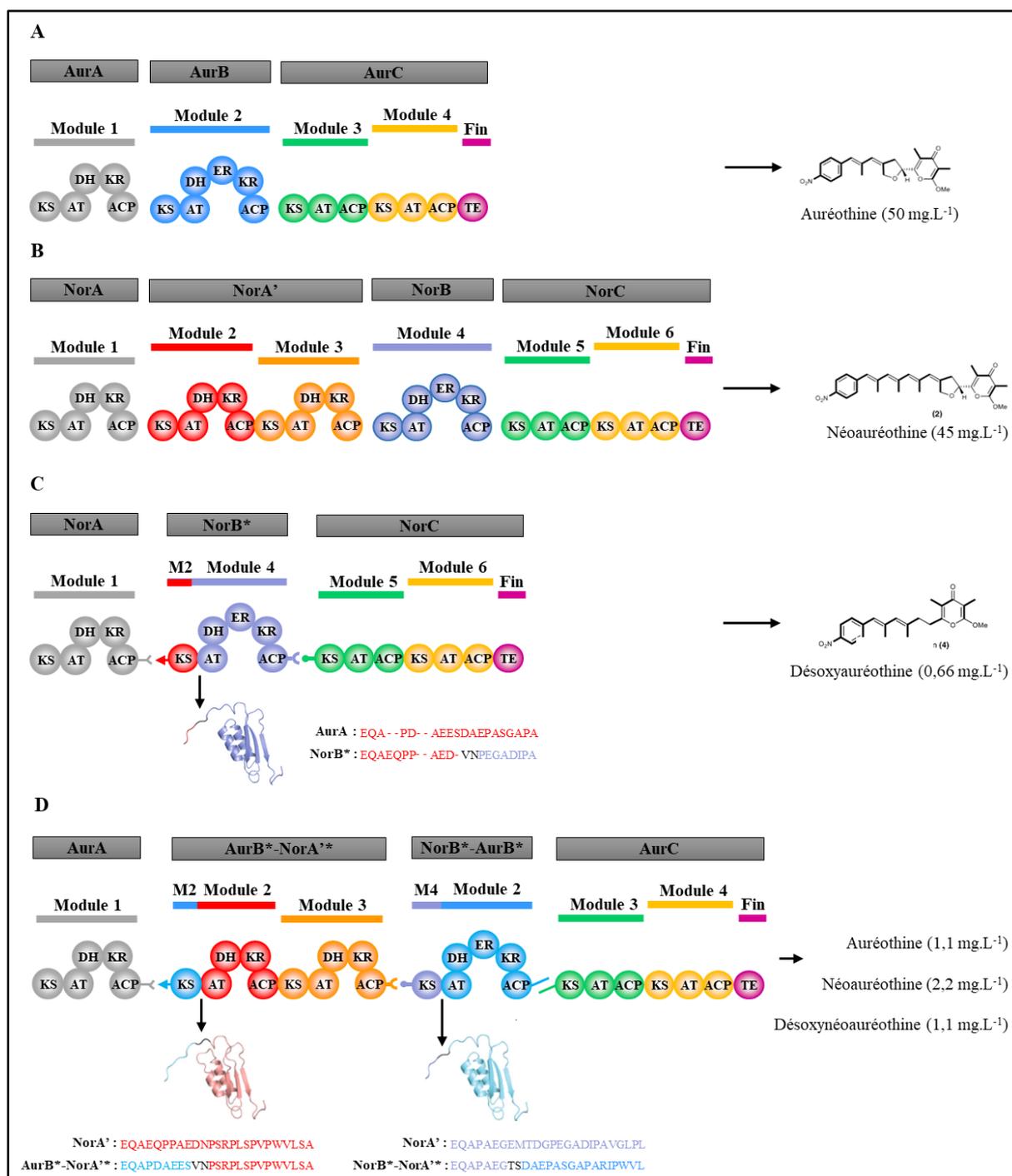
d'extension réalisé par les modules 2 et 3 de NorA') qui s'est montré inactif et ce, même en préservant l'interface ACP<sub>n</sub>/KS<sub>n+1</sub>. Ceci est probablement due à une incompatibilité entre ces deux domaines pour une translocation inter-modulaire favorable du polycétide qui peut être causé par l'absence de sélectivité du domaine KS du module 4 pour un substrat différent.

Une nouvelle construction a ensuite été générée, où le domaine KS du module 2 de NorA' a été fusionnée au module 3 de NorA' *via* la région KS-AT linker, permettant l'élimination du cycle d'extension réalisé par le module 2. La deuxième PKS obtenue a généré le produit attendu, l'homoauréothine. Cependant, des traces du produit parental néoauréothine ont aussi été générés, indiquant que le module 1 de NorA peut effectuer un cycle d'extension supplémentaire pour compenser l'éviction de l'activité d'extension du module 2 de NorA' (Traitcheva *et al.*, 2007).

La stratégie permettant d'échanger le module AT<sub>n</sub>-(KR-DH-ER)<sub>n</sub>-ACP<sub>n</sub>-KS<sub>n+1</sub> a fourni des résultats dans une autre étude du groupe de Hertweck. La délétion des gènes codant un module couplée à l'échange d'un domaine AT, a permis de faire évoluer la PKS responsable de la synthèse de l'auréothine vers la PKS responsable de la synthèse de la lutéoriculine. Cette expérience est le premier exemple de la conversion d'une PKS vers une autre proche évolutivement (Sugimoto *et al.*, 2014).

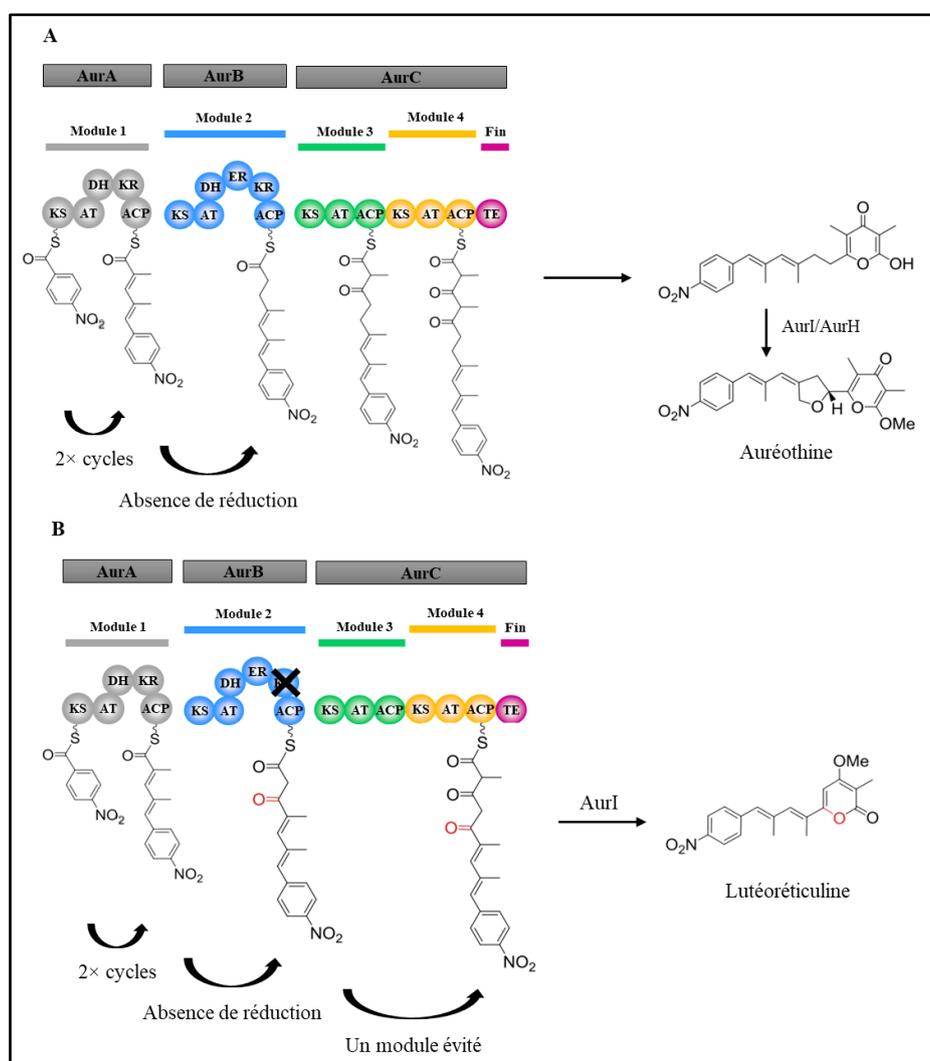
Le choix du point de fusion au niveau du linker KS<sub>n</sub>-AT<sub>n</sub> a ensuite été sondé pour comprendre le processus évolutif des PKS modulaires au travers des PKS NORS et AURS (Peng *et al.*, 2019a). L'évolution de la NORS vers la AURS (**Figure 38C**) a été testée par la création d'un module chimère (théoriquement celui correspondant à AurB) composé du domaine KS du module de 2 de NorA' et du module 4 de NorB *via* deux points de fusion différents au niveau du linker KS-AT. Le premier point de fusion est celui de leur étude initiale (Traitcheva *et al.*, 2007) qui correspond de près à celui qui avait été choisi lors de l'étude de Ranganathan, *et coll.* (Ranganathan *et al.*, 1999), et n'a pas permis d'obtenir le produit attendu. Le second point de fusion tient compte de la longueur du linker KS-AT pour être de la même taille que le module 2 de AurB, et a permis d'obtenir le produit attendu, mais avec un rendement 68 fois moindre que le composé parental (0,66 mg.L<sup>-1</sup> contre 45 mg.L<sup>-1</sup>). Ces résultats montrent que le choix du point de fusion dans le linker inter-modulaire peut s'avérer critique pour l'obtention d'une PKS chimère fonctionnelle. La perte majeure de rendement peut provenir du fait que, notamment, la taille du linker KS-AT chimère ne correspond pas exactement à celle du linker natif de la NORS (introduction des acides aminés en noir lors de l'ingénierie de NORS dans la **Figure 38C**).

L'évolution de la AURS vers la NORS a ensuite été testée (**Figure 38D**) par la création de deux modules chimères *via* l'introduction des deux modules de la sous-unité NorA' entre AurA et AurB (Peng *et al.*, 2019a). Le domaine KS du module 2 de AurB (compatible avec le domaine ACP<sub>1</sub> de Aur) a été fusionné au module 2 de NorA' et le domaine KS du module 4 de NorB (compatible avec le domaine ACP<sub>3</sub> de NorA') a été fusionné au module 2 de AurB. Les points de fusion ont respecté la taille des linkers KS-AT natifs, c'est-à-dire du module 2 de NorA' et du module 2 de AurB, respectivement. La PKS obtenue a permis de générer la néoauréothine, et montre que la PKS AURS a été convertie en PKS NORS. Cependant, le produit attendu a été obtenu avec un rendement 25 fois moindre que le composé parental auréothine (2,2 mg.L<sup>-1</sup> contre 50 mg.L<sup>-1</sup>). Le composé parental, qui est plus court que la néoauréothine, a été retrouvé parmi les produits générés, et est expliqué par le fait que le <sup>C</sup>DD d'AurA est capable de communiquer avec le module chimère NorB\*-AurB\* (**Figure 38D**). Cette deuxième preuve du choix du point de fusion dans le linker inter-modulaire apparaît identifiable pour des PKS proches dans l'évolution, mais doit être testé pour des systèmes plus éloignés. A l'instar de la PKS chimère précédemment décrite, la fusion des modules entre eux par génie génétique a aussi impliqué l'introduction d'acides aminés supplémentaires. Si les produits attendus ont pu être détectés dans ces études, le rendement est 25 à 68 fois moindre que le composé parental, indiquant que certes la PKS chimère est fonctionnelle, mais qu'elle reste moins efficace que la PKS parentale.



**Figure 38 : Evolution des PKS responsables de la synthèse de l'auréothine (AURS) et de la néoaureothine (NORS) (Adapté de Peng *et al.*, 2019a).** (A) Représentation de la PKS responsable de la synthèse de l'auréothine. (B) Représentation de la PKS responsable de la synthèse de la néoaureothine. (C) Evolution de la NORS vers la AURS. Le point de fusion dans le linker KS-AT du module 4 au sein de NorB\* est indiqué (en noir). (D) Evolution de la AURS vers la NORS. Les points de fusion dans le linker KS-AT du module 2 et du module 4 des PKS chimères sont indiqués (en noir). Les rendements des produits obtenus dans cette étude sont indiqués à droite de chaque PKS.

Chez cette même PKS responsable de la synthèse de l'auréothine (AURS) (**Figure 39A**), l'inactivation du domaine KR du second module de AURS a conduit à une erreur dans le transfert inter-modulaire de la chaîne du polycétide (Peng *et al.*, 2019b) (**Figure 39B**). En effet, le substrat a été transféré du module 2 de AURS directement au module 4, indiquant que le domaine KS du module 3 de AURS ne peut pas agir sur le substrat non réduit au niveau de la position C3. Au contraire, le domaine KS du module 4 possède une grande tolérance pour un substrat non natif. Cette expérience met en évidence un niveau supplémentaire de l'interface  $ACP_n/K_{S_{n+1}}$  qui engage une interaction non-covalente, et que le domaine KS peut être un point de contrôle qualité lors de la communication inter-modulaire qui engage deux-sous-unités.



**Figure 39 : Expérience d'inactivation d'un domaine KR dans la PKS responsable de la synthèse de l'auréothine (Adapté de Peng *et al.*, 2019b).** (A) Représentation de la PKS responsable de la synthèse de l'auréothine. (B) Le domaine KR du module 2 a été inactivé, avec pour conséquence un transfert inter-modulaire du module de 2 au module 4 de AURS. Le produit généré est la lutéoréticuline, et l'absence de réduction engendre les modifications indiquées en rouge.

L'ensemble de ces expériences ont montré qu'il est possible d'échanger des modules entre PKS proches évolutivement en exploitant un point de fusion dans le linker KS-AT. En outre, l'étude de Peng *et coll.* a révélé, pour deux expériences, l'importance de la longueur du linker KS-AT lorsqu'il est impliqué comme site de fusion pour des expériences d'échange de module. Cependant, ces études ont aussi montré que la préservation de l'interface ACP<sub>n</sub>/KS<sub>n+1</sub> peut conduire à des problèmes de reconnaissance entre sous-unité de PKS, même si la PKS chimère peut être fonctionnelle (Peng *et al.*, 2019a, 2019b). De plus, les rendements obtenus dans ces études restent bien en deçà des attentes de la biologie synthétique des PKS modulaires qui aspire à obtenir un dérivé de polycétide avec au moins le même rendement que le composé parental.

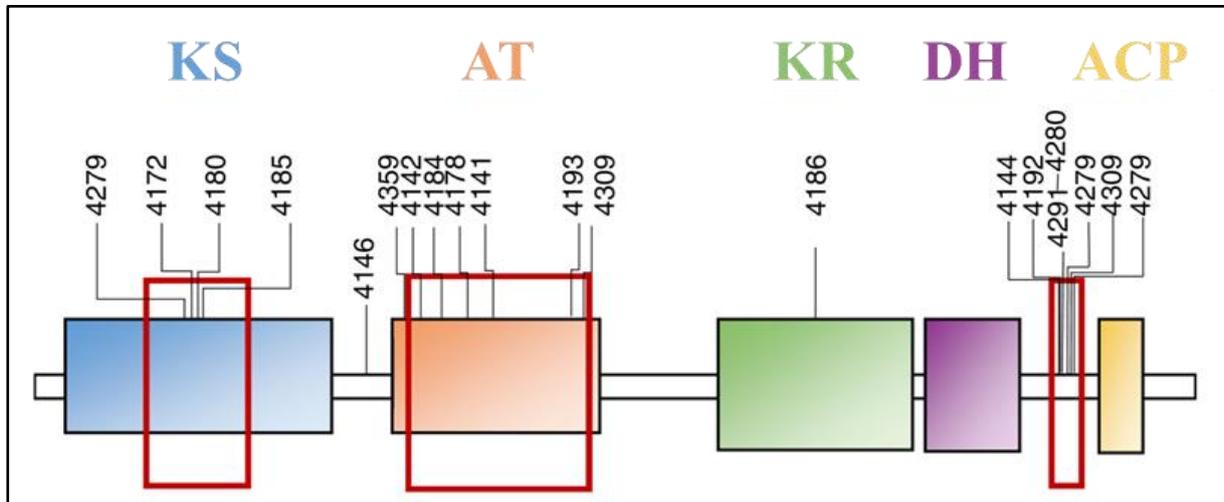
### c. Les mécanismes naturels de recombinaison pour générer de la diversité

La recombinaison homologue est un phénomène naturel qui repose sur l'échange de séquences d'ADN identiques, largement exploité dans les approches de génie génétique (Thomason *et al.*, 2014). En outre, ce phénomène serait prédominant dans les processus d'évolution des PKS modulaires (Nivina *et al.*, 2019). La possibilité d'utiliser ce phénomène naturel a d'abord été évaluée par bio-informatique, qui a suggéré que de nombreuses régions similaires au sein des modules de PKS modulaires de type *cis*-AT pourraient potentiellement être exploitées pour construire des PKS chimères et donc générer de nouvelles molécules (Starcevic *et al.*, 2011, 2012). Cette approche a été démontrée expérimentalement dans deux études majeures.

La première étude a exploité la recombinaison homologue chez la levure *Saccharomyces cerevisiae* dans le but de générer des groupes de gènes chimères entre la PKS DEBS et la PKS PIKS (Chemler *et al.*, 2015). Cette approche a permis d'obtenir des PKS chimères fonctionnelles, et les points de fusion identifiés sont situés dans le domaine AT et dans le linker KS-AT.

Une deuxième approche repose sur une méthode d'évolution accélérée, employée pour créer de la diversité pour la PKS RAPS et la PKS responsable de la synthèse de la tylosine (TYLS) (Wlodek *et al.*, 2017). Cette stratégie consiste à provoquer des recombinaisons du BGC associé aux PKS ciblées pour remplacer, ajouter ou éliminer des modules. Cette approche a permis, par exemple, de générer dix-sept nouvelles PKS RAPS fonctionnelles. L'analyse des séquences ADN des PKS chimères fonctionnelles générées montre que les sites de jonction se

situent généralement au sein des séquences conservées des domaines KS ou AT, ou dans le linker en amont du domaine ACP (**Figure 40**).



**Figure 40 : Analyses des sites de jonction dans un module de la PKS RAPS obtenus par l'approche d'évolution accélérée (Adapté de Wlodek *et al.*, 2017).** Les sites de recombinaisons sont indiqués par un cadre rouge. Le numéro des souches obtenues dans l'étude sont indiqués.

Les techniques basées sur la recombinaison homologue se sont montrées efficaces pour générer des PKS chimères fonctionnelles, bien qu'un effort de criblage important soit nécessaire pour tester leur viabilité. Cette approche est une avancée majeure pour la biologie synthétique des PKS modulaires, et ouvre de nombreuses possibilités de développements futurs. L'efficacité des sites de fusions comme les linkers inter-modulaires, les linkers KS-AT ou les DD pour la biologie synthétique des PKS modulaires repose sur leur homologie entre les systèmes PKS à modifier. Les points de fusion identifiés dans les études de Chemler *et coll.* et de Wlodek *et coll.*, certes spécifiques aux PKS générées dans ces études, pourraient aussi être exploités comme bornes alternatives pour manipuler des modules dans d'autres systèmes PKS lorsqu'ils engagent des régions conservées.

## 4. Vers une nouvelle définition d'un module

### A. Le modèle actuel

Un module de PKS modulaire de type *cis*-AT est classiquement défini comme s'étendant du domaine KS au domaine ACP (KS-AT-(KR-DH-ER)-ACP) (Nivina *et al.*, 2019). En

effet, de nombreuses sous-unités constituées d'un seul module contiennent cet ensemble de domaines, comme la PKS responsable de la synthèse de la stigmatelline qui est composée de neuf sous-unités monomodulaires (Gaitatzis *et al.*, 2002). Dans le contexte de l'ingénierie des PKS modulaires, les échanges de modules, selon ce modèle, préservent l'interaction KS-ACP intra-modulaire pour la réaction de condensation décarboxylative, mais introduisent une interface  $ACP_n/KS_{n+1}$  non-native pour le transfert inter-modulaire de la chaîne.

## B. Le modèle alternatif

Le modèle classique d'un module de PKS modulaire de type *cis*-AT a été récemment remis en cause par des analyses bio-informatiques (Zhang *et al.*, 2017). Pour cela, quatre des plus grands systèmes PKS modulaires de type *cis*-AT responsables de la synthèse d'aminopolyols (médiomycine, néomédiomycine, ECO-02301, et néotétrafibricine) ont été utilisés comme systèmes modèles pour étudier le comportement évolutif des domaines. Les données ont montré que le point de fusion se situe en amont du domaine AT, au niveau du linker KS-AT, ou en aval du domaine AT, au niveau du linker post-AT. Le module est alors redéfini comme  $AT_n-(KR-DH-ER)_n-ACP_n-KS_{n+1}$  ou  $(KR-DH-ER)_n-ACP_n-KS_{n+1}-AT_{n+1}$  (Zhang *et al.*, 2017). Ainsi, cette étude a révélé une co-évolution entre certains domaines KS et les domaines en amont dans la PKS. Cette observation a conduit à la proposition d'un modèle alternatif du module, alors défini comme s'étendant du domaine AT au domaine KS du module suivant ( $AT-(KR-DH-ER)-ACP-KS$ ) (Keatinge-Clay, 2017 ; Nivina *et al.*, 2019). Ce modèle implique que le domaine KS est associé au domaine ACP duquel il reçoit la chaîne de polycétide en cours de synthèse.

Un mécanisme possible qui explique cette co-migration est la spécificité de substrat du domaine KS. En effet, cette spécificité pourrait être conférée par le module en amont, et dépendrait du substrat sélectionné par le domaine AT ainsi que des modifications apportées aux positions en C2 et C3 par les domaines optionnels (Nguyen *et al.*, 2008). En outre, le domaine KS des PKS modulaires de type *trans*-AT possède une forte sélectivité pour son substrat, et il est possible de faire la corrélation entre un domaine KS et le domaine ACP en amont grâce à la nature de son substrat (Helfrich and Piel, 2016). Si ce phénomène est une hypothèse plausible chez les PKS de type *trans*-AT, et a seulement été montré dans un seul cas pour les PKS de type *cis*-AT (Peng *et al.*, 2019b), d'autres études ont montré que les interactions  $ACP_n/KS_{n+1}$  lors de l'étape de transfert inter-modulaire sont plus importantes que la spécificité pour le substrat, au moins en ce qui concerne les PKS chimères testées (Klaus *et al.*, 2016).

Néanmoins, les données décrites dans les parties précédentes montrent que les expériences d'échange de modules exploitant ces bornes alternatives ont mené à des résultats encore insatisfaisants au regard des attentes de la biologie synthétique des PKS modulaires (Ranganathan *et al.*, 1999 ; Peng *et al.*, 2019a ; Miyazawa *et al.*, 2020). En outre, selon ce modèle alternatif, les échanges de modules préservent certes l'interface ACP<sub>n</sub>/KS<sub>n+1</sub> pour le transfert inter-modulaire de la chaîne du polycétide, mais ils génèrent aussi une interface non-native entre les domaines KS et ACP du module.

### **C. Les dilemmes de l'ingénierie des modules**

Que ce soit pour le modèle classique ou alternatif, l'interface ACP<sub>n</sub>/KS<sub>n+1</sub> implique que les modules d'une PKS interagissent entre eux par des interactions covalentes (c'est-à-dire *via* un linker inter-modulaire), ou non-covalentes, ces dernières étant médiées par les DD. Le type d'interaction qui engage deux modules à échanger est une question centrale pour la biologie synthétique des PKS modulaires de type *cis*-AT. L'importance du linker inter-modulaire ainsi que des DD a été testée par des expériences d'échange de modules en utilisant le système PKS modèle DEBS (Tsuji *et al.*, 2001 ; Wu *et al.*, 2002). Le linker inter-modulaire natif a été montré comme essentiel dans une liaison covalente entre deux modules hybrides, et pour un transfert efficace, le linker du module donneur doit être conservé (Gokhale *et al.*, 1999 ; Tsuji *et al.*, 2001). Pour une liaison non-covalente entre deux modules, il a été montré que même en présence de DD compatibles, certains modules de la PKS DEBS sont sélectifs pour leur module n-1 naturel. Au contraire, les modules 2 et 6 de la PKS DEBS, qui sont connectés à leur module précédent par un linker inter-modulaire dans leur contexte natif, sont capables d'accepter un substrat non-naturel (Wu *et al.*, 2002). L'efficacité de l'emploi d'un linker inter-modulaire en lieu et place des DD a été montré par la fusion des sous-unités DEBS 1 avec DEBS 2, ou DEBS 2 avec DEBS 3 par un linker de deux acides aminés, ce qui a permis de maintenir et même d'améliorer la production d'érythromycine (Squire *et al.*, 2003). Placé dans le contexte de l'ingénierie des PKS modulaires, et nonobstant la spécificité de certains domaines KS pour leur substrat natif (Nguyen *et al.*, 2008 ; Peng *et al.*, 2019b), les modules connectés par un linker inter-modulaire n'auraient donc pas besoin d'être sélectifs entre eux pour garantir le transfert inter-modulaire du polycétide (Edwards *et al.*, 2014). L'interaction covalente permettrait alors de contraindre deux modules de PKS à interagir ensemble, dans un ordre donné, et ainsi d'augmenter la concentration locale de deux modules. En effet, des expériences d'ingénierie génétique sur la PKS PIKS ont montré que la séparation des modules 1 et 2 de la PKS PIKS réduit la production de pikromycine de plus de moitié (Yan *et al.*, 2009). La perte du linker

inter-modulaire obligerait les deux modules à interagir ensemble par une interaction non-covalente, ce qui aurait pu être rétabli par l'introduction de DD. Dans le cas de deux modules interagissant de manière non-covalente, la spécialisation de l'interface  $ACP_n/KS_{n+1}$  est parfois nécessaire pour contraindre ces deux sous-unités à interagir uniquement ensemble pour garantir la synthèse ordonnée qui a lieu au sein d'une PKS modulaire. Cette spécialisation est conférée certes par les DD, mais des contacts  $ACP_n/KS_{n+1}$  propre à l'interface peuvent être nécessaires, et ce type de module se prêterait moins aux expériences d'échange de modules. Pour déterminer les bases moléculaires des interactions  $ACP_n/KS_{n+1}$ , des expériences de mutagenèse dirigée et de modélisation moléculaire ont été menées en exploitant la PKS DEBS (Kapur *et al.*, 2010, 2012). L'hélice 1 du domaine  $ACP_2$  de la PKS DEBS a été identifiée comme une région conférant de la spécificité d'interaction pour le domaine  $KS_3$ . Cependant, ce résultat n'est pas extrapolable à toutes les interfaces  $ACP_n/KS_{n+1}$  qui impliquent une liaison non-covalente chez les PKS modulaires de type *cis*-AT. En effet, le modèle cryo-EM du  $PikAIII M_5$  fusionné au domaine *apo*- $ACP_4$  de  $PikAII$ , obtenu à basse résolution, suggère que l'hélice 1 du domaine ACP de  $PikAII$  ne serait pas en contact avec le domaine KS de  $PikAIII$  (Dutta *et al.*, 2014). Cette comparaison entre les PKS modèles DEBS et PIKS met en évidence que chaque module possède son propre contexte modulaire, et en ce qui concerne l'interface  $ACP_n/KS_{n+1}$ , les bases moléculaires des contacts peuvent être de différentes natures et restent à être approfondir.

Le contexte modulaire est clairement influencé par les interactions spécifiques du domaine ACP avec ses différents partenaires. Les modes de communication intra- et inter-modulaires impliquent vraisemblablement deux interfaces différentes entre les domaines ACP et KS impliqués (Whicher *et al.*, 2014), et les expériences d'échange de modules devraient perturber l'une de ces interfaces tout en préservant l'autre. Ce phénomène est un dilemme pour les expériences d'échange de modules, car il n'est pas possible de conserver ces deux interfaces. Pourtant, si l'une de ces interfaces est perturbée, le contexte modulaire reconstitué ne permet pas d'obtenir une PKS fonctionnelle. C'est pourquoi il n'est pas évident de définir les bornes d'un module de manière universelle comme le tendent à faire le modèle actuel et le modèle alternatif. En outre, la divergence des deux modèles montre qu'il faut différencier le module au niveau nucléotidique, et le module en tant que polypeptide fonctionnel (Klaus and Grninger, 2018 ; Nivina *et al.*, 2019). Une stratégie pour résoudre ce dilemme consiste à déterminer l'interface qui peut être perturbée, puis d'optimiser les contacts non-natifs entre les domaines KS et ACP par des approches de mutagenèse (Murphy *et al.*, 2016). En outre, les points de fusion naturels sont identifiables pour des modules de PKS proches évolutivement

(Zhang *et al.*, 2017 ; Peng *et al.*, 2019a), mais il apparaît plus difficile de modifier un module de PKS avec un autre plus éloigné dans l'évolution. Pourtant, ce type de modification est parfois nécessaire pour des modifications spécifiques, comme la modification ponctuelle de la stéréochimie d'un polycétide. En effet, les domaines KR de type réducteur et épimérisant sont rare et implique d'échanger des modules entre des PKS éloignées évolutivement. Les stratégies d'échange de modules entiers ou de modules chimères doivent être confrontées pour identifier une stratégie efficace pour la modification de la stéréochimie d'un polycétide.

**Dans le cadre de ma thèse, j'ai étudié l'effet du contexte modulaire sur le contrôle de la stéréochimie par des domaines spécifiques. De cette manière, j'ai pu déterminer qu'il était possible d'introduire des réactions de réduction et d'épimérisation *in vivo* par échange de module, permettant ainsi de démontrer que le remplacement de module pourrait être une stratégie pour manipuler la stéréochimie des polycétides et envisager une ingénierie efficace des PKS. Pour cela, j'ai utilisé le système biosynthétique modèle DEBS 1-TE et j'ai confronté deux stratégies décrites dans la littérature : la première considérant le modèle classique du module KS-AT-(KR-DH-ER)-ACP, et la deuxième respectant le modèle alternatif du module AT<sub>n</sub>-(KR-DH-ER)<sub>n</sub>-ACP<sub>n</sub>-KS<sub>n+1</sub>. Ces travaux se sont poursuivis au niveau de la PKS entière, en utilisant la DEBS comme système modèle.**

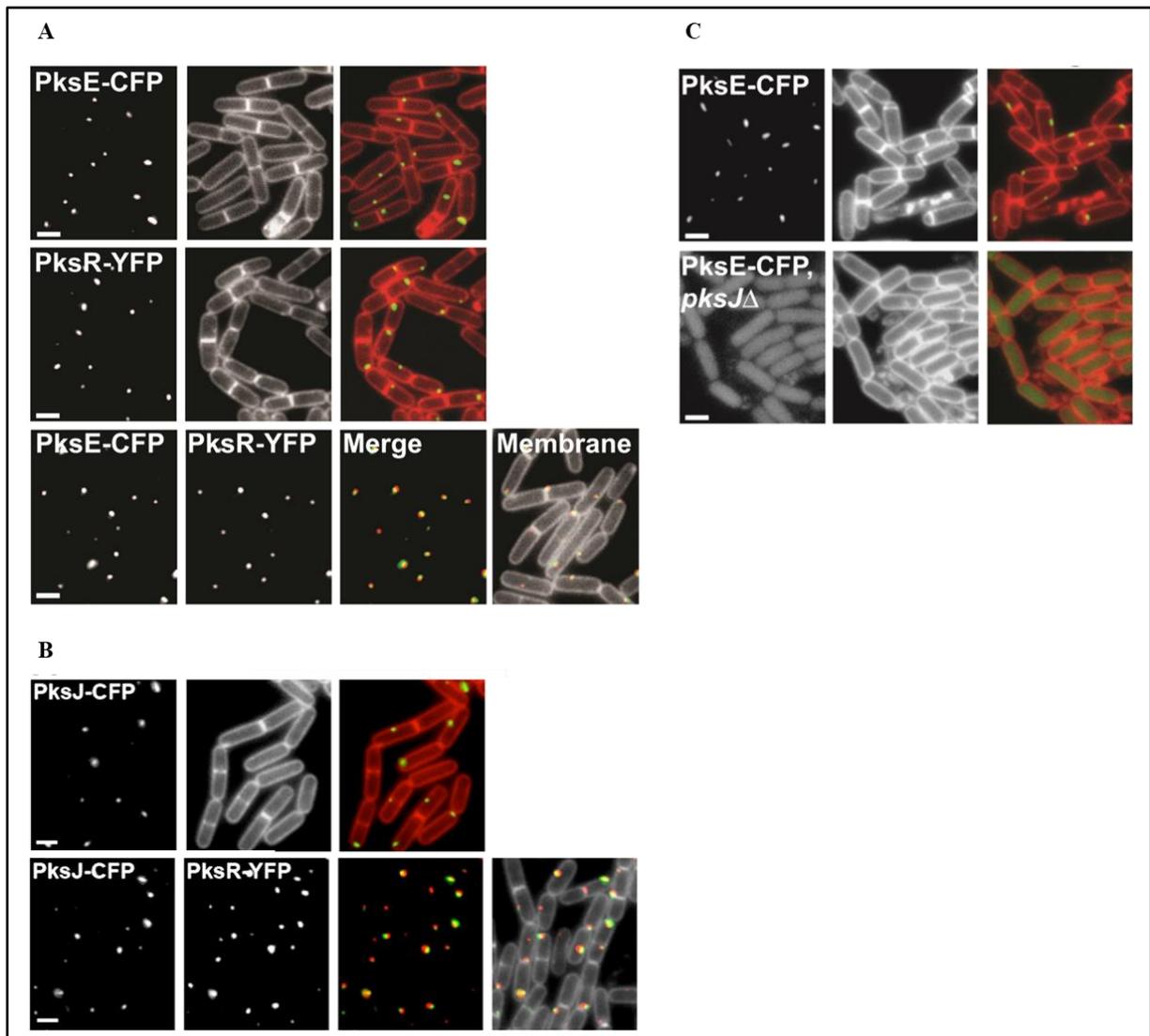
## V. L'organisation macromoléculaire des PKS modulaires

### 1. Le mégacomplexe chez *Bacillus subtilis*

Au niveau de la PKS entière, seule une étude s'est intéressée à obtenir des informations sur l'organisation macromoléculaire de ces systèmes multi-enzymatiques dans les cellules bactériennes (Straight *et al.*, 2007). Straight *et coll.* ont utilisé comme modèle d'étude la PKS de type *trans*-AT responsable de la synthèse de la bacillaene de *Bacillus subtilis*, décrite dans la section **II** de cet état de l'art. Des protéines chimères ont été générées en fusionnant les extrémités carboxy-terminales des sous-unités de la PKS et un domaine *trans*-AT (PksE) avec les protéines fluorescentes *cyan fluorescent protein* (CFP) ou *yellow fluorescent protein* (YFP). Les bactéries produisant ces protéines de fusion ont conservé leur activité inhibitrice contre *S. avermitilis*, montrant que les protéines de fusion restaient fonctionnelles. Puis, les différentes constructions ont été assujetties à des expériences de co-localisation par microscopie confocale.

En premier lieu, le domaine *trans*-AT PksE-CFP co-localise avec les sous-unités PksJ-YFP (PksK dans la **figure 11**) et PksR-YFP, mais de manière inattendue au niveau d'un locus unique dans la cellule et proche de la membrane (**Figure 41A**). En outre, les deux sous-unités PksJ et PksR, qui appartiennent au même système PKS, co-localisent dans la cellule bactérienne de la même manière (**Figure 41B**). Ces observations supportent le modèle selon lequel les PKS s'assembleraient en un seul mégacomplexe, et que ce mégacomplexe se juxtapose près de la membrane cellulaire.

Puis, Straight *et coll.* (Straight *et al.*, 2007) ont généré une souche de *B. subtilis* pour laquelle l'ORF codant la sous-unité PksJ a été délétée et la localisation du domaine *trans*-AT PksE a été étudiée. En l'absence de la sous-unité PksJ, le domaine discret PksE ne se localise plus en un seul locus unique dans la cellule, mais se retrouve diffus dans toute la bactérie (**Figure 41C**). La nécessité de la présence de la sous-unité PksJ pour la localisation en foci de PksE-CFP indique que la localisation des sous-unités détermine la localisation des enzymes *trans*-AT.

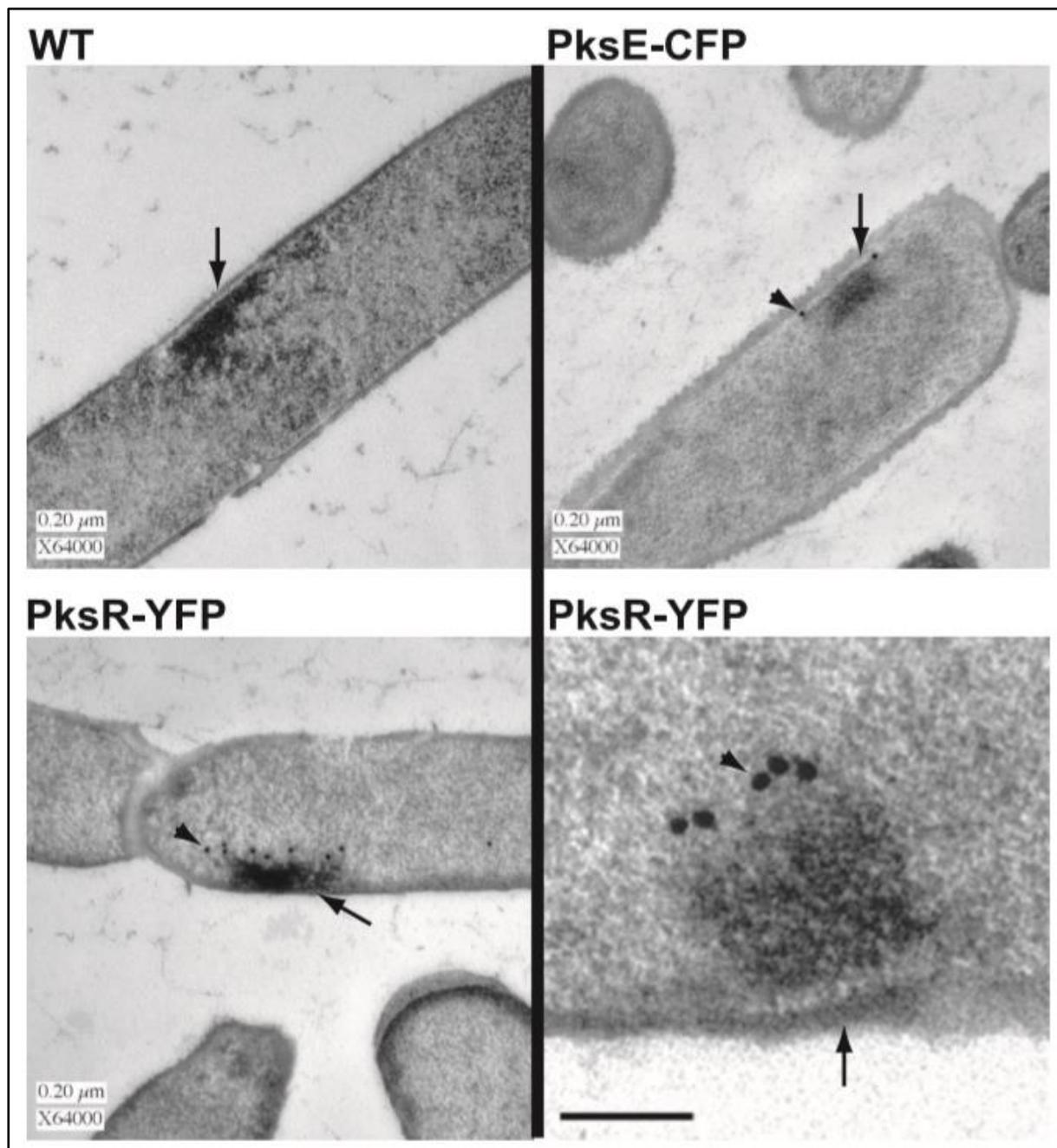


**Figure 41 : Localisation des sous-unités PksR et PksJ de la PKS responsable de la synthèse de la bacillaene et du domaine *trans*-AT PksE dans *B. subtilis* 3610 par microscopie confocale. (A) Le domaine *trans*-AT PksE-YFP (panel du dessus) et la sous-unité PksR (panel central) se localisent en un seul locus dans la bactérie. Les images à gauche correspondent aux signaux fluorescents des protéines de fusion. Les images au centre correspondent à la membrane marquée avec un colorant (TMA-DPH). Les images à droite correspondent à la superposition des images de gauche et du centre. Le domaine *trans*-AT PksE-CFP et PksR-YFP co-localisent en un seul locus (panel du dessous). (B) La sous-unité PksJ (panel du dessus) se localise en un seul locus dans la bactérie. Les sous-unités PksJ-CFP et PksR-YFP co-localisent en un seul locus (panel du dessous). (C) La sous-unité PksE-CFP se localise en un seul locus (panel du dessus), mais présente un signal diffus dans la bactérie en l'absence de la sous-unité PksJ (panel du dessous). L'échelle indiquée en bas à gauche de chaque image est de 1  $\mu$ m.**

Étant donné que la PKS responsable de la synthèse de la bacillaene se localise en un seul locus, Straight *et coll.* ont émis l'hypothèse que ce signal devrait être observé par la technique de microscopie électronique. Le développement de la microscopie électronique a permis d'obtenir des informations sur la localisation et la structure des macro-complexes biologiques

à très haute résolution. En outre, cette technique de microscopie peut être renforcée par une préparation particulière des échantillons biologiques, qui sont congelés presque instantanément, et conservés à la température de l'azote liquide pendant le stockage, de manière à figer l'état natif et donc de préserver la morphologie et la structure des échantillons (Nogales, 2018). L'apport de ce mode de préparation de l'échantillon couplé à la résolution de la microscopie électronique a permis de révéler de nombreuses nouvelles structures chez les bactéries (Dobro *et al.*, 2017). Grâce à cette approche, Straight *et coll.* ont révélé une masse dense en électrons dans le cytoplasme de *B. subtilis* qui n'avait jamais été observée auparavant.

Pour déterminer si la densité électronique observée dans une cellule sauvage de *B. subtilis* correspond au locus fluorescent de PKS observé dans les analyses de microscopie confocale, Straight *et coll.* ont utilisé un anticorps primaire anti-GFP (pouvant cibler les protéines YFP et CFP) et un anticorps secondaire conjugué avec des particules d'or. Cette expérience a montré qu'en effet, la sous-unité PksR-YFP et le domaine *trans*-AT PksE-CFP se localisent dans la zone dense en électrons dans le cytoplasme de *B. subtilis* (**Figure 42**), et que cette zone correspondrait à une structure protéique contenant au moins des protéines impliquées dans la synthèse de la bacillaene. Straight *et coll.* ont ensuite analysé une souche ne possédant plus la PKS responsable de la synthèse de la bacillaene. Les résultats montrent qu'en l'absence de cette PKS, la zone dense en électron n'est plus observée. Ces données supportent le fait que les systèmes PKS responsables de la synthèse de la bacillaene dans la bactérie *B. subtilis* s'organisent en un mégacomplexe possédant une taille comparable à un organelle, avec une masse moléculaire estimée de 10 à 100 MDa. De surcroît, les données de microscopie électronique indiquent que ce mégacomplexe ne semble pas être directement associé à la membrane.



**Figure 42 : Assemblage d'un mégacomplexe chez *B. subtilis* (adapté de Straight *et al.*, 2007).** Les images acquises par microscopie électronique à transmission montrent une densité électronique proche de la membrane plasmidique dans la souche sauvage NCIB3610, ainsi que les mutants PksR-YFP (sous-unité de la PKS synthétisant le bacillaene fusionnée à la YFP) et PksE-CFP (domaine *trans*-AT discret fusionné à la CFP). L'immunolocalisation par microscopie électronique des protéines PksR et PksE montre des particules d'or (correspondant aux anticorps secondaires (montré par les flèches courtes)) au sein des masses visibles dans les souches de *B. subtilis*. La flèche longue montre la paroi cellulaire. Echelle : 0,20 μm, à l'exception de l'image en bas à droite qui est de 0,1 μm.

Ces travaux pionniers dans l'étude de la localisation des PKS au sein de leur hôte bactérien ont donc mis en exergue qu'un système PKS modulaire de type *trans*-AT s'organise en un unique mégacomplexe, et que ce mégacomplexe se trouve proche de la membrane. Cette organelle serait une usine biosynthétique qui favoriserait la synthèse et l'export du métabolite. Dans ce contexte, l'association avec la membrane est supposée permettre le couplage entre la synthèse du métabolite et son export hors de la cellule (Simunovic *et al.*, 2003). Les bases moléculaires de l'association et son apport pour la bactérie restent cependant à déterminer.

Une hypothèse pouvant expliquer cette organisation particulière serait une association *via* des protéines membranaires qui pourrait servir de véritable plate-forme pour l'association des PKS avec la membrane (Straight *et al.*, 2007). Ce type d'association a déjà été observé entre une sous-unité de la PKS responsable de la synthèse d'un lipide de surface nécessaire pour la virulence de *Mycobacterium tuberculosis* et son transporteur associé (Jain and Cox, 2005). En outre, un domaine KR de la PKS de type II responsable de la synthèse de l'actinorhodin a été détecté avec la fraction membranaire de *Streptomyces coelicolor* par immuno-blotting (Xu *et al.*, 2008). Aussi, des protéines accessoires impliquées dans la synthèse de la mycolactone (biosynthétisée par une PKS de type I) ont été détectées par microscopie confocale dans la paroi cellulaire de *Mycobacterium ulcerans* et forment des îlots (Porter *et al.*, 2013), ce qui conforte l'hypothèse d'une usine biosynthétique. Par ailleurs, d'autres systèmes biosynthétiques proches des PKS ont été retrouvés dans les membranes des bactéries. Il a été décrit, par exemple, que les enzymes associées à des modifications post-traductionnelles lors de la maturation de RiPP, les lantibiotiques, s'organisent aussi comme des usines biosynthétiques à la surface des membranes bactériennes (Engelke *et al.*, 1992 ; Kiesau *et al.*, 1997). De la même façon, les enzymes responsables de la synthèse des sidérophores, incluant des NRPS, sont associés à la membrane (Hantash and Earhart, 2000) et s'organisent en mégacomplexe, appelé le sidérosome (Imperi and Visca, 2013).

Le rapprochement des PKS avec la membrane et la modification des peptides non ribosomiaux à la surface des membranes pour leur conférer leur bioactivité serait une stratégie conservée chez les bactéries visant à coupler la biosynthèse avec la sécrétion, évitant ainsi l'accumulation dans le cytoplasme de molécules potentiellement toxiques, telles que les antibiotiques et le sidérophore (Simunovic *et al.*, 2003 ; Straight *et al.*, 2007 ; Imperi and Visca, 2013). Il a été notamment montré que la linearmycine (un polycétide possédant des activités antifongique et antibiotique) peut être exporté sous formes de vésicules extracellulaires (Hoefler *et al.*, 2017), et que la biogénèse de ces vésicules dépend de la biosynthèse de la linearmycine. Le mode d'export des polycétides permettrait de conférer différentes

fonctions comme facteur de virulence, élément important de la physiologie membranaire de la bactérie, ou encore comme molécule de signalisation dans la communication microbienne.

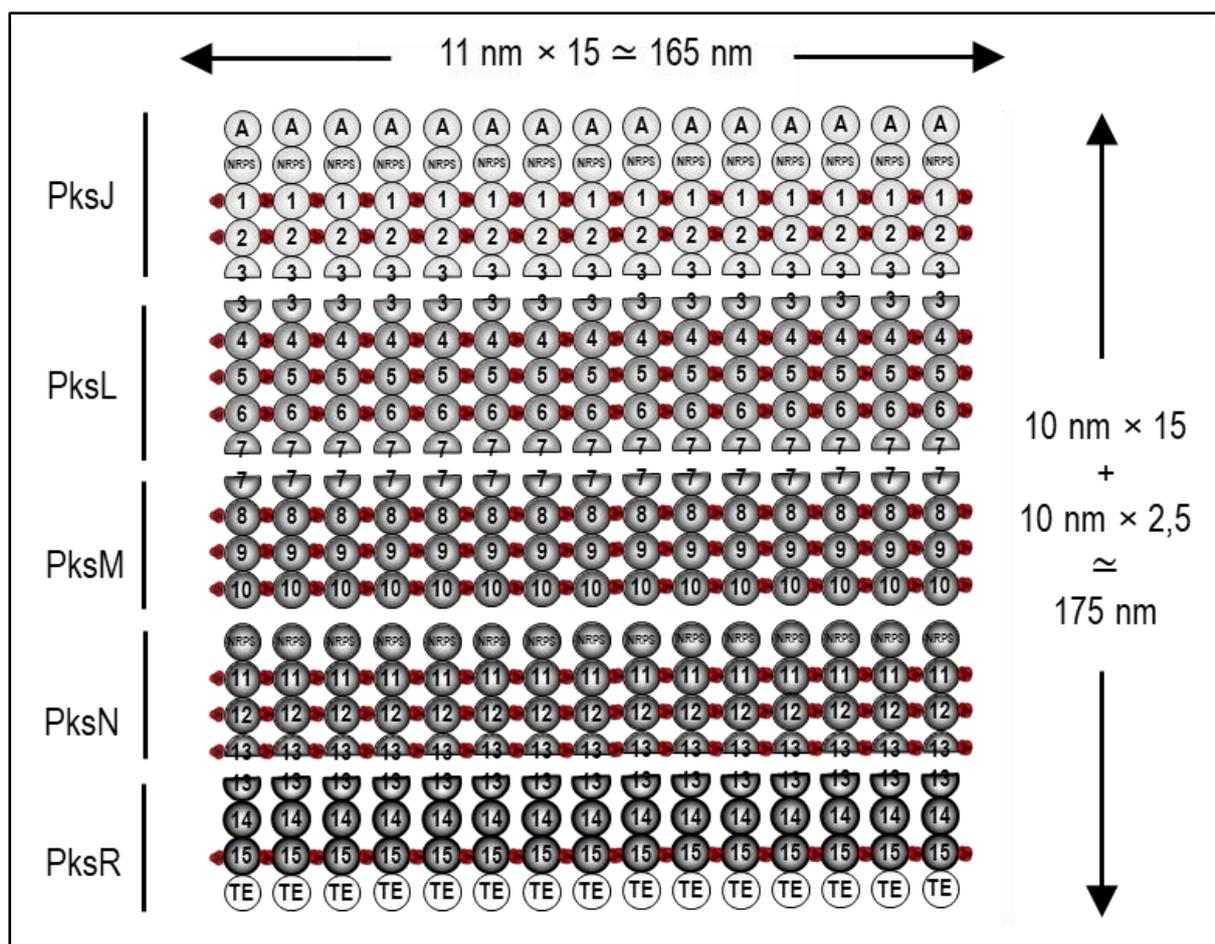
L'association en un mégacomplexe permettrait de concentrer les nombreuses protéines PKS, incluant les systèmes multi-enzymatiques, les enzymes agissant en *trans* comme les domaines *trans*-AT, ainsi que les précurseurs et les unités d'extension. En effet, les interactions entre les sous-unités de PKS sont faibles *in vitro* (Broadhurst *et al.*, 2003), ce qui pourrait avoir une importance physiologique. Le mégacomplexe permettrait alors de créer un environnement favorisant la communication entre les sous-unités de PKS en augmentant la concentration locale en sous-unités de PKS, ou encore de protéger les intermédiaires réactionnels.

## **2. Le modèle actuel du mode d'assemblage**

Une étude cristallographique ultérieure réalisée sur des domaines KS isolés a suggéré qu'un motif de trois hélices  $\alpha$  dans le linker KS-AT, nommé LINKS (de l'anglais *Laterally INteracting Ketosynthase Sequence*) serait responsable de l'assemblage des sous-unités PKS (Gay *et al.*, 2016). Ce motif a été observé dans 9 des 10 structures cristallographiques de domaines KS de PKS modulaires de type *trans*-AT, incluant des domaines KS impliqués dans la synthèse de la bacillaene. Les LINKS permettraient des interactions latérales entre domaines KS, favorisant alors un assemblage linéaire des PKS (**Figure 43**).

L'étude s'est ensuite intéressée à trouver dans d'autres domaines KS provenant de PKS de type *trans*-AT des motifs caractéristiques des LINKS par analyse bio-informatique. L'alignement de séquences provenant de 100 domaines KS a révélé que ce motif est très peu conservé, qu'il n'est pas présent au sein de tous les domaines KS, et que le mode d'interaction est unique pour chaque domaine KS (Gay *et al.*, 2016).

En effet, la réelle fonction des LINKS reste à élucider. Les interprétations ont été effectuées sur la bases de structures cristallographiques, et donc ces motifs pourraient être des artefacts cristallographiques, car les interactions étaient trop faibles pour être détectées par des méthodes d'étude des interactions en solution.



**Figure 43 : Représentation schématique de l'assemblage vertical de la PKS modulaire de type *trans*-AT responsable de la synthèse de la bacillaene (adapté de Gay *et al.*, 2016).**

Si de tels contacts sont critiques et si ce résultat est généralisable à plusieurs systèmes PKS – voire tous – alors ces données auront des conséquences importantes pour la conception des expériences de biologie synthétique. En effet, il sera alors nécessaire de maintenir la capacité des PKS modifiées à mettre en place ces associations. Néanmoins, les motifs LINKS ont été identifiés chez les PKS de type *trans*-AT et ne sont pas présents chez les PKS de type *cis*-AT (Gay *et al.*, 2016). Si les PKS de types *cis*-AT s'organisent aussi sous forme d'un mégacomplexe, alors d'autres bases moléculaires seraient engagées. Avec le recul, les résultats mitigés obtenus par l'ingénierie des PKS modulaire de type *cis*-AT sont peut-être imputables à la perturbation des contacts PKS/PKS ou des associations PKS/membrane, voire les deux.

**Dans le cadre de ma thèse, j'ai étudié le contexte macromoléculaire au niveau de la PKS entière. J'ai ainsi démontré que la formation de macrocomplexes est un phénomène pouvant être retrouvé chez les PKS de type *cis*-AT dans une bactérie du genre *Streptomyces*, le genre bactérien le plus prolifique en métabolites spécialisés. J'ai aussi**

*État de l'art*

**évalué la co-localisation d'une sous-unité avec le transporteur associé dans le but d'obtenir des informations sur le couplage entre la synthèse du polycétide et son export.**

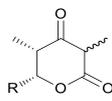
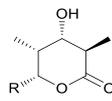
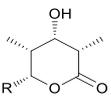
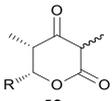
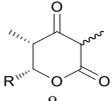
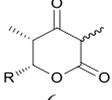
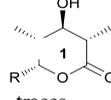
# **Objectifs**



Les polycétides réduits sont à la base de plusieurs médicaments à fortes valeurs thérapeutiques et économiques, tels que certains antibiotiques, immunosuppresseurs et agents anti-cancéreux. Par conséquent, il est d'un grand intérêt de trouver de nouveaux polycétides ou de modifier les structures existantes afin de pouvoir créer des analogues, dans le but d'augmenter les propriétés pharmacodynamiques et pharmacocinétiques de la molécule initiale. Dans ce contexte, une approche prometteuse consiste à modifier les enzymes responsables de leur synthèse – les polycétide synthases (PKSs) – par la biologie synthétique (Weissman, 2016). Cette approche vise à générer des analogues de structure de polycétides avec des rendements au moins similaires aux polycétides natifs. Pour que cette stratégie soit efficace, elle nécessite la compréhension en détail des relations structure/fonction de ces systèmes protéiques hautement complexes. Une des questions majeures qui reste posée est le rôle du contexte macromoléculaire sur le fonctionnement de ces systèmes. Cet aspect a été quasiment négligé pendant les tentatives précédentes d'ingénierie, ce qui peut expliquer, au moins en partie, le faible taux de réussite. Dans ce contexte, cette thèse avait pour objectifs d'aborder cette question à deux niveaux. Tout d'abord au niveau du module, pour déterminer l'effet du contexte modulaire sur le contrôle de la stéréochimie par des domaines spécifiques. Puis au niveau de la PKS entière, afin de déterminer si elle s'assemble en un mégacomplexe. Ces deux informations sont essentielles à la bonne conduite d'expériences de biologie synthétique sur les PKSs afin d'obtenir des dérivés de polycétides.

Le premier niveau auquel le contexte macromoléculaire peut jouer un rôle critique est celui du module, car la structure du module semble influencer le résultat stéréochimique lors de la synthèse. Lors des études précédentes, il a été suggéré par échange de domaine KR au sein du système bimodulaire DEBS 1-TE (présenté dans la **Figure 29** dans l'état de l'art (Kellenberger *et al.*, 2008)), que le contexte dans lequel la KR nouvellement introduite opère à une influence sur sa capacité à catalyser la réaction d'épimérisation (Annaval *et al.*, 2015). Plus spécifiquement, lorsque des KR de type A2 ont été introduites en lieu et place d'une KR de type A1, les produits escomptés ont été obtenus (introduction d'une épimérisation au niveau du groupe méthyle en C2), mais également des produits pour lesquels l'épimérisation n'a pas eu lieu (**Tableau 1**). Ce résultat suggère qu'un aspect du module hôte a empêché la réaction d'épimérisation. De plus, seulement des traces de produits attendus ont été obtenues à la suite de l'introduction des KR de type B2 dans le même contexte modulaire.

**Tableau 1 : Échange des domaines KR au sein de la PKS modèle DEBS 1-TE et analyse des produits obtenus (d'après Annaval *et al.*, 2015).** Les domaines KR montrés sont ceux ayant permis d'obtenir les résultats les plus significatifs lors de cette étude.

Domaine KR introduit	Type de KR	Produits et rendements (mg.L <sup>-1</sup> )		
Amph KR <sub>11</sub>	A2	 ≈ 3	 traces	 14
Olm KR <sub>5</sub>	A2	 52		
Las KR <sub>7</sub>	B2	 8		
Lkm KR <sub>1</sub>	B2	 6	 traces	

Deux conclusions se dégagent de ces études : i) le contexte modulaire dans lequel le domaine KR agit peut influencer la fidélité du stéréocontrôle (c'est-à-dire hors de leur contexte natif, les domaines KR de type A2 peuvent réduire les chaînes de polycétide non épimérisées) et, ii) les domaines KR de type B2 présentent une faible portabilité dans un contexte modulaire A1. Ces résultats ont incité à explorer le remplacement du module 2 de DEBS 1-TE par des modules contenant des domaines KR de type A2 et B2 provenant de systèmes PKS hétérologues, afin d'améliorer à la fois les rendements absolus et relatifs des lactones réduites et épimérisées. Dans le cadre de ce projet, mes objectifs étaient d'abord d'obtenir les constructions pour lesquelles le module 2 de la PKS modèle a été remplacé par un éventail d'autres modules. Puis, de transformer *Sac. erythraea* par conjugaison avec les constructions obtenues et permettre la production des lactones. Enfin, nous avons analysé la stéréochimie des produits obtenus par GC/MS et évalué l'influence du contexte modulaire sur le stéréocontrôle.

La deuxième partie de l'étude de l'influence du contexte modulaire sur le contrôle de la stéréochimie a été de déterminer si les résultats obtenus sur le système biosynthétique DEBS 1-TE pouvaient être exploités pour varier la stéréochimie d'un polycétide en modifiant la PKS entière DEBS. Mes objectifs ont été en premier lieu d'obtenir les constructions pour lesquelles le module 2 de la PKS DEBS a été remplacé par les modules ayant montré les résultats les plus significatifs lors de l'étude précédente. En second lieu, la transformation d'une souche de

*Sac. erythraea* était envisagée par les plasmides contenant les trois sous-unités de la DEBS, en incluant les constructions obtenues. Cette souche a notamment comme caractéristique de surproduire l'érythromycine, afin de permettre la production des analogues de la 6-désoxyérythronolide B. En dernier lieu, la caractérisation de la stéréochimie des produits obtenus a été effectuée, en collaboration cette fois-ci.

Au niveau de la PKS entière, il a été décrit qu'une PKS modulaire de type *trans*-AT (celle responsable de la synthèse de la bacillaene chez *B. subtilis*) s'assemblent pour former un mégacomplexe comparable à un organelle (taille estimée de 10 à 100 MDa), et que ce mégacomplexe se localise proche de la membrane cellulaire (**Figure 42 de l'état de l'art**) (Straight *et al.*, 2007). Cette proximité avec la membrane est supposée permettre le couplage entre la synthèse du métabolite et son export hors de la cellule. Une étude cristallographique ultérieure réalisée sur des domaines KS isolés a laissé suggérer à Keatinge-Clay *et coll.* (Gay *et al.*, 2016) qu'un motif de trois hélices  $\alpha$  suivant ces domaines (nommé « LINKS motif ») est responsable de l'assemblage des sous-unités PKS (**Figure 43 de l'état de l'art**), mais la base moléculaire de l'association avec la membrane reste à déterminer entièrement (Gay *et al.*, 2016). Ces observations nous ont incités à généraliser si la formation d'un mégacomplexe est un phénomène pouvant être retrouvé chez les PKS de type *cis*-AT dans une bactérie du genre *Streptomyces* qui est le genre bactérien le plus prolifique en métabolites spécialisés identifié à ce jour. Pour cela, la PKS responsable de la coelimycine (CPKS) de *S. coelicolor* (Pawlik *et al.*, 2007) a été utilisée comme modèle lors de cette étude. Concernant ce projet, mes objectifs ont été dans un premier temps de trouver les conditions permettant d'exprimer la CPKS, car le BGC *cpk* associé à cette PKS est cryptique dans des conditions normales de laboratoire (Bednarz *et al.*, 2019). Dans un second temps, pour pouvoir localiser les PKS au sein des cellules bactériennes, j'ai fusionné la première et/ou la dernière sous-unité ainsi que le transporteur associé, à une protéine fluorescente (Sun *et al.*, 1999 ; Schlimpert *et al.*, 2017) l'outil d'ingénierie génétique CRISPR/Cas9 (Cobb *et al.*, 2015; Tong *et al.*, 2015). Dans un troisième temps, les images de la localisation des protéines PKS et du transporteur associé ont été obtenus par microscopie à fluorescence, en collaboration avec le Dr Paula Yagüe (Université de Leiden, Pays-Bas, et Université d'Oviedo, Espagne).



# Résultats



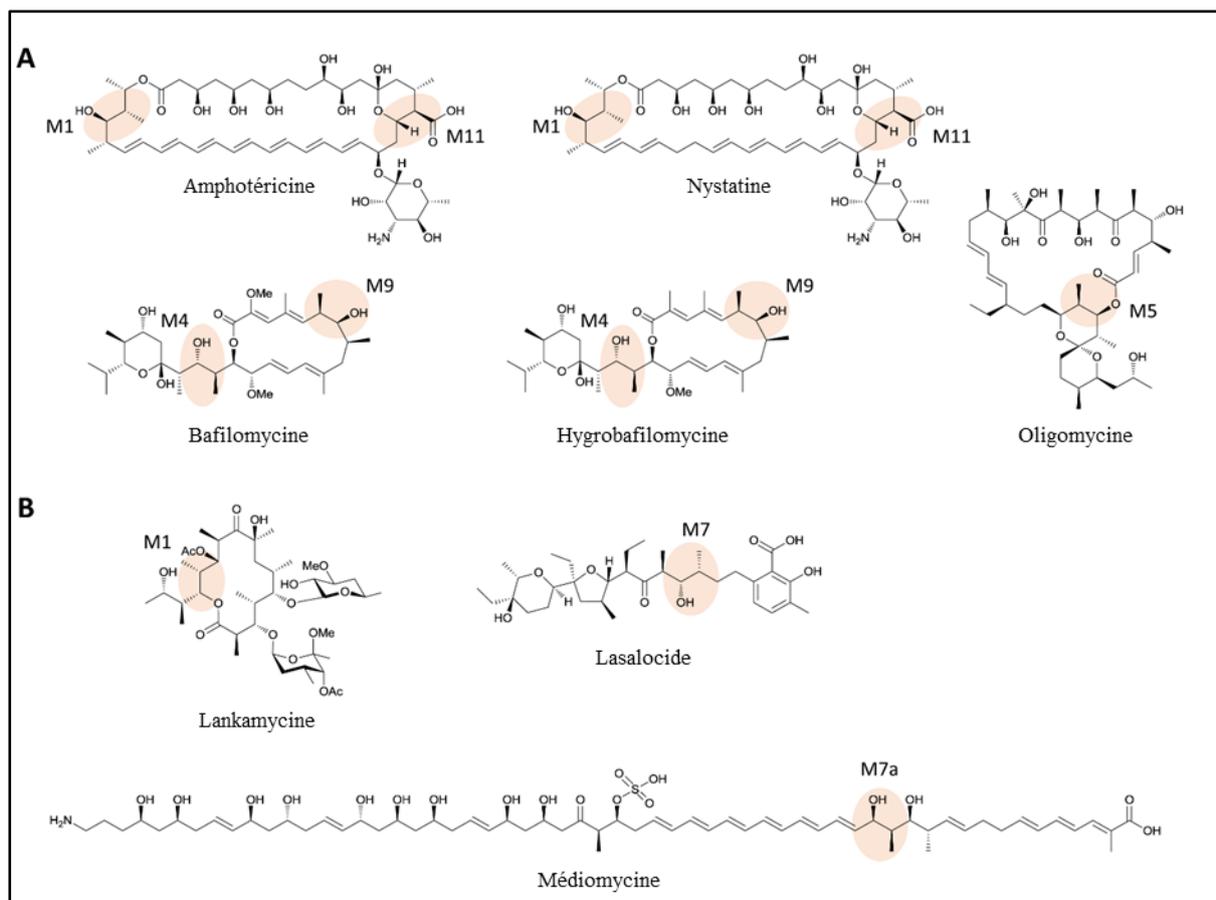
# I. Étude de l'influence du contexte modulaire sur le contrôle de la stéréochimie

## 1. Échange de modules dans un système biosynthétique modèle

### A. Résumé de l'étude

La première étape de cette étude a été de convertir le plasmide pMU3-DEBS 1-TE utilisé lors de l'étude de Annaval *et coll.* (Annaval *et al.*, 2015) en plasmide navette *E. coli* / *Sac. erythraea* / *Saccharomyces cerevisiae*. Celle-ci a été réalisé par le Dr Christophe Jacob avant mon arrivée au sein de l'équipe. Ce plasmide pMU3-DEBS 1-TE-URA3 est capable de se répliquer chez la levure et permet de cloner les séquences des modules étudiés par la technique de « Transformation-Associated Recombination » (TAR) (Kouprina and Larionov, 2016).

La deuxième étape a été de cloner l'ensemble des modules (**Figure 44A et 44B**) par recombinaison homologue chez la levure *Saccharomyces cerevisiae*, en accord avec les stratégies présentées dans la **Figure 1** de la publication n°1 en **Annexes**. En terme de bornes des modules, un module est défini classiquement comme s'étendant du domaine KS au domaine ACP. Pourtant, il a été démontré par génie génétique, que l'échange de module peut se faire, au moins dans certains cas, de façon plus efficace quand la jonction intermodulaire native (ACP<sub>n</sub>/KS<sub>n+1</sub>) est maintenue (Ranganathan *et al.*, 1999). Par conséquent, j'ai poursuivi deux stratégies d'échange de module : i) remplacement du module 2 entier (KS<sub>2</sub> → ACP<sub>2</sub>, construction de type A) (Gokhale *et al.*, 1998) ; et ii) échange de la région AT<sub>2</sub> → ACP<sub>2</sub> (construction de type B), préservant ainsi l'interface ACP<sub>1</sub>/KS<sub>2</sub>. De plus, le système modèle utilisé, DEBS 1-TE, a été construit de manière à relocaliser le domaine thioestérase (TE) en C-terminal de la première sous-unité de DEBS (DEBS 1). Pour ce faire, Rowe *et coll.* (Rowe *et al.*, 1998) ont utilisé le site de restriction *SacI* situé au milieu de la séquence nucléotidique du domaine ACP<sub>2</sub>. La conséquence est l'obtention d'une chimère ACP<sub>2</sub>/ACP<sub>6</sub> en lieu et place de l'ACP<sub>2</sub> au niveau du module 2 du DEBS 1-TE. Ainsi, les deux stratégies précédentes sont suivies soit en remplaçant l'ACP<sub>2/6</sub> en entier soit en ne remplaçant que l'ACP<sub>2</sub> *sensu stricto* (que jusqu'au site *SacI*, constructions de type C et D respectivement). Les alignements de séquences présentés en **Figure S2** de la publication n°1 en **Annexes** montrent les jonctions utilisées. Les plasmides recombinants obtenus ont ensuite été transformés par conjugaison dans la souche JC2 de *Sac. erythraea* dans le but de faire produire les différentes lactones.



**Figure 44 : Structures des polycétides dont les polycétide synthases (PKS) ont été utilisées comme sources de modules. (A) Sources des modules contenant un domaine KR de type A2. (B) Sources des modules contenant un domaine KR de type B2. En (A) et (B), les centres stéréogènes concernés et les modules qui les établissent sont indiqués (orange).**

La dernière étape a consisté en l'extraction de ces lactones puis en leur analyse par CPG-MS au Plateau d'Analyse Structurale et Métabolomique (PASM) de l'Ecole Nationale Supérieure d'Agronomie et des Industries Alimentaires (ENSAIA) de l'Université de Lorraine. Les standards dont j'ai disposé ont été obtenus après fermentation des souches *Sac. erythraea* JC2 pJLK25, 30 et 35 (Kellenberger *et al.*, 2008). L'identification des lactones **1a** et **3a** a été réalisée par comparaison avec un standard synthétique fourni par les docteurs S. Heinrich and F. Schulz (« Ruhr University Bochum », Allemagne). L'identification des lactones **2a** a été réalisée par comparaison avec un standard synthétique fourni par le Dr. A.T. Keatinge-Clay (« Texas University », Etats-Unis).

La souche *Sac. erythraea* JC2 est la souche bactérienne qui ne produit plus l'érythromycine car le groupe de gènes *ery* associé a été supprimé à l'exception de l'ORF correspondant au domaine TE. Cette dernière souche a été modifiée lors de l'étude de Kellenberger *et coll.* (Kellenberger *et al.*, 2008) par conjugaison bactérienne avec les plasmides

de la série pJLK. L'analyse des standards ainsi obtenus a permis d'attribuer à chacune des dix lactones possibles un temps de rétention (**Tableau 1**, publication n°1 en **Annexes**) associé à un spectre de masse en ionisation par impact électronique. Dans le but de quantifier les produits obtenus, j'ai réalisé une droite d'étalonnage au moyen de standards synthétiques fournis présentée en **Figure S3 et S4** de la publication n°1 en **Annexes**, sur laquelle j'ai pu reporter les abondances relatives obtenues au cours de mes analyses.

Pour les deux constructions dans lesquelles le module 2 de la PKS modèle a été remplacé par le onzième module de la PKS synthétisant l'amphotéricine (Amph M11) contenant un domaine KR de type A2, les analyses ont montré la formation des cétolactones **5**, des lactones **2** réduites par un domaine KR de type A au niveau du C3 et non épimérisées au niveau du C2, et des lactones **4** réduites par un domaine KR de type A au niveau du C3 et épimérisées au niveau du C2. Les lactones **4** réduites et épimérisées, constituent le produit majoritaire, avec un rendement de 36 mg.L<sup>-1</sup> et de 20,2 mg.L<sup>-1</sup> pour les constructions de type A et B respectivement. Ces résultats montrent une nette amélioration par rapport à l'étude consistant à échanger des domaines KR réducteurs et épimérisants seuls (Annaval *et al.*, 2015). En effet, un faible rendement de cétolactones **5** avait été observé (environ 3 mg.L<sup>-1</sup>), un bon rendement de 14 mg.L<sup>-1</sup> en lactones **4** réduites par un domaine KR de type A et épimérisées, ainsi que la présence de traces de lactones **2** réduites par un domaine KR de type A et non épimérisées. En utilisant les bornes classiques de module, le rendement global en lactone a augmenté de près de trois fois pour atteindre 48,1 mg.L<sup>-1</sup>, incluant 75,0 % des lactones **4** (36 mg.L<sup>-1</sup>). Les produits restants ont été répartis uniformément entre les cétolactones **5** observées précédemment (12,5 %) et les lactones **2** par un domaine KR de type A mais non épimérisées (12,5 %). Ainsi, la quantité totale de lactone réduite de l'une ou de l'autre stéréochimie a également été augmentée par rapport à l'expérience d'échange de KR (87,5 % contre 82,3 %). L'utilisation des bornes alternatives de module (construction de type B) a produit une augmentation d'environ 2 fois en rendement total en lactone par rapport à l'échange de KR seul (35,7 mg.L<sup>-1</sup>), mais a abouti à une plus grande proportion des lactones **2** (19 %), ainsi que de cétolactones **5** (24,4 %). Ainsi, ces résultats suggèrent que l'interaction native entre la KS<sub>n+1</sub> et l'ACP<sub>n+1</sub> au sein du module serait plus important que le maintien de l'interaction entre le domaine ACP<sub>n</sub> et le domaine KS<sub>n+1</sub>, afin d'obtenir le produit attendu.

L'analyse a été aussi effectuée avec les modules 1 et 11 de la PKS responsable de la synthèse de la nystatine (Nys M1 et Nys M11 respectivement), dont les domaines KR de type A2 partagent des pourcentages d'identité de séquence élevés avec leur équivalent chez la AMPHS (84,1% et 69,3%, respectivement). L'échange de ces modules, suivant les

## Résultats

constructions de type A et B, a permis de générer les lactones **4** attendues et, des deux modules, l'échange du module 11 de la PKS responsable de la synthèse de la nystatine (NYSS) a permis d'obtenir le rendement le plus élevé (**Tableau 2**, publication n°1 en **Annexes**). Ces résultats sont semblables à l'échange des domaines KR<sub>1</sub> et KR<sub>11</sub> en lieu et place de la KR<sub>2</sub> de la DEBS (Annaval *et al.*, 2015), où l'échange du domaine KR<sub>11</sub> était le plus productif. Ainsi, ces domaines KR qui interviennent à un stade plus tardif au sein de la PKS fonctionneraient avec plus d'efficacité dans le contexte du module de 2 de la DEBS. Encore une fois, l'échange de module complet dans chaque cas a produit des rendements plus élevés en lactones **4**, par rapport à l'échange de module alternatif (**Tableau 2**, publication n°1 en **Annexes**). De plus, la fidélité stéréochimique des deux modules de la NYSS s'est montré modérément plus élevée que pour l'échange du domaine KR<sub>11</sub> de la AMPHS (traces et 8,5% de lactones **2**, respectivement).

Pour la construction dans laquelle le module 2 de la PKS modèle a été remplacé par le cinquième module de la PKS responsable de la synthèse de l'oligomycine (Olm M5) contenant un domaine KR de type A2, nous avons mis en évidence la formation de cétolactones **5** mais pas de lactones réduites **4** et **2**. Ces résultats sont similaires à ceux obtenus lors des expériences précédentes (Annaval *et al.*, 2015), bien que nous observions une diminution de la quantité absolue des cétolactones produites. En effet, un bon rendement en cétolactones **5** avait été obtenu (environ 52 mg.L<sup>-1</sup>) contre 14,6 mg.L<sup>-1</sup> (Olm M<sub>5</sub>-A) et 6 mg.L<sup>-1</sup> (Olm M<sub>5</sub>-B) dans cette étude. Ces résultats indiquent que bien que la PKS est fonctionnelle (les cétolactones **5** sont produites) mais la KR du module 5 de la PKS synthétisant l'oligomycine (OLMS) n'est pas en mesure de réduire le polycétide en formation. Ce résultat semblerait indiquer une spécificité de substrat de la part de la KR du module 5 de la OLMS qui n'accepterait pas des substrats plus courts que son substrat natif, expliquant l'absence de réduction des cétolactones **5**. Néanmoins, le fait que l'échange de module entier ait permis d'obtenir un meilleur rendement en cétolactones **5** (14,6 mg.L<sup>-1</sup>) que l'échange de module partiel (6 mg.L<sup>-1</sup>) montre que cette première stratégie permet d'obtenir de meilleurs rendements.

Les échanges de modules provenant de la PKS responsable de la synthèse de la bafilomycine (BAFS) et leurs homologues chez la PKS responsable de la synthèse de l'hygrobafilomycine (HYGS), qui contiennent tous des domaines KR de type A2, n'ont pas permis d'obtenir les produits attendus. Dans le cas de l'échange de module entier, l'absence de produit peut provenir de la défaillance du transfert intermodulaire de l'intermédiaire entre le module 1 de la DEBS et le module inséré. Dans le cas des échanges de modules partiels, qui ont conservé l'interface native ACP<sub>1</sub>/KS<sub>2</sub> du DEBS 1-TE, le problème s'est probablement posé

au stade de la réaction de condensation décarboxylative entre le domaine KS<sub>2</sub> de la DEBS et le domaine ACP du système PKS hétérologue.

Concernant l'échange de module contenant un domaine KR de type B2, pour les deux constructions dans lesquelles le module 2 de la PKS modèle a été remplacé par le premier module de la PKS responsable de la synthèse de la lankamycine (Lkm M1), nous avons mis en évidence la formation des cétolactones **5**, des lactones **3** réduites par un domaine KR de type B au niveau du C3 et non épimérisées au niveau du C2 et des lactones **1** réduites par un domaine KR de type B et épimérisées. Le rendement en lactones réduites a été sensiblement amélioré par rapport à l'échange de domaine KR seul, pour lequel les cétolactones **5** prédominaient (6 mg.L<sup>-1</sup>) et seulement des traces de lactones réduites **1** ont été observées (Annaval *et al.*, 2015). Contrairement à l'échange de module contenant des domaines KR de type A2, le rendement en lactone **1** était supérieur lors de l'échange de module partiel de type B au lieu du module complet de type A (18,7 mg.L<sup>-1</sup> soit 75 % de lactones **1**, contre 6,2 mg.L<sup>-1</sup> soit 56 % de lactones **1**, **Tableau 2** de la publication n°1 en **Annexes**), bien que des lactones réduites par un domaine KR de type B et non épimérisées aient été également observées. De plus, les lactones **3** n'ont pas été observées lors de l'échange de module entier de type A, ce qui suggère que l'échange incluant la KS<sub>1</sub> de la LKMS améliore la fidélité du contrôle de la stéréochimie par le module. Le plus faible rendement obtenu pour la construction Lkm M1-A peut-être dû à une moindre efficacité du transfert intermodulaire de l'intermédiaire.

Cependant, lorsque le septième module de la lasalocide synthase (Las M7) possédant aussi un domaine KR de type B2, est introduit à la place du second module au sein du DEBS 1-TE, aucun produit n'a été détecté. Pourtant, lors de nos travaux précédents, l'échange du domaine KR<sub>7</sub> de la LASS avait permis de produire des cétolactones **5** (8 mg.L<sup>-1</sup>) et des traces de lactones **1** réduites par un domaine KR de type B au niveau du C3 et épimérisées au niveau du C2. Cette absence de produits a aussi été observée pour le premier module de la septième sous-unité de la PKS responsable de la médiomycine, qui contient un domaine KR de type B2. Comme pour l'échange de module provenant de la BAFS et de la HYGS, ces PKS chimères se sont montrées non-fonctionnelles, et ce même en maintenant l'une des interfaces natives ACP<sub>1</sub>-KS<sub>2</sub> ou KS<sub>2</sub>-ACP<sub>2</sub>.

Concernant les échanges de modules en remplaçant l'ACP<sub>2</sub> jusqu'au site *SacI*, les PKS chimères générées étaient non-fonctionnelles. De manière surprenante, le système parental DEBS 1-TE, qui reste fonctionnel malgré la présence d'un domaine ACP hybride, possède une forte activité et permet d'obtenir le produit attendu avec une grande fidélité stéréochimique (Kellenberger *et al.*, 2008). Nonobstant ce dernier phénomène, ces résultats montrent la nécessité de conserver les domaines ACP intacts dans des expériences d'échanges de module.

**B. Publication n°1**

L'ensemble de ces données ont abouti à la rédaction d'un article intitulé « Manipulating polyketide stereochemistry by exchange of polyketide synthase modules » paru dans *Chemical Communications* (DOI : 10.1039/d0cc05068g ; facteur d'impact : 5,996) qui est fourni en **Annexes**.



## Résultats

Nous avons ainsi exploré la possibilité d'introduire une épimérisation du groupement méthyl en C3 de la 6-désoxyérythronolide B (6-deB), et aussi une inversion de la configuration du groupement hydroxyle en C2 (il est à noter que les enzymes post-PKS associées ne sont pas introduites dans cette étude). Pour cela, les modules ayant montrés les meilleurs rendements lors de l'étude précédente (Massicard *et al.*, 2020) ont été employés, c'est-à-dire le onzième module en entier (KS → ACP) contenant un domaine KR de type A2 de la AMPHS pour introduire une épimérisation en C2 (**Figure 45B**) et le module 1 partiel (AT → ACP) contenant un domaine KR de type B2 de la LKMS pour modifier le sens de la réduction en C3 et introduire une épimérisation en C2 (**Figure 45C**). Dans ce cadre, nous collaborons avec Isomerase (Chesterford Research Park, Cambridge) qui nous ont fourni une souche de *Sac. erythraea* possédant une plateforme avec 4 sites *attB* (TG1, SV1,  $\Phi$ C31 et  $\Phi$ BT1) (Fayed *et al.*, 2015) en lieu et place du BGC *ery*. Cette plateforme peut être exploitée avec 3 plasmides correspondant à chacune des 3 sous-unités et un plasmide correspondant aux gènes associés aux enzymes post-PKS, séparément exprimés en utilisant le promoteur natif de *ery* (*eryAlp*) dans les plasmides avec différentes intégrases (TG1, SV1, phiC31 et phiBT1).

En premier lieu, le second module de la PKS a été remplacé soit par le onzième module de la AMPHS, soit par le premier module de la LKM. Les modules ont été introduits de manière à préserver l'interaction native entre les DD, bien qu'une interface ACP / KS<sub>3</sub> hétérologue et donc non native ait évidemment été introduite. Le plasmide portant la séquence ADN du second module de DEBS a été modifié dans un premier temps par la technique de « PCR-targeting » (Gust *et al.*, 2003) pour l'échanger avec la cassette de sélection/contre-sélection TetA-SacB (Li *et al.*, 2013) (Don de Donald Court, « Frederick National Laboratory for Cancer Research, Leidos Biomedical Research », Frederick, Maryland. Puis dans un second temps, la cassette doit être remplacée par le onzième module de la AMPHS ou le premier module de la LKMS. Malheureusement, à ce jour, cette modification n'a pas encore abouti et est toujours en cours.

La modification par échange de module de la DEBS introduit une interface non-native entre la KS<sub>3</sub> et le domaine ACP<sub>2</sub>, et la mutation de l'acide aminé alanine en position 154 (notation basée sur la structure du fragment KS<sub>3</sub>-AT<sub>3</sub> de la DEBS (Tang *et al.*, 2007)) du domaine KS en tryptophane a été décrit comme augmentant la tolérance du domaine KS<sub>3</sub> pour des substrats non-natifs (Murphy *et al.*, 2016). C'est pourquoi le domaine KS<sub>3</sub> de la DEBS a été modifié par mutagénèse.

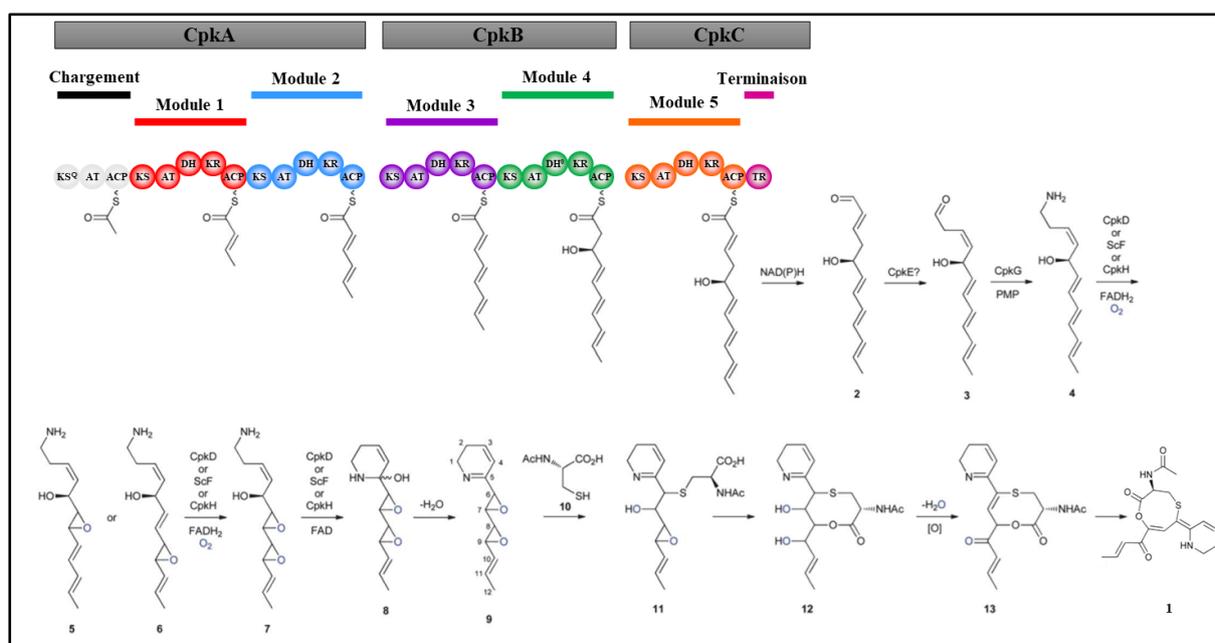
Les plasmides recombinants ainsi obtenus pourront être ensuite transformés par conjugaison dans la souche *Sac. erythraea* dans le but de faire produire les différents dérivés.

## II. Étude de la localisation des systèmes PKS au sein des hôtes bactériens

### 1. Activation du BGC *cpk* et visualisation de la coelimycine P2

Au cours des 50 dernières années, *S. coelicolor* a été le modèle majeur pour l'étude de la production des antibiotiques et de son contrôle (Heul *et al.*, 2018). Parmi le large répertoire en métabolites spécialisés de ce *Streptomyces*, *S. coelicolor* A3 (2) synthétise 4 composés antimicrobiens : la coelimycine A (CPKA, précurseur des coelimycines P1 et P2), l'antibiotique dépendant du calcium (CDA), l'undécylprodigiosine de couleur rouge (RED) et l'actinorhodin de couleur bleue (ACT) (Liu *et al.*, 2013). Le groupe de gènes *cpk*, associé à une PKS de type I de type *cis*-AT, a été identifié pour la première fois en 1997 par la technique d'hybridation de sonde ADN *via* le domaine AT spécifique du malonyl-CoA (Kuczek *et al.*, 1997). En 2002, le génome de *S. coelicolor* a été publié (Bentley *et al.*, 2002) permettant notamment d'annoter le groupe de gène *cpk* (Pawlik *et al.*, 2007). Les produits issus de ce groupe de gènes sont visualisés à partir de 2010, sous forme de pigments jaunes sécrétés dans le milieu de culture (yCpk, de l'anglais *yellow coelimycin*) (Gottelt *et al.*, 2010; Pawlik *et al.*, 2007), plus tard identifiés comme coelimycine P1 et P2 (Gomez-Escribano *et al.*, 2012), ainsi qu'un produit sans couleur mais possédant une propriété antimicrobienne (abCpk) (Gottelt *et al.*, 2010), identifié plus tard en tant que coelimycine A (Challis, 2014).

L'expression des gènes *cpk* au tout début de la phase de transition de croissance est un événement précoce du passage métabolique du métabolisme primaire au métabolisme secondaire (Nielsen *et al.*, 2016). Nous avons alors commencé par trouver un milieu permettant de visualiser la coelimycine, indiquant une expression du groupe de gènes *cpk* et l'activité de la CPKS (**Figure 46**), à l'aide de la souche FM145 de *S. coelicolor* (**Tableau 2**). Cette souche a pour particularité de posséder une auto-fluorescence fortement réduite (Willemse and van Wezel, 2009), et ainsi elle est un candidat idéal pour étudier des protéines faiblement exprimées par microscopie à fluorescence.

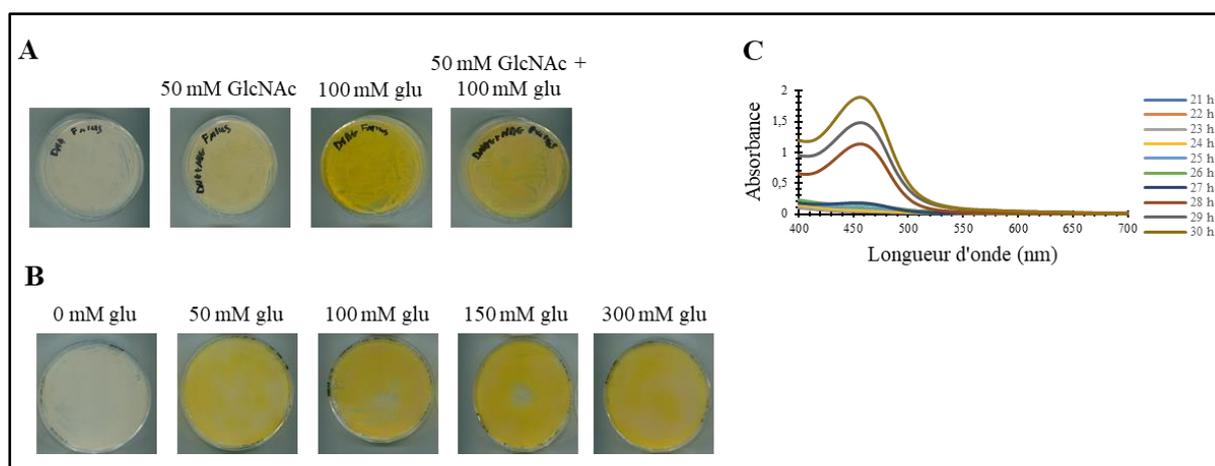


**Figure 46 : Représentation de la PKS responsable de la synthèse de la coelimycine (d'après Gomez-Escribano *et al.*, 2012).**

Légende : KS : cétylsynthase, KS<sup>Q</sup> : domaine cétylsynthase avec le résidu Cys du site actif remplacé par une glutamine, AT : acyltransférase, ACP : *acyl carrier protein*, DH : déshydratase, KR : cétyroréductase, TR : thioester réductase. Le domaine DH dans le module 4 est prédit pour être inactif. Les structures prédites des intermédiaires liés aux domaines ACP de chaque module de la PKS sont indiquées. La configuration du groupe hydroxyle C5 en **2** a été déduite de l'analyse de séquence du domaine KR du module 4.

Nous avons choisi d'activer le BGC *cpk* en modifiant uniquement le milieu de culture, car nous souhaitons conserver le métabolisme et la physiologie d'origines du *Streptomyces*, dans le but d'étudier l'organisation macromoléculaire de la CPKS dans son contexte natif. La production du pigment jaune associée à la coelimycine était possible dans les milieux DNA, DNA supplémenté en L-glutamate (Gottelt *et al.*, 2010) et M79 (Pawlik *et al.*, 2010). Nous n'avons pas observé de production du composé jaune pour le milieu minimal, même additionné de L-glutamate. En accord avec Gomez-Escribano *et coll.* (Gomez-Escribano *et al.*, 2012), la coelimycine observée correspond à l'adduit réactionnel entre l'intermédiaire **8** montré dans la **Figure 46** et le glutamate, correspondant alors à la coelimycine P2. Puisque la production du composé jaune était maximale sur DNA supplémenté avec du L-glutamate (**Figure 47A**), ce milieu a été utilisé pour le reste de l'étude. Il a été noté que la *N*-acétylglucosamine (GlcNAc), connue pour favoriser la production d'antibiotiques chez *S. coelicolor* (Rigali *et al.*, 2008), n'a pas amélioré la production de coelimycine P2 dans nos conditions (**Figure 47A**). Initialement, la concentration de L-glutamate utilisée était de 325 mM (Gottelt *et al.*, 2010), mais nous avons montré que la production de coelimycine P2 apparaissait à partir de 50 mM de L-glutamate (**Figure 47B**). La souche FM145 de *S. coelicolor* a été aussi

cultivée à différents temps de culture (21 h à 30 h) dans la version liquide du milieu de production de la coelimycine P2 (DNB + 100 mM d'acide glutamique). Les spectres d'absorption (400–700 nm) des surnageants de culture ont été déterminés en utilisant le milieu DNB supplémenté en acide glutamique comme blanc. Comme pour le milieu solide, la production de coelimycine P2 était possible à partir de 27 h avec un maximum d'absorption de 460 nm (**Figure 47C**). La visualisation de coelimycine P2 en milieu de production liquide et solide est cohérente avec les études précédentes (Gottelt *et al.*, 2010; Pawlik *et al.*, 2010).



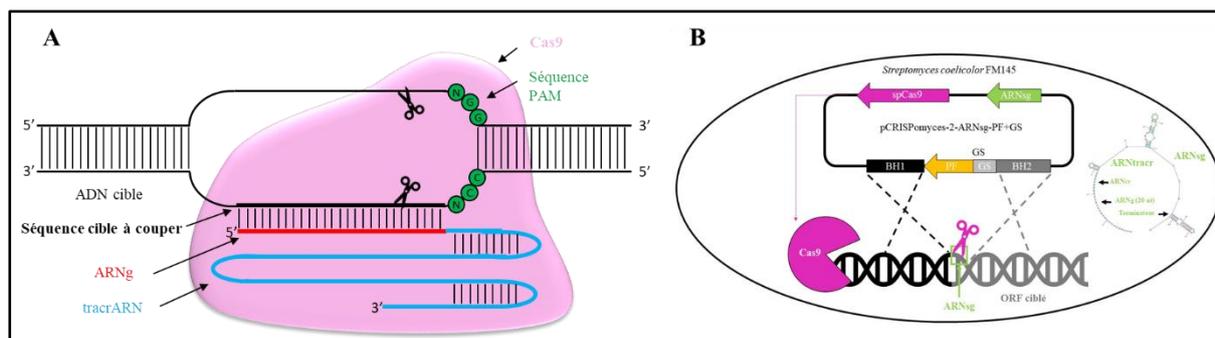
**Figure 47 : Production du pigment jaune associé à la coelimycine P2 sur DNA. (A)** Production en milieu solide DNA supplémenté ou non avec de l'acide L-glutamique (glu) et/ou de la *N*-acétylglucosamine (GlcNAc). **(B)** Production en milieu solide DNA supplémenté avec plusieurs concentrations d'acide glutamique. **(C)** Production en milieu liquide DNB supplémenté avec 100 mM d'acide glutamique à différents temps de culture. Pour quantifier de manière relative la production de pigment jaune, les spectres d'absorption ( $A_{400-700}$ ) ont été déterminés.

## 2. Construction des protéines de fusion

### A. Insertion des protéines fluorescentes par la technique CRISPR-Cas9

Pour pouvoir localiser les PKS au sein des cellules bactériennes, la première et/ou la dernière sous-unité de la CPKS a été fusionnée à une protéine fluorescente *via* un linker protéique flexible de sept glycines. Nous avons supposé que le marquage des extrémités la CPKS ne devraient pas affecter les associations entre les sous-unités consécutives médiées par les domaines de docking (Broadhurst *et al.*, 2003). Nous avons employé deux protéines fluorescentes qui sont connues pour fonctionner chez les *Streptomyces* – mCherry (Schlimpert *et al.*, 2017) et eGFP (Celler *et al.*, 2016). La création des fusions a été effectuée avec l’outil d’ingénierie CRISPR/Cas9, qui a été employé récemment pour modifier les génomes de différentes espèces de *Streptomyces* (Cobb *et al.*, 2015 ; Tong *et al.*, 2015).

La première étape a été de concevoir deux ARNsg grâce au logiciel en ligne CRISPy-web (Blin *et al.*, 2016) pour guider la protéine Cas9 vers son site cible. Chaque ARN guide a été assemblé dans le pCRISPomyces-2 par la technique « Golden Gate » (Engler *et al.*, 2008). La seconde étape a été d’adapter le plasmide pCRISPomyces-2 pour pouvoir utiliser la technique CRISPR/Cas9 afin de modifier l’ADN génomique de la souche bactérienne (**Figure 48A**) et ainsi, d’insérer la séquence correspondant à la protéine fluorescente au niveau du locus ciblé. Pour cela, la stratégie adoptée a été de construire une matrice par PCR de chevauchement contenant deux régions d’homologie de part et d’autre du site de coupure de l’enzyme Cas9 (**Figure 48B**). Entre ces deux régions, se trouve la séquence correspondant à la protéine fluorescente qui pourra donc être insérée dans le génome de *S. coelicolor* lors de la réparation de l’ADN par recombinaison homologue au niveau du site cible. Afin d’éviter que la protéine Cas9 effectue une coupure sur la plasmide contenant la matrice, un gène synthétique a été créé en modifiant les codons de la séquence correspondant à l’ARNcr contenu dans l’ARNsg. Les différents fragments d’ADN qui composent la matrice d’édition ont été assemblés par PCR de chevauchement puis insérée dans le vecteur pCRISPomyces-2 par clonage de type Gibson (Gibson *et al.*, 2009). Les jonctions matrice/vecteur et ARNsg/vecteur ainsi que toutes les régions amplifiées par PCR ont été contrôlées par séquençage.



**Figure 48 : Stratégie pour modifier le génome de *S. coelicolor* FM145 par la technique CRISPR/Cas9.** (A) Représentation schématique du complexe d'édition du génome CRISPR/Cas9. Ce complexe est composé de la protéine Cas9 de *Streptococcus pyogenes*, de la fusion entre l'ARNcr et l'ARNtracr appelé alors l'ARNsg, et de l'ADN cible. (B) Représentation schématique de la stratégie d'édition du génome de *S. coelicolor* FM145. La matrice pour la recombinaison homologue après la cassure double brin médiée par la protéine Cas9 active est composée de deux régions d'homologies (BH1 en noir et BH2 en gris), qui bordent l'ORF correspondant à la protéine fluorescente (en orange). La nature des codons correspondant à la cible de l'ARNsg a été modifiée pour empêcher la protéine Cas9 de couper le plasmide dans l'hôte *Streptomyces* au moyen d'un gène synthétique (GS). La structure de l'ARNsg est également indiquée.

La troisième étape a été de transférer par conjugaison bactérienne les plasmides pCRISPOmyces-2 génétiquement modifiés dans la souche FM145 de *S. coelicolor*. Les exconjugants présentant une résistance à l'apramycine ont été confirmés par repiquages sur milieu sélectif. La quatrième étape a été de vérifier si la modification génétique souhaitée a bien eu lieu au sein du chromosome de *S. coelicolor*. Pour cela, les exconjugants ont été testés par PCR sur colonie en utilisant des oligodésoxyribonucléotides spécifiques de l'ADN génomique de *S. coelicolor* afin de ne pas amplifier le plasmide. Puis, lorsque les exconjugants comportaient la modification souhaitée, le pCRISPOmyces-2 a été éliminé comme décrit par Cobb *et coll.*, (Cobb *et al.*, 2015). Les exconjugants ne pouvant plus croître sur milieu supplémenté en apramycine, montrant alors la perte du pCRISPOmyces-2, ont été mis en culture dans le but d'extraire l'ADN génomique. De cette façon, les différentes modifications souhaitées dans le chromosome des exconjugants de *Streptomyces coelicolor* ont pu être contrôlées par séquençage.

Les mutants générés au cours de cette étude sont présentés dans le **Tableau 2**. Les jonctions du site d'insertion de chaque mutant ont été contrôlées par séquençage. En plus des sous-unités de PKS, nous avons choisi de construire une fusion entre l'extrémité carboxy-terminale du potentiel transporteur CpkF et la protéine fluorescente eGFP. CpkF a été identifié comme étant l'hypothétique transporteur de type MFS (Major Facilitator Superfamily) codé au sein du groupe de gènes *cpk* (Pawlik *et al.*, 2007). Si la formation et la localisation d'un éventuel

## Résultats

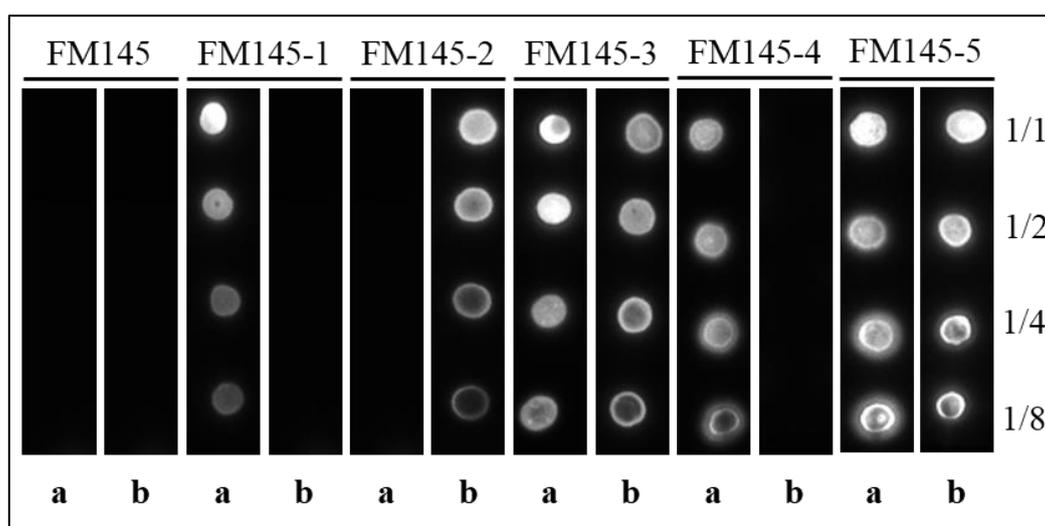
mégacomplexe chez *Streptomyces* sont cohérentes avec l'assemblage d'une usine de biosynthèse (Straight *et al.*, 2007), alors celui-ci est potentiellement couplé à des machines de sécrétion dédiées pour éviter l'accumulation de composés toxiques dans le cytoplasme. Si nous observons une organisation macromoléculaire en mégacomplexe de la CPKS chez *Streptomyces coelicolor*, et qu'il se trouve proche de la membrane comme dans l'observation initiale chez *Bacillus* (Straight *et al.*, 2007), alors l'étude de la localisation du transporteur CpkF avec le mégacomplexe permettra d'obtenir des informations sur le couplage entre la synthèse et l'export du métabolite.

**Tableau 2 : Dérivés de la souche FM145 de *S. coelicolor* obtenus ou en cours pour l'étude de la localisation d'un mégacomplexe.**

Souche	Génotype	Description	Source
FM145	/	FM145 montre une diminution d'environ trois fois ( $2,9 \pm 0,3$ sur la base d'au moins 100 hyphes) en autofluorescence dans le canal vert et de plus de cinq fois ( $7 \pm 1$ ) dans le canal rouge par rapport à M145	(Willemse and van Wezel, 2009)
FM145-1	<i>mCherry-cpkA</i>	Insertion de <i>mCherry</i> à l'extrémité 5' de l'ORF <i>cpkA</i>	Ce projet
FM145-2	<i>cpkC-eGFP</i>	Insertion de <i>eGFP</i> à l'extrémité 3' de l'ORF <i>cpkC</i>	Ce projet
FM145-3	<i>mCherry-cpkA</i> ; <i>cpkC-eGFP</i>	Insertion de <i>mCherry</i> à l'extrémité 5' de <i>cpkA</i> et insertion de <i>eGFP</i> à l'extrémité 3' de <i>cpkC</i>	Ce projet
FM145-4	<i>cpkF-mCherry</i>	Insertion de <i>mCherry</i> à l'extrémité 3' de <i>cpkF</i>	Ce projet
FM145-5	<i>cpkC-eGFP</i> ; <i>cpkF-mCherry</i>	Insertion de <i>mCherry</i> à l'extrémité 5' de <i>cpkA</i> et insertion de <i>mCherry</i> à l'extrémité 3' de <i>cpkF</i>	Ce projet
FM145-6	$\Delta cpkC$	Délétion de <i>cpkC</i>	Ce projet
FM145-7	<i>mCherry-cpkA</i> ; $\Delta cpkC$	Insertion de <i>mCherry</i> à l'extrémité 5' de <i>cpkA</i> et délétion de <i>cpkC</i>	Ce projet
FM145-8	$\Delta cpkA$	Délétion de <i>cpkA</i>	Ce projet
FM145-9	<i>cpkC-eGFP</i> ; $\Delta cpkA$	Insertion de <i>eGFP</i> à l'extrémité 3' de <i>cpkC</i> et délétion de <i>cpkA</i>	Ce projet
FM145-10	$\Delta cpkA$ ; <i>cpkF-mCherry</i>	Délétion de <i>cpkA</i> et insertion de <i>mCherry</i> à l'extrémité 3' de <i>cpkF</i>	Ce projet
FM145-11	$\Delta cpkF$	Délétion de <i>cpkC</i>	Ce projet
FM145-12	$\Delta cpkF$ ; <i>cpkC-eGFP</i>	Délétion de <i>cpkC</i> et insertion de <i>mCherry</i> à l'extrémité 3' de <i>cpkF</i>	Ce projet

## B. Contrôle de l'expression et de la fonction des sous-unités chimériques

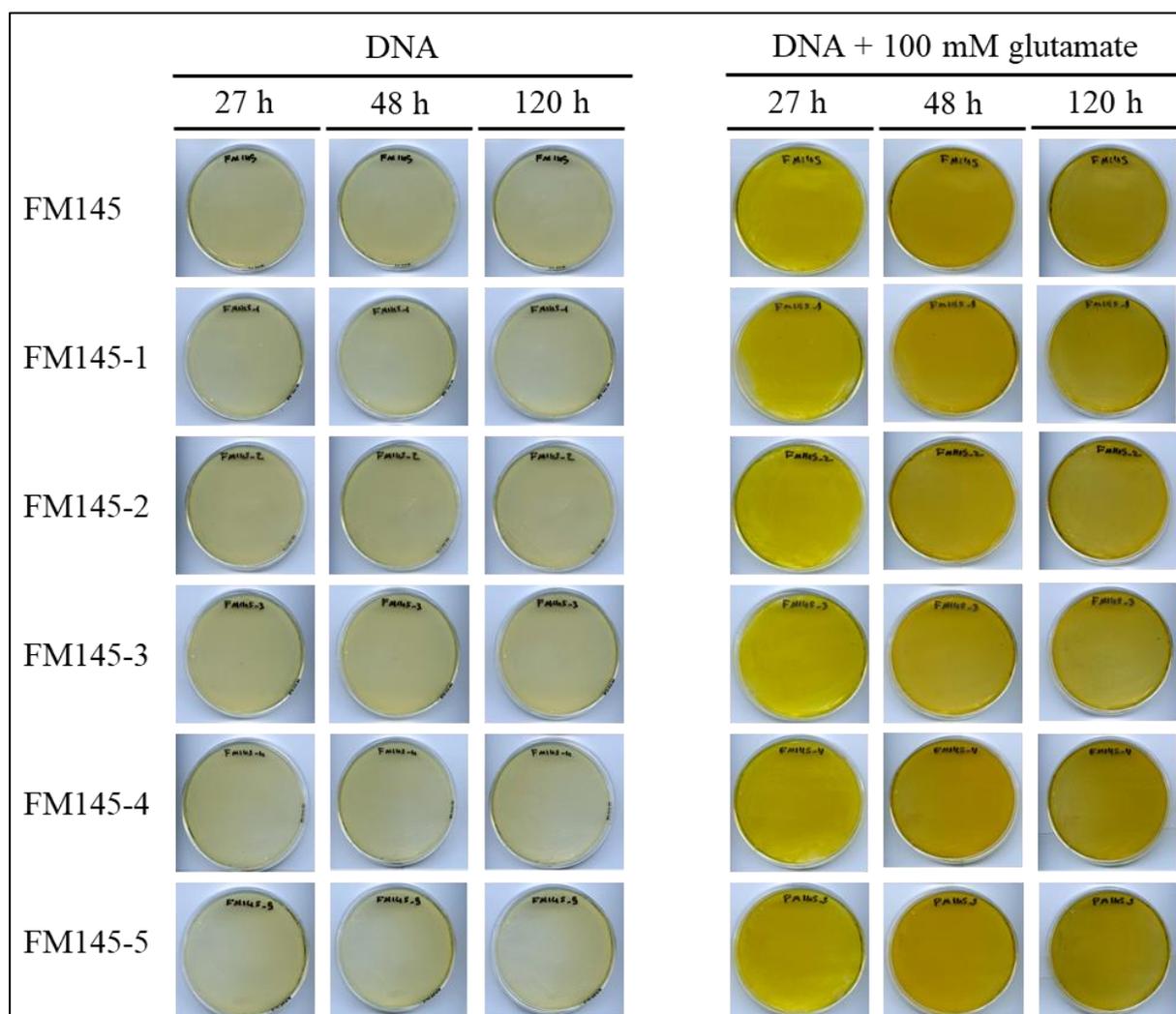
Dans les différentes constructions générées, l'expression des différents gènes du BGC *cpk*, incluant les gènes fusionnés à l'ORF des protéines fluorescentes mCherry ou eGFP, sont soumis à leur promoteur natif. C'est pourquoi leur détection par des techniques d'immunoblotting est nécessaire afin de contrôler si les modifications génétiques n'ont pas engendré de perturbation aux différents niveaux de leur régulation transcriptionnelle et traductionnelle. Cependant, les sous-unités de polycétide synthases sont de très grande taille (CpkA : 476,05 kDa, CpkC : 229,83 kDa), ce qui rend leur analyse par la technique de Western-blot difficile. Ainsi, les analyses ont été effectuées dans un premier temps par la technique de Dot-blot en utilisant des anticorps anti-protéines fluorescentes, permettant alors de détecter la protéine fluorescente. La **Figure 49** montre donc la spécificité de l'anticorps dirigé contre la protéine fluorescente, permettant de confirmer la présence de mCherry chez la souche FM145-1, la présence de l'eGFP chez la souche FM145-2, la présence de mCherry et de l'eGFP chez la souche FM145-3, la présence de mCherry dans la souche FM145-4 et la présence de mCherry et de l'eGFP chez la souche FM145-5. Cette technique a ainsi permis de montrer que la protéine fluorescente dont l'ORF a été inséré par la technique de CRISP/Cas9 était bien exprimée dans les souches dérivées de *S. coelicolor* FM145, mais elle n'a pas permis de savoir si les protéines fluorescentes étaient bien fusionnées en contrôlant les tailles des différentes protéines de fusion générées lors de cette étude.



**Figure 49 : Contrôle de l'expression des protéines de fusion.** Dot-blot utilisant les souches FM145-1, FM145-2, FM145-3, FM145-4 et FM145-5 pour contrôler la spécificité des anticorps vis-à-vis de leur protéine fluorescente respective. Les dilutions des extraits bruts sont indiquées sur la droite et renseignent sur le rapport signal / quantité de protéine. Les pistes **a** représentent le traitement avec l'anticorps anti-mCherry (1/4000<sup>ème</sup>). Les pistes **b** représentent le traitement avec l'anticorps anti-eGFP (1/3000<sup>ème</sup>).

## Résultats

La production de coelimycine P2 nécessite que la CPKS soit fonctionnelle, ainsi les mutants de la souche FM145 de *Streptomyces* ont aussi été testé sur le milieu DNA supplémenté en glutamate (**Figure 50**). De cette manière, nous avons contrôlé si malgré l'insertion des protéines fluorescentes, la CPKS restait fonctionnelle. Comme observé précédemment, la production a été optimale dans le milieu DNA supplémenté en glutamate. Nous avons montré que la CPKS reste fonctionnelle malgré l'insertion d'une ou de deux protéines fluorescentes à l'une des extrémités (ou les deux) de la CPKS, ainsi qu'à l'extrémité C-terminale du transporteur associé CpkF. La quantification de la production de coelimycine P2 pour les souches contenant les fusions en comparaison à la souche FM145 permettra de confirmer la conservation de l'efficacité des CPKS de fusion.



**Figure 50 : Production de la coelimycine P2 sur le milieu DNA modifié par l'ajout de glutamate.** La souche FM145 et ses dérivées ont été ensemencées à raison de  $10^8$  spores et incubées à 30 °C. La production du pigment jaune a été enregistrée aux temps indiqués.

### 3. Localisation sous-cellulaire des protéines de PKS

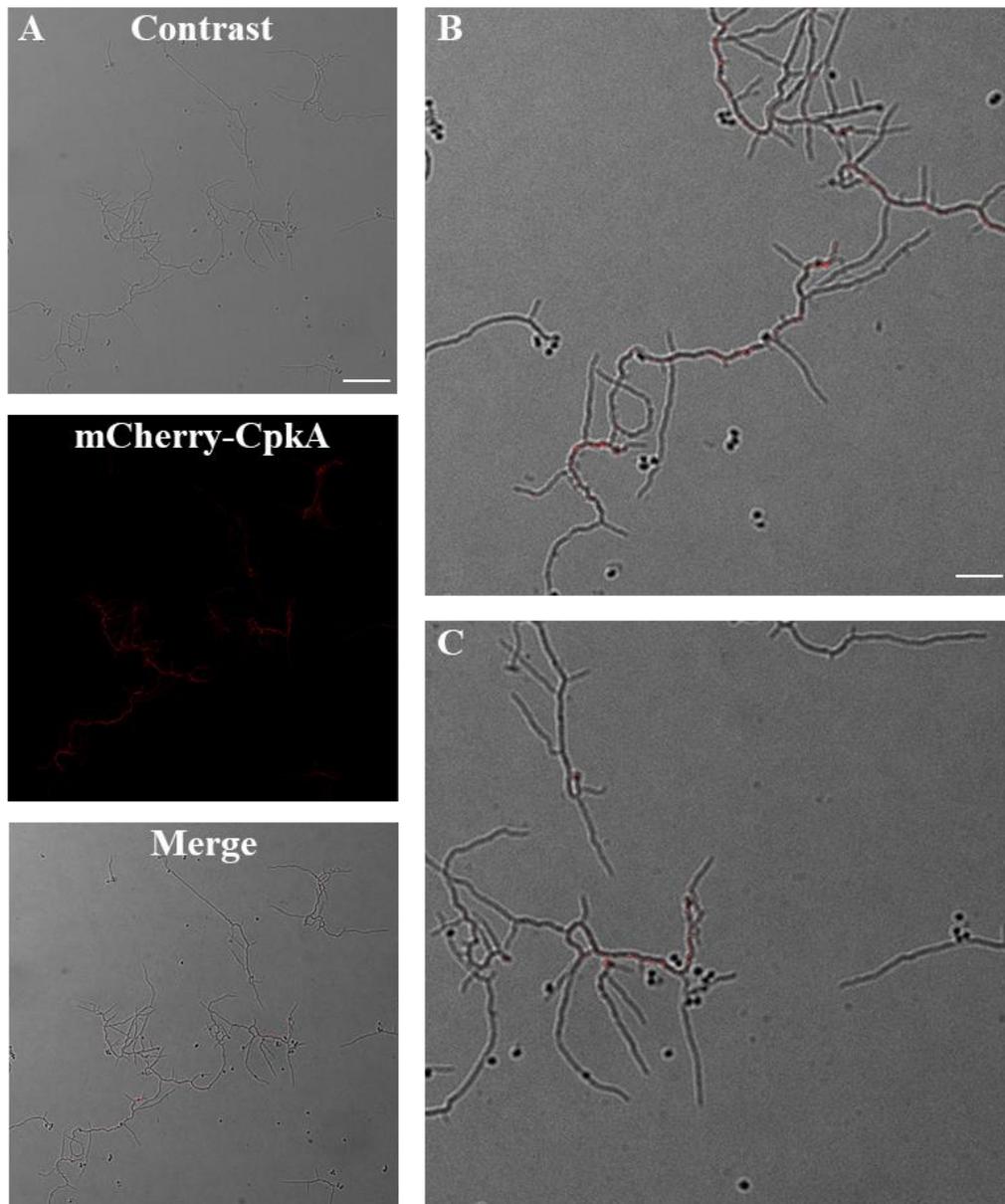
#### A. CpkA et CpkC localisent en loci dans les hyphes de *S. coelicolor*

A la suite des contrôles des souches contenant les protéines de fusion, nous avons obtenu des images de la localisation des sous-unités de fusion de la CPKS par microscopie à fluorescence en collaboration avec la Dr Paula Yagüe (Institut de Biologie de l'Université de Leiden, Pays-Bas et Faculté de Médecine de Université d'Oviedo, Espagne). La production des polycétides étant liée à la phase de croissance des cellules de *Streptomyces* (classiquement aux phases tardives (Heul *et al.*, 2018)) nous avons d'abord analysé la fluorescence liée aux protéines PKS dans les cellules à différents temps de culture, afin de déterminer le temps pour lequel le maximum de fluorescence est observé. Pour cela, des images ont tout d'abord été acquises par microscopie à contraste interférentiel (DIC en anglais), suivi par des images fluorescentes. La superposition des deux images nous a donné la localisation des signaux fluorescents au sein des cellules.

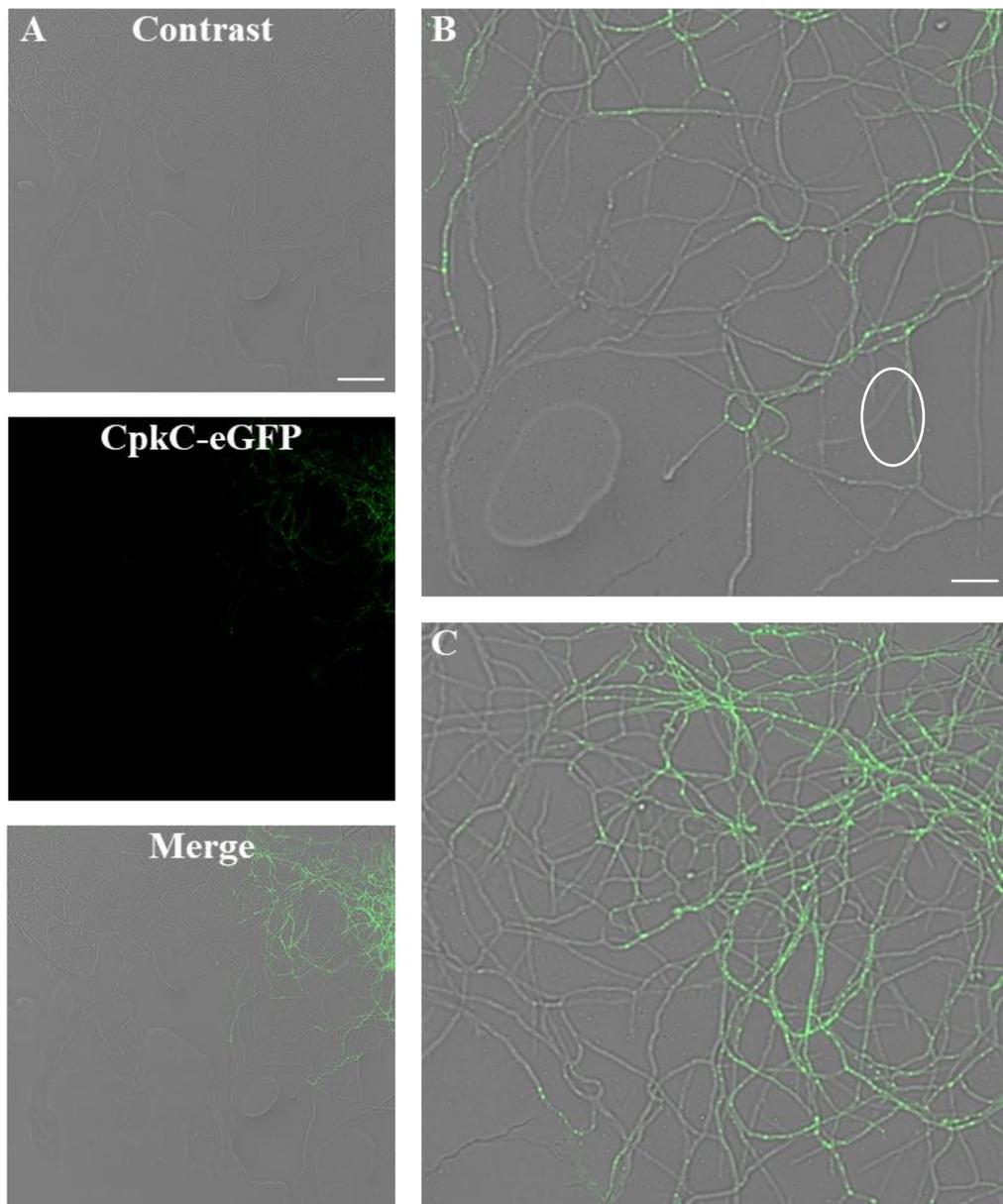
Les souches exprimant les protéines de fusion ont été cultivées jusqu'à 24 h, ce qui correspondait à la croissance des hyphes aériens et à la production de coelimycine P2. Nous avons d'abord analysé les fusions mCherry-CpkA et CpkC-eGFP de manière individuelle dans les souches FM145-1 et FM145-2 respectivement. Comme montré dans les **Figures 51** et **52**, de fortes intensités de fluorescence ont été observées dans les hyphes de *Streptomyces*. L'analyse de l'ensemble des échantillons traités a montré que 10 % des hyphes observés ont exprimé les fusions de protéines, et que les loci de fluorescence ont été observés tout le long des hyphes. La localisation des sous-unités fusionnées individuellement dans les souches FM145-1 et-2 est similaire à celle de la PKS responsable de la synthèse de la bacillaene (Straight *et al.*, 2007). En effet, la majorité des cellules de *Bacillus* avaient montré un seul locus dans 69 % des cas, une minorité (6 %) avait possédé deux loci par cellule, et 25 % des cellules avaient révélé une absence de signaux fluorescents. Une explication possible de la faible quantité d'hyphes exprimant Cpk a été fournie récemment par une étude ayant proposé que la synthèse et l'export des polycétides, si ce n'est le cas des métabolites spécialisés en général, est coûteux en terme d'énergie et de ressources métaboliques, et que seulement une fraction des hyphes au sein d'une colonie de *Streptomyces* les produisent (Zhang *et al.*, 2020). Il est à noter que le bruit de fond de fluorescence lors de la détection de mCherry (excitation à 522 nm, émission entre 586 nm et 789 nm) était très important et donc que le signal correspondant à la fusion mCherry-CpkA était difficile à détecter. Néanmoins, nous avons attribuer avec certitudes les signaux fluorescents observés à nos protéines de fusion (Sun *et al.*, 1999 ; Schlimpert *et al.*, 2017), et la taille moyenne des foci observés n'a pas pu être déterminée car la majorité des foci observés

## Résultats

possèdent une taille inférieure à 400 nm (c'est-à-dire en dessous du seuil de confiance pour la microscopie à fluorescence). Toutefois, les foci les plus grands observés possèdent une taille de 600 nm. Malgré les difficultés liées aux analyses, nous avons aussi observé qu'un signal diffus ne correspondant pas au bruit de fond de fluorescence pouvait être observé. En effet, ce type de signal a été observé uniquement dans les hyphes exprimant mCherry-CpkA et CpkC-eGFP. Ceci peut s'expliquer par la possibilité de la traduction des protéines de fusion et de leur acheminement vers les loci de fluorescence (comme encerclé dans la **Figure 52B**).



**Figure 51 : Localisation de la sous-unité CpkA dans *S. coelicolor* FM145.** (A) La fusion mCherry-CpkA se localise à un endroit discret. (en haut) Contraste des dérivés de *S. coelicolor* FM145-1, fluorescence (au centre) de la fusion mCherry-CpkA, superposition (en bas) du contraste et de la fluorescence. (B et C) Agrandissement du contraste et de la fluorescence. Échelle (barre blanche) : (A) 10  $\mu$ m, (B-C) 3  $\mu$ m.



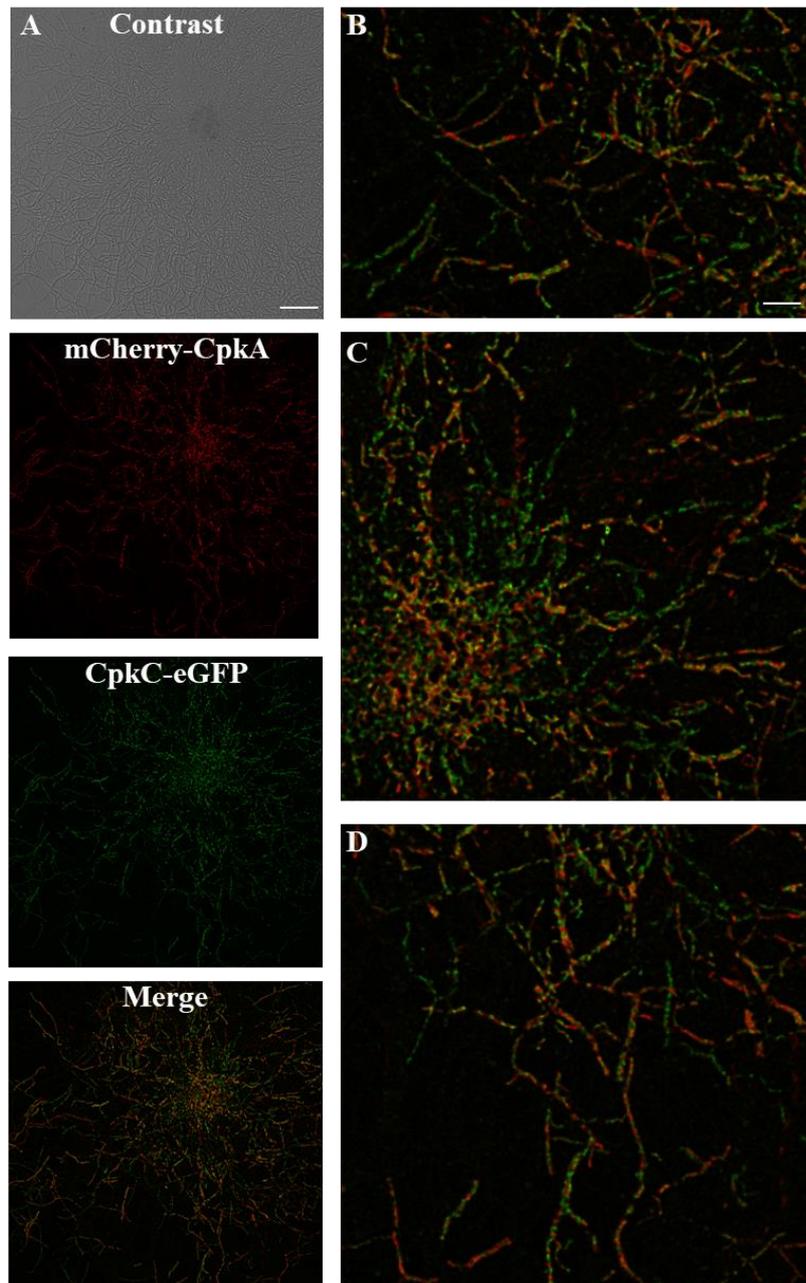
**Figure 52 : Localisation de la sous-unité CpkC dans *S. coelicolor* FM145.** La fusion CpkC-eGFP se localise à un endroit discret. (en haut) Contraste des dérivés de *S. coelicolor* FM145-2, fluorescence (au centre) de la fusion CpkC-eGFP, superposition (en bas) du contraste et de la fluorescence. (B et C) Agrandissement du contraste et de la fluorescence. En cercle est montré un exemple de signal diffus détecté dans les hyphes bactériens ne correspondant pas à de l'auto-fluorescence. Échelle (barre blanche) : (A) 10  $\mu\text{m}$ , (B-C) 3  $\mu\text{m}$ .

## B. Colocalisation de CpkA avec CpkC

Nous avons ensuite comparé la localisation de CpkA et CkC dans la souche FM145-3 contenant à la fois les fusions mCherry-CpkA et CpkC-eGFP. Nous avons aussi observé que 10 % des hyphes de *Streptomyces* ont exprimées l'une ou l'autre des sous-unités chimériques. Lorsque les deux protéines ont pu être observées dans un même hyphe, nous avons constaté que les sous-unités CpkA et CpkC se succèdent sans se chevaucher, indiquant que les sous-unités ne se colocalisent pas (Figure **53A-D**). Il est à noter qu'il n'a pas été possible de quantifier la colocalisation entre les deux sous-unités de la CPKS, car l'important bruit de fond détecté n'a pas permis d'être suffisamment discriminant pas rapport au signal correspondant à mCherry. Néanmoins, à la résolution de la microscopie confocale (0,2  $\mu\text{m}$ ), et en état des connaissances de la taille du mégacomplexe de PKS de type *trans*-AT observé chez *B. subtilis* (0,1  $\mu\text{m}$ ), nous aurions dû observer une parfaite colocalisation entre les sous-unités de PKS de types *cis*-AT si les PKS s'associaient entre-elles pour former un mégacomplexe. Il est possible que les deux signaux fluorescents n'aient pas pu être détecté en même temps dans les hyphes de *Streptomyces*. Toutefois, à ce stade de l'étude, nous ne pouvons pas aussi exclure que la fusion entre les sous-unités de PKS avec une protéine fluorescente perturbe l'assemblage entre les PKS. De plus dans nos données, les signaux correspondant à l'eGFP et à la mCherry peuvent être séparé de plusieurs micromètres dans les hyphes de *Streptomyces*, et même être séparé par des parois transversales (Yagiie *et al.*, 2016). Ces observations ne sont pas en adéquation avec la fonction de la CPKS qui impliquerait que les deux sous-unités fonctionnent ensemble par l'interaction entre plusieurs sous-unités *via* les DD. Aussi, nous ne pouvons pas exclure que les protéines fluorescentes peuvent perturber la localisation des sous-unités de PKS, en particulier pour la sous-unité CpkC fusionnée à l'eGFP. Concernant cette dernière protéine fluorescente, il a été décrit que son caractère dimérique peut provoquer la formation d'agrégats protéiques dans les cellules bactériennes (Margolin, 2012; Swulius and Jensen, 2012). Dans le but de clarifier nos observations, nous devons tester des mutants pour lesquels l'une ou l'autre sous-unité de la CPKS est délétée (**Tableau 2**). Ainsi, nous obtiendrons les images par microscopie confocale pour observer si la localisation en foci des sous-unités individuelles est maintenue ou si elle nécessite la présence de l'autre sous-unité de PKS pour s'organiser. Aussi, nous changerons la protéine fluorescente eGFP par une protéine fluorescente monomérique, afin de savoir si les artefacts de fluorescence sont dû à l'eGFP ou que les assemblages sont perturbés par la présence des deux sous-unités fusionnées à des protéines fluorescences.

Nonobstant les difficultés d'interprétation liées aux analyses, l'ensemble des résultats qui concernent la localisation des sous-unités CpkA et CpkC suggèrent une organisation

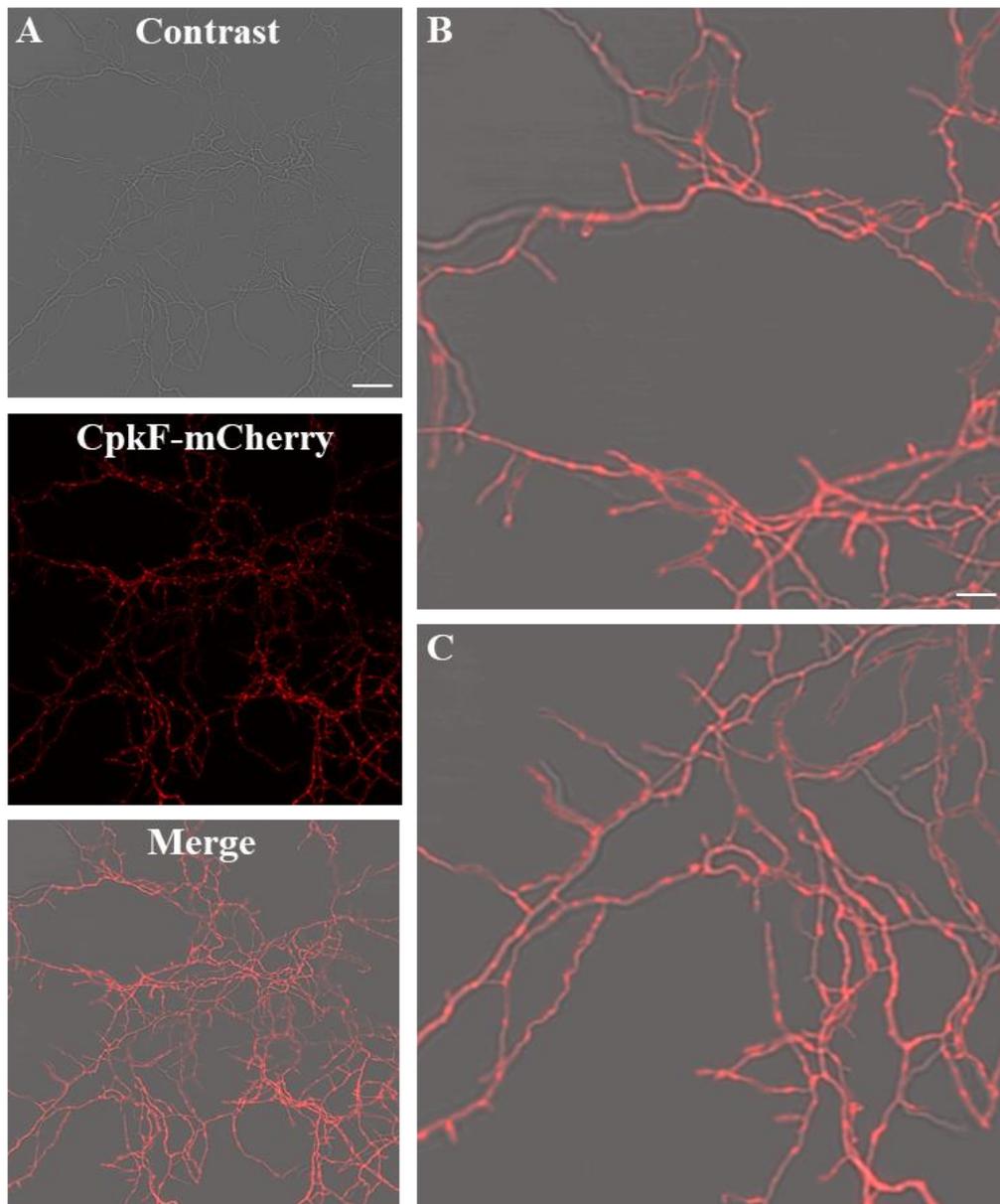
macromoléculaire différente de cette PKS de type *cis*-AT que celle observée pour la PKS de type *trans*-AT chez *B. subtilis*. En effet, sur la base du modèle de l'organisation macromoléculaire de la PKS modulaire de type *trans*-AT observée chez *B. subtilis*, l'introduction de protéines fluorescentes de part et d'autre de la CPKS ne devrait perturber un assemblage vertical de cette PKS de type *cis*-AT.



**Figure 53 : Colocalisation des sous-unités CpkA et CpkC dans *S. coelicolor* FM145.** (A) Les fusions mCherry-CpkA et CpkC-eGFP ne semblent pas se colocaliser. (en haut) Contraste de *S. coelicolor* FM145-3. (au centre) Fluorescence des protéines de fusion comme indiqué ci-dessus. (en bas) Les images mCherry-CpkA et CpkC-eGFP sont fusionnées pour montrer la colocalisation aux points de chevauchement en jaune (B à D) Agrandissement de la superposition des fluorescences. Échelle (barre blanche) : (A) 10  $\mu$ m, (B-D) 3  $\mu$ m.

### C. Colocalisation de CpkC avec CpkF

La visualisation du transporteur dans la souche FM145-4 contenant la fusion CpkF-mCherry a montré une localisation en un seul foyer dans les hyphes (**Figure 54**). Cette observation est similaire aux sous-unités CpkA et CpkC, à la seule grande différence que cette protéine ait été observées dans l'ensemble des hyphes bactériens. De plus, l'expression de CpkF-mCherry comparé à mCherry-CpkC semble plus importante, car il a été plus facile de détecter le signal correspondant au transporteur que la sous-unité de PKS. La distribution de CpkF est à première vue surprenante car si seulement certains hyphes expriment la CPKS, nous aurions pu nous attendre à ce même résultat pour le transporteur associé. Toutefois, il est à noter que la régulation de *cpkA*, *cpkC* et *cpkF* est soumise au même mécanisme de régulation transcriptionnelle (Bednarz *et al.*, 2019), mais *cpkA* et *cpkC* n'appartiennent pas à la même unité transcriptionnelle que *cpkF* (Chen *et al.*, 2016), ce qui suggèrent une régulation traductionnelle différente. La localisation en foci de CpkF-mCherry renforce l'observation faite pour Cherry-CpkA, car pour une même protéine fluorescente monomérique utilisée, la localisation était différente. De plus, l'organisation en loci de mCherry-CpkA est en accord avec l'observation effectuée pour CpkC-eGFP, qui est certes fusionnée à une protéine fluorescence à caractère dimérique, mais qui possède une localisation biologiquement relevant. Un mutant comportant une délétion de l'une ou l'autre sous-unité de PKS est en cours de construction afin de savoir si le transporteur CpkF conserve cette localisation au sein des hyphes bactériens (**Tableau 2**).

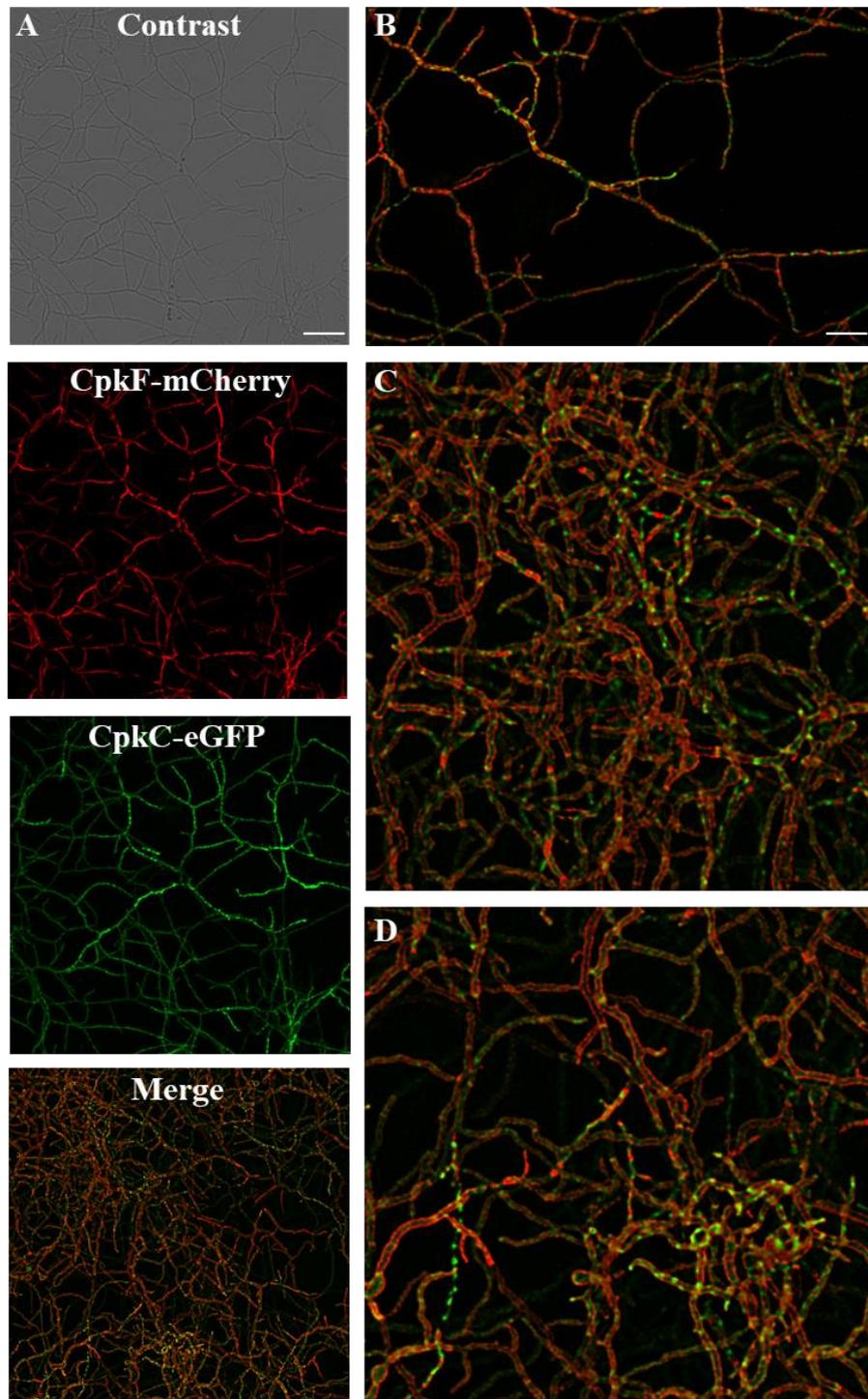


**Figure 54 : Localisation du transporteur CpkF dans *S. coelicolor* FM145.** La fusion CpkF-mCherry se localise à un endroit discret. (en haut) Contraste des dérivés de *S. coelicolor* FM145-4, fluorescence (au centre) de la fusion CpkC-eGFP, superposition (en bas) du contraste et de la fluorescence. (B et C) Agrandissement du contraste et de la fluorescence. Échelle (barre blanche) : (A) 10  $\mu\text{m}$ , (B-C) 3  $\mu\text{m}$ .

Nous avons ensuite comparé la localisation de CpkF et CkC dans la souche FM145-5 contenant à la fois les fusions CpkF-mCherry et CpkC-eGFP. Nous avons d'abord observé un changement d'organisation du transporteur dans certains hyphes bactériens. En effet, les signaux correspondant ont parfois changé d'une localisation en locus pour une localisation tout le long des membranes des hyphes de *Streptomyces* (**Figure 55A-D**). Seulement certains hyphes comportaient des loci de fluorescences correspondant à CpkF. Néanmoins, les résultats montrent que les deux signaux ne se colocalisent pas, mais se succèdent le long des hyphes bactériens. Un mutant contenant la délétion du transporteur CpkF permettra de savoir si, dans

## Résultats

un premier temps, l'accumulation de coelimycine est cytotoxique, et si non, les sous-unités de PKS conservent leur organisation macromoléculaire.



**Figure 55 : Colocalisation du transporteur CpkF et de la sous-unité CpkC dans *S. coelicolor* FM145.** (A) (en haut) Contraste de *S. coelicolor* FM145-5. (au centre) Fluorescence des protéines de fusion comme indiqué ci-dessus. (en bas) Les images CpkF-mCherry et CpkC-eGFP sont fusionnées pour montrer la colocalisation aux points de chevauchement en jaune (B à D) Agrandissement de la superposition des fluorescences. Échelle (barre blanche) : (A) 10  $\mu\text{m}$ , (B-D) 3  $\mu\text{m}$ .

### III. Ingénierie de la biosynthèse des produits naturels

#### 1. Contexte du chapitre

Au cours de ma thèse, mes différentes études de l'organisation du contexte macromoléculaire ont nécessité de mettre en application des méthodes de génie génétique dans le but de modifier le génome de *Streptomyces*. Pour cela, j'ai employé diverses technologies d'édition du génome décrites dans la littérature scientifique qui ont été employées, entre autres, pour l'activation des BGC silencieux sous les conditions normales de laboratoire et la production des produits naturels qui leurs sont associés dans l'hôte natif ou de manière hétérologue (Laureti *et al.*, 2011 ; Cobb *et al.*, 2015). Aussi, il est possible de déléter des BGC entièrement ou partiellement, afin de pourvoir, entre autre, étudier la fonction des gènes associés. Plusieurs excellentes revues scientifiques couvrent les avancées des progrès récents de la biologie synthétique pour l'ingénierie du génome des *Actinomycètes* (Palazzotto *et al.*, 2019; Zhao *et al.*, 2020), mais aucune ne permet de mettre en place une stratégie claire et efficace pour obtenir une modification génétique souhaitée en fonction des outils disponibles.

Ce chapitre a été co-écrit avec ma collègue Li Su (doctorante) et mes directeurs de thèse Pr. Kira Weissman et Dr. Christophe Jacob et est destiné aux jeunes chercheurs (étudiants / doctorants / post-doctorants) qui souhaiteraient modifier le génome d'un *Streptomyces*. Ainsi, nous couvrons les dernières technologies d'édition du génome, et nous discutons aussi de leurs avantages et de leurs inconvénients. Des exemples provenant de notre laboratoire serviront à illustrer les différents protocoles.

Dans la première partie, nous décrivons les différentes possibilités de modification génétique couramment utilisées en laboratoire. L'accent est mis tout particulièrement sur les techniques CRISPR/Cas9, qui, aujourd'hui, sont les méthodes les plus rapides et efficaces quel que soit l'objectif de modification génétique souhaité.

La deuxième partie décrit dans un second temps la conception de l'ARNs guide en prenant en compte les effets hors des cibles de coupure. Dans un second temps, nous décrivons la préparation de la matrice d'édition en tenant compte de l'objectif de modification génétique souhaité.

Dans la troisième partie, nous décrivons différents protocoles qui détaillent la préparation des spores de *Streptomyces*, leur transformation par conjugaison bactérienne et le criblage des exconjugants.

## **2. Chapitre d'ouvrage n°1**

L'ensemble de ces méthodes va aboutir à la rédaction d'un chapitre d'ouvrage intitulé « Engineering modular polyketide biosynthesis in *Streptomyces*: a practical guide » dans la série « Engineering Natural Product Biosynthesis: Methods and Protocols » dans « Methods in Molecular biology ».

# **Discussion et perspectives**



## L'échange de module pour modifier la stéréochimie d'un polycétide

Les résultats de ce chapitre ont montré que l'échange de module est une meilleure stratégie que l'échange de domaine KR seul pour modifier la stéréochimie du polycétide aux positions C3 et C2. Dans ce contexte, nous avons identifié trois modules contenant des domaines KR de type A2 et un contenant un domaine KR de type B2 qui donnent à la fois des rendements globaux et relatifs encourageant des produits souhaités, ce qui en fait des atouts de choix dans la boîte à outils de l'ingénierie des PKS.

Les résultats obtenus de l'échange de module contenant des domaines KR de type A2 suggèrent qu'afin d'améliorer les rendements en lactones réduites, le module complet doit être échangé, maintenant ainsi l'interaction entre le domaine KS et le domaine ACP du même module pour une condensation décarboxylative native (Gokhale *et al.*, 1999). Ce maintien semble être plus important que l'interaction entre le domaine ACP du module précédent et le domaine KS du module suivant pour le transfert de la chaîne d'un module à l'autre. En effet, lorsque le contact entre le domaine ACP<sub>n</sub> et le domaine KS<sub>n+1</sub> a été maintenu, comme le suggère l'étude de Ranganathan *et coll.* (Ranganathan *et al.*, 1999), ce qui implique que l'interaction entre le domaine KS<sub>2</sub> de la DEBS et le domaine ACP introduit soit non native, celui-ci a conduit à une baisse apparente de la quantité de produit réduit et épimérisé ou non, voire de cétolactone. Néanmoins, cette observation est plus mitigée pour les échanges de module contenant des KR de type B2, pour qui l'échange partiel a fourni un meilleur rendement. Ce résultat a été seulement observé pour un seul module contenant un domaine KR de type B2, et n'a pas pu être évalué pour les PKS chimères non-fonctionnelles. De plus, malgré un rendement plus faible, l'échange de module entier avait permis d'obtenir une plus grande fidélité stéréochimique du produit attendu.

De toute évidence, les effets négatifs relatifs à l'interaction non-native entre les domaines ACP et KS lors du transfert de l'intermédiaire et de son extension dans les PKS chimériques semblent différer en fonction de l'origine des modules hybrides. Le contexte modulaire du module source serait prépondérant pour des expériences d'échange de modules et l'obtention d'une PKS chimère fonctionnelle. En effet, tous les modules fonctionnels de ce chapitre sont liés de manière covalente à des modules en amont au sein de leur sous-unité native. Dans de tels contextes, ils sont contraints par le domaine ACP du module donneur et n'ont donc apparemment pas évolué pour discriminer les substrats provenant des domaines ACP non-natifs. Au contraire, tous les modules échangés qui se sont montrés non fonctionnels sont situés aux extrémités N-terminales de leur sous-unité native. Dans ces cas, le transfert intermodulaire

du substrat impose des contraintes de spécificité plus strictes sur le module, à la fois en termes d'interactions protéine-protéine (ACP / KS) et domaine-substrat.

En termes de bornes de modules, les résultats obtenus dans ce chapitre ne s'inscrivent donc pas dans le nouveau paradigme de la définition d'un module (Keatinge-Clay, 2017), qui aurait les bornes alternatives  $AT_n-(KR-DH-ER)_n-ACP_n-KS_{n+1}$ . En effet, cette observation se base sur plusieurs publications qui donnent comme information que les domaines KS des systèmes PKS *cis*-AT agissent comme les gardiens du transfert intermodulaire de la chaîne d'un polycétide en extension. L'étude de Ranganathan *et coll.* (Ranganathan *et al.*, 1999) a montré que les PKS chimères étaient plus actives lorsque le domaine KS natif est conservé lors d'expérience d'échange de module. Cependant lors de cette étude, aucun module incluant des domaines KR épimérisants n'avait été introduit. Nonobstant le fait que cette définition d'un module ne permet pas d'expliquer comment des sous-unités composées d'un seul module ont évolué, cette définition met en exergue la fonction du domaine KS comme gardien de la synthèse (Klaus *et al.*, 2020).

Le second domaine KS (pourtant appartenant à un module interne) de la DEBS, celui qui fait l'objet d'un échange dans notre étude, est connu pour fortement favoriser la stéréochimie de l'intermédiaire défini par la KR de type B1 du module 1 de DEBS (Weissman, 2017a). Ce niveau de spécificité stéréochimique domaine-substrat a aussi été observé pour le second domaine KS de la PKS synthétisant la nanchangmycine (appartenant à un module interne), qui sélectionne entre deux épimères interconvertis par l'activité d'épimérisation du domaine KR du module précédent (Xie *et al.*, 2017a, 2017b). En ce qui concerne les domaines KS situés en N-terminal des sous-unités, le second domaine KS de la PKS synthétisant l'auréothine (à l'extrémité N-terminale de la seconde sous-unité) sélectionne une chaîne de polycétide générée après deux cycles d'élongation par le même module plutôt qu'une chaîne de polycétide qui a été allongée une fois par le même module (Traitcheva *et al.*, 2007). Chez cette même PKS, l'inactivation du domaine KR du second module de AURS a conduit à une erreur dans le transfert inter-modulaire de la chaîne du polycétide (Peng *et al.*, 2019b). En effet, le substrat a été transféré du module 2 de AURS directement au module 4, indiquant que le domaine KS du module 3 de AURS (à l'extrémité N-terminale de la troisième sous-unité) ne peut pas agir sur le substrat non réduit au niveau de la position C3 (**Figure 39** de l'état de l'art). Au contraire, le domaine KS du module 4 (appartenant à un module interne à la troisième sous-unité) possède une grande tolérance pour le substrat natif. Ces expériences ont mis en évidence un niveau supplémentaire de l'interface  $ACP_n/KS_{n+1}$  qui engage une interaction non-covalente, et que le domaine KS est un point de contrôle qualité de la stéréochimie de l'intermédiaire lors de la

communication inter-modulaire qui engage deux sous-unités. L'ensemble de ces résultats indiquent que le domaine KS qui suit un module « classique » devrait être conservé, en tout cas s'il est situé sur une autre chaîne polypeptidique. De façon contradictoire, les résultats de la **publication n°1** montre que ce type d'échange de module ne permet pas d'obtenir les meilleurs rendement et même de générer une PKS chimère fonctionnelle.

Lors de ce chapitre, il n'a pas été possible de comparer l'effet de la région « linker » intermodulaire qui a été maintenue dans les stratégies d'échange de module entier ou partiel. En effet, l'interaction entre le domaine KS<sub>2</sub> de DEBS 1 et le domaine ACP introduit aurait pu ne pas être maintenue malgré la conservation de la région « linker » intermodulaire, tel que pour les constructions AT→ACP. D'autre part, la capacité du domaine KR de Amph M11 à réduire un substrat tricétidique à 6 carbones alors que son substrat naturel possède 24 carbones provient probablement des 76 % d'identité qu'elle partage avec le domaine KR de Amph M1. Par conséquent, ces deux domaines KR sont susceptibles de présenter une spécificité de substrat proche, leur permettant de fonctionner sur des chaînes de tailles différentes de celles de leur contexte d'origine. Cette observation a été confirmée avec des modules homologues chez la NYSS. Ce n'est évidemment pas le cas pour le domaine KR de la PKS OLMS qui posséderait une spécificité de substrat plus élevée, ce qui expliquerait pourquoi, même si la construction dans cette étude est fonctionnelle, la KR n'agit pas.

Pris ensemble, les résultats soutiennent que les modules internes aux sous-unités sont intrinsèquement plus adaptés à l'échange de module par génie génétique dans de nouveaux contextes covalents que leurs homologues situés à l'extrémité N-terminale. Fort de cette information, l'étude s'est poursuivie avec cette fois-ci la DEBS dans le but de modifier la stéréochimie de la 6-désoxyérythronolide B. Pour cela, le second module de la DEBS est en cours d'échange avec les modules ayant permis d'obtenir les meilleurs rendements en lactones réduites et épimérisées. En termes de bornes des modules, nous avons choisi de conserver le module entier du onzième module de la AMPHS (Amph M11-A) et le premier module partiel de la LKMS (Lkm M1-B). Cependant, l'introduction de ces module génèrent une interface non-native entre le Amph M11-A et le Lkm M1-B et le domaine KS<sub>3</sub> de la DEBS. Ce dernier domaine KS est attendu pour fortement discriminer le substrat provenant du module 2 (Murphy *et al.*, 2016). Dans le but de contourner la spécificité du domaine KS<sub>3</sub> de la DEBS lors du transfert intermodulaire, il serait possible de conserver le domaine KS<sub>12</sub> de la AMPHS ou le domaine KS<sub>2</sub> de la LKMS et de générer un troisième module de la DEBS hybride. Mais ce type d'échange introduirait aussi une interface non-native, cette fois-ci entre le domaine KS<sub>12</sub> de la AMPHS ou le domaine KS<sub>2</sub> de la LKMS et l'ACP<sub>3</sub> de la DEBS, qui pourrait conduire à une

baisse de l'activité catalytique de condensation décarboxylative qui dépend de ces deux domaines KS non natifs et ACP<sub>3</sub> natif.

Une stratégie alternative serait alors de modifier la spécificité de la KS<sub>3</sub> par mutagenèse dirigée. En effet, il a déjà été observé qu'une mutation ponctuelle dans le domaine KS<sub>3</sub> de la DEBS (à l'extrémité N-terminale de la seconde sous-unité de la DEBS) permet d'améliorer sa spécificité pour des substrats non-natifs (Murphy *et al.*, 2016). Cette mutation se situe dans une boucle au niveau de l'interface de dimérisation du domaine KS, dans une région faiblement conservée qui fait face à la triade catalytique du domaine KS au sein des PKS de type *cis*-AT. La DEBS devra donc être modifiée par échange du module 2 et par mutagenèse dirigée du domaine KS<sub>3</sub>. Lorsque les plasmides génétiquement modifiés seront disponibles, ils seront transformés par conjugaison dans la souche surproductrice de *Sac. erythraea*. Puis, les différentes souches génétiquement modifiées seront mises en culture dans le but de savoir si des produits sont synthétisés. Une fois extraits et purifiés par chromatographie, les composés seront analysés par RMN et par spectrométrie de masse dans le but d'identifier et de caractériser les produits obtenus. Pour cela, une collaboration avec Andreas Kirsching (Université d'Hanover) est envisagée.

Pour aller plus loin dans l'amélioration des rendements en lactones réduites et épimérisées, des expériences de mutagenèse sur le domaine KS pourront être effectuées. En effet, notre étude a permis notamment de mettre en lumière que des PKS chimères peuvent être générées par l'échange de module tout en gardant leurs fonctionnalités d'extension et de modification du polycétide. La préservation de l'interface ACP<sub>n</sub> et KS<sub>n+1</sub> pourra être affinée par des expériences de mutagenèse, dans le but d'améliorer les contacts protéines-protéines notamment (Klaus *et al.*, 2016).

Aussi, les modules mis en évidence (Amph M11, Nys M1 et M11 et Lkm M1) pourront aussi être testés dans un autre contexte modulaire que le deuxième module de la DEBS. Par exemple, à la place d'un autre module de la DEBS, ou dans un système modèle basé sur un autre système PKS. De cette manière, l'efficacité de ces modules en qualité de véritables biobriques pour la biologie synthétique des PKS modulaires pourra être démontrée. Ces résultats seront encourageant pour la biologie synthétique des PKS modulaires et inciteront à utiliser ces approches de biologie de synthèse pour générer les médicaments de demain.

## Localisation d'un mégacomplexe de PKS *cis*-AT chez *S. coelicolor*

Dans cette étude, nous avons utilisé des fusions de protéines fluorescentes pour déterminer la localisation des sous-unités CpkA et CpkC de la CPKS ainsi que du transporteur associé CpkF dans l'hôte natif *S. coelicolor*. Ces fusions ont été effectuées au moyen de l'outil d'ingénierie génétique CRISPR/Cas9, qui nous a permis d'insérer les séquences codant des protéines fluorescentes mCherry et eGFP en 5' de *cpkA* ou en 3' de *cpkF* et en 3' de *cpkC* respectivement. Cette technique de modification du génome a présenté les avantages d'être bien plus facile à acquérir et à appliquer que les nucléases à doigts de zinc (ZFN, de l'anglais *Zinc finger nuclease*) ou les nucléases proches des activateurs transcriptionnels (TALEN, de l'anglais *Transcription activator-like effector nuclease*), de ne pas laisser une cicatrice (comme l'aurait fait la technique de *PCR-targeting* (Gust *et al.*, 2003)) et d'effectuer des modifications génétiques avec plus d'efficacité (Yasue *et al.*, 2014). Cette technique a largement été employée pour modifier les génomes de *Streptomyces* dans le but d'effectuer essentiellement des délétions, des remplacements de gènes ou des mutations (Zhao *et al.*, 2020). En l'état actuel des connaissances, nous rapportons pour la première fois l'utilisation de l'outil CRISPR/Cas9 pour observer la localisation de protéines par microscopie à fluorescence dans *Streptomyces*.

Concernant la localisation des sous-unités CpkA et CpkC de façon individuelle, les résultats de cette étude ont montré que chez *Streptomyces coelicolor*, les sous-unités de la PKS de type *cis*-AT CPKS semblent s'assembler pour former des macrocomplexes. Cette organisation macromoléculaire avait déjà été observée pour les sous-unités de la PKS de type *trans*-AT chez *Bacillus subtilis* (Straight *et al.*, 2007) et en comparaison avec nos résultats, ce phénomène suggère que cette organisation macromoléculaire des sous-unités de PKS est un mécanisme conservé chez les PKS de type I modulaires et entre ces deux genres bactériens. De plus, l'expression relative des sous-unités de PKS par certains hyphes dans les échantillons testés témoigne d'une variabilité phénotypique au sein d'une colonie de *Streptomyces* (Zhang *et al.*, 2020). En tenant compte qu'une souche de *Streptomyces* contient en moyenne une vingtaine de BGC associés à la synthèse de métabolites spécialisés (van Keulen and Dyson, 2014), ce différentiel d'expression des PKS dans une colonie de *Streptomyces* permettrait d'optimiser l'expression d'un ou de quelques BGC seulement. Ainsi, plusieurs antibiotiques pourraient être produits de manière optimale, de sorte que l'ensemble de la colonie bénéficie de l'avantage sur son environnement qu'ils procurent (van der Meij *et al.*, 2017). Concernant la coelimycine, le BGC *cpk* est soumis à une régulation croisée avec d'autres voies biosynthétiques (Bednarz *et al.*, 2019). En effet, la régulation de ce BGC est soumise à l'action d'un activateur transcriptionnel ScbA et de deux répresseurs transcriptionnels ScbR et ScbR2.

L'étude de leur cible de fixation dans le génome de *Streptomyces* a montré que ces régulateurs transcriptionnels pouvaient avoir d'autres cibles que le BGC *cpk* (Li *et al.*, 2015). De plus, il a été montré que ScbR2 interagit avec des polycétides comme l'actinorhodin et l'undécylprodigiosine (Xu *et al.*, 2010 ; Wang *et al.*, 2014). Cela témoigne que les BGC dans *S. coelicolor* influencent la régulation transcriptionnelle entre eux, privilégiant l'expression que de certains BGC dans les hyphes de *Streptomyces*, et qu'au sein d'un même hyphe, plusieurs BGC peuvent être exprimés au cours du temps.

À propos de la localisation des sous-unités CpkA et CpkC entre elles, nous n'avons pas trouvé d'évidence d'une organisation macromoléculaire de la CPKS comparable à celle de la PKS modulaire de type *trans*-AT chez *B. subtilis* (Straight *et al.*, 2007). En l'état des résultats présentés dans ce manuscrit, les difficultés rencontrées au cours de nos analyses ne nous ont pas permis de savoir si l'assemblage des PKS est perturbé par la présence des deux sous-unités de fusion en même temps, ou si le caractère dimérique de l'eGFP modifie la localisation de CpkC en provoquant la formation d'agrégats protéiques (Margolin, 2012). Ainsi, des expériences supplémentaires seront nécessaires, en modifiant notamment l'eGFP avec une protéine fluorescente monomérique. Aussi, une autre approche est envisagée pour la localisation des sous-unités de la CPKS, en utilisant cette fois-ci non pas des protéines fluorescentes, mais des anticorps monoclonaux contre des régions spécifiques de la CPKS.

En ce qui concerne le mécanisme d'assemblage d'un éventuel mégacomplexe de PKS modulaires, le modèle de l'assemblage du mégacomplexe observé chez *B. subtilis* via les domaines LINKS (Gay *et al.*, 2016), qui est caractéristique des PKS modulaires de type *trans*-AT, ne permet pas d'expliquer comment les sous-unités de PKS modulaires de type *cis*-AT pourraient s'assembler. En effet, ces motifs structuraux ne sont pas présent dans les domaines KS de la CPKS et de manière plus générale chez les PKS modulaires de type *cis*-AT, et si une telle organisation macromoléculaire existe, alors d'autres bases moléculaires seraient engagées.

Ces informations de la localisation des PKS modulaires sont cruciales pour le développement des outils de biologie synthétique, et il sera nécessaire d'explorer dans d'autres *Actinomycètes* (comme la DEBS chez *Sac. erythraea* ou encore la LKMS chez *S. rochei*) si les complexes observés sont maintenus. Aussi, ce type d'étude devra être mené chez *E. coli*, car c'est le système d'expression hétérologue le plus utilisé pour l'expression, la purification et l'obtention des structures des protéines PKS. Si l'ensemble de ces résultats ainsi que ceux obtenus par Straight *et coll.* (Straight *et al.*, 2007) montrent l'organisation macromoléculaire de PKS modulaires dans des cellules bactériennes fixées, alors des expériences de « live imaging » permettraient d'obtenir des informations sur la dynamique d'assemblage de ces

macrocomplexes. Dans le cas où le marquage des PKS avec des protéines fluorescentes ne biaise pas la localisation des macrocomplexes observés, les protéines de fusion générées serviront à élucider les mécanismes moléculaires de l'assemblage de ces complexes et tout particulièrement dans le cadre d'interactions protéines-protéine avec des techniques comme le FRET (de l'anglais *Fluorescence Resonance Energy Transfer*).

Le fait de retrouver cette organisation quaternaire des machineries de biosynthèse des produits naturels, comme les polycétides (Straight *et al.*, 2007) ou les sidérophores (Imperi and Visca, 2013) suggère, comme l'avaient proposé ces études, que la biogénèse localisée sur la membrane et/ou l'activation de peptides non ribosomiaux est une stratégie bactérienne conservée visant à coupler la biosynthèse avec la sécrétion, évitant ainsi l'accumulation cytoplasmique de molécules potentiellement toxiques, comme les antibiotiques et les sidérophores. Ainsi, il apparaît nécessaire d'identifier les protéines associées à l'export de ces produits naturels ainsi que les bases moléculaires de la localisation des systèmes enzymatiques associés à leur synthèse à la membrane. Initialement, Straight *et coll.* (Straight *et al.*, 2007) avait proposé que l'assemblage du mégacomplexe de PKS modulaires de type *trans*-AT à la membrane de la bactérie pouvait se faire *via* l'association entre une portion de la PKS par un domaine membranaire, que ce soit par une interaction directe ou avec une protéine membranaire. Ce type d'observation avait été déjà effectué pour d'autres systèmes multi-modulaires (Simunovic *et al.*, 2003 ; Jain and Cox, 2005), et notamment *via* un transporteur membranaire.

Ainsi, nous avons cherché à connaître la localisation du transporteur associé à la CPKS, CpkF, en générant une souche dans laquelle ce transporteur transmembranaire a été fusionné avec une protéine fluorescente. Nous avons observé que le transporteur CpkF présentent le même modèle d'organisation que les sous-unités de la CPKS, mais qu'il pouvait être visualisé en plus grande abondance dans les hyphes de *Streptomyces*. Puis, nous nous sommes intéressés à la colocalisation entre le transporteur et la sous-unité CpkC. En terme de colocalisation, nous ne pouvons toujours pas exclure que le caractère dimérique de l'eGFP perturbe sa colocalisation avec le transporteur CpkF. Néanmoins, en ce qui concerne la distribution de CpkF et de CpkC dans les hyphes de *Streptomyces*, nos résultats ont montré que le transporteur semble s'organiser indépendamment de la PKS, et cela suggère que lorsque le polycétide est synthétisé, il diffuse dans le cytoplasme pour rejoindre le transporteur et être exporté. La distribution du transporteur tout le long des hyphes permettrait de maximiser l'export du polycétide en évitant son accumulation dans le cytoplasme lors du phénomène de diffusion. Cette hypothèse est renforcée par la distribution relative de CpkF et de CpkC. En effet, si

seulement certains hyphes bactériens expriment un antibiotique, c'est l'ensemble de la colonie qui doit pouvoir résister à son action. Ainsi, l'ensemble des hyphes d'une colonie exprime le transporteur associé, conférant ainsi un avantage conséquent sur l'environnement pour le développement de la colonie de *Streptomyces*. Le transporteur CpkF pourrait permettre aussi l'export d'autres produits naturels. Cependant, ce transporteur, dont la séquence codante est contenue dans le BGC *cpk* (Pawlik *et al.*, 2007), pourrait aussi être spécifique uniquement de la coelimycine *via* des interactions spécifiques (Romano *et al.*, 2018). L'absence de données supplémentaires sur les transporteurs contenus dans les BGC témoigne que l'étude de ces transporteurs transmembranaires et/ou des mécanismes de résistance chez *Streptomyces* reste pauvrement décrite (Severi and Thomas, 2019) et reste à être élucidée.

Les bases moléculaires de l'assemblage du mégacomplexe et du couplage entre la synthèse du polycétide et son export sont à l'étude (**Tableau 2**), et permettront notamment de savoir si la présence du transporteur est essentielle pour l'organisation macromoléculaire de la CPKS. Pour obtenir des informations plus approfondies sur les complexes visualisés, nous envisageons de continuer notre collaboration avec les équipes des Pr Gilles P. van Wezel et Arianne Briegel (Université de Leiden, Pays Bas) qui sont spécialisés dans l'étude de l'organisation des cellules de *Streptomyces*. Ce groupe est récemment parvenu à révéler des détails structuraux fins concernant des assemblages membranaires en 3D à une résolution du nanomètre chez la bactérie *S. albus* par une approche de microscopie lumineuse corrélée à la cryo-microscopie électronique (cryo-CLEM, de l'anglais *Cryo-correlative light and electron microscopy*) (Celler *et al.*, 2016; Yagüe *et al.*, 2016).

# **Matériels et méthodes**



## I. Matériel biologique

### 1. Souches de bactéries et de levure

***Escherichia coli* DH5 $\alpha$**  : *supE44*,  $\Delta$ *lacU169* ( $\Phi$ 80 *lacZ* $\Delta$ M15), *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*. Cette souche est utilisée pour les transformations de routine. Elle permet la propagation des plasmides et est notamment caractérisée par un taux de transformation élevé.

***Escherichia coli* ET12567 (pUZ8002)** :  $F^-$ , *dam-13::Tn9*, *dcm-6*, *hsdM*, *hsdS*. Cette souche est utilisée pour la transformation par conjugaison de *Sac. erythraea* et de *S. coelicolor*. Elle possède le plasmide pUZ8002 contenant le gène de résistance à la kanamycine et l'opéron *tra* codant des protéines de transfert d'ADN plasmidique (MacNeil *et al.*, 1992).

***Escherichia coli* BW25113 (pIJ790)** :  $F^-$ , *DE(araD-araB)567*, *lacZ4787(del)::rrnB-3*, *LAM-*, *rph-1*, *DE(rhaDrhaB)568*, *hsdR514*. Elle est transformée par le plasmide pIJ790 possédant le gène de résistance au chloramphénicol, le réplicon thermosensible pSG5 permettant la perte du plasmide et les gènes *gam*, *bet* et *exo* impliqués dans le mécanisme de recombinaison homologue. Cette souche est utilisée comme hôte dans les expériences de recombinaison entre un plasmide et un produit PCR approprié (Gust *et al.*, 2003).

***Escherichia coli* T-SACK** : W3110 *araD*, *tetA-sacB-amp fliC*, *cat arG::Tn5*. Cette souche a été utilisée comme matrice pour l'amplification de la cassette *tetA-sacB* (Li *et al.*, 2013). Cette souche a été fournie par le Dr Court (National Cancer Institute, Frederick).

***Saccharomyces cerevisiae* 23.344c** : *mat $\alpha$* , *ura3*. Cette souche de *Saccharomyces cerevisiae* dérive de la souche  $\Sigma$ 1278b (Bechet *et al.*, 1970). Cette souche a été utilisée lors de l'étape de clonage, par recombinaison homologue, des divers modules dans le plasmide pMU3-DEBS 1-TE-URA3.

***Saccharopolyspora erythraea* JC2** :  $\Delta$ *eryA*. Cette souche est modifiée au niveau de son génome par délétion de la région *eryA* codant la polycétide synthase synthétisant l'érythromycine A (DEBS) à l'exception de la région d'*eryAIII* codant le domaine thioestérase (Rowe *et al.*, 1998).

***Saccharopolyspora erythraea* JC2 (pJLK 25)** :  $\Delta$ *eryA*. Cette souche *Sac. erythraea* dérive de la souche JC2 (Rowe *et al.*, 1998). Cette souche est transformée par le plasmide intégratif pJLK25 et a été utilisée pour la production de lactones (Kellenberger *et al.*, 2008).

***Saccharopolyspora erythraea* JC2 (pJLK 30)** :  $\Delta$ *eryA*. Cette souche *Sac. erythraea* dérive de la souche JC2 (Rowe *et al.*, 1998). Cette souche est transformée par le plasmide intégratif pJLK30 et a été utilisée pour la production de lactones (Kellenberger *et al.*, 2008).

***Saccharopolyspora erythraea* JC2 (pJLK 35)** :  $\Delta$ eryA. Cette souche *Sac. erythraea* dérive de la souche JC2 (Rowe *et al.*, 1998). Cette souche est transformée par le plasmide intégratif pJLK35 et a été utilisée pour la production de lactones (Kellenberger *et al.*, 2008).

***Saccharopolyspora erythraea* ISOM-5202** :  $\Delta$ eryA. Cette souche de *Sac. erythraea* a la particularité de surproduire l'érythromycine et possède une plate-forme avec 4 sites *attB* (TG1, SV1, phiC31 et phiBT1) (Fayed *et al.*, 2015) en lieu et place du groupe de gènes *ery*. Cette souche a été fournie par Isomerase (Chesterford Research Park, Cambridge).

***Saccharopolyspora erythraea* ISOM-5738** :  $\Delta$ eryA. Cette souche de *Sac. erythraea* ISOM-5202 contient les pièces du puzzle de gènes associés à la synthèse de la désoxyérythronolide B (Fayed *et al.*, 2015) (3 plasmides correspondant à chacune des 3 sous-unités). Cette souche a été fournie par Isomerase (Chesterford Research Park, Cambridge).

***Saccharopolyspora erythraea* ISOM-5744** :  $\Delta$ eryA. Cette souche de *Sac. erythraea* ISOM-5202 contient toutes les pièces du puzzle de gènes associés à la synthèse de l'érythromycine (Fayed *et al.*, 2015) (3 plasmides correspondant à chacune des 3 sous-unités et un plasmide correspondant aux gènes associés aux enzymes post-PKS). Cette souche a été fournie par Isomerase (Chesterford Research Park, Cambridge).

***Kitasatospora setae* DSM 43861** : Cette souche produisant la bafilomycine a été obtenue du DSMZ (« Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH », Allemagne).

***Streptomyces avermitilis* ATTC 31267** : Cette souche produisant l'oligomycine a été obtenue du DSMZ (« Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH », Allemagne).

***Streptomyces blastmyceticus* DSMZ 40029** : Cette souche produisant la médiomycine a été obtenue du DSMZ (« Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH », Allemagne).

***Streptomyces varsoviensis* DSMZ 40346** : Cette souche produisant l'hygrobafilomycine a été obtenue du DSMZ (« Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH », Allemagne).

***Streptomyces lasaliensis* NRRL 3382** : Cette souche produisant la lasalocide a été fournie par le Professeur Leadlay (« University of Cambridge », Royaume-Uni).

***Streptomyces neyagawaensis* ATCC 27449** : Cette souche produisant la concanamycine a été fournie par le Professeur Leadlay (« University of Cambridge », Royaume-Uni).

*Streptomyces noursei* **DSM 40635** : Cette souche produisant la nystatine a été obtenue du DSMZ (« Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH », Allemagne).

*Streptomyces nodosus* **ATCC 14899** : Cette souche produisant l'amphotéricine a été fournie par le Professeur Caffrey (« University College Dublin », Irlande).

*Streptomyces coelicolor* **M145** : Cette souche de *Streptomyces* modèle a été fournie par le Professeur Aigle (Université de Lorraine, France).

*Streptomyces coelicolor* **FM145** : Cette souche est dérivée de *S. coelicolor* M145 et présente une auto-fluorescence réduite, favorisant la visualisation des protéines présentes en faible abondance (Willemse and van Wezel, 2009). Cette souche a été fournie par le Professeur van Wezel (« University of Leiden », Pays-Bas).

Les micro-organismes utilisés ont été conservés à  $-80\text{ °C}$  dans 20 % de glycérol.

## 2. Vecteurs plasmidiques

**pMU3-DEBS 1-TE** : Plasmide navette *E. coli* / *Sac. erythraea* (**Figure 56**). L'ADN codant DEBS 1-TE (Cortes *et al.*, 1995) est placé sous le contrôle du promoteur *actI* qui est lui-même sous le contrôle de l'activateur transcriptionnel *actIII-ORF4* (Rowe *et al.*, 1998 ; Annaval *et al.*, 2015).

**pFL38** : Plasmide navette *E. coli* / *Saccharomyces cerevisiae* (**Figure 57**). Ce plasmide, qui dérive du plasmide pUC19, est centromérique et porte le marqueur de sélection de levure URA3 ainsi qu'une séquence de réplication autonome (ARS) chromosomique associée au centromère VI (Bonneaud *et al.*, 1991).

**pBF21** : Plasmide navette *E. coli* / *Sac. erythraea* (**Figure 58A**). L'ADN codant la première sous-unité de la DEBS, *eryAI*, est placé sous le contrôle du promoteur *eryAIP* (Fayed *et al.*, 2015).

**pBF23** : Plasmide navette *E. coli* / *Sac. erythraea* (**Figure 58B**). L'ADN codant la troisième sous-unité de la DEBS, *eryAIII*, est placé sous le contrôle du promoteur *eryAIP* (Fayed *et al.*, 2015).

**pBF25** : Plasmide navette *E. coli* / *Sac. erythraea* (**Figure 58C**). L'ADN codant la première sous-unité de la DEBS, *eryAII*, est placé sous le contrôle du promoteur *eryAIP* (Fayed *et al.*, 2015).

**pBF30** : Plasmide navette *E. coli* / *Sac. erythraea* (**Figure 58D**). L'ADN codant les enzymes post-PKS associées à la DEBS est placé sous le contrôle du promoteur *eryAIP* (Fayed *et al.*, 2015).

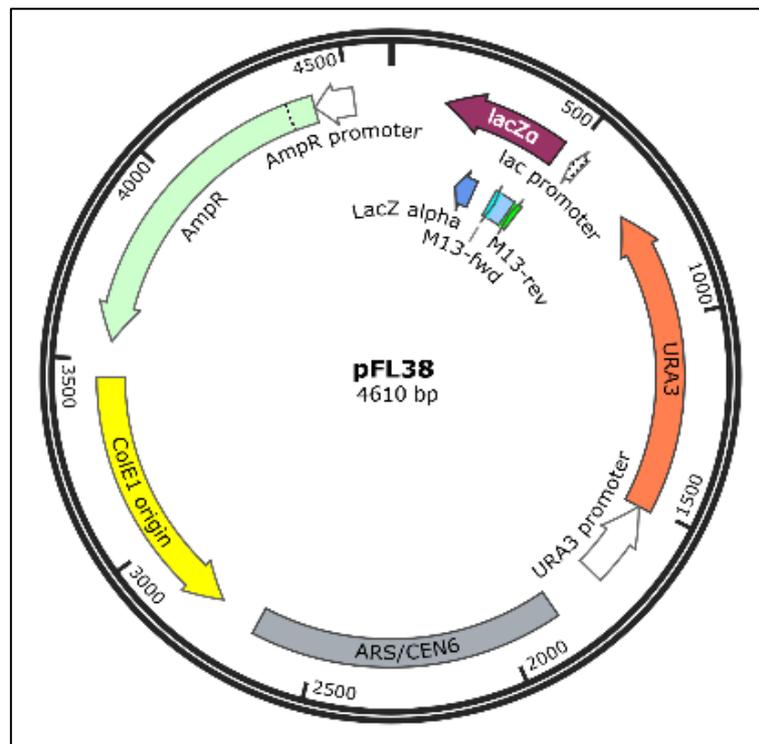
**pCRISPomyces-2** : Plasmide navette *E. coli* / *Streptomyces* (**Figure 59**). Ce plasmide a été utilisé pour modifier l'ADN génomique de *Streptomyces coelicolor* par la technique de génie génétique CRISPR/Cas9 (Cobb *et al.*, 2015).

**pCRISPR-Cas9** : plasmide navette *E. coli* / *Streptomyces* (**Figure 60**). Ce plasmide a été utilisé pour modifier l'ADN génomique de *Streptomyces coelicolor* par la technique de génie génétique CRISPR/Cas9 (Tong *et al.*, 2015).

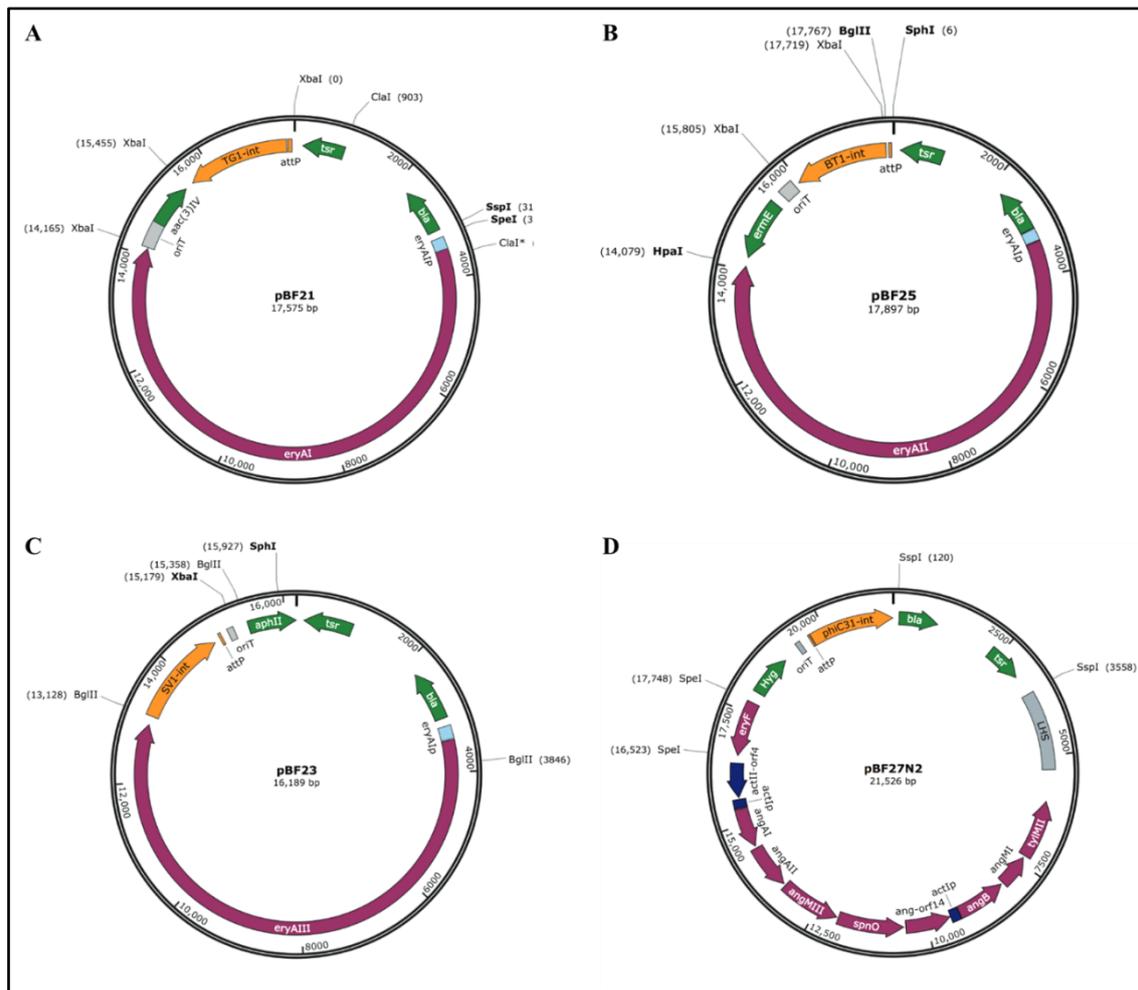
**pSS172** : Plasmide navette *E. coli* / *Streptomyces* (**Figure 61**) contenant le gène codant la protéine fluorescente mCherry (Schlimpert *et al.*, 2017).

**pIJ8630** : Plasmide navette *E. coli* / *Streptomyces* (**Figure 62**) contenant le gène codant la protéine fluorescente eGFP (Sun *et al.*, 1999).





**Figure 57 : Carte du plasmide pFL38 (Bonneaud *et al.*, 1991).** Ce vecteur cargo possède le gène de résistance à l’ampicilline et l’origine de répllication *colE1* d’*E. coli*. Ce plasmide a été construit à partir du plasmide pUC19 où deux sites *AluI* ont été modifiés. Le site 629 a été remplacé par un site *BglIII* et le site 747 par un site *ClaI*. Le marqueur de sélection de levure URA3 a été cloné dans le site *BglIII* et la séquence de répllication autonome (ARS) chromosomique associée au centromère VI dans le site *ClaI*. Une cassette d’insertion multiple est présente en amont d’une cassette *LacZ*.



**Figure 58 : Carte des plasmides codant la DEBS et les enzymes post-PKS associées (Fayed et al., 2015).** Ces vecteurs d'expression possèdent l'origine de réplication *colE1* d'*E. coli* et l'origine de transfert *oriT*.

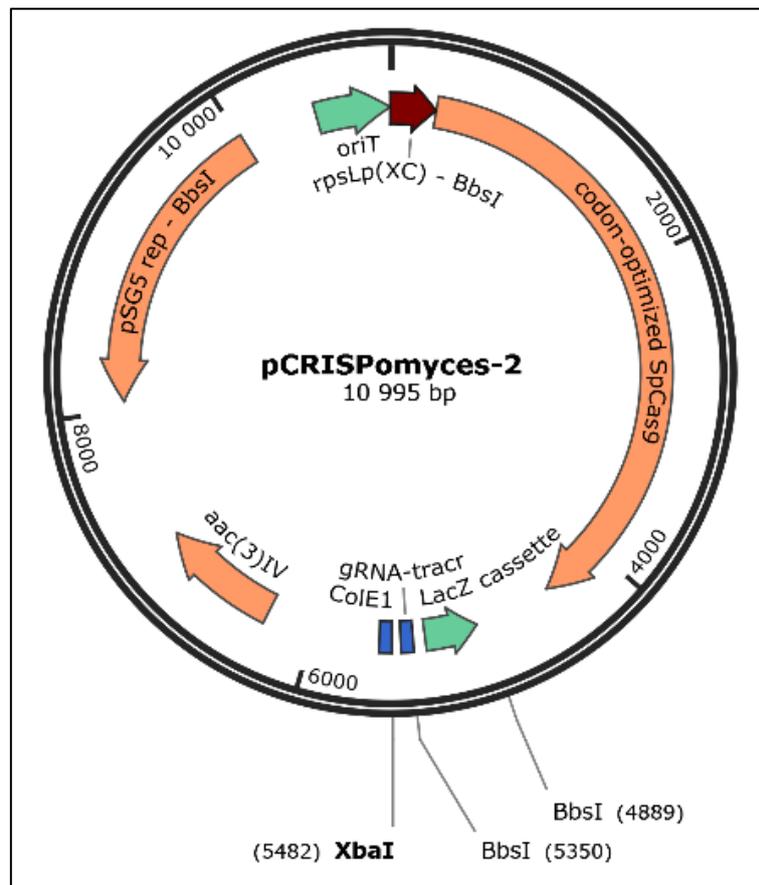
(A) Le vecteur d'expression pBF21 possède le gène de résistance à l'apramycine, le site d'intégration *attP* et l'*attP* TG1 intégrase du bactériophage TG1 et le gène codant la première sous-unité de la DEBS *eryAI*.

(B) Le vecteur d'expression pBF25 possède le gène de résistance à la kanamycine, le site d'intégration *attP* et l'*attP* SV1 intégrase du bactériophage SV1 et le gène codant la deuxième sous-unité de la DEBS *eryAII*.

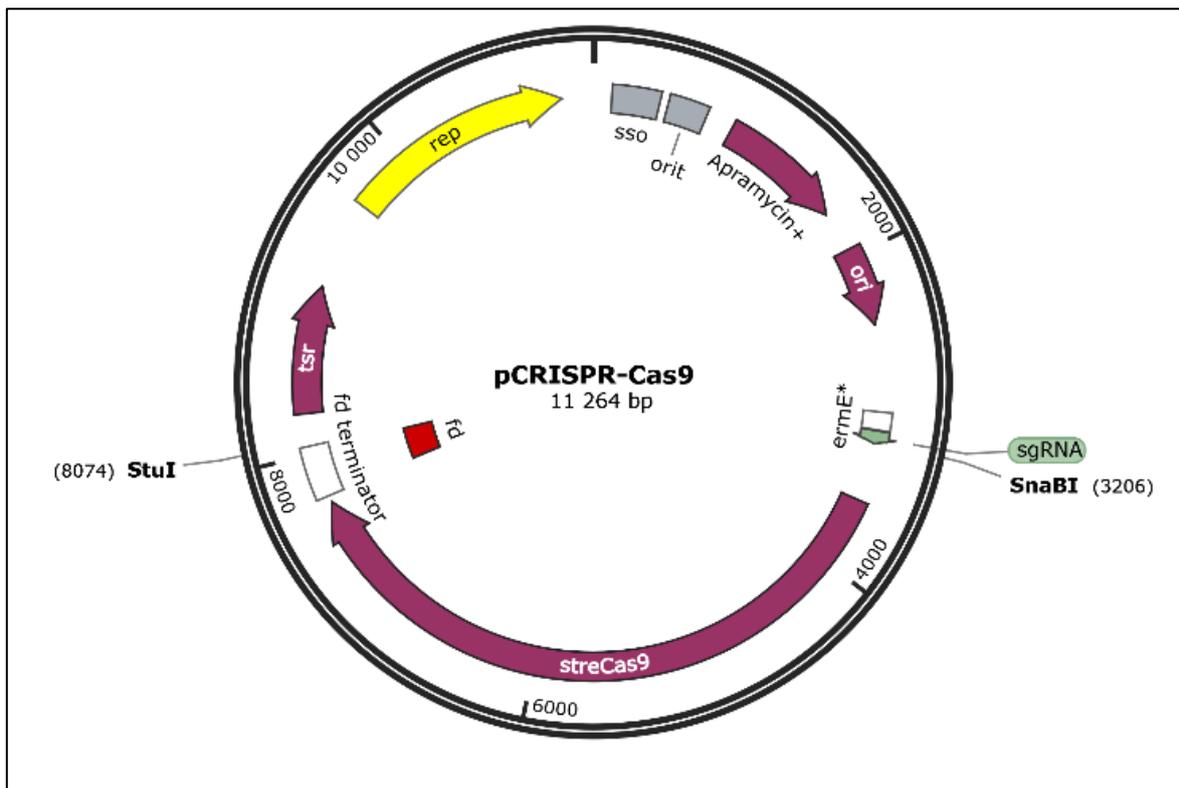
(C) Le vecteur d'expression pBF23 possède le gène de résistance à l'érythromycine, le site d'intégration *attP* et l'*attP* TBT1 intégrase du bactériophage TBT1 et le gène codant la troisième sous-unité de la DEBS *eryAIII*.

(D) Le vecteur d'expression pBF27N2 possède le gène de résistance à l'hygromycine, le site d'intégration *attP* et l'*attP*  $\phi$ C31 intégrase du bactériophage  $\phi$ C31 et les gènes codant les enzymes post-PKS associées à la modification du cœur aglycone du 6-désoxyérythronolide B.

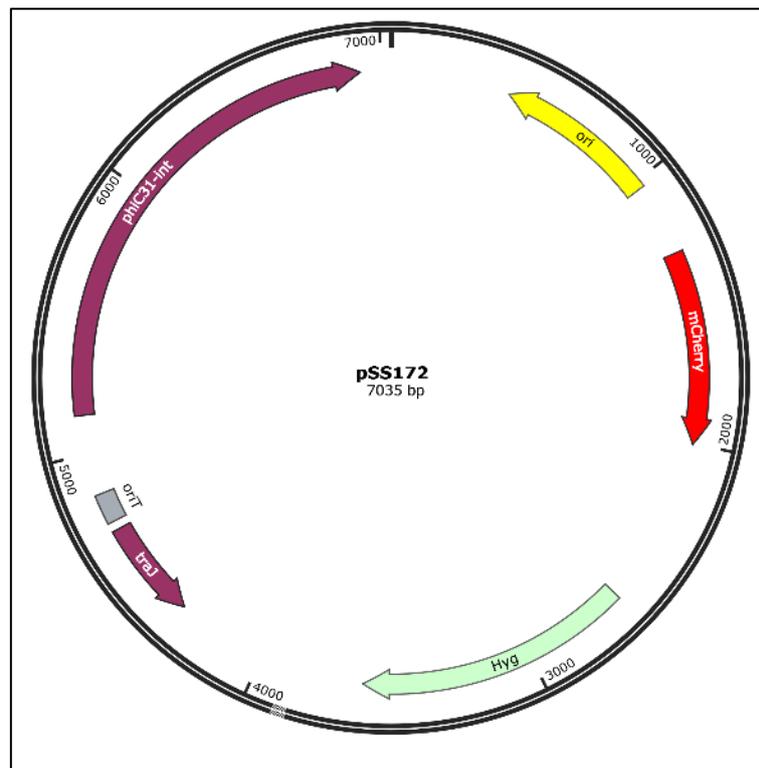
Les gènes codant EryAI, EryAII, EryAIII et les enzymes post-PKS sont sous le contrôle du promoteur *eryAIP*.



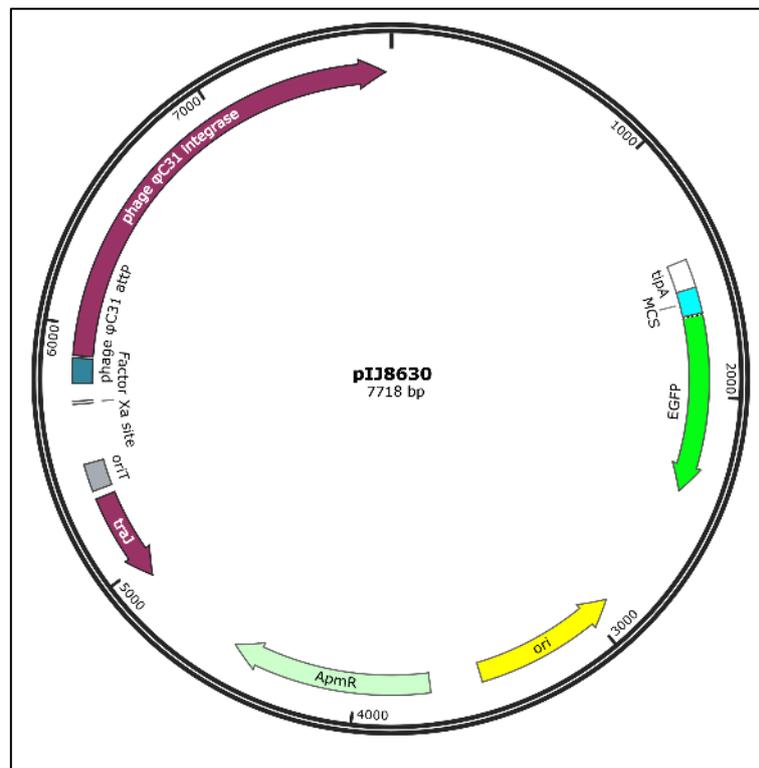
**Figure 59 : Carte du plasmide pCRISPomyces-2 (Cobb *et al.*, 2015).** Ce vecteur possède le gène de résistance à l’apramycine, l’origine de répliation *colE1* d’*E. coli*, l’origine de transfert *oriT* et l’origine de répliation sensible à la température pSG5 dans le but de pouvoir éliminer le plasmide des *Streptomyces* recombinants. L’ARNcr et l’ARNtracr sont fusionnés en un seul ARNsg, qui est plus efficace et le plus utilisé pour une édition à une seule cible. Le vecteur contient la cassette *LacZ* qui est soumise au promoteur fort *Pgapdh*. Le gène codant la protéine Cas9, où les codons ont été optimisés pour le genre *Streptomyces* (qui possède un pourcentage en GC allant jusqu’à 73 % au sein de leur génome), est soumis au promoteur fort *rpsLp* (XC) de la protéine ribosomale 30S de *Xylanimonas cellulosilytica*. Les terminateurs des phages fd et  $\lambda$  ont été utilisés pour les cassettes d’expression.



**Figure 60 : Carte du plasmide pCRISPR-Cas9 (Tong *et al.*, 2015).** Ce vecteur possède le gène de résistance à l'apramycine, l'origine de réplication *colE1* d'*E. coli*, l'origine de transfert *oriT* et l'origine de réplication sensible à la température pSG5 dans le but de pouvoir éliminer le plasmide des *Streptomyces* recombinants. L'ARNcr et l'ARNtracr sont fusionnés en un seul ARNsg qui est soumis au promoteur *ErmE\**. Le gène codant la protéine Cas9 est soumis au promoteur inductible *tipA* qui est lui-même sous le contrôle de l'activateur transcriptionnel TipA. Le gène codant ce dernier est présent dans le génome de *Streptomyces coelicolor* et son activité dépend de la présence de la thiostreptone. Les terminateurs des phages fd et λ ont été utilisés pour les cassettes d'expression.



**Figure 61 : Carte du plasmide pSS172 (Schlimpert *et al.*, 2017).** Ce vecteur d'expression contient le gène de résistance à l'hygromycine, l'origine de répllication *colE1*, l'origine de transfert *oriT* et le site d'intégration *attP* et l'*attP*  $\phi$ C31 intégrase du bactériophage  $\phi$ C31. Ce plasmide a été construit à partir du plasmide pMS82, par de la délétion du promoteur du gène codant l'apramycine, et par l'insertion du gène codant la protéine fluorescente mCherry qui contient les codons optimisés pour le gène *Streptomyces*.



**Figure 62 : Carte du plasmide pIJ8630 (Sun *et al.*, 1999).** Ce vecteur d'expression possède le gène de résistance à l'apramycine, l'origine de réplication *colE1*, l'origine de transfert *oriT* et le site d'intégration *attP* et l'*attP* φC31 intégrase du bactériophage φC31. Le gène codant la protéine fluorescente eGFP de *Aequorea victoria*, où les codons ont été optimisés pour le genre *Streptomyces*, est fusionné à une cassette d'insertion qui est soumise au promoteur inductible *tipA* qui est lui-même sous le contrôle de l'activateur transcriptionnel TipA.

### 3. Cosmide

**Cosmide 14F1** : Cosmide contenant les BGC des PKS responsables de la synthèse de la lankamycine et de la lankacidine. Il a été fourni par le Docteur Arakawa (Université d'Hiroshima, Japon).

## II. Milieux de cultures des micro-organismes

### 1. Milieux de culture d'*E. coli*

**LB (*Lysogeny Broth*)** : Ce milieu est composé de tryptone ( $10 \text{ g.L}^{-1}$ ), d'extrait de levure ( $5 \text{ g.L}^{-1}$ ) et de NaCl ( $5 \text{ g.L}^{-1}$ ). Le pH est ajusté à 7,5 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Pour les milieux solides, de l'agar est ajouté à raison de  $17 \text{ g.L}^{-1}$  de milieu de culture avant stérilisation. Le milieu LB a été utilisé pour la croissance d'*E. coli*.

**2YT** : Ce milieu est composé de tryptone ( $16 \text{ g.L}^{-1}$ ), d'extrait de levure ( $5 \text{ g.L}^{-1}$ ) et de NaCl ( $5 \text{ g.L}^{-1}$ ). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Pour les milieux solides, de l'agar est ajouté à raison de  $17 \text{ g.L}^{-1}$  de milieu de culture avant stérilisation. Le milieu 2YT a été utilisé pour la croissance d'*E. coli*.

**SOB (*Super Optimal Broth*)** : Ce milieu est composé de tryptone ( $20 \text{ g.L}^{-1}$ ), d'extrait de levure ( $5 \text{ g.L}^{-1}$ ), de NaCl ( $0,5 \text{ g.L}^{-1}$ ), de KCl ( $0,186 \text{ g.L}^{-1}$ ), de  $\text{MgCl}_2$  ( $2 \text{ g.L}^{-1}$ ) et de  $\text{MgSO}_4$  ( $2,4 \text{ g.L}^{-1}$ ). Le pH est ajusté à 7 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu SOB a été utilisé pour la croissance d'*E. coli*.

**SOC (*Super Optimal broth with Catabolite repression*)** : Ce milieu est composé de tryptone ( $20 \text{ g.L}^{-1}$ ), d'extrait de levure ( $5 \text{ g.L}^{-1}$ ), de NaCl ( $0,5 \text{ g.L}^{-1}$ ), de KCl ( $0,186 \text{ g.L}^{-1}$ ), de  $\text{MgCl}_2$  ( $2 \text{ g.L}^{-1}$ ), de  $\text{MgSO}_4$  ( $2,4 \text{ g.L}^{-1}$ ) et de glucose ( $3,6 \text{ g.L}^{-1}$ ). Le pH est ajusté à 7 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu SOC a été utilisé pour la régénération des bactéries compétentes après transformation.

## 2. Milieux de culture de *Saccharomyces cerevisiae*

**Milieu complet :** Ce milieu est composé d'extrait de levure ( $10 \text{ g.L}^{-1}$ ), de bactopeptone ( $20 \text{ g.L}^{-1}$ ) et de D-glucose ( $20 \text{ g.L}^{-1}$ ). Le pH est ajusté à 7,5 par ajout de NaOH et le milieu est stérilisé par autoclave pendant 20 min à 1 bar. Pour les milieux solides, de l'agar est ajouté à raison de  $17 \text{ g.L}^{-1}$  de milieu de culture avant stérilisation. Le milieu complet a été utilisé pour la croissance de *Saccharomyces cerevisiae*.

**Milieu de sélection ura- :** Une première solution permettant de constituer le milieu de sélection est composée d'arginine ( $0,8 \text{ g.L}^{-1}$ ), de lysine ( $1,2 \text{ g.L}^{-1}$ ), de méthionine ( $0,8 \text{ g.L}^{-1}$ ), de phénylalanine ( $2 \text{ g.L}^{-1}$ ), de thréonine ( $8 \text{ g.L}^{-1}$ ), de tyrosine ( $1,2 \text{ g.L}^{-1}$ ), d'adénine ( $0,8 \text{ g.L}^{-1}$ ), d'histidine ( $0,8 \text{ g.L}^{-1}$ ), de leucine ( $2,4 \text{ g.L}^{-1}$ ), de tryptophane ( $0,8 \text{ g.L}^{-1}$ ) et d'agar ( $20 \text{ g.L}^{-1}$ ). Une deuxième solution est composée de « yeast nitrogen base » sans acides aminés ( $6,7 \text{ g.L}^{-1}$ ). Une troisième solution est composée de D-glucose ( $20 \text{ g.L}^{-1}$ ). Les trois solutions sont autoclavées pendant 20 min à 1 bar puis mélangées. Le milieu de sélection a été utilisé pour la sélection des levures.

### 3. Milieux de culture de *Sac. erythraea* et de *S. coelicolor*

**ABB13** : Ce milieu est composé de soytone (5 g.L<sup>-1</sup>), d'amidon soluble (5 g.L<sup>-1</sup>), de CaCO<sub>3</sub> (3 g.L<sup>-1</sup>), de MOPS (2,1 g.L<sup>-1</sup>) et d'agar (17 g.L<sup>-1</sup>). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Du MgCl<sub>2</sub> (0,952 g.L<sup>-1</sup>), du FeSO<sub>4</sub> (0,012 g.L<sup>-1</sup>) et de la thiamine (5 g.L<sup>-1</sup>) sont ajoutés extemporanément à la solution. Le milieu ABB13 a été utilisé pour la croissance en milieu solide des souches de *Sac. erythraea* JC2.

**DNA (Difco Nutrient Agar)** : Ce milieu est composé d'extrait de bœuf (3 g.L<sup>-1</sup>), de peptone (5 g.L<sup>-1</sup>) et d'agar (15 g.L<sup>-1</sup>). Le pH est ajusté à 6,8 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Lorsque nécessaire, de l'acide glutamique est ajouté à raison de 325 mM après stérilisation. Le milieu DNA a été utilisé pour la production et la détection de coelimycine P2 en milieu solide par *S. coelicolor*.

**GMTSB** : Ce milieu est composé de TSB (30 g.L<sup>-1</sup>), de glucose (10 g.L<sup>-1</sup>) et de MES (21,3 g.L<sup>-1</sup>). Le pH est ajusté à 6 par ajout de HCl et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu GMTSB a été utilisé pour la transformation par conjugaison de *Sac. erythraea* JC2.

**ISP2** : Ce milieu est composé d'extrait de levure (4 g.L<sup>-1</sup>), d'extrait de malt (10 g.L<sup>-1</sup>), de glucose (4 g.L<sup>-1</sup>) et d'agar (20 g.L<sup>-1</sup>). Le pH est ajusté à 7,3 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu ISP2 a été utilisé pour la mise en culture des différentes espèces de *Streptomyces*.

**MYG** : Ce milieu est composé d'extrait de malt (10 g.L<sup>-1</sup>), d'extrait de levure (4 g.L<sup>-1</sup>) et de glucose (4 g.L<sup>-1</sup>). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Pour les milieux solides, de l'agar est ajouté à raison de 10 g.L<sup>-1</sup> de milieu de culture avant stérilisation. Le milieu MYG a été utilisé pour la mise en culture de *S. coelicolor*.

**M65** : Ce milieu est composé de glucose (4 g.L<sup>-1</sup>), d'extrait de levure (4 g.L<sup>-1</sup>) et d'extrait de malt (10 g.L<sup>-1</sup>). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Pour les milieux solides, de l'agar et du CaCO<sub>3</sub> sont ajoutés avant stérilisation à raison de 17 g.L<sup>-1</sup> et de 2 g.L<sup>-1</sup>, respectivement. Le milieu M65 a été utilisé pour la mise en culture des bactéries *Kitasatospora setae*, *Streptomyces noursei* et *Streptomyces varsoviensis*.

**M79 sans glucose** : Ce milieu est composé de peptone (1 g.L<sup>-1</sup>), d'extrait de levure (0,2 g.L<sup>-1</sup>), de caséine (0,2 g.L<sup>-1</sup>), de NaCl (0,6 g.L<sup>-1</sup>) et d'agar (15 g.L<sup>-1</sup>). Le pH est ajusté à 7,5 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le

milieu M79 sans glucose a été utilisé pour la production et la détection de coelimycine P2 en milieu solide par *S. coelicolor*.

**M252** : Ce milieu est composé d'amidon soluble (10 g.L<sup>-1</sup>), de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (2 g.L<sup>-1</sup>), de MgSO<sub>4</sub> (1 g.L<sup>-1</sup>), de K<sub>2</sub>HPO<sub>4</sub> (1 g.L<sup>-1</sup>), de NaCl (1 g.L<sup>-1</sup>) et de CaCO<sub>3</sub> (2 g.L<sup>-1</sup>). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Des éléments traces (1 mL) sont ajoutés extemporanément à la solution (FeSO<sub>4</sub> (0,1 g.100 mL<sup>-1</sup>), MnCl<sub>2</sub> (0,1 g.100 mL<sup>-1</sup>), et ZnSO<sub>4</sub> (0,1 g.100 mL<sup>-1</sup>)). Pour les milieux solides, de l'agar (15 g.L<sup>-1</sup>) est ajouté. Le milieu M252 a été utilisé pour la mise en culture de la souche de *S. blastmyceticus*.

**R2T20** : Ce milieu est composé de saccharose (206 g.L<sup>-1</sup>), de K<sub>2</sub>SO<sub>4</sub> (0,25 g.L<sup>-1</sup>), d'extrait de levure (6,5 g.L<sup>-1</sup>), de tryptone (5 g.L<sup>-1</sup>) et d'agar (22 g.L<sup>-1</sup>). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Du glucose (10 g.L<sup>-1</sup>), du Tris (3,028 g.L<sup>-1</sup>), du KH<sub>2</sub>PO<sub>4</sub> (0,025 g.L<sup>-1</sup>), du NaOH (0,1 g.L<sup>-1</sup>), du CaCl<sub>2</sub> (5,549 g.L<sup>-1</sup>), et du MgCl<sub>2</sub> (4,76 g.L<sup>-1</sup>) sont ajoutés extemporanément à la solution. Le milieu R2T20 a été utilisé pour la mise en culture de *Sac. erythraea* JC2.

**SM3** : Ce milieu est composé de glucose (4 g.L<sup>-1</sup>), de MD30 (40 g.L<sup>-1</sup>), de farine de soja (20 g.L<sup>-1</sup>), de mélasse de betteraves (2,4 g.L<sup>-1</sup>), de K<sub>2</sub>HPO<sub>4</sub> (2,5 g.L<sup>-1</sup>), de CaCO<sub>3</sub> (2 g.L<sup>-1</sup>) et d'agar (17 g.L<sup>-1</sup>). Le pH est ajusté à 7 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu SM3 a été utilisé pour la production de métabolites secondaires en milieu solide par *Sac. erythraea* JC2.

**SFM (Soya Flour Mannitol)** : Ce milieu est composé de farine de soja torréfiée (20 g.L<sup>-1</sup>), de mannitol (20 g.L<sup>-1</sup>), d'agar (20 g.L<sup>-1</sup>) et de château-la-pompe. Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu SFM a été utilisé pour la croissance bactérienne de *S. coelicolor*.

**TSB (Tryptic Soy Broth)** : Ce milieu est composé de tryptone (17 g.L<sup>-1</sup>), de soytone (3 g.L<sup>-1</sup>), de glucose (2,5 g.L<sup>-1</sup>), de NaCl (5 g.L<sup>-1</sup>) et de K<sub>2</sub>HPO<sub>4</sub> (2,5 g.L<sup>-1</sup>). Le pH est ajusté à 7,2 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Le milieu TSB a été utilisé pour la croissance bactérienne de *Sac. erythraea* et de *S. coelicolor*.

**TWM** : Ce milieu est composé de glucose (5 g.L<sup>-1</sup>), de tryptone (5 g.L<sup>-1</sup>), de saccharose (10 g.L<sup>-1</sup>) et d'extrait de levure (2,5 g.L<sup>-1</sup>). Le pH est ajusté à 7,1 par ajout de NaOH et le milieu est stérilisé par autoclavage pendant 20 min à 1 bar. Pour les milieux solides, de l'agar est ajouté à raison de 17 g.L<sup>-1</sup> de milieu de culture avant stérilisation. Le milieu TWM a été utilisé pour la mise en culture de *Sac. erythraea* JC2.

### *Matériels et méthodes*

Pour assurer la pression de sélection, les milieux ont été supplémentés en acide nalidixique ( $50 \mu\text{g.L}^{-1}$ ), en ampicilline ( $100 \mu\text{g.L}^{-1}$ ), en apramycine ( $50 \mu\text{g.L}^{-1}$ ), en chloramphénicol ( $25 \mu\text{g.L}^{-1}$ ), en kanamycine ( $50 \mu\text{g.L}^{-1}$ ), en hygromycine ( $10 \mu\text{g.L}^{-1}$ ), en spectinomycine ( $50 \mu\text{g.L}^{-1}$ ), en tétracycline ( $15 \mu\text{g.L}^{-1}$ ) ou en thiostreptone ( $50 \mu\text{g.mL}^{-1}$ ), en fonction des besoins.

### III. Solutions tampon

**Tampon de bleu de dépôt pour gel d'agarose (6X)** : Tris-HCl 10 mM pH 8,0, EDTA 1 mM, glycérol 25 % (v/v), bleu de bromophénol 0,12 % (p/v).

**Tampon de bleu de dépôt pour gel de protéines dénaturant** : Tris-HCl 80 mM pH 6,8, SDS 3% (v/v), glycérol 10 % (v/v), bleu de bromophénol 0,1 % (p/v), DTT 100 mM.

**Tampon de migration en gel polyacrylamide** : Tris-HCl 50 mM pH 8,5, glycine 0,2 M, SDS 0,1 % (v/v).

**Tampon de migration en gel d'agarose** : TrisBase 40 mM, EDTA 1 mM. Le pH est ajusté à pH 8,0 par ajout d'acide acétique.

**Tampon de lyse de *Streptomyces coelicolor*** : Tris-HCl 0,05 M pH 8.

**Tampon TBST** : Tris-HCl 20 mM pH 7,5, NaCl 150 mM, Tween-20 0,1 % (v/v).

**Tampon de blocage pour l'immuno-empreinte** : TBST, 0,1 % Tween-20 0,1 % (v/v), lait écrémé en poudre.

**Tampon de transfert pour western blot** : Tris-HCl 25 mM pH 8,5, glycine 0,2 M, éthanol 10 %.

## IV. Construction des plasmides

### 1. Clonage de la cassette URA3-ARS-CEN6 dans le plasmide pMU3-DEBS 1-TE

#### A. Préparation de la cassette URA3-ARS-CEN6

La séquence nucléotidique codant le marqueur de sélection de levure (URA3) ainsi que la séquence de répllication autonome (ARS) chromosomique associée au centromère VI a été amplifiée par PCR à partir du plasmide pFL38 (Bonneaud *et al.*, 1991). Les oligodésoxyribonucléotides utilisés pour cette PCR ont été choisis de manière à ce que leurs extrémités 5' (environ 15 bases) soient identiques aux séquences situées de part et d'autre du site de reconnaissance de l'enzyme de restriction *PciI*, sur le plasmide pMU3-DEBS 1-TE, puis le site de reconnaissance de l'enzyme *PciI* (5'-A↓CATGT-3', Thermo Fisher Scientific) a été ajouté, et en 3' une quinzaine de bases spécifiques du fragment à amplifier a été incluse.

L'enzyme utilisée pour les réactions de PCR est l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific). Les oligodésoxyribonucléotides, les conditions de réaction et le programme utilisé sont indiqués dans le **Tableau 3**. La présence et la taille de l'ADN amplifié ont été contrôlées par dépôt sur gel d'agarose 0,9 %. L'ADN amplifié a été purifié grâce au kit NucleoSpin® « Gel and PCR Clean-up » (Macherey-Nagel). Cette extraction est basée sur l'utilisation de colonnes de silice sur lesquelles l'ADN va se lier en présence de sels chaotropiques.

**Tableau 3 : Séquence nucléotidique des désoxyribonucléotides, composition du mélange réactionnel et conditions expérimentales utilisées pour la réaction de PCR de la cassette URA3-ARS-CEN6.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
Cassette URA3-ARS-CEN6	B925	direct	CCTTTTGCTCACATGtttagtttgcggccg
	B926	indirect	GCAGGAAAGAACATGtagtgagcatcgattag

Composants	Volume
Matrice	50 ng
dNTP ( 1,25 mM chaque)	12 µL
Oligodésoxyribonucléotides sens direct (10 mM)	1 µL
Oligodésoxyribonucléotides sens indirect (10 mM)	1 µL
Tampon Phusion HF (5X)	10 µL
ADN polymérases haute-fidélité (1 U.µL <sup>-1</sup> )	0,6 µL
H <sub>2</sub> O	qsp 50 µL
Volume total	50 µL

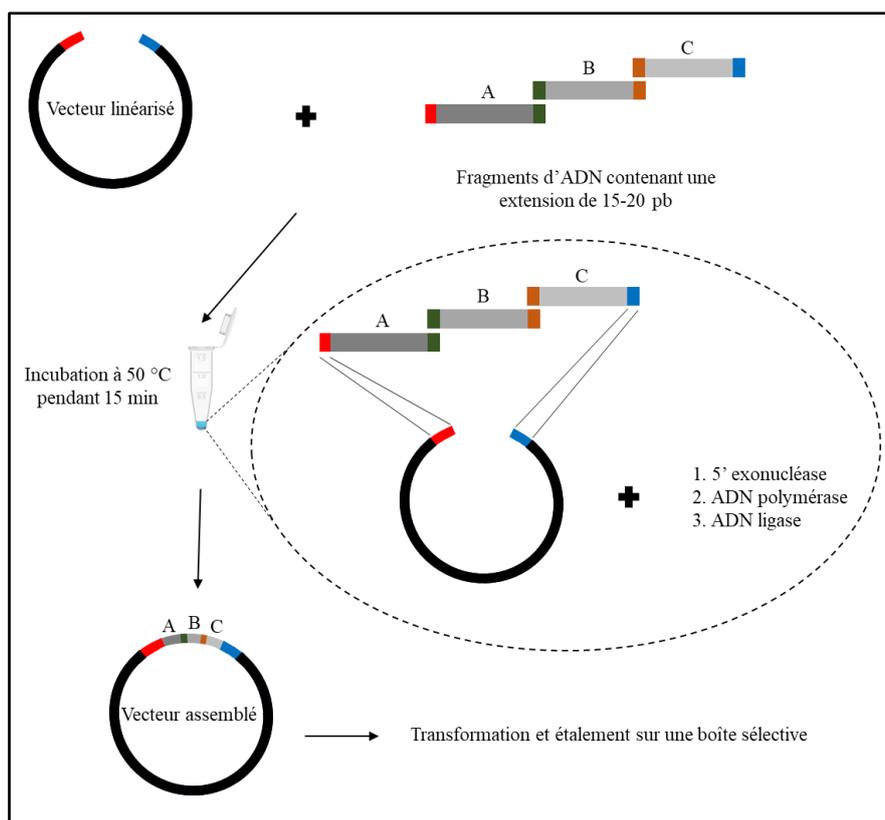
Phase	Durée	Température	Cycles
Dénaturation initiale	2 min	98 °C	1
Dénaturation	10 s	98 °C	30
Hybridation	30 s	56 °C	
Elongation	1 min	72 °C	
Elongation finale	10 min	72 °C	
Conservation	∞	20 °C	1

### B. Préparation du vecteur pMU3-DEBS 1-TE

Le vecteur pMU3-DEBS 1-TE a été linéarisé par l'enzyme de restriction *Pci*I (5'-A↓CATGT-3', Thermo Fisher Scientific) pendant 1,5 h à 37 °C à raison de 5 unités d'enzyme par µg d'ADN, puis traité avec 1 unité de phosphatase alcaline (Fast-AP, Thermo Fisher Scientific) par µg d'ADN pendant 2 h à 37 °C. Après migration sur gel d'agarose 0,9 %, le vecteur linéarisé et déphosphorylé a été purifié comme décrit dans la **partie IV.1.A.**

### C. Assemblage de type Gibson

L'assemblage de la cassette URA3-ARS-CEN6 et du plasmide pMU3-DEBS 1-TE a été réalisé à l'aide du kit « In-Fusion® HD Cloning » (Clontech). Ce kit est basé sur l'assemblage de Gibson (Gibson *et al.*, 2009) comme montré **Figure 63**. Pour cela, 100 ng de plasmide linéarisé et purifié a été mélangé avec le fragment d'ADN à insérer à raison d'un rapport molaire de 1:2, et 1 µL de Mix In-Fusion a été ajouté pour un volume finale de 10 µL. Puis, le mélange a été incubé à 50 °C pendant 15 min. Une fraction aliquote du produit précédent a été transformée dans des cellules d'*E. coli* DH5α compétentes par choc thermique. Les cellules ont été étalées sur boîte LB contenant de l'apramycine et incubées une nuit à 37 °C. Quelques clones obtenus sont mis en culture et une minipréparation d'ADN plasmidique est ensuite réalisée. La présence de la cassette URA3-ARS-CEN6 dans le plasmide pMU3-DEBS 1-TE a été confirmée par séquençage et le plasmide recombinant nommé pMU3-DEBS 1-TE-URA3.



**Figure 63 : Principe de la méthode Gibson (Gibson *et al.*, 2009).** Cette technique consiste en une réaction réalisée dans des conditions isothermes en utilisant trois activités enzymatiques : une 5'-exonucléase qui va permettre d'obtenir des régions complémentaires simple-brins, une polymérase qui va combler les écarts après l'appariement des régions simple-brins complémentaires, et une ADN ligase qui va permettre la formation des liaisons phosphodiester.

## 2. Clonage des séquences codant les modules de PKS étudiés par recombinaison *in vivo*

### A. Préparation des fragments d'ADN

Les séquences nucléotidiques codant les modules de diverses PKS provenant de différentes espèces de *Streptomyces* ont été amplifiées par PCR à partir d'ADN génomique ou d'ADN cosmétique. Les oligodésoxyribonucléotides utilisés pour ces PCR ont été choisis de manière à ce que leurs extrémités 5' (environ 50 bases) permettent une recombinaison homologue avec le vecteur pMU3-DEBS 1-TE-URA3 (séquences en majuscule dans le **Tableau 4**). Leurs extrémités 3', quant à elles, s'hybrident de manière spécifique sur le génome des *Streptomyces* ou sur l'ADN cosmétique de manière à amplifier les modules désirés (en minuscule). L'enzyme utilisée pour les réactions de PCR est l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific). Les oligodésoxyribonucléotides, les conditions de réaction et le programme utilisé sont indiqués dans le **Tableau 4**. L'ADN amplifié a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A**.

**Tableau 4 : Séquence nucléotidique des oligodésoxyribonucléotides, composition du mélange réactionnel et conditions expérimentales utilisées pour l'amplification des fragments d'ADN codant les modules de PKS.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
Amp11_KS	C393	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCAtcggcatgagctgccgtacccggcgga
Amp11_Linkers-AT	C392	direct	GGGAGTGGTCGCCCGCCGCGGACG GGGTGCGCCGGGCAGGTGTGTCGtc cttcggcatcagcgggtaccaacgccac
Amp11_ACP complet	C391	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCcagttcggagtggaggaagccggccagggc
Amp11_ACP SacI	C632	indirect	GCCTGCTGGAGCTGGTTGCGCAGC CCGACCGCGGTCAGCGAGTCGAAG CCGAcgtcgcggaaggcgcgctggtcggtag
Baf4_KS	D494	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgtctcgatggcctgccgtacccggcggc
Baf4_Linkers-AT	D493	direct	GGGGTGGTCGCCCGCCGCGGACGG GGTGCGCCGGGCAGGTGTGTCGgcc ttcgggtcagcggcaccatgccacctg
Baf4_ACP complet	D492	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCcaggccggtctccaggcggccggcgagcgc

Baf9_KS	D491	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgtggcgatgagctgccgcttccccggcggg
Baf9_Linkers-AT	D490	direct	GGGAGTGGTCGCCCCCGCGGACG GGGTGCGCCGGGCAGGTGTGTGTCGtc gttcggcatgagcggcaccaacgcgcacgtc
Baf9_ACP complet	D489	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCcagttcggcgggatctccccggcggcgc
Con4_KS	C397	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgcgatggcctgccgcttccccgcgggc
Con4_Linkers-AT	C394	direct	GGGAGTGGTCGCCCCCGCGGACG GGGTGCGCCGGGCAGGTGTGTGTCG gcgttcggcatgagcggcaccaacgccccat
Con4_ACP complet	C396	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCcaaccctcgtacaggtggtcggcggggc
Con4_ACP <i>SacI</i>	C628	indirect	GCCTGCTGGAGCTGGTTGCGCAGC CCGACCGCGGTCAGCGAGTCGAAG CCGAcggcgggagaacgcctggtcggcggggac
Hyg 4_KS	D488	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgtctccatgggctgccgctaccccggtggc
Hyg 4_Linkers-AT	D487	direct	GGGAGTGGTCGCCCCCGCGGACG GGGTGCGCCGGGCAGGTGTGTGTCGg cgttcggcgtcagcggcaccaacgccccac
Hyg 4_ACP complet	D486	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCgaggagggcgcccaggtggcggggcggggc
Hyg 4_KS	D485	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgcggtcgtctcgatggcctgccgcttcccc
Hyg 9_Linkers-AT	D484	direct	GGGAGTGGTCGCCCCCGCGGACG GGGTGCGCCGGGCAGGTGTGTGTCGtc cttcggcatgagcggcaccaacgccccac
Hyg 9_ACP complet	D483	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCgagttccgtgccgatgtcggggacagggc
Med 7a_KS	D497	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgtcgcgatgagctgccgctacccggggggc
Med 7a_Linkers-AT	D496	direct	GGGAGTGGTCGCCCCCGCGGACG GGGTGCGCCGGGCAGGTGTGTGTCGtc cttcggctcagcggcaccaacgccccac
Med 7a_ACP complet	D495	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCgagctcggcggcaggtgctggcgagcac
Nys 1_KS	D501	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCgtcggaatgggctgccgacttccccggcggc

Nys 1_Linkers KS-AT	D500	direct	GGGAGTGGTCGCCCGCCGCGGACG GGGTGCGCCGGGCAGGTGTGTCGtc gttcggcatcagcggcaccaacgcgcac
Nys 1_ACP complet	D481	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCcgcgcgtcgttctcaggaacgcggcgagcgc
Nys 11_KS	D499	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCcatcgcatggcctgccgttccccggcggc
Nys 11_Linkers KS-AT	D498	direct	GGGAGTGGTCGCCCGCCGCGGACG GGGTGCGCCGGGCAGGTGTGTCGtc gttcggcatcagcgggtaccaacgccat
Nys 11_ACP complet	D482	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCcagctcgcgtgcagataggcgggcgagcgc
Olm 5_KS	C408	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATC atcgtcgcatggcctgccgtacccccggcggc
Olm 5_Linkers KS-AT	C407	direct	GGGAGTGGTCGCCCGCCGCGGACG GGGTGCGCCGGGCAGGTGTGTCG gcgttcggagtcagcggcaccaacgcacat
Olm 5_ACP complet	C406	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TC gaggccgcgggcccaggaaatccgccagttg
Olm 5_ACP SacI	C631	indirect	GCCTGCTGGAGCTGGTTGCGCAGC CCGACCGCGGTCAGCGAGTCGAAG CCGAgtccttgagcgcggggcggggtcgac
Las 7_KS	C404	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATC ggcattggcctgccggtgccccggtggc
Las 7_Linkers KS-AT	C403	direct	GGGAGTGGTCGCCCGCCGCGGACG GGGTGCGCCGGGCAGGTGTGTCG gccttcggcatcagcggcaccaacgccac
Las7_ACP complet	C402	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TC gagtccggcccaggtggcgtaccaccgc
Las7_ACP SacI	C629	indirect	GCCTGCTGGAGCTGGTTGCGCAGC CCGACCGCGGTCAGCGAGTCGAAG CCGAgttcccgaaggcacgcgtcggtcgac
Lkm 1_KS	C400	direct	AGCAGGCGGCACCGGCGACCACG GCCCCGTCGACGAGCCGATCGCG ATCcgctcgcatggcctgacggctgccccggcggg
Lkm 1_Linkers KS-AT	C409	direct	GGGAGTGGTCGCCCGCCGCGGACG GGGTGCGCCGGGCAGGTGTGTCGtc gaagaggcaccacgcgacgtccggggcgcc
Lkm 1_ACP complet	C399	indirect	TAGCCGTCGCGAAGAGCGCTGCTC GCTTCCCGGGCGGGAGTCCCGCTG TCgagttcggcctccaggcgcagggcgagcgc
Lkm 1_ACP SacI	C630	indirect	GCCTGCTGGAGCTGGTTGCGCAGC CCGACCGCGGTCAGCGAGTCGAAG CCGAgttcctgtaacgcccgggtccgggggac

Composants	Volume
Matrice	50 ng
dNTP ( 1,25 mM chaque)	12 µL
Oligodésoxyribonucléotides sens direct (10 mM)	1 µL
Oligodésoxyribonucléotides sens indirect (10 mM)	1 µL
DMSO (100 %)	5 µL
Tampon Phusion GC (5X)	10 µL
ADN polymérase haute-fidélité (1 U.µL <sup>-1</sup> )	0,6 µL
H <sub>2</sub> O	qsp 50 µL
Volume total	50 µL

Phase	Durée	Température	Cycles
Dénaturation initiale	2 min	98 °C	1
Dénaturation	30 s	98 °C	35
Hybridation	30 s	88 °C	
Elongation	4 min	72 °C	
Elongation finale	10 min	72 °C	
Conservation	∞	20 °C	1

### B. Préparation du vecteur pMU3-DEBS 1-TE-URA3

Le vecteur pMU3-DEBS 1-TE-URA3 a été linéarisé par l'enzyme de restriction *AvrII* (5'-C↓CTAGG-3', Thermo Fisher Scientific) pendant 3 h à 37 °C à raison de 5 unités d'enzyme par µg d'ADN, puis traité avec 1 unité de phosphatase alcaline (Fast-AP, Thermo Fisher Scientific) par µg d'ADN pendant 2 h à 37 °C. Le vecteur linéarisé et déphosphorylé a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A**.

### C. Transformation des cellules de levure

La transformation de la souche 23.344c de *Saccharomyces cerevisiae* a été effectuée selon la méthode de Gietz *et coll.* (Gietz *et al.*, 1992).

A partir d'une préculture, une culture de levures a été réalisée dans 100 mL de milieu complet à 30 °C sous agitation. Cette dernière a été arrêtée lorsque la DO<sub>600 nm</sub> a atteint 0,5–0,7. Après une centrifugation de 5 min à 4000 g, le culot a été repris dans 1 mL d'eau stérile et le tout a été centrifugé quelques secondes. Cette étape a été répétée une seconde fois dans le but d'éliminer les traces de milieu riche. Le culot a été ensuite repris dans 1 mL d'une solution d'AcLi/TE (0,1 M acétate de lithium, 10 mM Tris-HCl, 1 mM EDTA (pH 7,5)). Après une centrifugation de quelques secondes, le culot a été suspendu dans 0,5 mL d'AcLi/TE. À ce stade, les cellules de levure étaient compétentes pour la transformation.

Le mélange de transformation se compose de 50  $\mu\text{L}$  de suspension de levures compétentes dans lesquels sont ajoutés dans l'ordre, 5  $\mu\text{L}$  d'ADN de sperme de saumon soniqué ( $10 \text{ mg} \cdot \text{mL}^{-1}$ ) et préalablement dénaturé par chauffage (10 min à  $96 \text{ }^\circ\text{C}$ ), 0,5 à 1  $\mu\text{g}$  d'ADN (le plasmide pMU3-DEBS 1-TE-URA3 préalablement préparé et le fragment d'ADN à cloner) et 300  $\mu\text{L}$  d'une solution de 40 % PEG/AcLi/TE (40 % PEG 4000, 0,1 M acétate de lithium, 10 mM Tris-HCl, 1mM EDTA (pH 7,5)). Après chaque ajout, la suspension a été mélangée délicatement. Le tout a été ensuite incubé à  $30 \text{ }^\circ\text{C}$  pendant 30 min sous agitation avant d'être soumis à un choc thermique à  $42 \text{ }^\circ\text{C}$  pendant 15 min. Les levures ont été alors lavées dans 1 mL d'eau stérile, suspendues dans 1 mL de milieu complet et incubées à  $30 \text{ }^\circ\text{C}$  pendant 4 h sous agitation. Après centrifugation pendant 8 s, les cellules ont été étalées sur un milieu de sélection et incubées pendant 3 jours à  $30 \text{ }^\circ\text{C}$ .

#### **D. Isolement des plasmides recombinants**

L'ADN total des cellules de levure a été extrait comme décrit dans la **partie V.2.**, puis dialysé avant d'être transformé par électroporation comme décrit dans la **partie VI.5.** dans la souche DH5 $\alpha$  d'*E. coli*. Après sélection sur milieu LB contenant de l'apramycine, l'ADN plasmidique de 12 clones a été extrait. Les plasmides ainsi obtenus ont été ensuite analysés par restriction enzymatique et séquençage.

### 3. Clonage de la cassette *tetA-sacB* par *PCR-targeting*

#### A. Préparation de la cassette *tetA-sacB*

La séquence nucléotidique codant la cassette *tetA-sacB* a été amplifiée par PCR sur colonie à partir de la souche T-SACK d'*E. coli*. Les oligodésoxyribonucléotides utilisés pour cette PCR ont été choisis de manière à ce que leurs extrémités 5' (environ 39 bases) permettent une recombinaison homologe avec le vecteur pBF21 (séquences en majuscules dans le **Tableau 5**). Leurs extrémités 3' (en minuscule), quant à elles, s'hybrident de manière spécifique sur le génome de la souche T-SACK d'*E. coli*.

L'enzyme utilisée pour la réaction de PCR sur colonie l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific). Les oligodésoxyribonucléotides utilisés, les conditions de réaction et le programme utilisé sont indiqués dans le **Tableau 5**. L'ADN amplifié a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A**.

**Tableau 5 : Séquence ADN des oligodésoxyribonucléotides, composition du mélange réactionnel et conditions expérimentales utilisées pour l'amplification de la cassette *tetA-sacB*.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
Cassette TetA-SacB	D856	direct	CCGGCGACCACGGCCCCCGTTCGACG AGCCGATCGCGATCatcaagggaactgtc catat
	D855	indirect	GGCGAGGGCCGACGGCGCCTCCCCC CGGACCTCGGTGCCctctaattttgtgacactc ta

Composants	Volume
Matrice	1 µL
dNTP ( 1,25 mM chaque)	4 µL
Oligodésoxyribonucléotides sens direct (10 mM)	1 µL
Oligodésoxyribonucléotides sens indirect (10 mM)	1 µL
Tampon Phusion HF (5X)	2,5 µL
ADN polymérase haute-fidélité (1 U.µL <sup>-1</sup> )	0,5 µL
H <sub>2</sub> O	10
Volume total	25 µL

Phase	Durée	Température	Cycles
Dénaturation initiale	2 min	95 °C	1
Dénaturation	30 s	95 °C	30
Hybridation	30 s	50 °C	
Elongation	4 min	72 °C	
Elongation finale	10 min	72 °C	
Conservation	∞	20 °C	1

## B. Recombinaison entre le plasmide pBF21 et la cassette tetA-sacB par PCR-targeting

Le plasmide pBF21 a été transformé par électroporation comme décrit dans la **partie VI.2.** dans la souche BW25133 (pIJ790) d'*E. coli*. La sélection des transformants a été effectuée sur milieu LB contenant du chloramphénicol et de l'apramycine.

La recombinaison entre le pBF21 et la cassette *tetA-sacB* a été réalisée par la technique de *PCR-targeting* (Gust *et al.*, 2003). Cette technique exploite l'activité de trois protéines : la protéine Gam qui inhibe l'activité nucléase des complexes protéiques RecBCD et SbcCD afin de prévenir de la dégradation de l'ADN linéaire, la protéine Exo dont l'activité 5'→3' exonucléase dégrade l'extrémité 5' de l'ADN double-brin pour générer des régions complémentaires simples-brins et la protéine Béta qui protège l'ADN simple brin généré par l'activité de l'exonucléase et promeut son hybridation avec l'ADN complémentaire simple-brin. Ces protéines sont exprimées à partir du plasmide pIJ790 contenu dans la souche BW25113 d'*E. coli* et sont soumis à l'activateur transcriptionnel *araC* inductible à l'arabinose (Gust *et al.*, 2003).

Un clone de la souche BW25133 (pIJ790) (pBF21) a été mis en culture dans 10 mL de LB contenant du chloramphénicol et de l'apramycine. Un volume de 500 µL de cette préculture a été utilisé pour inoculer 35 mL de SOB (sans MgSO<sub>4</sub>) contenant du chloramphénicol, de l'apramycine et du L-arabinose (10 mM). Lorsque la culture a atteint une DO<sub>600 nm</sub> de 0,4, les cellules ont été rendues électro-compétentes comme décrit dans la **partie VI.2.** Cent ng de la cassette *tetA-sacB* a été transformée dans les cellules d'*E. coli* précédemment rendues compétentes par électroporation. Les cellules ont été étalées sur boîte LB contenant de la tétracycline et incubées une nuit à 37 °C. Quelques clones obtenus ont été mis en culture et une minipréparation d'ADN plasmidique a été ensuite réalisée. La présence de la cassette *tetA-sacB* dans le plasmide pBF21 a été confirmée par séquençage et le plasmide recombinant nommé pBF21-tetA-sacB.

#### 4. Clonage du module 11 de l'AMPHS par *PCR-targeting*

##### A. Préparation du fragment d'ADN

La séquence nucléotidique codant le module 11 de l'AMPHS a été amplifiée par PCR partir de la souche *S. nodosus*. Les oligodésoxyribonucléotides utilisés pour cette PCR ont été choisis de la même manière que pour la cassette *tetA-sacB*, à la seule différence que leurs extrémités 3' s'hybrident de manière spécifique sur le génome de la souche *S. nodosus*. L'enzyme utilisée pour les réactions de PCR est l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific). Les oligodésoxyribonucléotides utilisés sont indiqués dans le **Tableau 6**. Les conditions de réaction ainsi que le programme utilisé ont été décrits dans le **Tableau 4**. L'ADN amplifié a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV4.A**.

**Tableau 6 : Séquence ADN des oligodésoxyribonucléotides pour la réaction de PCR du fragment ADN du onzième module de la AMPHS.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
Module 11 de AMPHS	C393	direct	CCGGCGACCACGGCCCCCGTCGACG AGCCGATCGCGATCcatcgcatgagctgccgc taccggcgga
	E339	indirect	GGCGAGGGCCGACGGCGCCTCCCC CGGACCTCGGTGCCcagttcgagtgaggaa gccggccagggc

##### B. Recombinaison entre le fragment de onzième module de AMPHS et le pBF21-tetA-sacB par *PCR-targeting*

Le plasmide pBF21-tetA-sacB a été transformé par électroporation comme décrit dans la **partie VI.2**, dans la souche BW25133 (pIJ790) d'*E. coli*. La sélection des transformants a été effectuée sur milieu LB contenant du chloramphénicol, de l'apramycine et de la tétracycline.

Le fragment d'ADN du 11<sup>ème</sup> module de la AMPHS a été introduite dans la souche pBF21-tetA-sacB comme décrit dans la **partie VI.2**. Après la transformation, les cellules ont été mises en culture dans 10 mL de LB contenant de l'apramycine à 37 °C. Après sélection, l'ADN plasmidique de 10 clones a été extrait.

### C. Isolement des plasmides recombinants

Le mélange d'ADN plasmidique a été digéré ensuite par l'enzyme de restriction *NdeI* (5'-CA↓TATG-3', Thermo Fisher Scientific) pendant 1,5 h à 37 °C à raison de 5 unités d'enzyme par µg d'ADN puis traité avec 1 unité de phosphatase alcaline (Fast-AP, Thermo Fisher Scientific) par µg d'ADN pendant 2 h à 37 °C. Cette dernière digestion permet d'éliminer les plasmides pBF21-tetA-sacB possédant un site *NdeI* que ne possèdent pas les plasmides ayant recombinés le fragment d'ADN du onzième module de l'AMPHS. Une fraction aliquote du produit précédent a été transformée dans des cellules d'*E. coli* DH5α compétentes par choc thermique. Les cellules ont été étalées sur boîte LB contenant de l'apramycine et incubées une nuit à 37 °C. Quelques clones obtenus sont mis en culture et une minipréparation d'ADN plasmidique est ensuite réalisée. La présence du onzième module de l'AMPHS dans le plasmide pBF21 a été confirmée par séquençage et le plasmide recombinant nommé pBF21-AMPHS<sub>module11</sub>.

## 5. Clonage des matrices d'édition et des séquences espaceurs dans le plasmide pCRISPomyces-2

### A. Assemblage des matrices d'édition par PCR de chevauchement

Les séquences nucléotidiques des différentes matrices d'édition, chacune composée de deux régions homologues aux séquences nucléotidiques en amont et en aval du protospaceur, ont été amplifiées par PCR à partir de l'ADN génomique de *S. coelicolor*.

Pour amplifier les fragments d'ADN et dans un second temps les assembler, les oligodésoxyribonucléotides utilisés pour ces PCR ont été choisis de manière à ce que leurs extrémités 5' (environ 15 bases) génèrent des régions complémentaires avec la séquence à fusionner (en majuscule). Leurs extrémités 3', quant à elles, s'hybrident de manière à amplifier spécifiquement les fragments d'ADN désirés à partir du génome de *S. coelicolor* ou des plasmides pSS172 ou pIJ8630 (en minuscule).

L'enzyme utilisée pour les réactions de PCR est l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific) et les oligodésoxyribonucléotides sont présentés dans le **Tableau 7**. Les conditions de réaction et le programme utilisé sont indiqués dans le **Tableau 4**. La température d'hybridation a été adaptée pour chaque couple d'oligodésoxyribonucléotides. Lorsque les fragments d'ADN devaient être fusionnés entre eux, le temps d'hybridation a été augmenté à 45 s. La présence et la taille de l'ADN amplifié ont été contrôlées par dépôt sur gel d'agarose 0,9 %. L'ADN amplifié a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A**.

**Tableau 7 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'assemblage des matrices d'édition et pour le séquençage du pCRISPomyces-2.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
cpkA-RH-amont	D451	direct	TGCCGCCGGGCGTTTTTATcggtgaa cagcagggcggtg
	D450	indirect	GTGTCAATCTGAACCTCCTCgccgaga acgccactacggagg
cpkA-gène synthétique	D449	direct	TCCGTAGTGGCGTTCTCGGCgaggag gttcagattgaca
	D444	indirect	GCATGGACGAGCTGTACAAGggcggc ggcggcggcggcggcgtgtcaccagtgaaaataa
mCherry-cpkA	D445	direct	CCAGGAAAGGGCAGCCGAGCatggtc tccaagggcgag
	D446	indirect	CCGCCGCCGCCGCCGCCcttgtaca gctcgtccatgcc
cpkA-RH-aval	D447	direct	TCCTCGCCCTTGAGACCATgctcggc tgcccttct

	D448	indirect	TTACGGTTCCTGGCCTCTAGcgtgttga tggagatgtgccc
cpkC-RH-amont	D786	direct	GCCGGGCGTTTTTTATCTAGcaacttcg gcgtcatcacc
	D785	indirect	GAGCTGTACAAGTAAccaccgcacgega aagc
cpkC-eGFP	D784	direct	TTCGCGTGCGGTGGGttactgtacagctcgt ccatgcc
	D783	indirect	GGCGGCGGCGGCGGCgtgagcaagggcg aggagc
cpkC-gène synthétique	D782	direct	CCTTGCTCACgccgcgccgccgccgccccc ccggcccggggaagtg
	D781	indirect	GGTGGCGGCGCTGGAaggcacgggcata ggcatacc
cpkC-RH-aval	D778	direct	TCCTCGCCCTTGAGACCATccgcgga cccgaaggagatc
	D776	indirect	TTACGGTTCCTGGCCTCTAGgcactgc ctggtccgagg
cpkF-RH-amont	E199	direct	GCCGGGCGTTTTTTATCTAGtactccac cagccccgagatg
	E210	indirect	CGCGGGGATTCCGGCGGGAGcgggtgt cggcgatcccgg
cpkF-gène synthétique	E209	direct	CTCCGGGATCGCCGACACCGctcccgc cggaatccccgagg
	E208	indirect	TCCTCCTCGCCCTTGAGACgccccg ccgccgccgccggcgctggtgacgggctcttc
cpkF-mCherry	E207	direct	GCGGCGGCGGCGGCGGCGGCgtctcc aagggcgaggag
	E206	indirect	GGTCGGTCATGTCCGTCCCTcacttgt acagctcgtccat
cpkF-RH-aval	E205	direct	TGGACGAGCTGTACAAGTGAagggac cgacatgaccacca
	E204	indirect	TTACGGTTCCTGGCCTCTAGcttcagcc ccagccgggtg
$\Delta$ cpkA-RH- amont	E852	direct	GCCGGGCGTTTTTTATCTAGgccggcgg gcgtcggagagc
	E851	indirect	CCAGGAAAGGGCAGCCGAGCccggg caccggacagcgacac
$\Delta$ cpkA-RH-aval	E850	direct	TGTCGCTGTCCGGTGCCCGGgctcggc tgcccttctggc
	E849	indirect	TTACGGTTCCTGGCCTCTAGcttctggc agatcgccggcg

## B. Préparation du vecteur

Le vecteur pCRISPomyces-2 a été linéarisé par l'enzyme de restriction *Xba*I (5'-T↓CTAGA-3') (Thermo Fisher Scientific) pendant 1 h à 37 °C à raison d'une unité d'enzyme par  $\mu$ g d'ADN, puis traité avec 1 unité de phosphatase alcaline (Fast-AP, Thermo Fisher

Scientific) par  $\mu\text{g}$  d'ADN pendant 2 h à 37 °C. Le vecteur linéarisé et déphosphorylé a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A**.

### C. Ligation de la matrice d'édition et du plasmide par assemblage de type Gibson

L'assemblage des matrices d'édition et du plasmide pCRISPomyces-2 a été effectué en utilisant le kit In-Fusion<sup>®</sup> HD Cloning (Ozyme, France), dont le principe repose sur l'assemblage de Gibson (Gibson *et al.*, 2009) comme décrit précédemment (**partie IV.1.C**). La présence de la matrice d'édition dans le plasmide pCRISPomyces-2 a été confirmée par séquençage.

### D. Identification des séquences spaceurs

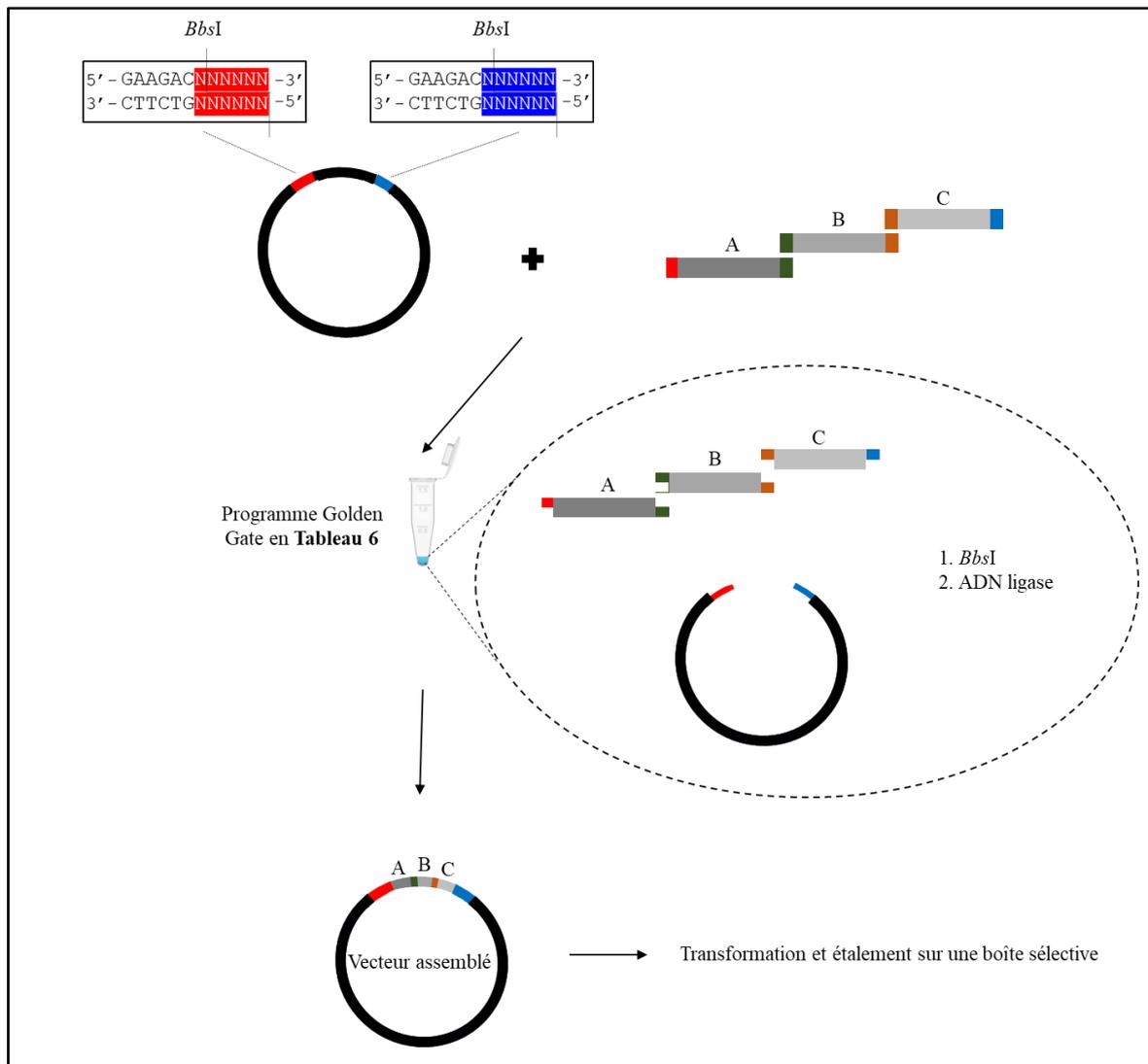
Les séquences spaceurs (séquences en minuscule dans le **Tableau 8**) ont été déterminés grâce au logiciel en ligne CRISPy-web (Blin *et al.*, 2016). Celles-ci ont été choisies de telle manière que la séquence soit préférentiellement sur le brin non-codant, comporte des purines à son extrémités 3', soit suivie de la séquence PAM (5'-NGG-3') et soit unique dans le génome bactérien (Cobb *et al.*, 2015).

**Tableau 8 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'amplification des séquences spaceurs.** Les séquences en majuscules correspondent à l'extension en 5' pour le clonage de type Golden Gate.

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
cpkA-ARNsg-1	D455	direct	ACGCtgctacggaattgttctca
	D454	indirect	AAACgtgagaacaattccgtagca
cpkA-ARNsg-2	D453	direct	ACGCcagcagattgaggttcacac
	D452	indirect	AAACgtgtgaacctcaatctgctg
cpkC-ARNsg-1	D774	direct	ACGCcatacttctcgaacagctcc
	D772	indirect	AAACggagctgttcgagaagtatg
cpkC-ARNsg-2	D771	direct	ACGCgaatcccgataccggttccc
	D770	indirect	AAACgggaaccggtatcgggattc
cpkF-ARNsg-1	E203	direct	ACGCggccgggatacccggcgag
	E202	indirect	AAACctccggcggtatcccggcc
cpkF-ARNsg-2	E201	direct	ACGCgggctcggcgtctcgggtg
	E200	indirect	AAACcaccggagagcggcgagccc

### E. Insertion des séquences spaceur par assemblage de type Golden Gate

Le clonage de type Golden Gate (Engler *et al.*, 2008), dont le principe est montré **Figure 64** permet d'assembler simultanément et directionnellement un ou plusieurs fragments d'ADN. Pour assembler les séquences spaceurs avec le plasmide pCRISPiomyces-2 contenant la matrice d'édition, les séquences spaceurs ont été conçu de manière à inclure une extension à leur extrémité 5' qui est complémentaire à celle générée par l'activité de l'enzyme *BbsI* dont le site de restriction est contenu dans le pCRISPomyces-2 (séquences en majuscule dans le **Tableau 8**). L'appariement des oligodésoxyribonucléotides constitutifs de la séquence spaceur a été réalisé en mélangeant 5 µL de chaque oligodésoxyribonucléotides dans 90 µL de tampon HEPES (30 mM, pH 7,8) et en effectuant une dénaturation par un chauffage à 95 °C pendant 5 min suivie d'un refroidissement progressif (0,1 °C.s<sup>-1</sup>) jusqu'à 4 °C. Pour assembler la séquence spaceur ainsi générée avec le pCRISPomyces-2 contenant la matrice d'édition par le clonage de type Golden Gate, le plasmide pCRISPomyces-2 contenant la matrice d'édition a été linéarisé par l'enzyme de restriction *BbsI* (5'-GAAGAC(N2)↓-3') (NEB, NewEnglandBiolabs), permettant alors d'insérer la séquence spaceur par ligation grâce à l'action d'une ADN ligase (NEB, NewEnglandBiolabs). Le mélange réactionnel et le programme utilisés sont présentés dans le **Tableau 9**. Par la suite, 50 µL de bactéries compétentes (souche DH5α d'*E. coli*) ont été transformées par 3 µL du mélange réactionnel comme décrit dans la **partie VI.1**. Le criblage des clones positifs ayant incorporés le plasmide pCRISPomyces-2 contenant la séquence spaceur a été réalisé par étalement des bactéries sur boîtes de gélose LB contenant l'apramycine, et recouvertes de 10 µL d'IPTG (500 mM) et de 40 µL de bluo-gal (20 mg.mL<sup>-1</sup>) nécessaires à la réalisation du criblage blanc/bleu. Ces boîtes ont été incubées une nuit à 37 °C. Quelques clones blanc obtenus sont mis en culture et une minipréparation d'ADN plasmidique est ensuite réalisée. La présence de la séquence spaceur dans le plasmide pCRISPomyces-2 contenant la matrice d'édition a été vérifiée par séquençage.



**Figure 64 : Principe de la méthode Golden Gate (Engler *et al.*, 2008).** Cette technique présente les avantages suivants : i) la séquence de l'extrémité saillante générée n'est pas dictée par l'endonucléase de restriction – aucune séquence supplémentaire n'est donc ajoutée, ii) la séquence des extrémités saillantes, spécifique à chaque fragment, permet l'assemblage simultané de plusieurs fragments dans le bon ordre, et iii) le site de restriction est éliminé du produit ligaturé, ce qui permet de réaliser la digestion et la ligature simultanément.

**Tableau 9 : Mélange réactionnel et programme utilisé pour l'assemblage de la séquence espaceur avec le plasmide pCRISPomyces-2 contenant la matrice d'édition.**

Composants	Volume	Phase	Durée	Température	Cycles
pCRISPomyces-2	100 ng	Golden Gate	10 min	37 °C	10
Oligodésoxyribonucléotides chevauchants (dilué 10 fois)	0,3 µL		10 min	16 °C	
Tampon ADN ligase T4 (NEB, NewEnglandBiolabs) (10X)	2 µL		5 min	50 °C	
<i>Bbs</i> I (NEB, NewEnglandBiolabs) (10 U.µL <sup>-1</sup> )	1 µL		20 min	65 °C	
ADN ligase T4 (NEB, NewEnglandBiolabs) (400 U.µL <sup>-1</sup> )	1 µL		∞	20 °C	
H <sub>2</sub> O	qsp 20 µL	Conservation	∞	20 °C	1
Volume total	20 µL				

## 6. Clonage des matrices d'édition et des séquences espaceurs dans le plasmide pCRISPR-Cas9

### A. Préparation des matrices d'édition

L'amplification des fragments d'ADN et leur fusion dans le but d'assembler les matrices d'édition ont été effectués comme dans la **partie IV.5.A**. L'enzyme utilisée pour les réactions de PCR est l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific) et les oligodésoxyribonucléotides sont présentés dans le **Tableau 10**. Les conditions de réaction et le programme utilisé sont indiqués dans le **Tableau 4**. La température d'hybridation a été adaptée pour chaque couple d'oligodésoxyribonucléotides. Lorsque les fragments d'ADN devaient être fusionnés entre eux, le temps d'hybridation a été augmenté à 45 s. La présence et la taille de l'ADN amplifié ont été contrôlées par dépôt sur gel d'agarose 0,9 %. L'ADN amplifié a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A**.

**Tableau 10 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'assemblage des matrices d'édition et le séquençage du pCRISPR-Cas9.** Les séquences en majuscules et en italique correspondent au site favorisant le recrutement de l'enzyme de restriction. Les séquences en majuscules et soulignées correspondent au site de restriction inséré. Les séquences en minuscules correspondent à la borne du génome de *Streptomyces* utilisée pour amplifier le fragment d'ADN d'intérêt.

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
$\Delta$ cpkC-RH-amont	F413	direct	<i>TTTCCGAGGCCT</i> gaccgtcgaccatctgtacgc
	F412	indirect	CACCCGGCTCTCAGGAGAAGtgtggctacggctacggcttg
$\Delta$ cpkC-RH-aval	F411	direct	AAGCCGTAGCCGTAGCCACAAttctcctgagagccgggtgg
	F410	indirect	<i>TTTCCGAGGCCT</i> gcacgagctgacgaaggacgt
$\Delta$ cpkC-RH-amont	F409	direct	<i>TTTCCGAGGCCT</i> acctcagcttccccgaggtcg
	F408	indirect	GGTCGGTCATGTTCGGTCCCTtcagtgggtctgccggtacgc
$\Delta$ cpkC-RH-aval	F407	direct	CGTACCGGCAGACCCACTGAagggacgcacatgaccgacca
	F406	indirect	<i>TTTCCGAGGCCT</i> aggctcttggccagcgtgtag
Contrôle-insertion de la matrice d'édition	E815	direct	ATAGAAAGGAACA ACTAAAGG

Chaque matrice d'édition a été ensuite linéarisée par l'enzyme de restriction *StuI* (5'-AGG↓CCT-3') (Thermo Fisher Scientific) pendant 1 h à 37 °C à raison d'une unité d'enzyme par µg d'ADN. Le fragment d'ADN ainsi digéré a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans le **partie IV.1.A.**

### **B. Linéarisation du plasmide par *StuI***

Le vecteur pCRISPR-Cas 9 a été linéarisé par l'enzyme de restriction *StuI* (5'-AGG↓CCT-3') (Thermo Fisher Scientific) pendant 1 h à 37 °C à raison d'une unité d'enzyme par µg d'ADN, puis traité avec 1 unité de phosphatase alcaline (Fast-AP, Thermo Fisher Scientific) par µg d'ADN pendant 2 h à 37 °C. Le vecteur linéarisé et déphosphorylé a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A.**

### **C. Ligation de la matrice d'édition et du plasmide pCRISPR-Cas9**

La formation des liaisons phosphodiester entre les extrémités 5' phosphate de l'insert et 3' hydroxyle du vecteur est catalysée par l'ADN ligase du bactériophage T4 (Thermo Fisher Scientific) à raison d'une unité d'enzyme par µg d'ADN. La ligation est réalisée en présence d'un excès molaire d'insert de trois fois par rapport au vecteur linéarisé. Les mélanges de ligation ont été placés à température ambiante pendant une nuit. Une fraction aliquote du produit précédent a été transformée dans des cellules d'*E. coli* DH5α compétentes par choc thermique. Les cellules ont été étalées sur boîte LB contenant de l'apramycine et incubées une nuit à 37 °C. Quelques clones obtenus sont mis en culture et une minipréparation d'ADN plasmidique est ensuite réalisée. La présence de la matrice d'édition dans le plasmide pCRISPR-Cas9 a été confirmée par séquençage.

### **D. Identification et préparation des séquences spaceurs**

Les séquences spaceurs ont été choisies comme décrit dans la **partie IV.5.D.** Les séquences spaceurs ont été amplifiées de manière à inclure à leur extrémité 5' un site de restriction *SnaBI*. L'enzyme utilisée pour les réactions de PCR est l'ADN polymérase Phusion HOT START II (Thermo Fisher Scientific). Les oligodésoxyribonucléotides, les conditions de réaction et le programme utilisé sont indiqués dans les **Tableaux 11 et 4.** La température d'hybridation a été adaptée pour chaque couple d'oligodésoxyribonucléotides. La présence et la taille de l'ADN amplifié ont été contrôlées par dépôt sur gel d'agarose 0,9 %. L'ADN

amplifié a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A.**

Chaque séquence spaceur a été ensuite digérée par l'enzyme de restriction *Sna*BI (5'-TAC↓GTA-3') (Thermo Fisher Scientific) pendant 1 h à 37 °C à raison d'une unité d'enzyme par µg d'ADN. Le fragment d'ADN ainsi digéré a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A.**

**Tableau 11 : Séquence nucléotidique des oligodésoxyribonucléotides utilisés pour l'amplification des séquences spaceurs.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
ΔcpkC-ARNsg-1	F402	direct	TTTCCGCCATGGGGGAACCGGTAT CGGGATTCggttttagagctagaaatagc
ΔcpkC-ARNsg-2	F405	direct	TTTCCGCCATGGGGAGCTGTTCGA GAAGTATGgttttagagctagaaatagc
ΔcpkF-ARNsg-1	F404	direct	CATGCCATGGGGCCGGGATACCCG CCGGAGgttttagagctagaaatagc
ΔcpkF-ARNsg-2	F403	direct	CATGCCATGGGGGCTCGCCGCTCT CGGGTGgttttagagctagaaatagc
ARNsg	E758	indirect	ACGCCTACGTAAAAAAGCACCGA CTCGGTGCC

### E. Linéarisation du plasmide par *Sna*BI

Le vecteur pCRISPR-Cas 9 a été linéarisé par l'enzyme de restriction *Sna*BI (5'-TAC↓GTA-3') (Thermo Fisher Scientific) pendant 1 h à 37 °C à raison d'une unité d'enzyme par µg d'ADN, puis traité avec 1 unité de phosphatase alcaline (Fast-AP, Thermo Fisher Scientific) par µg d'ADN pendant 2 h à 37 °C. Le vecteur linéarisé et déphosphorylé a été purifié après migration sur gel d'agarose 0,9 % comme décrit dans la **partie IV.1.A.**

### F. Ligation de la séquence spaceur et du plasmide

La réaction de ligation entre le pCRISPR-Cas9 et la séquence spaceur a été effectuée comme décrit dans la **partie IV.5.C.** La présence de la séquence spaceur dans le plasmide pCRISPR-Cas9 a été confirmée par séquençage.

## V. Extraction d'ADN génomique, d'ADN plasmidique et de fragments d'ADN

### 1. Extraction de l'ADN génomique de différentes espèces de *Streptomyces*

L'extraction a été réalisée à l'aide du kit Wizard® Genomic DNA Purification (Promega). Le protocole se décompose en cinq étapes : les cellules préalablement cultivées en milieu liquide ont été centrifugées à 16000 g pendant 2 minutes à température ambiante et le surnageant a été éliminé. Le culot a été repris en présence d'EDTA (50 mM), puis les cellules ont été lysées et les ARN contaminants ont été dégradés par l'ajout de RNase (60 µg.mL<sup>-1</sup>). Ensuite, les protéines ont été précipitées et éliminées par centrifugation. L'ADN génomique contenu dans le surnageant a été précipité par l'ajout d'une solution d'isopropanol et lavé par l'ajout d'une solution d'éthanol à 70 %. L'éthanol résiduel a été séché par centrifugation et l'ADN génomique présent dans le culot a été repris dans de l'eau stérile, avant d'être dosé à 260 nm en utilisant un appareil Nanodrop2000C (Thermo Fisher Scientific) et l'intégrité de l'ADN génomique a été contrôlée par migration sur gel d'agarose 0,8 %. L'ADN génomique extrait a été conservé à -20 °C.

### 2. Extraction de l'ADN total de cellules de levure

L'équivalent de deux anses de platine de cellules de levure a été suspendu dans 200 µL d'eau stérile et 200 µL de billes de verre (425–600 µm) ont été ajoutés. Afin de permettre l'extraction de l'ADN chromosomique et plasmidique, 300 µL d'un mélange phénol-chloroforme-alcool isoamylique (25:24:1) ont été ajoutés et le mélange a été soumis au vortex pendant 10 min, puis centrifugé à 16000 g pendant 5 min. La phase aqueuse a été récupérée, additionnée de 200 µL de diéthyl éther et le mélange a été traité comme précédemment. La phase aqueuse a été récupérée et incubée à température ambiante afin d'éliminer les traces de diéthyl éther. Les extraits d'ADN total de cellules de levure ont été conservés à -20 °C jusqu'à la transformation d'*E. coli*.

### 3. Extraction d'ADN plasmique par midipréparation et minipréparation

La midipréparation d'ADN a été utilisée pour la préparation d'une grande quantité de plasmide à partir de 200 mL de cultureensemencée par 2 mL de préculture. L'ADN plasmidique a été purifié à l'aide du kit NucleoBond® Xtra Midi / Maxi (Marcherey Nagel).

### *Matériels et méthodes*

Après avoir été élué avec de l'eau, l'ADN plasmidique a été dosé à 260 nm et conservé à  $-20\text{ }^{\circ}\text{C}$ .

La minipréparation d'ADN a été utilisée pour la préparation rapide d'une faible quantité d'ADN dans un but analytique, à partir de 5 mL de culture. L'ADN plasmidique a été purifié à l'aide du kit Nucleospin<sup>®</sup> plasmid (Marcherey Nagel). Après avoir été élué avec de l'eau, l'ADN plasmidique a été dosé à 260 nm et conservé à  $-20\text{ }^{\circ}\text{C}$ .

## VI. Préparation et transformation des cellules bactériennes

### 1. Préparation et transformation de cellules compétentes d'*E. coli*

La préparation de cellules compétentes d'*E. coli* a été réalisée pour pouvoir transformer les souches d'*E. coli* DH5 $\alpha$  et ET12567 (pUZ8002). Une culture bactérienne a été réalisée dans 100 mL de milieu LB à 37 °C sous agitation. Cette dernière a été arrêtée en phase exponentielle de croissance lorsque la DO<sub>600 nm</sub> a atteint 0,5. La culture a été centrifugée à 4 °C à 4500 g pendant 15 min, et le culot a été repris dans une solution de TfbI (CaCl<sub>2</sub> 10 mM, KCl 100 mM, MnCl<sub>2</sub> 50 mM, glycérol 15 % (pH 7,0)). Après une incubation de 10 min dans la glace et une nouvelle centrifugation à 4 °C à 4500 g pendant 15 min, le culot a été repris dans 4 mL de solution TfbII (MOPS 10 mM, CaCl<sub>2</sub> 75 mM, KCl 10 mM, glycérol 15 % (pH 7,0)). Des fractions ont été aliquotées et rapidement congelées par immersion dans de l'azote liquide, puis stockées à -80 °C.

La transformation a été réalisée en mélangeant 100 ng de plasmide et 50  $\mu$ L de bactéries compétentes. Après incubation sur glace pendant 30 min, le mélange a été soumis à un choc thermique de 45 s à 42 °C. Les cellules ont été ensuite régénérées pendant 1 h à 37 °C en milieu LB avant d'être étalées sur boîte LB contenant l'antibiotique approprié au plasmide utilisé et incubées une nuit à 37 °C.

### 2. Préparation et transformation des cellules électro-compétentes d'*E. coli*

La préparation de cellules électro-compétentes est effectuée pour pouvoir transformer les souches d'*E. coli* DH5 $\alpha$  et BW25113 (pIJ790). Pour cela, deux cultures bactériennes ont été réalisées dans 250 mL de milieu LB à 37 °C sous agitation. Ces dernières ont été arrêtées en phase exponentielle de croissance lorsque la DO<sub>600 nm</sub> a atteint 0,5. Les cultures ont été refroidies dans la glace pendant 10 min, puis centrifugées à 4 °C à 4500 g pendant 15 min. Les culots ont été lavés successivement deux fois dans une solution de glycérol 5 % glacé et une fois dans une solution de glycérol 10 % glacé. Le culot a été finalement repris dans une solution de glycérol 10 % glacé, puis des fractions ont été aliquotées et rapidement congelées par immersion dans de l'azote liquide et stockées à -80°C.

La transformation a été réalisée en mélangeant 10  $\mu$ L d'ADN total de levure, préalablement dialysé sur une membrane d'esters de cellulose contenant des pores d'une dimension de 0,025  $\mu$ m pendant 30 min (afin d'éliminer les traces de sels), et 50  $\mu$ L de bactéries électro-compétentes. Le mélange a été placé dans une cuve d'électroporation glacée dont les

électrodes sont espacées de 0,2 cm. L'ensemble a été placé dans l'appareil d'électroporation (Eppendorf Eporator<sup>®</sup>) puis soumis à un choc électrique de 2500 V pendant environ 5 ms. Les cellules transformées ont été reprises immédiatement dans 1 mL de LB froid et régénérées pendant 1 h à 37 °C, avant d'être étalées sur boîte LB contenant l'antibiotique approprié au plasmide utilisé, et incubées une nuit à 37 °C.

### **3. Préparation des spores et transfert des plasmides d'*E. coli* à *S. coelicolor* par conjugaison**

#### **A. Extraction des spores**

La préparation des spores est effectuée pour pouvoir transformer par conjugaison les souches de *S. coelicolor* et de *Sac. erythraea*. Pour cela, une boîte de milieu de sporulation, SFM pour *S. erythraea*, ABB13 pour *Sac. erythraea*, a étéensemencée à raison de 10<sup>6</sup> spores. La boîte de sporulation a été enveloppée dans de l'aluminium et placée à 30 °C pendant 5 jours. Lorsque la sporulation était suffisante, les spores ont été récoltées en déposant 10 mL d'eau stérile filtrée sur la gélose et en frottant légèrement avec un coton tige stérile. La suspension a été ensuite filtrée au moyen d'un système de filtration contenant du coton cardé stérile, puis placée dans un tube Falcon de 50 mL et finalement conservée à 4 °C. Pour une conservation supérieure à deux mois, la suspension de spore a été mélangée avec une solution de glycérol 40 % v/v et stockées à -80°C.

#### **B. Préparation, transformation par conjugaison de *S. coelicolor* et criblage des exconjuguants**

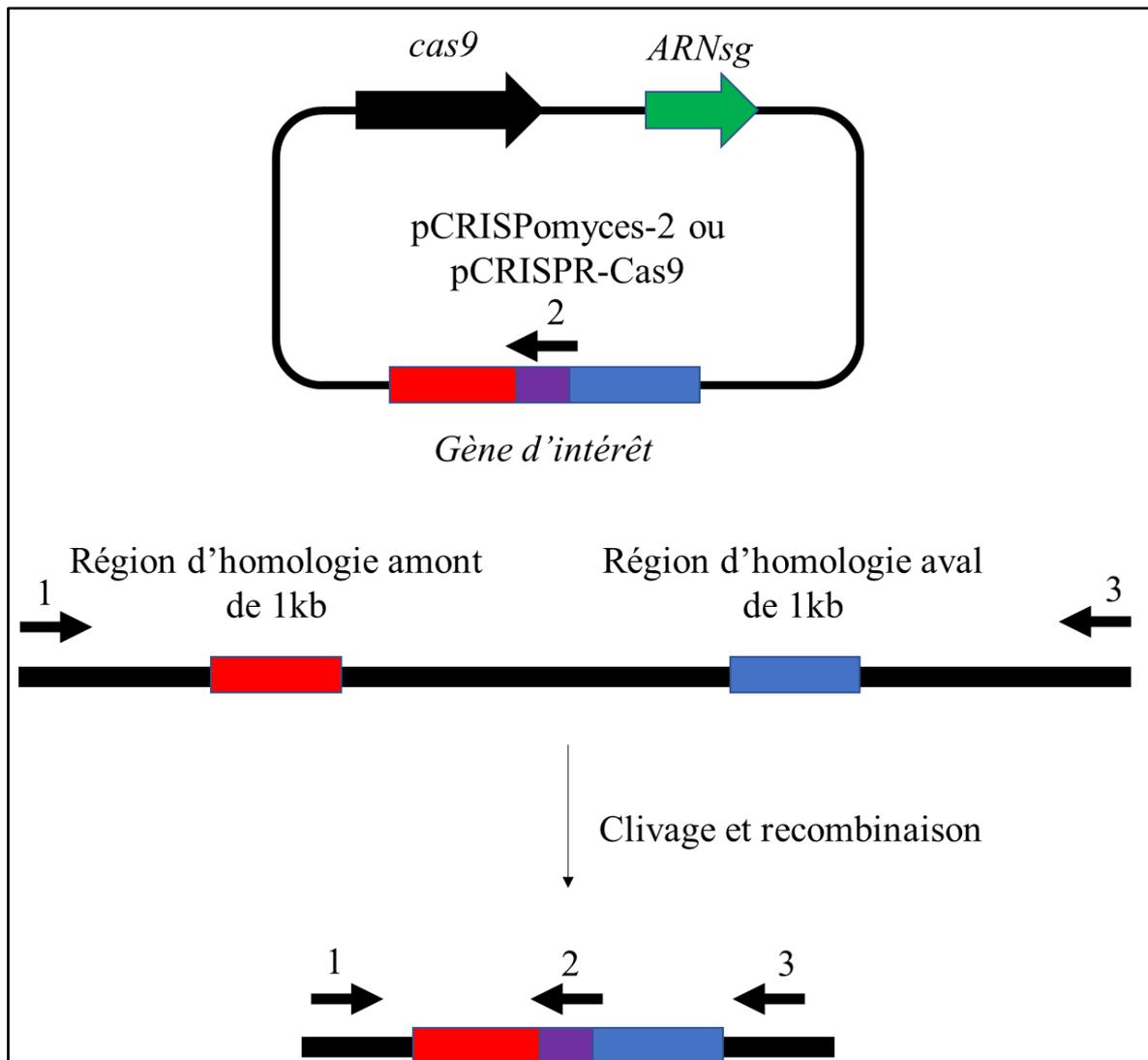
##### **a. Transformation par conjugaison de *S. coelicolor***

Les plasmides génétiquement modifiés pCRISPomyces-2 ou pCRISPR-Cas9 ont chacun été introduits dans la souche ET12567 d'*E. coli* contenant le plasmide pUZ8002 par électroporation et transférés à la souche A3(2) FM145 de *S. coelicolor* par conjugaison. À la suite de la sélection sur boîtes de gélose LB contenant de l'apramycine, un unique clone de la souche ET12567 (pUZ8002) d'*E. coli* a permis d'ensemencer une pré-culture de 5 mL de milieu LB, additionnée d'apramycine, de chloramphénicol et de kanamycine, qui a été incubée sous agitation à 37 °C pendant une nuit. Un volume de 1 mL de la pré-culture a été prélevé pour inoculer 25 mL de milieu LB supplémenté avec de l'apramycine du chloramphénicol et de la kanamycine. La culture, incubée sous agitation à 37 °C jusqu'à une concentration cellulaire de

0,3 DO<sub>600 nm</sub>, a été centrifugée à 1780 g pendant 12 minutes à température ambiante et le culot a été repris dans 25 mL de milieu LB. Cette dernière étape a été répétée deux fois et le culot a été repris dans 2,5 mL de milieu 2YT. En parallèle, plusieurs réplicats de 500 µL de milieu 2YT inoculés par une suspension de 10<sup>6</sup> spores de la souche de A3(2) FM145 de *S. coelicolor* ont été chauffés à 50 °C pendant 10 minutes et centrifugés à 4000 g pendant 1 minute, avant que les culots ne soient chacun repris à l'aide de 500 µL de bactéries *E. coli*. Les suspensions ont été directement étalées sur boîtes de gélose SFM contenant du MgCl<sub>2</sub> (10mM) et placées à 30 °C pendant 16 heures. Ensuite, de l'apramycine et de l'acide nalidixique ont été ajoutés sur les boîtes afin de sélectionner les exconjugants et d'éliminer les bactéries de la souche ET12567 (pUZ8002) d'*E. coli*, respectivement. Les boîtes ont de nouveau été incubées à 30 °C durant 3 jours. Les exconjugants obtenus ont été repiqués sur boîtes de gélose SFM contenant du MgCl<sub>2</sub> (10 mM) et de l'apramycine, et placées 30°C pendant 3 jours. Pour les cellules de *S. coelicolor* transformées avec les plasmides pCRISPR-Cas9 génétiquement modifiés, le milieu SFM contenant du MgCl<sub>2</sub> est supplémenté avec de la thiostreptone à raison de 1 µg.L<sup>-1</sup>.

### b. Géotypage des exconjugants par PCR sur colonies

Pour cribler l'événement de modification de l'ADN génomique de *S. coelicolor*, 0,1 cm<sup>2</sup> du mycélium d'une centaine de colonies a été prélevé et placé ensuite dans 10 µL de DMSO. Le mélange a été incubé pendant 10 min à 100°C, puis centrifugé pendant 10 s à 16000 g. Un µL de la préparation a été utilisé en tant que matrice lors de la PCR sur colonies. Pour s'assurer que le produit PCR ait été amplifié à partir du chromosome plutôt que du plasmide recombinant, les oligodésoxyribonucléotides présentés dans les **Tableaux 9** ont été conçus de sorte à s'apparier environ 500 pb en amont et en aval des régions d'homologie de la cassette de délétion sur le chromosome bactérien (**Figure 65**). Les conditions de réaction et le programme sont indiqués dans le **Tableau 12**. L'enzyme utilisée pour la réaction de PCR sur colonie l'ADN polymérase DreamTaq (Thermo Fisher Scientific). La température d'hybridation a été ajustée pour chaque couple d'oligodésoxyribonucléotides. La présence et la taille de l'ADN amplifié ont été contrôlées par dépôt sur gel d'agarose 0,9 %.



**Figure 65 : Stratégie pour le criblage de l'événement de modification génétique souhaité.** La flèche 1 désigne l'oligodésoxyribonucléotide à utiliser dans le sens direct. La flèche 2 désigne l'oligodésoxyribonucléotide à utiliser dans le sens indirect dans le cas de l'insertion d'un gène. La flèche 3 désigne l'oligodésoxyribonucléotide à utiliser dans le sens indirect dans le cas de la délétion d'un gène.

**Tableau 12 : Composition du mélange réactionnel et conditions expérimentales utilisées pour les réactions de PCR sur colonies dans le cadre du criblage des événements de modification dans l'ADN génomique des *Streptomyces*.**

Nom	Nomenclature ESCOLA	Sens	Séquence ADN (5'→3')
Contrôle- mCherry-cpkA- amont	E016	direct	acgccccaggactccagcagc
	E015	indirect	gaggacggcgccctgaagggc
Contrôle- mCherry-cpkA- aval	E014	direct	ttggagccgtacatgaactgc
	E013	indirect	caggttggtcgggaacttgta
Contrôle-cpkC- eGFP-amont	D769	direct	atcacggggcggtggttacg
	D768	indirect	ctaccagcagaacacccccat
Contrôle- cpkC- eGFP -aval	D767	direct	gccagggcacgggcagctt
	D766	indirect	ccacgctcgtcttcgaccatccg
Contrôle-cpkF- mCherry-amont	E483	direct	aggcctacgtcaagcaccggg
	E482	indirect	cccgcaccgtctcggggaagt
Contrôle- cpkF- mCherry -aval	E483	direct	atgaacttcgaggacggcggc
	E482	indirect	cgatgaggcaggagaaggcgt
Contrôle - $\Delta$ cpkA	E848	direct	accacgtgcgcgttggtgccg
	E847	indirect	atcgccgctcgatcggcgag
Contrôle - $\Delta$ cpkC	F132	direct	atcaactacgccgacgccgac
	F133	indirect	tccgaactggaccggctggag
Contrôle - $\Delta$ cpkF	E848	direct	ggtaccgcgccaacaacctcc
	E847	indirect	tggtcaccgctcgacgtcg

Composants	Volume
Matrice	1 $\mu$ L
dNTP ( 1,25 mM chaque)	4 $\mu$ L
Oligodésoxyribonucléotides sens direct (10 mM)	1 $\mu$ L
Oligodésoxyribonucléotides sens direct (10 mM)	1 $\mu$ L
DMSO (100 %)	5 $\mu$ L
Tampon Dream Taq (10X)	0,5 $\mu$ L
Dream Taq (1,25 U. $\mu$ L <sup>-1</sup> )	0,5 $\mu$ L
H <sub>2</sub> O	qsp 25 $\mu$ L
Volume total	25 $\mu$ L

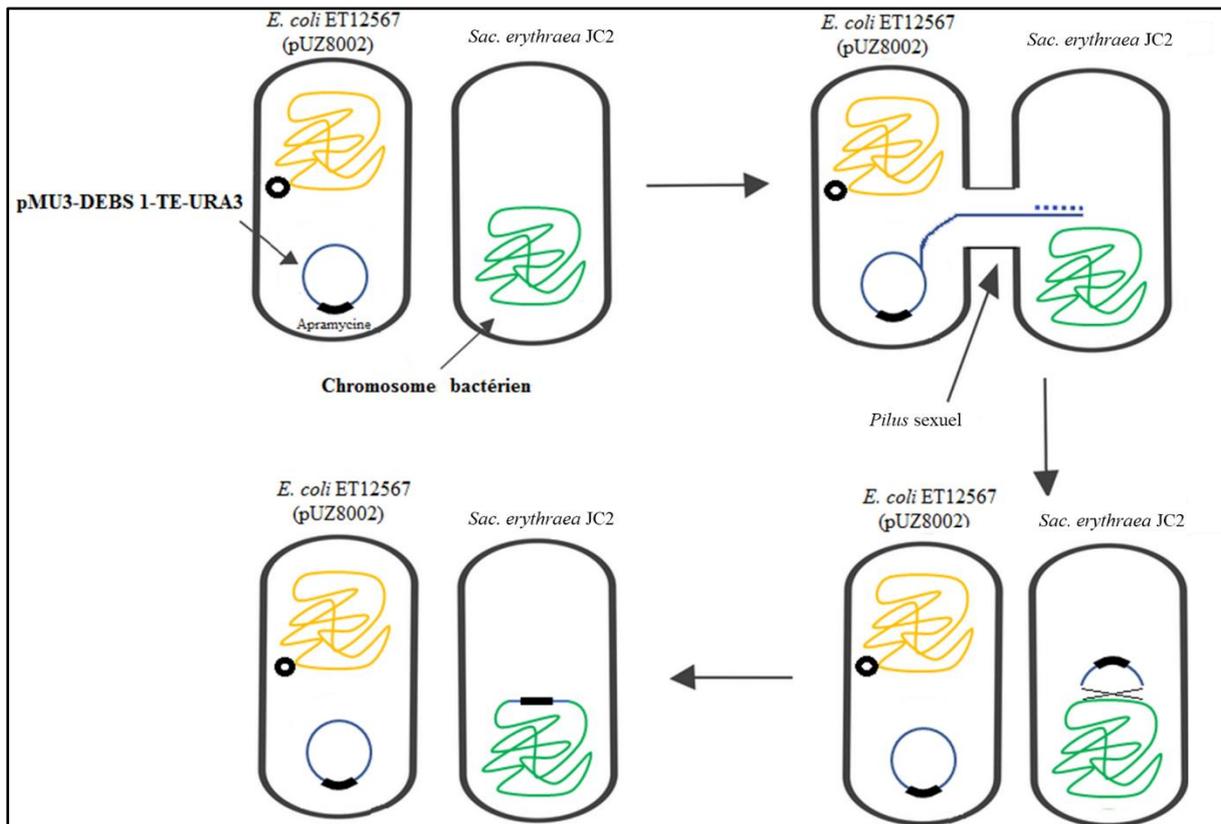
Phase	Durée	Température	Cycles
Dénaturation initiale	2 min	95 °C	1
Dénaturation	30 s	95 °C	30
Hybridation	30 s	50 °C	
Elongation	4 min	72 °C	
Elongation finale	10 min	72 °C	
Conservation	$\infty$	20 °C	1

### c. Elimination des plasmides génétiquement modifiés

L'élimination des plasmides a été effectuée en exploitant le réplicon thermosensible pSG5 contenu dans les plasmides pCRISPomyces-2 et pCRISPR-Cas9. Pour cela, une culture à 37 °C pendant 3 jours a été effectuée, suivie par un ensemencement sur des boîtes sélectives et non sélectives pour confirmer la restauration de la sensibilité à l'apramycine.

### C. Transformation par conjugaison et culture de *Sac. erythraea* JC2

Les différents plasmides génétiquement modifiés ou non pMU3-DEBS 1-TE, pBF21, pBF23, pBF25 et pBF30 ont été utilisés pour transformer la souche ET12567 (pUZ8002) d'*E. coli* rendue compétente. Après sélection sur milieu LB contenant de l'apramycine, de la kanamycine et du chloramphénicol, les clones ont été mis en culture dans 3,5 mL de milieu 2YT contenant de l'apramycine, de la kanamycine et du chloramphénicol à 37 °C pendant une nuit. Un volume de 1 mL de cette préculture a été utilisé pour inoculer 25 mL de LB contenant du chloramphénicol, de la kanamycine et de l'apramycine. Lorsque la culture a atteint une  $DO_{600\text{ nm}}$  de 0,3, elle a été centrifugée pendant 12 min à 1780 g et le culot a été repris dans 25 mL de 2YT. Cette dernière étape a été répétée deux fois et le culot a été repris dans 2,5 mL de 2YT. De manière concomitante, 10 mL de GMTSB inoculés par 200 µL de mycélium de *Sac. erythraea* JC2 ont été cultivés pendant deux jours à 30 °C. Ensuite, 1 mL de culture de *Sac. erythraea* JC2 a été centrifugé à 16000 g pendant 1 min et le culot obtenu a été repris dans 0,5 mL de 2YT. Ces 500 µL ont été ajoutés aux 500 µL de culture d'*E. coli*, puis ce mélange de bactéries a été étalé sur boîte ABB13. Le principe de cette conjugaison est montré **Figure 66**. Les boîtes ont ensuite été placées pendant 24 h à 30 °C. Puis de l'apramycine et de l'acide nalidixique ont été étalées sur la boîte qui a été ensuite incubée à 30 °C pendant 12 jours. L'acide nalidixique a pour fonction d'éliminer les cellules d'*E. coli* ET12567 (pUZ8002) et l'apramycine permet de maintenir une pression de sélection sur les cellules de *Sac. erythraea* JC2 transformées par conjugaison par le plasmide recombinant. Huit clones bactériens obtenus ont été repris dans 30 µL d'eau, puis étalés sur boîte ABB13 contenant de l'apramycine et de l'acide nalidixique et placés à 30 °C pendant 13 jours.



**Figure 66 : Représentation schématique du protocole de conjugaison entre des cellules d'*E. coli* ET12567 (pUZ8002) et de *Sac. erythraea* JC2, ainsi que l'événement de modification génétique.** Les cellules d'*E. coli* ET 12567 (pUZ8002) compétentes pour la conjugaison sont transformées avec le vecteur pMU3-DEBS 1-TE-URA3 recombinant ou non. En jaune est représenté l'ADN génomique d'*E. coli* ET12567. En noir est représenté le plasmide pUZ8002 possédant l'opéron *tra* codant des protéines de transfert d'ADN plasmidique. En bleu est représenté le plasmide pMU3-DEBS 1-TE-URA3 possédant le gène de résistance à l'apramycine qui est symbolisé par la cassette en noir. En vert est représenté l'ADN génomique de *Sac. erythraea* JC2. Le vecteur pMU3-DEBS 1-TE-URA3 est transféré aux cellules de *Sac. erythraea* JC2 par conjugaison. Les cellules de *Sac. erythraea* JC2 transformées sont sélectionnées par ajout d'apramycine et d'acide nalidixique. Le vecteur n'est pas capable de se répliquer à l'intérieur des cellules de *Sac. erythraea* JC2 et s'intègre dans le génome de la bactérie.

## VII. Production des métabolites spécialisés

### 1. Production de coelimycine

#### A. Production de coelimycine en milieu solide

Cent  $\mu\text{L}$  contenant  $10^7$  spores de la souche de *S. coelicolor* FM145 et de ses dérivés ont été étalés sur le milieu DNA supplémenté ou non en acide glutamique. Puis la boîte a été enveloppée dans de l'aluminium pour la conserver dans le noir et incubée à 30 °C pendant 27 heures. Après incubation, la production du pigment jaune a été enregistrée.

#### B. Production de coelimycine en milieu liquide

Cette méthode est adaptée de Pawlik *et coll.* (Pawlik *et al.*, 2010). Cent  $\mu\text{L}$  contenant  $10^8$  de la souche *S. coelicolor* ont étéensemencés dans 50 ml de DNB supplémenté ou non en acide glutamique dans un erlenmeyer de 250 mL (contenant un ressort en acier inoxydable) et incubés sous agitation à 220 tr / min à 30 °C pendant 27 h jusqu'à une  $\text{DO}_{450 \text{ nm}}$  de 1,2. Puis, 1 mL de culture a été centrifugé à 4000 g. Pour quantifier la production de pigment jaune, le surnageant a été utilisé pour mesurer l'absorbance  $A_{400-700 \text{ nm}}$  au moyen d'un spectrophotomètre. Pour déterminer la croissance de la culture bactérienne, le culot bactérien a été repris dans du tampon de lyse de *Streptomyces* et l'absorbance  $A_{450 \text{ nm}}$  a été mesurée au moyen d'un spectrophotomètre.

### 2. Production, extraction, et analyses des lactones

#### A. Production de lactones

Un  $\text{cm}^2$  du mycélium obtenu a été repris dans 200  $\mu\text{L}$  de TSB puis inoculé dans 10 mL de TSB contenant de l'apramycine puis placé à 30 °C pendant 48 h. Trois cents  $\mu\text{L}$  de cette culture ont été étalés sur boîte SM3 contenant de l'apramycine puis incubés à 30 °C pendant 13 jours. Cinq cents  $\mu\text{L}$  de cette même culture ont été placés dans 20 % de glycérol et conservés à -80 °C.

#### B. Production des standards

Les quatre diastéréoisomères de la lactone dont la configuration varie seulement au niveau des carbones C2 et C3, ainsi que la cétolactone, sont produits grâce à des cellules de *Sac. erythraea* JC2 transformées par les plasmides pJLK25, pJLK30 et pJLK35 codant les PKS

qui synthétisent ces produits et obtenues par l'équipe du Professeur Leadlay (Kellenberger *et al.*, 2008). Ces cellules ont été mises en culture à 30 °C pendant 48 h dans 10 mL de TWM contenant de la thiostreptone. Trois cents µL de la culture ont été ensuite déposés sur boîte SM3 contenant de la thiostreptone et incubés à 30 °C pendant 13 jours. Les lactones produites ont été extraites puis analysées par chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC/MS) sur le Plateau d'Analyse Structurale et Métabolomique (PASM) de l'École Nationale Supérieure d'Agronomie et des Industries Alimentaires (ENSAIA) de Nancy.

### **C. Extraction des lactones**

Deux cm<sup>2</sup> de tapis bactérien provenant de la boîte SM3 contenant de l'apramycine (50 µg.mL<sup>-1</sup>) ont été prélevés stérilement et placés dans un tube Eppendorf. Un mélange de 1,2 mL d'acétate d'éthyle et de 20 µL d'acide formique a été ajouté, puis le tout a été incubé à 50 °C pendant 15 min. Après être mélangé pendant 30 min à l'aide d'un vortex, l'ensemble a été centrifugé 1 min à 16000 g. Un mL du surnageant a été récupéré et évaporé à l'aide d'un concentrateur sous vide (Concentrator plus, Eppendorf) pendant 30 min. Le culot a été repris dans 100 µL d'acétate d'éthyle et l'extrait a été analysé par GC/MS.

### **D. Analyse des lactones**

Les lactones ont été analysées par GC/MS sur un appareil 7890A-5975C (Agilent Technologies) équipé avec une colonne capillaire J&W DB 5 MS (30 m × 0,25 mm ; 0,25 µm ; Agilent Technologies). Les températures de l'injecteur et du détecteur ont été réglées à 250 °C. Le gaz vecteur est l'hélium à 1 mL.min<sup>-1</sup>. Le volume d'injection a été de 4 µL. Le profil de température du four est : 2 min à 40 °C après injection puis augmentation de 10 °C par min, jusqu'à atteindre 250 °C, puis 25 °C par min de hausse jusqu'à atteindre 300 °C et la température a été maintenue à 300 °C pendant 2 min. Les produits ont été analysés par spectrométrie de masse où les ions subissent une fragmentation dans une source à impact électronique.

## VIII. Détection des sous-unités de la CPKS par immuno-empreinte

### 1. Préparation des échantillons

Les souches de *S. coelicolor* ont été cultivées dans du milieu M79 et les cellules ont été récoltées à différents temps, entre 17 h et 24 h de culture. La croissance bactérienne a été suivie à DO<sub>450 nm</sub>. Au temps de culture désiré, 10 mL de culture ont été centrifugés pendant 5 min à 4000 g. Puis, le culot a été lavé dans du tampon de lyse de *S. coelicolor* et repris dans 500 µL de ce même tampon. Les cellules ont été ensuite lysées par sonication (Branson) à raison de 5 × 1 min, puissance 4 dans la glace et le lysat a été centrifugé pendant 40 min à 4 °C à 16000 g. Les protéines du surnageant ont été précipitées en ajoutant 3 volumes d'acétone froid, puis le mélange a été brièvement mélangé par vortex pour être ensuite incubé à -20 °C pendant 1 h. L'extrait a été centrifugé à 4 °C pendant 10 min à 16 000 g, et le culot a été finalement repris dans 100 µL de tampon de lyse de *S. coelicolor*.

### 2. Analyse par Dot Blot

Une fraction aliquote (5 µL) de l'extrait total a été déposée sur une membrane de nitrocellulose contenant des pores d'une dimension de 0,45 µm (Biorad). Pour la détection des protéines, la membrane a été bloquée dans le tampon de blocage pendant 1 h à température ambiante sous agitation. Après 3 lavages dans le tampon TBST, la membrane a été incubée pendant une nuit à 4 °C avec les anticorps primaires anti-GFP (abcam, 1/3000<sup>ème</sup>) ou anti-mCherry (GenTex, 1/4000<sup>ème</sup>) sous agitation. Après 3 lavages dans le tampon TBST, la membrane a été incubée pendant 1 h à température ambiante en présence de l'anticorps secondaire adapté (IgG anti-souris pour anti-mCherry, IgG anti-chèvre pour anti-GFP) (Thermo Fisher Scientific) couplé à l'HRP (*Horse Radish Peroxidase*) dilué au 1/10000<sup>ème</sup> dans du TBST / lait (5 %) sous agitation. Après 3 nouveaux lavages, les protéines d'intérêt sont révélées par chimioluminescence (kit ECL<sup>TM</sup> Prime Western Blotting Detection Reagent, GE Healthcare) et détectées par le système Fusion Solo (Vilbert Laurmat) et le logiciel FusionCapt Advanced.

### 3. Analyse par Western Blot

Dix µL d'extrait total ont été déposés sur un gel de polyacrylamide dénaturant à 6 % et la migration électrophorétique a été effectuée pendant 1 h 30 min à 150 V et 50 mA. Les protéines sont transférées ensuite en milieu liquide sur une membrane de polyfluorure de vinylidène

contenant des pores d'une dimension de 0,45  $\mu\text{m}$  (Immobilon-P, Merck) dans le tampon de transfert pendant 1 h 30 min à 100 V et 450 mA. La détection des protéines a été effectuée ensuite comme pour l'analyse par Dot Blot.

## IX. Localisation des sous-unités de la CPKS chez *S. coelicolor* par microscopie

### 1. Préparation des échantillons

Dix  $\mu\text{L}$  d'une suspension de  $10^7$  spores ont été étalés sur le milieu DNA contenant 10 mM de glutamate, puis une lamelle a été insérée avec un angle de  $45^\circ$  dans la gélose dans le but d'isoler les hyphes aériens des *Streptomyces* (**Figure 67**). Puis la boîte a été enveloppée dans de l'aluminium pour la conserver dans le noir et incubée à  $30^\circ\text{C}$  pendant 27 heures.



**Figure 67 : Exemple de préparation d'un échantillon pour l'analyse par microscopie.** La souche FM 145 (WT sur la boîte) a étéensemencée à la base d'une lamelle à raison de  $10^7$  spores sur le milieu DNA contenant 10 mM de glutamate. A droite,  $5\ \mu\text{L}$  de la suspension ont étéensemencés comme contrôle de la production de coelimycine P2.

### 2. Acquisition des images par microscopie

Des échantillons ont été observés en utilisant un microscope confocal à balayage laser inversé Leica TCS SP8 utilisant des détecteurs hybrides. Le signal fluorescent correspondant à la protéine eGFP a été analysé à des longueurs d'onde de 488 pour l'excitation et de 500 à 548 nm pour l'émission. Le signal fluorescent correspondant à la protéine mCherry a été analysé à des longueurs d'onde de 522 nm pour l'excitation et de 586-789 nm pour l'émission. Pour analyser la co-localisation, le mode de balayage séquentiel a été utilisé pour éviter la diaphonie. Les images ont été traitées à l'aide du logiciel Fiji-Image J.

# **Références bibliographiques**



1. Acuner Ozbabacan, S.E., Engin, H.B., Gursoy, A., and Keskin, O. (2011). Transient protein-protein interactions. *Protein Eng. Des. Sel.* *24*, 635–648.
2. Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B.O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R.E., *et al.* (2002). The Calicheamicin Gene Cluster and Its Iterative Type I Eneidyne PKS. *Science* *297*, 1173–1176.
3. Alekseyev, V.Y., Liu, C.W., Cane, D.E., Puglisi, J.D., and Khosla, C. (2007). Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Sci.* *16*, 2093–2107.
4. Ames, B.D., Nguyen, C., Bruegger, J., Smith, P., Xu, W., Ma, S., Wong, E., Wong, S., Xie, X., Li, J.W.-H., *et al.* (2012). Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* *109*, 11144–11149.
5. Andersson, D.I., and Hughes, D. (2014). Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* *12*, 465–478.
6. Annavaal, T., Paris, C., Leadlay, P.F., Jacob, C., and Weissman, K.J. (2015a). Evaluating ketoreductase exchanges as a means of rationally altering polyketide stereochemistry. *Chembiochem* *16*, 1357–1364.
7. Arnison, P.G., Bibb, M.J., Bierbaum, G., Bowers, A.A., Bugni, T.S., Bulaj, G., Camarero, J.A., Campopiano, D.J., Challis, G.L., Clardy, J., *et al.* (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* *30*, 108–160.
8. Arsic, B., Barber, J., Čikoš, A., Mladenovic, M., Stankovic, N., and Novak, P. (2018). 16-membered macrolide antibiotics: a review. *Int. J. Antimicrob. Agents* *51*, 283–298.
9. Arthur, C.J., Szafranska, A., Evans, S.E., Findlow, S.C., Burston, S.G., Owen, P., Clark-Lewis, I., Simpson, T.J., Crosby, J., and Crump, M.P. (2005). Self-malonylation is an intrinsic property of a chemically synthesized Type II polyketide synthase acyl carrier protein. *Biochemistry* *44*, 15414–15421.
10. Baerga-Ortiz, A., Popovic, B., Siskos, A.P., O’Hare, H.M., Spitteller, D., Williams, M.G., Campillo, N., Spencer, J.B., and Leadlay, P.F. (2006). Directed mutagenesis

- alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase. *Chem. Biol.* *13*, 277–285.
11. 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., *et al.* (2019). Structural insights into dehydratase substrate selection for the borrelidin and fluvirucin polyketide synthases. *J. Ind. Microbiol. Biotechnol.* *46*, 1225–1235.
  12. Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J.P., Klenk, H.-P., Clément, C., Ouhdouch, Y., and van Wezel, G.P. (2016). Taxonomy, physiology, and natural products of actinobacteria. *Microbiol. Mol. Biol. Rev.* *80*, 1–43.
  13. Bechet, J., Greenson, M., and Wiame, J.M. (1970). Mutations affecting the repressibility of arginine biosynthetic enzymes in *Saccharomyces cerevisiae*. *Eur. J. Biochem.* *12*, 31–39.
  14. Bedford, D., Jacobsen, J.R., Luo, G., Cane, D.E., and Khosla, C. (1996). A functional chimeric modular polyketide synthase generated *via* domain replacement. *Chem. Biol.* *3*, 827–831.
  15. Bednarz, B., Kotowska, M., and Pawlik, K.J. (2019). Multi-level regulation of coelimycin synthesis in *Streptomyces coelicolor* A3(2). *Appl. Microbiol. Biotechnol.* *103*, 6423–6434.
  16. Beld, J., Lee, D.J., and Burkart, M.D. (2015). Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Mol Biosyst* *11*, 38–59.
  17. 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., *et al.* (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* *417*, 141–147.
  18. Bills, G.F., and Gloer, J.B. (2016). Biologically active secondary metabolites from the fungi. *Microbiol Spectr* *4*.
  19. Blin, K., Pedersen, L.E., Weber, T., and Lee, S.Y. (2016). CRISPy-web: An online resource to design sgRNAs for CRISPR applications. *Synthetic and Systems Biotechnology* *1*, 118–121.

20. Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G.Y., Labouesse, M., Minvielle-Sebastia, L., and Lacroute, F. (1991). A family of low and high copy replicative, integrative and single-stranded *S. cerevisiae*/*E. coli* shuttle vectors. *Yeast* 7, 609–615.
21. Bonnett, S.A., Whicher, J.R., Papireddy, K., Florova, G., Smith, J.L., and Reynolds, K.A. (2013). Structural and stereochemical analysis of a modular polyketide synthase ketoreductase domain required for the generation of a *cis*-alkene. *Chem. Biol.* 20, 772–783.
22. Bretschneider, T., Heim, J.B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013). Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* 502, 124–128.
23. Broadhurst, R.W., Nietlispach, D., Wheatcroft, M.P., Leadlay, P.F., and Weissman, K.J. (2003). The structure of docking domains in modular polyketide synthases. *Chem. Biol.* 10, 723–731.
24. Buchholz, T.J., Geders, T.W., Bartley, F.E., Reynolds, K.A., Smith, J.L., and Sherman, D.H. (2009). Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem. Biol.* 4, 41–52.
25. Butler, M.S., Robertson, A.A.B., and Cooper, M.A. (2014). Natural product and natural product derived drugs in clinical trials. *Nat Prod Rep* 31, 1612–1661.
26. Butler, M.S., Blaskovich, M.A., and Cooper, M.A. (2017). Antibiotics in the clinical pipeline at the end of 2015. *J. Antibiot.* 70, 3–24.
27. Caffrey, P. (2003). Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. *Chembiochem* 4, 654–657.
28. Carmody, M., Byrne, B., Murphy, B., Breen, C., Lynch, S., Flood, E., Finnan, S., and Caffrey, P. (2004). Analysis and manipulation of amphotericin biosynthetic genes by means of modified phage KC515 transduction techniques. *Gene* 343, 107–115.
29. Celler, K., Koning, R.I., Willemsse, J., Koster, A.J., and van Wezel, G.P. (2016). Cross-membranes orchestrate compartmentalization and morphogenesis in *Streptomyces*. *Nat Commun* 7.

30. Challis, G.L. (2014). Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J Ind Microbiol Biotechnol* 41, 219–232.
31. Challis, G.L., and Hopwood, D.A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc. Natl. Acad. Sci. U.S.A.* 100 Suppl 2, 14555–14561.
32. Chemler, J.A., Tripathi, A., Hansen, D.A., O’Neil-Johnson, M., Williams, R.B., Starks, C., Park, S.R., and Sherman, D.H. (2015). Evolution of efficient modular polyketide synthases by homologous recombination. *J. Am. Chem. Soc.* 137, 10603–10609.
33. Chen, A., Re, R.N., and Burkart, M.D. (2018). Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions. *Nat Prod Rep* 35, 1029–1045.
34. Chen, S., Xue, Y., Sherman, D.H., and Reynolds, K.A. (2000). Mechanisms of molecular recognition in the pikromycin polyketide synthase. *Chem. Biol.* 7, 907–918.
35. Chen, S., Zheng, G., Zhu, H., He, H., Chen, L., Zhang, W., Jiang, W., and Lu, Y. (2016). Roles of two-component system AfsQ1/Q2 in regulating biosynthesis of the yellow-pigmented coelimycin P2 in *Streptomyces coelicolor*. *FEMS Microbiology Letters* 363.
36. Chen, X.-H., Vater, J., Piel, J., Franke, P., Scholz, R., Schneider, K., Koumoutsis, A., Hitzeroth, G., Grammel, N., Strittmatter, A.W., *et al.* (2006). Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42. *J Bacteriol* 188, 4024–4036.
37. Chen, X.H., Koumoutsis, A., Scholz, R., Eisenreich, A., Schneider, K., Heinemeyer, I., Morgenstern, B., Voss, B., Hess, W.R., Reva, O., *et al.* (2007). Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Nat. Biotechnol.* 25, 1007–1014.
38. Cheng, Y.-Q., Tang, G.-L., and Shen, B. (2003). Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3149–3154.
39. Cimermancic, P., Medema, M.H., Claesen, J., Kurita, K., Wieland Brown, L.C., Mavrommatis, K., Pati, A., Godfrey, P.A., Koehrsen, M., Clardy, J., *et al.* (2014).

- Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158, 412–421.
40. Clardy, J., and Walsh, C. (2004). Lessons from natural molecules. *Nature* 432, 829–837.
  41. Cobb, R.E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas System. *ACS Synth Biol* 4, 723–728.
  42. Cogan, D.P., Li, X., Sevillano, N., Mathews, I.I., Matsui, T., Craik, C.S., and Khosla, C. (2020). Antibody probes of module 1 of the 6-deoxyerythronolide B synthase reveal an extended conformation during ketoreduction. *J. Am. Chem. Soc.* 142, 14933–14939.
  43. Cortes, J., Haydock, S.F., Roberts, G.A., Bevitt, D.J., and Leadlay, P.F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* 348, 176–178.
  44. Cortes, J., Wiesmann, K.E., Roberts, G.A., Brown, M.J., Staunton, J., and Leadlay, P.F. (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* 268, 1487–1489.
  45. Cragg, G.M., Grothaus, P.G., and Newman, D.J. (2014). New horizons for old drugs and drug leads. *J. Nat. Prod.* 77, 703–723.
  46. Davies, J. (2013). Specialized microbial metabolites: functions and origins. *J. Antibiot.* 66, 361–364.
  47. Davies, J., Spiegelman, G.B., and Yim, G. (2006). The world of subinhibitory antibiotic concentrations. *Curr. Opin. Microbiol.* 9, 445–453.
  48. Del Vecchio, F., Petkovic, H., Kendrew, S.G., Low, L., Wilkinson, B., Lill, R., Cortés, J., Rudd, B.A.M., Staunton, J., and Leadlay, P.F. (2003). Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol.* 30, 489–494.
  49. Demain, A.L. (2014). Importance of microbial natural products and the need to revitalize their discovery. *J. Ind. Microbiol. Biotechnol.* 41, 185–201.
  50. Devereaux, A.L., Mercer, S.L., and Cunningham, C.W. (2018). DARK Classics in chemical neuroscience: morphine. *ACS Chem Neurosci* 9, 2395–2407.

51. Dhakal, D., Sohng, J.K., and Pandey, R.P. (2019). Engineering actinomycetes for biosynthesis of macrolactone polyketides. *Microb. Cell Fact.* *18*, 137.
52. Dobro, M.J., Oikonomou, C.M., Piper, A., Cohen, J., Guo, K., Jensen, T., Tadayon, J., Donermeyer, J., Park, Y., Solis, B.A., *et al.* (2017). Uncharacterized bacterial structures revealed by electron cryotomography. *J Bacteriol* *199*.
53. Dodge, G.J., Maloney, F.P., and Smith, J.L. (2018). Protein-protein interactions in “*cis*-AT” polyketide synthases. *Nat Prod Rep* *35*, 1082–1096.
54. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J., and Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* *252*, 675–679.
55. Donadio, S., McAlpine, J.B., Sheldon, P.J., Jackson, M., and Katz, L. (1993). An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl. Acad. Sci. U.S.A.* *90*, 7119–7123.
56. Dorival, J., Annaval, T., Risser, F., Collin, S., Roblin, P., Jacob, C., Gruez, A., Chagot, B., and Weissman, K.J. (2016). Characterization of Intersubunit Communication in the virginiamycin *trans*-acyl transferase polyketide synthase. *J. Am. Chem. Soc.* *138*, 4155–4167.
57. Dunn, B.J., Cane, D.E., and Khosla, C. (2013). Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase. *Biochemistry* *52*, 1839–1841.
58. Dutta, S., Whicher, J.R., Hansen, D.A., Hale, W.A., Chemler, J.A., Congdon, G.R., Narayan, A.R.H., Håkansson, K., Sherman, D.H., Smith, J.L., *et al.* (2014). Structure of a modular polyketide synthase. *Nature* *510*, 512–517.
59. Dutton, C.J., Gibson, S.P., Goudie, A.C., Holdom, K.S., Pacey, M.S., Ruddock, J.C., Bu’Lock, J.D., and Richards, M.K. (1991). Novel avermectins produced by mutational biosynthesis. *J. Antibiot.* *44*, 357–365.
60. Edwards, A.L., Matsui, T., Weiss, T.M., and Khosla, C. (2014). Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J. Mol. Biol.* *426*, 2229–2245.

61. Eng, C.H., Yuzawa, S., Wang, G., Baidoo, E.E.K., Katz, L., and Keasling, J.D. (2016). Alteration of polyketide stereochemistry from anti to *syn* by a ketoreductase domain exchange in a type I modular polyketide synthase subunit. *Biochemistry* 55, 1677–1680.
62. Engelke, G., Gutowski-Eckel, Z., Hammelmann, M., and Entian, K.D. (1992). Biosynthesis of the lantibiotic nisin: genomic organization and membrane localization of the NisB protein. *Appl Environ Microbiol* 58, 3730–3743.
63. Engler, C., Kandzia, R., and Marillonnet, S. (2008). A One Pot, One Step, Precision cloning method with high throughput capability. *PLoS One* 3.
64. Fayed, B., Ashford, D.A., Hashem, A.M., Amin, M.A., Gazayerly, O.N.E., Gregory, M.A., and Smith, M.C.M. (2015). Multiplexed integrating plasmids for engineering of the erythromycin gene cluster for expression in *Streptomyces* spp. and combinatorial biosynthesis. *Appl. Environ. Microbiol.* 81, 8402–8413.
65. Feitelson, J.S., Malpartida, F., and Hopwood, D.A. (1985). Genetic and biochemical characterization of the red gene cluster of *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* 131, 2431–2441.
66. Feng, Z., Kallifidas, D., and Brady, S.F. (2011). Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. *Proc Natl Acad Sci U S A* 108, 12629–12634.
67. Fernández-Moreno, M.A., Martínez, E., Boto, L., Hopwood, D.A., and Malpartida, F. (1992). Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J. Biol. Chem.* 267, 19278–19290.
68. Fischbach, M.A., and Walsh, C.T. (2009). Antibiotics for emerging pathogens. *Science* 325, 1089–1093.
69. Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br J Exp Pathol* 10, 226–236.
70. Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. (2002). The Biosynthesis of the aromatic Myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* 277, 13082–13090.

71. Garg, A., Khosla, C., and Cane, D.E. (2013). Coupled methyl group epimerization and reduction by polyketide synthase ketoreductase domains. Ketoreductase-catalyzed equilibrium isotope exchange. *J. Am. Chem. Soc.* *135*, 16324–16327.
72. Garg, A., Xie, X., Keatinge-Clay, A., Khosla, C., and Cane, D.E. (2014). Elucidation of the cryptic epimerase activity of redox-inactive ketoreductase domains from modular polyketide synthases by tandem equilibrium isotope exchange. *J. Am. Chem. Soc.* *136*, 10190–10193.
73. Gay, D.C., Wagner, D.T., Meinke, J.L., Zogzas, C.E., Gay, G.R., and Keatinge-Clay, A.T. (2016). The LINKS motif zippers *trans*-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *J. Struct. Biol.* *193*, 196–205.
74. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* *6*, 343–345.
75. Giddings, L.-A., and Newman, D.J. (2013). Microbial natural products: molecular blueprints for antitumor drugs. *J. Ind. Microbiol. Biotechnol.* *40*, 1181–1210.
76. Gietz, D., St Jean, A., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* *20*, 1425.
77. Gokhale, R.S., Lau, J., Cane, D.E., and Khosla, C. (1998). Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. *Biochemistry* *37*, 2524–2528.
78. Gokhale, R.S., Tsuji, S.Y., Cane, D.E., and Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science* *284*, 482–485.
79. Gomez-Escribano, J.P., and Bibb, M.J. (2011). Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* *4*, 207–215.
80. Gomez-Escribano, J.P., Song, L., Fox, D.J., Yeo, V., Bibb, M.J., and 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. *Chem. Sci.* *3*, 2716–2720.

81. Gottelt, M., Kol, S., Gomez-Escribano, J.P., Bibb, M., and Takano, E. (2010). Deletion of a regulatory gene within the *cpk* gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology (Reading, Engl.)* *156*, 2343–2353.
82. Goudie, A.C., Evans, N.A., Gration, K.A., Bishop, B.F., Gibson, S.P., Holdom, K.S., Kaye, B., Wicks, S.R., Lewis, D., and Weatherley, A.J. (1993). Doramectin--a potent novel endectocide. *Vet. Parasitol.* *49*, 5–15.
83. Grenson, M. (1969). The utilization of exogenous pyrimidines and the recycling of uridine-5'-phosphate derivatives in *Saccharomyces cerevisiae*, as studied by means of mutants affected in pyrimidine uptake and metabolism. *Eur. J. Biochem.* *11*, 249–260.
84. Gullberg, E., Albrecht, L.M., Karlsson, C., Sandegren, L., and Andersson, D.I. (2014). Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *MBio* *5*, e01918-01914.
85. Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* *100*, 1541–1546.
86. Hacker, C., Cai, X., Kegler, C., Zhao, L., Weickhmann, A.K., Wurm, J.P., Bode, H.B., and Wöhnert, J. (2018). Structure-based redesign of docking domain interactions modulates the product spectrum of a rhabdopeptide-synthesizing NRPS. *Nat Commun* *9*, 4366.
87. Hagen, A., Poust, S., Rond, T. de, Fortman, J.L., Katz, L., Petzold, C.J., and Keasling, J.D. (2016). Engineering a polyketide synthase for *in vitro* production of adipic acid. *ACS Synth Biol* *5*, 21–27.
88. Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* *5*, R245–R249.
89. Hansen, D.A., Koch, A.A., and Sherman, D.H. (2017). Identification of a thioesterase bottleneck in the pikromycin pathway through full-module processing of unnatural pentaketides. *J Am Chem Soc* *139*, 13450–13455.
90. Hantash, F.M., and Earhart, C.F. (2000). Membrane association of the *Escherichia coli* enterobactin synthase proteins EntB/G, EntE, and EntF. *J Bacteriol* *182*, 1768–1773.

91. Helfrich, E.J.N., and Piel, J. (2016). Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep* 33, 231–316.
92. Helfrich, E.J., Reiter, S., and Piel, J. (2014). Recent advances in genome-based polyketide discovery. *Current Opinion in Biotechnology* 29, 107–115.
93. Herbst, D.A., Townsend, C.A., and Maier, T. (2018). The architectures of iterative type I PKS and FAS. *Nat Prod Rep* 35, 1046–1069.
94. Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A. (2007). Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* 24, 162–190.
95. Heul, H.U. van der, Bilyk, B.L., McDowall, K.J., Seipke, R.F., and Wezel, G.P. van (2018). Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Nat. Prod. Rep.* 35, 575–604.
96. Hitchman, T.S., Crosby, J., Byrom, K.J., Cox, R.J., and Simpson, T.J. (1998). Catalytic self-acylation of type II polyketide synthase acyl carrier proteins. *Chemistry & Biology* 5, 35–47.
97. Hoefler, B.C., Stubbendieck, R.M., Josyula, N.K., Moisan, S.M., Schulze, E.M., and Straight, P.D. (2017). A Link between linearmycin biosynthesis and extracellular vesicle genesis connects specialized metabolism and bacterial membrane physiology. *Cell Chem Biol* 24, 1238-1249.e7.
98. Hong, H., Appleyard, A.N., Siskos, A.P., Garcia-Bernardo, J., Staunton, J., and Leadlay, P.F. (2005). Chain initiation on type I modular polyketide synthases revealed by limited proteolysis and ion-trap mass spectrometry. *FEBS J.* 272, 2373–2387.
99. Hopwood, D.A., and Wright, H.M. (1983). CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* 129, 3575–3579.
100. Horsman, M.E., Hari, T.P.A., and Boddy, C.N. (2016). Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat Prod Rep* 33, 183–202.
101. Ikeda, H., Kazuo, S., and Omura, S. (2014). Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J. Ind. Microbiol. Biotechnol.* 41, 233–250.

102. Imperi, F., and Visca, P. (2013). Subcellular localization of the pyoverdine biogenesis machinery of *Pseudomonas aeruginosa*: A membrane-associated “siderosome.” *FEBS Letters* 587, 3387–3391.
103. Imran, M., Rauf, A., Abu-Izneid, T., Nadeem, M., Shariati, M.A., Khan, I.A., Imran, A., Orhan, I.E., Rizwan, M., Atif, M., *et al.* (2019). Luteolin, a flavonoid, as an anticancer agent: A review. *Biomed. Pharmacother.* 112, 108612.
104. Jain, M., and Cox, J.S. (2005). Interaction between polyketide synthase and transporter suggests coupled synthesis and export of virulence lipid in *M. tuberculosis*. *PLoS Pathog.* 1, e2.
105. Jenner, M., Kosol, S., Griffiths, D., Prasongpholchai, P., Manzi, L., Barrow, A.S., Moses, J.E., Oldham, N.J., Lewandowski, J.R., and Challis, G.L. (2018). Mechanism of intersubunit ketosynthase-dehydratase interaction in polyketide synthases. *Nat. Chem. Biol.* 14, 270–275.
106. Jin, C., Gibani, M.M., Pennington, S.H., Liu, X., Ardrey, A., Aljayyousi, G., Moore, M., Angus, B., Parry, C.M., Biagini, G.A., *et al.* (2019). Treatment responses to azithromycin and ciprofloxacin in uncomplicated *Salmonella typhi* infection: A comparison of clinical and microbiological data from a controlled human infection Model. *PLoS Negl Trop Dis* 13.
107. Kalkreuter, E., Pan, G., Cepeda, A.J., and Shen, B. (2020). Targeting bacterial genomes for natural product discovery. *Trends in Pharmacological Sciences* 41, 13–26.
108. Kao, C.M., Katz, L., and Khosla, C. (1994). Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science* 265, 509–512.
109. Kao, C.M., McPherson, M., McDaniel, R.N., Fu, H., Cane, D.E., and Khosla, C. (1998). Alcohol stereochemistry in polyketide backbones is controlled by the  $\beta$ -ketoreductase domains of modular polyketide synthases. *J. Am. Chem. Soc.* 120, 2478–2479.
110. Kapur, S., Chen, A.Y., Cane, D.E., and Khosla, C. (2010). Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22066–22071.

111. Kapur, S., Lowry, B., Yuzawa, S., Kenthirapalan, S., Chen, A.Y., Cane, D.E., and Khosla, C. (2012). Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation. *Proc. Natl. Acad. Sci. U.S.A.* *109*, 4110–4115.
112. Katz, L., and Baltz, R.H. (2016). Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* *43*, 155–176.
113. Keatinge-Clay, A. (2008). Crystal structure of the erythromycin polyketide synthase dehydratase. *J. Mol. Biol.* *384*, 941–953.
114. Keatinge-Clay, A.T. (2007). A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem. Biol.* *14*, 898–908.
115. Keatinge-Clay, A.T. (2012). The structures of type I polyketide synthases. *Nat Prod Rep* *29*, 1050–1073.
116. Keatinge-Clay, A.T. (2017). Polyketide synthase modules redefined. *Angew. Chem. Int. Ed. Engl.* *56*, 4658–4660.
117. Kellenberger, L., Galloway, I.S., Sauter, G., Böhm, G., Hanefeld, U., Cortés, J., Staunton, J., and Leadlay, P.F. (2008a). A polylinker approach to reductive loop swaps in modular polyketide synthases. *Chembiochem* *9*, 2740–2749.
118. van Keulen, G., and Dyson, P.J. (2014). Production of specialized metabolites by *Streptomyces coelicolor* A3(2). *Adv. Appl. Microbiol.* *89*, 217–266.
119. Khosla, C., Tang, Y., Chen, A.Y., Schnarr, N.A., and Cane, D.E. (2007). Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem.* *76*, 195–221.
120. Kiesau, P., Eikmanns, U., Gutowski-Eckel, Z., Weber, S., Hammelmann, M., and Entian, K.D. (1997). Evidence for a multimeric subtilin synthetase complex. *J. Bacteriol.* *179*, 1475–1481.
121. Kim, J., Chang, J.H., Kim, E.-J., and Kim, K.-J. (2014). Crystal structure of (R)-3-hydroxybutyryl-CoA dehydrogenase PhaB from *Ralstonia eutropha*. *Biochem. Biophys. Res. Commun.* *443*, 783–788.
122. Kitsche, A., and Kalesse, M. (2013). Configurational assignment of secondary hydroxyl groups and methyl branches in polyketide natural products through bioinformatic analysis of the ketoreductase domain. *Chembiochem* *14*, 851–861.

123. Klaus, M., and Grninger, M. (2018). Engineering strategies for rational polyketide synthase design. *Nat Prod Rep* 35, 1070–1081.
124. Klaus, M., Ostrowski, M.P., Austerjost, J., Robbins, T., Lowry, B., Cane, D.E., and Khosla, C. (2016). Protein-protein interactions, not substrate recognition, dominate the turnover of chimeric assembly line polyketide synthases. *J. Biol. Chem.* 291, 16404–16415.
125. Klaus, M., Buyachuihan, L., and Grninger, M. (2020). Ketosynthase domain constrains the design of polyketide synthases. *ACS Chem. Biol.*
126. Koryakina, I., Kasey, C., McArthur, J.B., Lowell, A.N., Chemler, J.A., Li, S., Hansen, D.A., Sherman, D.H., and Williams, G.J. (2017). Inversion of extender unit selectivity in the erythromycin polyketide synthase by acyltransferase domain engineering. *ACS Chem. Biol.* 12, 114–123.
127. Kosol, S., Gallo, A., Griffiths, D., Valentic, T.R., Masschelein, J., Jenner, M., de Los Santos, E.L.C., Manzi, L., Sydor, P.K., Rea, D., *et al.* (2019). Structural basis for chain release from the enacyloxin polyketide synthase. *Nat Chem* 11, 913–923.
128. Kouprina, N., and Larionov, V. (2016). Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* 125, 621–632.
129. Kuczek, K., Pawlik, K., Kotowska, M., and Mordarski, M. (1997). *Streptomyces coelicolor* DNA homologous with acyltransferase domains of type I polyketide synthase gene complex. *FEMS Microbiol Lett* 157, 195–200.
130. Kulkarni, S.S., and Cantó, C. (2015). The molecular targets of resveratrol. *Biochim. Biophys. Acta* 1852, 1114–1123.
131. Kwan, D.H., and Leadlay, P.F. (2010). Mutagenesis of a modular polyketide synthase enoylreductase domain reveals insights into catalysis and stereospecificity. *ACS Chem. Biol.* 5, 829–838.
132. Kwan, D.H., Sun, Y., Schulz, F., Hong, H., Popovic, B., Sim-Stark, J.C.C., Haydock, S.F., and Leadlay, P.F. (2008). Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases. *Chem. Biol.* 15, 1231–1240.

133. Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C., and Walsh, C.T. (1996). A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem. Biol.* 3, 923–936.
134. Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G.L., and Aigle, B. (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6258–6263.
135. Li, Y., and Rebuffat, S. (2019). The manifold roles of microbial ribosomal peptide-based natural products in physiology and ecology. *J. Biol. Chem.* jbc.REV119.006545.
136. Li, X., Thomason, L.C., Sawitzke, J.A., Costantino, N., and Court, D.L. (2013a). Positive and negative selection using the tetA-sacB cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res* 41, e204.
137. Li, X., Wang, J., Li, S., Ji, J., Wang, W., and Yang, K. (2015). ScbR- and ScbR2-mediated signal transduction networks coordinate complex physiological responses in *Streptomyces coelicolor*. *Sci Rep* 5.
138. Li, X., Sevillano, N., La Greca, F., Deis, L., Liu, Y.-C., Deller, M.C., Mathews, I.I., Matsui, T., Cane, D.E., Craik, C.S., *et al.* (2018). Structure-Function Analysis of the Extended Conformation of a Polyketide Synthase Module. *J. Am. Chem. Soc.* 140, 6518–6521.
139. Li, X.-T., Thomason, L.C., Sawitzke, J.A., Costantino, N., and Court, D.L. (2013b). Positive and negative selection using the tetA-sacB cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res.* 41, e204.
140. Linares, J.F., Gustafsson, I., Baquero, F., and Martinez, J.L. (2006). Antibiotics as intermicrobial signaling agents instead of weapons. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19484–19489.
141. Liu, C., Yuan, M., Xu, X., Wang, L., Keatinge-Clay, A.T., Deng, Z., Lin, S., and Zheng, J. (2018). Substrate-bound structures of a ketoreductase from amphotericin modular polyketide synthase. *J. Struct. Biol.* 203, 135–141.
142. Lobanovska, M., and Pilla, G. (2017). Penicillin's Discovery and Antibiotic Resistance: Lessons for the Future? *Yale J Biol Med* 90, 135–145.

143. Lohman, J.R., Ma, M., Osipiuk, J., Nocek, B., Kim, Y., Chang, C., Cuff, M., Mack, J., Bigelow, L., Li, H., *et al.* (2015). Structural and evolutionary relationships of “AT-less” type I polyketide synthase ketosynthases. *Proc. Natl. Acad. Sci. U.S.A.* *112*, 12693–12698.
144. Lucas, X., Senger, C., Erxleben, A., Grüning, B.A., Döring, K., Mosch, J., Flemming, S., and Günther, S. (2013). StreptomeDB: a resource for natural compounds isolated from *Streptomyces* species. *Nucleic Acids Res.* *41*, D1130-1136.
145. MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., and MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* *111*, 61–68.
146. Maier, T., Leibundgut, M., and Ban, N. (2008). The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* *321*, 1315–1322.
147. Margolin, W. (2012). The Price of Tags in Protein Localization Studies. *J Bacteriol* *194*, 6369–6371.
148. Marsden, A.F., Caffrey, P., Aparicio, J.F., Loughran, M.S., Staunton, J., and Leadlay, P.F. (1994). Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. *Science* *263*, 378–380.
149. Marsden, A.F., Wilkinson, B., Cortés, J., Dunster, N.J., Staunton, J., and Leadlay, P.F. (1998). Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* *279*, 199–202.
150. Massicard, J.-M., Soligot, C., Weissman, K., and Jacob, C. (2020). Manipulating polyketide stereochemistry by exchange of polyketide synthase modules. *Chem. Commun.*
151. McAlpine, J.B., Bachmann, B.O., Pirae, M., Tremblay, S., Alarco, A.-M., Zazopoulos, E., and Farnet, C.M. (2005). Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example. *J. Nat. Prod.* *68*, 493–496.
152. McCranie, E.K., and Bachmann, B.O. (2014). Bioactive oligosaccharide natural products. *Nat Prod Rep* *31*, 1026–1042.

153. McPherson, M., Khosla, C., and Cane, D.E. (1998). Erythromycin biosynthesis: the  $\beta$ -ketoreductase domains catalyze the stereospecific transfer of the 4-pro-*S* hydride of NADPH. *J. Am. Chem. Soc.* *120*, 3267–3268.
154. Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A., Weber, T., Takano, E., and Breitling, R. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* *39*, W339-346.
155. van der Meij, A., Worsley, S.F., Hutchings, M.I., and van Wezel, G.P. (2017). Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiol. Rev.* *41*, 392–416.
156. Menzella, H.G., Reid, R., Carney, J.R., Chandran, S.S., Reisinger, S.J., Patel, K.G., Hopwood, D.A., and Santi, D.V. (2005). Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* *23*, 1171–1176.
157. Miyanaga, A., Ouchi, R., Ishikawa, F., Goto, E., Tanabe, G., Kudo, F., and Eguchi, T. (2018). Structural basis of protein-protein interactions between a *trans*-acting acyltransferase and acyl carrier protein in polyketide disorazole biosynthesis. *J. Am. Chem. Soc.* *140*, 7970–7978.
158. Miyazawa, T., Hirsch, M., Zhang, Z., and Keatinge-Clay, A.T. (2020). An *in vitro* platform for engineering and harnessing modular polyketide synthases. *Nat Commun* *11*.
159. Mohammadi, S., Jafari, B., Asgharian, P., Martorell, M., and Sharifi-Rad, J. (2020). Medicinal plants used in the treatment of Malaria: A key emphasis to *Artemisia*, *Cinchona*, *Cryptolepis*, and *Tabebuia* genera. *Phytother Res.*
160. Moncrieffe, M.C., Fernandez, M.-J., Spitteller, D., Matsumura, H., Gay, N.J., Luisi, B.F., and Leadlay, P.F. (2012). Structure of the glycosyltransferase EryCIII in complex with its activating P450 homologue EryCII. *J Mol Biol* *415–20*, 92–101.
161. Moretto, L., Vance, S., Heames, B., and Broadhurst, R.W. (2017). Dissecting how modular polyketide synthase ketoreductases interact with acyl carrier protein-attached substrates. *Chem. Commun. (Camb.)* *53*, 11457–11460.

162. Moretto, L., Heylen, R., Holroyd, N., Vance, S., and Broadhurst, R.W. (2019). Modular type I polyketide synthase acyl carrier protein domains share a common N-terminally extended fold. *Sci Rep* 9, 2325.
163. Morita, H., Wong, C.P., and Abe, I. (2019). How structural subtleties lead to molecular diversity for the type III polyketide synthases. *J. Biol. Chem.* 294, 15121–15136.
164. Mukherjee, S., Seshadri, R., Varghese, N.J., Eloë-Fadrosch, E.A., Meier-Kolthoff, J.P., Göker, M., Coates, R.C., Hadjithomas, M., Pavlopoulos, G.A., Paez-Espino, D., *et al.* (2017). 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nat Biotechnol* 35, 676–683.
165. Murphy, A.C., Hong, H., Vance, S., Broadhurst, R.W., and Leadlay, P.F. (2016). Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chem. Commun. (Camb.)* 52, 8373–8376.
166. Nappi, J., Soldi, E., and Egan, S. (2019). Diversity and distribution of bacteria producing known secondary metabolites. *Microb. Ecol.* 78, 885–894.
167. Newman, D.J., and Cragg, G.M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* 79, 629–661.
168. Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., and Piel, J. (2008). Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat. Biotechnol.* 26, 225–233.
169. Nielsen, M.L., Isbrandt, T., Petersen, L.M., Mortensen, U.H., Andersen, M.R., Hoof, J.B., and Larsen, T.O. (2016). Linker flexibility facilitates module exchange in fungal hybrid PKS-NRPS engineering. *PLoS ONE* 11, e0161199.
170. Nivina, A., Yuet, K.P., Hsu, J., and Khosla, C. (2019). Evolution and diversity of assembly-line polyketide synthases. *Chem. Rev.* 119, 12524–12547.
171. Nogales, E. (2018). Cryo-EM. *Curr. Biol.* 28, R1127–R1128.
172. Nooren, I.M.A., and Thornton, J.M. (2003). Structural characterisation and functional significance of transient protein-protein interactions. *J. Mol. Biol.* 325, 991–1018.

173. O'Brien, R.V., Davis, R.W., Khosla, C., and Hillenmeyer, M.E. (2014). Computational identification and analysis of orphan assembly-line polyketide synthases. *J. Antibiot.* *67*, 89–97.
174. Ochi, K. (2007). From microbial differentiation to ribosome engineering. *Biosci. Biotechnol. Biochem.* *71*, 1373–1386.
175. Olesen, S.W., Barnett, M.L., MacFadden, D.R., Brownstein, J.S., Hernández-Díaz, S., Lipsitch, M., and Grad, Y.H. (2018). The distribution of antibiotic use and its association with antibiotic resistance. *Elife* *7*.
176. Oliynyk, M., Brown, M.J., Cortés, J., Staunton, J., and Leadlay, P.F. (1996). A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.* *3*, 833–839.
177. Ostrowski, M.P., Cane, D.E., and Khosla, C. (2016). Recognition of acyl carrier proteins by ketoreductases in assembly line polyketide synthases. *J. Antibiot.* *69*, 507–510.
178. Pacey, M.S., Dirlam, J.P., Geldart, R.W., Leadlay, P.F., McArthur, H.A., McCormick, E.L., Monday, R.A., O'Connell, T.N., Staunton, J., and Winchester, T.J. (1998). Novel erythromycins from a recombinant *Saccharopolyspora erythraea* strain NRRL 2338 pIG1. I. Fermentation, isolation and biological activity. *J. Antibiot.* *51*, 1029–1034.
179. Palaniappan, N., Kim, B.S., Sekiyama, Y., Osada, H., and Reynolds, K.A. (2003). Enhancement and selective production of phoslactomycin B, a protein phosphatase IIa inhibitor, through identification and engineering of the corresponding biosynthetic gene cluster. *J. Biol. Chem.* *278*, 35552–35557.
180. Palazzotto, E., Tong, Y., Lee, S.Y., and Weber, T. (2019). Synthetic biology and metabolic engineering of actinomycetes for natural product discovery. *Biotechnol. Adv.* *37*, 107366.
181. Patel, P.S., Huang, S., Fisher, S., Pirnik, D., Aklonis, C., Dean, L., Meyers, E., Fernandes, P., and Mayerl, F. (1995). Bacillaene, a novel inhibitor of procaryotic protein synthesis produced by *Bacillus subtilis*: production, taxonomy, isolation, physico-chemical characterization and biological activity. *J. Antibiot.* *48*, 997–1003.

182. Paul, V.J., Frautschy, S., Fenical, W., and Neilson, K.H. (1981). Antibiotics in microbial ecology: Isolation and structure assignment of several new antibacterial compounds from the insect-symbiotic bacteria *Xenorhabdus spp.* *J. Chem. Ecol.* *7*, 589–597.
183. Pawlik, K., Kotowska, M., Chater, K.F., Kuczek, K., and Takano, E. (2007). A cryptic type I polyketide synthase (cpk) gene cluster in *Streptomyces coelicolor* A3(2). *Arch. Microbiol.* *187*, 87–99.
184. Pawlik, K., Kotowska, M., and Kolesiński, P. (2010). *Streptomyces coelicolor* A3(2) produces a new yellow pigment associated with the polyketide synthase Cpk. *J. Mol. Microbiol. Biotechnol.* *19*, 147–151.
185. Peltier, K. (1992). The new macrolactones: clarithromycin (Biaxin) and azithromycin (Zithromax). *Conn Med* *56*, 371–374.
186. Peng, H., Ishida, K., Sugimoto, Y., Jenke-Kodama, H., and Hertweck, C. (2019a). Emulating evolutionary processes to morph aureothin-type modular polyketide synthases and associated oxygenases. *Nat Commun* *10*, 3918.
187. Peng, H., Ishida, K., and Hertweck, C. (2019b). Loss of single-domain function in a modular assembly line alters the size and shape of a complex polyketide. *Angew. Chem. Int. Ed. Engl.* *58*, 18252–18256.
188. Petković, H., Sandmann, A., Challis, I.R., Hecht, H.-J., Silakowski, B., Low, L., Beeston, N., Kuscer, E., Garcia-Bernardo, J., Leadlay, P.F., *et al.* (2008). Substrate specificity of the acyl transferase domains of EpoC from the epothilone polyketide synthase. *Org. Biomol. Chem.* *6*, 500–506.
189. Piasecki, S.K., Zheng, J., Axelrod, A.J., Detelich, M.E., and Keatinge-Clay, A.T. (2014). Structural and functional studies of a *trans*-acyltransferase polyketide assembly line enzyme that catalyzes stereoselective  $\alpha$ - and  $\beta$ -ketoreduction. *Proteins* *82*, 2067–2077.
190. Piel, J. (2002). A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc Natl Acad Sci U S A* *99*, 14002–14007.

191. Pieper, R., Luo, G., Cane, D.E., and Khosla, C. (1995). Cell-free synthesis of polyketides by recombinant erythromycin polyketide synthases. *Nature* 378, 263–266.
192. Porter, J.L., Tobias, N.J., Pidot, S.J., Falgner, S., Tuck, K.L., Vettiger, A., Hong, H., Leadlay, P.F., and Stinear, T.P. (2013). The cell wall-associated mycolactone polyketide synthases are necessary but not sufficient for mycolactone biosynthesis. *PLoS ONE* 8, e70520.
193. Ranganathan, A., Timoney, M., Bycroft, M., Cortés, J., Thomas, I.P., Wilkinson, B., Kellenberger, L., Hanefeld, U., Galloway, I.S., Staunton, J., *et al.* (1999). Knowledge-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: a route to simple statin analogues. *Chem. Biol.* 6, 731–741.
194. Ray, L., and Moore, B.S. (2016). Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Nat Prod Rep* 33, 150–161.
195. Reeves, C.D., Murli, S., Ashley, G.W., Piagentini, M., Hutchinson, C.R., and McDaniel, R. (2001). Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* 40, 15464–15470.
196. Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D.V., Hutchinson, C.R., and McDaniel, R. (2003). A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry* 42, 72–79.
197. Richter, C.D., Nietlispach, D., Broadhurst, R.W., and Weissman, K.J. (2008). Multienzyme docking in hybrid megasynthetases. *Nat. Chem. Biol.* 4, 75–81.
198. Risser, F., Collin, S., Dos Santos-Morais, R., Gruez, A., Chagot, B., and Weissman, K.J. (2020). Towards improved understanding of intersubunit interactions in modular polyketide biosynthesis: docking in the enacyloxin IIa polyketide synthase. *J. Struct. Biol.* 107581.
199. Robbins, T., Liu, Y.-C., Cane, D.E., and Khosla, C. (2016). Structure and mechanism of assembly line polyketide synthases. *Curr. Opin. Struct. Biol.* 41, 10–18.
200. Romano M., Fusco G., Choudhury H.G., Mehmood S., Robinson C.V., Zirah S., Hegemann J.D., Lescop E., Marahiel M.A., Rebuffat S., De Simone A., and Beis K.

- (2018). Structural basis for natural product selection and export by bacterial abc transporters. *ACS Chem Biol.* *13*, 1598–1609.
201. Romero, D., Traxler, M.F., López, D., and Kolter, R. (2011). Antibiotics as signal molecules. *Chem. Rev.* *111*, 5492–5505.
202. Röttig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011). NRSPredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* *39*, W362-367.
203. Rowe, C.J., Cortés, J., Gaisser, S., Staunton, J., and Leadlay, P.F. (1998). Construction of new vectors for high-level expression in actinomycetes. *Gene* *216*, 215–223.
204. Rowe, C.J., Böhm, I.U., Thomas, I.P., Wilkinson, B., Rudd, B.A., Foster, G., Blackaby, A.P., Sidebottom, P.J., Roddis, Y., Buss, A.D., *et al.* (2001). Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chem. Biol.* *8*, 475–485.
205. Rowinsky, E.K., and Donehower, R.C. (1995). Paclitaxel (taxol). *N. Engl. J. Med.* *332*, 1004–1014.
206. Rudd, B.A., and Hopwood, D.A. (1979). Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* *114*, 35–43.
207. Scaglione, J.B., Akey, D.L., Sullivan, R., Kittendorf, J.D., Rath, C.M., Kim, E.-S., Smith, J.L., and Sherman, D.H. (2010). Biochemical and structural characterization of the tautomycetin thioesterase: analysis of a stereoselective polyketide hydrolase. *Angew. Chem. Int. Ed. Engl.* *49*, 5726–5730.
208. Schlimpert, S., Wasserstrom, S., Chandra, G., Bibb, M.J., Findlay, K.C., Flärdh, K., and Buttner, M.J. (2017). Two dynamin-like proteins stabilize FtsZ rings during *Streptomyces* sporulation. *Proc Natl Acad Sci U S A* *114*, E6176–E6183.
209. Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001). Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* *413*, 814–821.
210. Seow, K.T., Meurer, G., Gerlitz, M., Wendt-Pienkowski, E., Hutchinson, C.R., and Davies, J. (1997). A study of iterative type II polyketide synthases, using bacterial genes

- cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J Bacteriol* *179*, 7360–7368.
211. Severi E., and Thomas G.H. (2019). Antibiotic export: transporters involved in the final step of natural product production. *Microbiology* *165*, 805–818.
212. Shama, G. (2016). La Moisissure et la Bactérie: Deconstructing the fable of the discovery of penicillin by Ernest Duchesne. *Endeavour* *40*, 188–200.
213. Shatz, A., Bugie, E., and Waksman, S. (1944). Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proc Soc Exp Biol Med* *55*, 66–69.
214. Sherman, D.H., Koch, A.A., Schmidt, J., Lowell, A., Hansen, D., Coburn, K., and Chemler, J. (2020). Probing selectivity and creating structural diversity through hybrid polyketide synthases. *Angew. Chem. Int. Ed. Engl.*
215. Sherman, D.H., Koch, A.A., Schmidt, J., Lowell, A., Hansen, D., Coburn, K., and Chemler, J. Probing selectivity and creating structural diversity through hybrid polyketide synthases. *Angewandte Chemie International Edition n/a*.
216. Shimizu, Y., Ogata, H., and Goto, S. (2017). Type III polyketide synthases: functional classification and phylogenomics. *Chembiochem* *18*, 50–65.
217. Shrestha, B.M. (2017). Two decades of tacrolimus in renal transplant: basic science and clinical evidences. *Exp Clin Transplant* *15*, 1–9.
218. Simunovic, V., Gherardini, F.C., and Shimkets, L.J. (2003). Membrane localization of motility, signaling, and polyketide synthetase proteins in *Myxococcus xanthus*. *J. Bacteriol.* *185*, 5066–5075.
219. Siskos, A.P., Baerga-Ortiz, A., Bali, S., Stein, V., Mamdani, H., Spitteller, D., Popovic, B., Spencer, J.B., Staunton, J., Weissman, K.J., *et al.* (2005). Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. *Chem. Biol.* *12*, 1145–1153.
220. Squire, C.M., Goss, R.J.M., Hong, H., Leadlay, P.F., and Staunton, J. (2003). Catalytically active tetramodular 6-deoxyerythronolide B synthase fusion proteins. *Chembiochem* *4*, 1225–1228.

221. Starcevic, A., Diminic, J., Zucko, J., Elbekali, M., Schlosser, T., Lisfi, M., Vukelic, A., Long, P.F., Hranueli, D., and Cullum, J. (2011). A novel docking domain interface model predicting recombination between homoeologous modular biosynthetic gene clusters. *J. Ind. Microbiol. Biotechnol.* 38, 1295–1304.
222. Starcevic, A., Wolf, K., Diminic, J., Zucko, J., Ruzic, I.T., Long, P.F., Hranueli, D., and Cullum, J. (2012). Recombinatorial biosynthesis of polyketides. *J Ind Microbiol Biotechnol* 39, 503–511.
223. Stassi, D.L., Kakavas, S.J., Reynolds, K.A., Gunawardana, G., Swanson, S., Zeidner, D., Jackson, M., Liu, H., Buko, A., and Katz, L. (1998). Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering. *Proc Natl Acad Sci U S A* 95, 7305–7309.
224. Straight, P.D., Fischbach, M.A., Walsh, C.T., Rudner, D.Z., and Kolter, R. (2007). A singular enzymatic megacomplex from *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 305–310.
225. Sugimoto, Y., Ding, L., Ishida, K., and Hertweck, C. (2014). Rational design of modular polyketide synthases: morphing the aureothin pathway into a luteoreticulin assembly line. *Angew. Chem. Int. Ed. Engl.* 53, 1560–1564.
226. Sun, J., Kelemen, G.H., Fernández-Abalos, J.M., and Bibb, M.J. (1999). Green fluorescent protein as a reporter for spatial and temporal gene expression in *Streptomyces coelicolor* A3(2). *Microbiology*, 145, 2221–2227.
227. Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., and Schulz, F. (2013). Enzyme-directed mutasynthesis: a combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem. Biol.* 8, 443–450.
228. Süßmuth, R.D., and Mainz, A. (2017). Nonribosomal peptide synthesis-principles and prospects. *Angew. Chem. Int. Ed. Engl.* 56, 3770–3821.
229. Swulius, M.T., and Jensen, G.J. (2012). The helical MreB cytoskeleton in *Escherichia coli* MC1000/pLE7 is an artifact of the N-terminal yellow fluorescent protein tag. *J Bacteriol* 194, 6382–6386.

230. Tang, L., Fu, H., and McDaniel, R. (2000). Formation of functional heterologous complexes using subunits from the picromycin, erythromycin and oleandomycin polyketide synthases. *Chem. Biol.* *7*, 77–84.
231. Tang, Y., Kim, C.-Y., Mathews, I.I., Cane, D.E., and Khosla, C. (2006). The 2.7-Ångstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 11124–11129.
232. Tang, Y., Chen, A.Y., Kim, C.-Y., Cane, D.E., and Khosla, C. (2007). Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem. Biol.* *14*, 931–943.
233. Taylor, S.D., and Palmer, M. (2016). The action mechanism of daptomycin. *Bioorg. Med. Chem.* *24*, 6253–6268.
234. Tholl, D. (2015). Biosynthesis and biological functions of terpenoids in plants. *Adv. Biochem. Eng. Biotechnol.* *148*, 63–106.
235. Thomas, I., Martin, C.J., Wilkinson, C.J., Staunton, J., and Leadlay, P.F. (2002). Skipping in a hybrid polyketide synthase. Evidence for ACP-to-ACP chain transfer. *Chem. Biol.* *9*, 781–787.
236. Thomason, L.C., Sawitzke, J.A., Li, X., Costantino, N., and Court, D.L. (2014). Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* *106*, 1.16.1-39.
237. Tong, Y., Charusanti, P., Zhang, L., Weber, T., and Lee, S.Y. (2015). CRISPR-Cas9 based engineering of Actinomycetal genomes. *ACS Synth Biol* *4*, 1020–1029.
238. Torres, J.P., and Schmidt, E.W. (2019). The biosynthetic diversity of the animal world. *J. Biol. Chem.* *294*, 17684–17692.
239. Traitcheva, N., Jenke-Kodama, H., He, J., Dittmann, E., and Hertweck, C. (2007). Non-colinear polyketide biosynthesis in the aureothin and neo-aureothin pathways: an evolutionary perspective. *Chembiochem* *8*, 1841–1849.
240. Tran, H.D.M., Tran, Y.H., Tran, T.D., Jong, M., Coroneo, M., and Sankaridurg, P. (2018). A review of myopia control with atropine. *J Ocul Pharmacol Ther* *34*, 374–379.

241. Tran, L., Tosin, M., Spencer, J.B., Leadlay, P.F., and Weissman, K.J. (2008). Covalent linkage mediates communication between ACP and TE domains in modular polyketide synthases. *Chembiochem* 9, 905–915.
242. Tran, L., Broadhurst, R.W., Tosin, M., Cavalli, A., and Weissman, K.J. (2010). Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. *Chem. Biol.* 17, 705–716.
243. Tsai, S.-C. (Sheryl), and Ames, B.D. (2009). Structural enzymology of polyketide synthases. *Methods Enzymol* 459, 17–47.
244. Tsai, S.C., Miercke, L.J., Krucinski, J., Gokhale, R., Chen, J.C., Foster, P.G., Cane, D.E., Khosla, C., and Stroud, R.M. (2001). Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14808–14813.
245. Tsai, S.-C., Lu, H., Cane, D.E., Khosla, C., and Stroud, R.M. (2002). Insights into channel architecture and substrate specificity from crystal structures of two macrocycle-forming thioesterases of modular polyketide synthases. *Biochemistry* 41, 12598–12606.
246. Tsuji, S.Y., Cane, D.E., and Khosla, C. (2001). Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. *Biochemistry* 40, 2326–2331.
247. Udvary, D.W., Zeigler, L., Asolkar, R.N., Singan, V., Lapidus, A., Fenical, W., Jensen, P.R., and Moore, B.S. (2007). Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10376–10381.
248. Vogt, T. (2010). Phenylpropanoid biosynthesis. *Mol Plant* 3, 2–20.
249. Waksman, S.A., and Lechevalier, H.A. (1949). Neomycin, a new antibiotic active against streptomycin-resistant bacteria, including tuberculosis organisms. *Science* 109, 305–307.
250. Wallace, T.C., and Giusti, M.M. (2015). Anthocyanins1. *Adv Nutr* 6, 620–622.
251. Wang, B., Guo, F., Dong, S.-H., and Zhao, H. (2019). Activation of silent biosynthetic gene clusters using transcription factor decoys. *Nat. Chem. Biol.* 15, 111–114.

252. Wang, W., Ji, J., Li, X., Wang, J., Li, S., Pan, G., Fan, K., and Yang, K. (2014). Angucyclines as signals modulate the behaviors of *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* *111*, 5688–5693.
253. Wawrik, B., Kerkhof, L., Zylstra, G.J., and Kukor, J.J. (2005). Identification of Unique Type II Polyketide Synthase Genes in Soil. *Appl Environ Microbiol* *71*, 2232–2238.
254. Wawrik, B., Kutliev, D., Abdivasievna, U.A., Kukor, J.J., Zylstra, G.J., and Kerkhof, L. (2007). Biogeography of Actinomycete communities and type II polyketide synthase genes in soils collected in New Jersey and central Asia. *Appl Environ Microbiol* *73*, 2982–2989.
255. Weber, T., Rausch, C., Lopez, P., Hoof, I., Gaykova, V., Huson, D.H., and Wohlleben, W. (2009). CLUSEAN: a computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. *J. Biotechnol.* *140*, 13–17.
256. Weissman, K.J. (2009). Introduction to polyketide biosynthesis. *Meth. Enzymol.* *459*, 3–16.
257. Weissman, K.J. (2015). The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* *11*, 660–670.
258. Weissman, K.J. (2016). Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat Prod Rep* *33*, 203–230.
259. Weissman, K.J. (2017). Polyketide stereocontrol: a study in chemical biology. *Beilstein J Org Chem* *13*, 348–371.
260. Weissman, K.J., Timoney, M., Bycroft, M., Grice, P., Hanefeld, U., Staunton, J., and Leadlay, P.F. (1997). The molecular basis of Celmer's rules: the stereochemistry of the condensation step in chain extension on the erythromycin polyketide synthase. *Biochemistry* *36*, 13849–13855.
261. Weissman, K.J., Hong, H., Popovic, B., and Meersman, F. (2006). Evidence for a protein-protein interaction motif on an acyl carrier protein domain from a modular polyketide synthase. *Chem. Biol.* *13*, 625–636.

262. Westhoff, S., van Leeuwe, T.M., Qachach, O., Zhang, Z., van Wezel, G.P., and Rozen, D.E. (2017). The evolution of no-cost resistance at sub-MIC concentrations of streptomycin in *Streptomyces coelicolor*. *ISME J* 11, 1168–1178.
263. Whicher, J.R., Dutta, S., Hansen, D.A., Hale, W.A., Chemler, J.A., Dosey, A.M., Narayan, A.R.H., Håkansson, K., Sherman, D.H., Smith, J.L., *et al.* (2014). Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* 510, 560–564.
264. Wiesmann, K.E., Cortés, J., Brown, M.J., Cutter, A.L., Staunton, J., and Leadlay, P.F. (1995). Polyketide synthesis *in vitro* on a modular polyketide synthase. *Chem. Biol.* 2, 583–589.
265. Willemse, J., and van Wezel, G.P. (2009). Imaging of *Streptomyces coelicolor* A3(2) with reduced autofluorescence reveals a novel stage of FtsZ localization. *PLoS One* 4.
266. Willemse, J., Ruban-Ośmialowska, B., Widdick, D., Celler, K., Hutchings, M.I., van Wezel, G.P., and Palmer, T. (2012). Dynamic localization of Tat protein transport machinery components in *Streptomyces coelicolor*. *Journal of Bacteriology* 194, 6272–6281.
267. Witkowski, A., Joshi, A.K., and Smith, S. (2002). Mechanism of the beta-ketoacyl synthase reaction catalyzed by the animal fatty acid synthase. *Biochemistry* 41, 10877–10887.
268. 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., *et al.* (2017). Diversity oriented biosynthesis *via* accelerated evolution of modular gene clusters. *Nat Commun* 8, 1206.
269. Wohlleben, W., Mast, Y., Stegmann, E., and Ziemert, N. (2016). Antibiotic drug discovery. *Microb Biotechnol* 9, 541–548.
270. Wong, F.T., Jin, X., Mathews, I.I., Cane, D.E., and Khosla, C. (2011). Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. *Biochemistry* 50, 6539–6548.

271. Wright, L.F., and Hopwood, D.A. (1976). Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* 95, 96–106.
272. Wu, N., Cane, D.E., and Khosla, C. (2002). Quantitative analysis of the relative contributions of donor acyl carrier proteins, acceptor ketosynthases, and linker regions to intermodular transfer of intermediates in hybrid polyketide synthases. *Biochemistry* 41, 5056–5066.
273. Xie, X., and Cane, D.E. (2018). pH-Rate profiles establish that polyketide synthase dehydratase domains utilize a single-base mechanism. *Org. Biomol. Chem.* 16, 9165–9170.
274. Xie, X., Garg, A., Keatinge-Clay, A.T., Khosla, C., and Cane, D.E. (2016). Epimerase and Reductase Activities of Polyketide Synthase Ketoreductase Domains Utilize the Same Conserved Tyrosine and Serine Residues. *Biochemistry* 55, 1179–1186.
275. Xie, X., Garg, A., Khosla, C., and Cane, D.E. (2017a). Elucidation of the cryptic methyl group epimerase activity of dehydratase domains from modular polyketide synthases using a tandem modules epimerase assay. *J. Am. Chem. Soc.* 139, 9507–9510.
276. Xie, X., Garg, A., Khosla, C., and Cane, D.E. (2017b). Mechanism and stereochemistry of polyketide chain elongation and methyl group epimerization in polyether biosynthesis. *J. Am. Chem. Soc.* 139, 3283–3292.
277. Xu, G., Wang, J., Wang, L., Tian, X., Yang, H., Fan, K., Yang, K., and Tan, H. (2010). “Pseudo”  $\gamma$ -butyrolactone receptors respond to antibiotic signals to coordinate antibiotic biosynthesis. *J Biol Chem* 285, 27440–27448.
278. Xu, X.-P., Wang, Z.-J., Fan, K.-Q., Wang, S.-L., Jia, C.-J., Han, H., Ramalingam, E., and Yang, K.-Q. (2008). Localization of the ActIII actinorhodin polyketide ketoreductase to the cell wall. *FEMS Microbiol. Lett.* 287, 15–21.
279. Yagüe, P., Willemse, J., Koning, R.I., Rioseras, B., López-García, M.T., Gonzalez-Quiñonez, N., Lopez-Iglesias, C., Shliaha, P.V., Rogowska-Wrzesinska, A., Koster, A.J., *et al.* (2016). Subcompartmentalization by cross-membranes during early growth of *Streptomyces* hyphae. *Nature Communications* 7, 12467.

280. Yan, J., Gupta, S., Sherman, D.H., and Reynolds, K.A. (2009). Functional dissection of a multimodular polypeptide of the pikromycin polyketide synthase into monomodules by using a matched pair of heterologous docking domains. *Chembiochem* *10*, 1537–1543.
281. Yasue, A., Mitsui, S.N., Watanabe, T., Sakuma, T., Oyadomari, S., Yamamoto, T., Noji, S., Mito, T., and Tanaka, E. (2014). Highly efficient targeted mutagenesis in one-cell mouse embryos mediated by the TALEN and CRISPR/Cas systems. *Scientific Reports* *4*, 5705.
282. Yuzawa, S., Deng, K., Wang, G., Baidoo, E.E.K., Northen, T.R., Adams, P.D., Katz, L., and Keasling, J.D. (2017). Comprehensive *in vitro* analysis of acyltransferase domain exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synth Biol* *6*, 139–147.
283. Zaynab, M., Fatima, M., Abbas, S., Sharif, Y., Umair, M., Zafar, M.H., and Bahadar, K. (2018). Role of secondary metabolites in plant defense against pathogens. *Microb. Pathog.* *124*, 198–202.
284. Zeng, J., Wagner, D.T., Zhang, Z., Moretto, L., Addison, J.D., and Keatinge-Clay, A.T. (2016). Portability and structure of the four-helix bundle docking domains of *trans*-acyltransferase modular polyketide synthases. *ACS Chem. Biol.* *11*, 2466–2474.
285. Zhang, L., Hashimoto, T., Qin, B., Hashimoto, J., Kozono, I., Kawahara, T., Okada, M., Awakawa, T., Ito, T., Asakawa, Y., *et al.* (2017). Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolyol polyketides. *Angew. Chem. Int. Ed. Engl.* *56*, 1740–1745.
286. Zhang, L., Ji, J., Yuan, M., Feng, Y., Wang, L., Deng, Z., Bai, L., and Zheng, J. (2018). Stereospecificity of Enoylreductase Domains from Modular Polyketide Synthases. *ACS Chem. Biol.* *13*, 871–875.
287. Zhang, Z., Du, C., de Barsey, F., Liem, M., Liakopoulos, A., van Wezel, G.P., Choi, Y.H., Claessen, D., and Rozen, D.E. (2020). Antibiotic production in *Streptomyces* is organized by a division of labor through terminal genomic differentiation. *Science Advances* *6*, eaay5781.
288. Zhao, Y., Li, G., Chen, Y., and Lu, Y. (2020). Challenges and Advances in Genome Editing Technologies in *Streptomyces*. *Biomolecules* *10*.

*Références bibliographiques*

289. Zheng, J., and Keatinge-Clay, A.T. (2011). Structural and functional analysis of C2-type ketoreductases from modular polyketide synthases. *J. Mol. Biol.* *410*, 105–117.
290. Zheng, J., Taylor, C.A., Piasecki, S.K., and Keatinge-Clay, A.T. (2010). Structural and functional analysis of A-type ketoreductases from the amphotericin modular polyketide synthase. *Structure* *18*, 913–922.
291. Zheng, J., Gay, D.C., Demeler, B., White, M.A., and Keatinge-Clay, A.T. (2012). Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nat. Chem. Biol.* *8*, 615–621.
292. Zheng, J., Piasecki, S.K., and Keatinge-Clay, A.T. (2013). Structural studies of an A2-type modular polyketide synthase ketoreductase reveal features controlling  $\alpha$ -substituent stereochemistry. *ACS Chem. Biol.* *8*, 1964–1971.
293. Ziegler, J., and Facchini, P.J. (2008). Alkaloid biosynthesis: metabolism and trafficking. *Annu Rev Plant Biol* *59*, 735–769.

# **Annexes**





# Manipulating polyketide stereochemistry by exchange of polyketide synthase modules†

Cite this: DOI: 10.1039/d0cc05068g

Jean-Malo Massicard,<sup>a</sup> Claire Soligot,<sup>bc</sup> Kira J. Weissman<sup>id</sup>\*<sup>a</sup> and Christophe Jacob<sup>id</sup>\*<sup>a</sup>

Received 24th July 2020,  
Accepted 30th August 2020

DOI: 10.1039/d0cc05068g

rsc.li/chemcomm

**A key goal of modular polyketide synthase (PKS) engineering is to alter polyketide stereochemistry. Here we report that exchanging whole PKS modules is a more productive approach than swapping individual ketoreductase (KR) domains for introducing rare 'A2' and 'B2' stereochemistry into model polyketides, and identify four modular 'biobricks' for such synthetic biology efforts.**

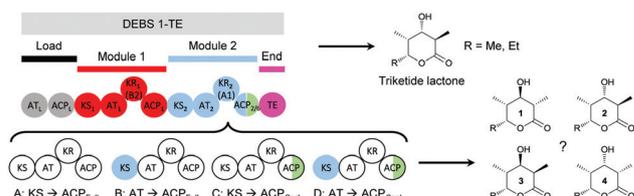
The complex polyketide natural products of bacteria are heavily exploited in both human and veterinary medicine.<sup>1,2</sup> Their exquisite target specificity and high binding affinity<sup>3</sup> result from the density and defined stereochemistry of their diverse functional groups. In particular, hydroxyl substituents participate in directional hydrogen bonding interactions and act as attachment points for further functionality, while methyl and other acyl groups can impose important steric constraints. Together, these control elements impart conformational preferences to the large, typically macrocyclic molecules, giving them a defined shape while maintaining overall backbone flexibility.<sup>4</sup> This intimate link between molecular architecture and activity makes stereochemistry an attractive target for manipulation in polyketide analoguing efforts.

To this end, efforts have focused on genetically engineering the multienzymes responsible for building polyketide skeletons, the modular polyketide synthases (PKSs).<sup>5</sup> Remarkably, each step in PKS-catalysed chain assembly is carried out by a dedicated enzymatic domain: an acyl transferase (AT) selects an extender unit and transfers it onto an acyl carrier protein (ACP), upon which it is joined to the growing chain by a ketosynthase (KS) domain. The intermediate can then undergo stereospecific tailoring of the newly-established functionality at the  $\alpha$ - and  $\beta$ -positions by ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains. It is thus possible, at least in principle, to modify the stereochemistry of a particular functional group

by targeting the domain responsible for its formation. In this context, the KR domains are notable, as they carry out reduction of the  $\beta$ -keto groups formed upon chain extension to yield secondary hydroxyls of defined stereochemistry (3*S* or 3*R*, respectively, when C2 > C4; the stereochemistry is reversed when the priorities are inverted). Such KRs are referred to as A1 and B1.<sup>6</sup> Certain KRs also catalyse epimerisation at the adjacent  $\alpha$ -methyl centre<sup>7</sup> generating the 2*S* configuration (A2 and B2 KRs, respectively).<sup>8</sup> Thus, in theory, engineering of type 2 KRs could result in modification at two adjacent stereocentres.

Indeed, knowledge of these different types of KRs has enabled attempts to modify polyketide stereochemistry by KR domain exchange within engineered, mini PKSs.<sup>9–11</sup> These systems include DEBS 1-TE<sup>12</sup> derived from the erythromycin PKS, in which the chain terminating TE has been fused to the end of DEBS 1, giving rise to a triketide  $\delta$ -lactone incorporating two methyl and two hydroxyl stereocentres (Fig. 1). In earlier work, the A1 KR domain of the second chain extension module was substituted by a panel of B1, A2 and B2 KR domains.<sup>9</sup> While introduction of the B1 KRs resulted in the predicted product 3 (Fig. 1), only traces of the anticipated products 4 and 1 were obtained by swapping with A2 or B2 KRs, respectively.

In follow-up experiments,<sup>10</sup> we evaluated an extensive pool of A2 and B2 KRs. Globally, the nine tested A2-type KRs produced mixtures of epimerised and non-epimerised lactones (4 and 2, respectively), as well as substantial quantities of



**Fig. 1** Engineering strategies (A–D) used to replace module 2 of the model PKS DEBS 1-TE by heterologous modules containing A2 and B2 KRs. Key: AT, acyl transferase; ACP, acyl carrier protein; KS, ketosynthase; KR, ketoreductase; TE, thioesterase.

<sup>a</sup> Université de Lorraine, CNRS, IMoPA, F-54000 Nancy, France.

E-mail: kira.weissman@univ-lorraine.fr, christophe.jacob@univ-lorraine.fr

<sup>b</sup> Université de Lorraine, UR AFPA, USC 340 INRAE, F-54000 Nancy, France

<sup>c</sup> Université de Lorraine, PASM, F-54000 Nancy, France

† Electronic supplementary information (ESI) available: Materials and methods, four supplementary figures and two supplementary tables. See DOI: 10.1039/d0cc05068g

ketolactone **5**. Notably, however, two KR (amphotericin (Amph) KRs **1** and **11**) yielded **4** as the major product, and in respectable yield (7–14 mg L<sup>-1</sup>). Discouragingly, all six B2 KRs produced at best traces of the expected epimerised and reduced lactone **1**, although the modified modules were still active for chain extension, producing ketolactone **5**. Two conclusions emerged from this work: (i) the KR modular environment can influence the fidelity of stereocontrol (*i.e.* out of their native context, epimerising A2-type KRs can reduce non-epimerised chains); and, (ii) B2 KRs exhibit low portability into an A1 modular context. These results prompted us to explore whether swapping module 2 of DEBS 1-TE for intact modules containing A2 and B2 KR from heterologous PKS systems (Fig. S1, ESI<sup>†</sup>), could improve both the absolute and relative yields of reduced lactones **1** and **4**.

As target modules, we chose those harbouring four KRs tested previously<sup>10</sup> which gave a range of results: Amph<sup>13</sup> module 11 (M11) (A2 KR, predominantly **4**), oligomycin (Olm)<sup>14</sup> M5 (A2 KR, exclusively **5**), lankamycin (Lkm)<sup>15</sup> M1 (B2 KR, largely **5**, but also traces of **1**), and lasalocid (Las)<sup>16</sup> M7 (B2 KR, exclusively **5**). For the A2 KRs, we also tested bafilomycin (Bfm) M4 and M9<sup>17</sup> (and M4 and M9 of the closely-related hygrobafilomycin (HbaA) PKS<sup>18</sup>), and nystatin (Nys) M1 and M11 (Fig. S1, ESI<sup>†</sup>),<sup>19</sup> which had not previously been examined. We also studied one additional B2 KR-containing module, mediomycin (Med) M7a (Fig. S1, ESI<sup>†</sup>),<sup>20</sup> as this was one of the few novel modules of this type described in the literature when this project was initiated. The selected modules are present in different contexts: Amph/Nys M1 and M11, Olm M5 and Lkm M1 are linked covalently to an upstream module, analogous to the situation within the parental DEBS 1-TE system. In contrast, Las M7, Bfm/HbaA M4 and M9, and Med M7a are located at the N-termini of their respective subunits, and thus communicate with upstream modules non-covalently.<sup>21</sup> It was thus of interest to investigate the functional consequences of linking such modules covalently to an upstream module, as this question has not yet been explored comprehensively.<sup>22–25</sup>

PKS modules have been classically defined as running from KS to ACP, and encompassing all intervening domains (KS–AT–(DH–ER–KR)–ACP).<sup>26</sup> Many mono-module subunits contain precisely this complement of domains, notably including Las M7 and two other subunits from the Las system.<sup>16</sup> According to this definition, a module includes the KS and ACP domains which partner each other during chain extension. Module exchanges based on such classical boundaries thus preserve the key condensation interaction, while introducing a non-native ACP/KS chain transfer junction. This model has been challenged by phylogenetic analysis<sup>20</sup> which revealed co-evolutionary relationships between certain KSs and the upstream processing domains. This observation has led recently to a proposed redefinition of module borders (AT–(DH–ER–KR)–ACP–KS),<sup>26,27</sup> in which the KS is associated with the ACP domain from which it receives the extending chain. Indeed, it has been reported that ACP/KS interactions during intermodular substrate translocation are more important than substrate recognition in determining turnover efficiency, at least in chimeric synthases.<sup>28</sup> Furthermore, several published module swap experiments were only successful or produced substantial gains in yield, when employing modules with these redefined boundaries.<sup>25,29,30</sup> Nevertheless, module exchanges

carried out in this way also generate a non-native interaction, that between the KS and ACP domains participating in chain extension.

Given the uncertainty surrounding which domains constitute a functional module and the degree to which non-native KS/ACP interactions impact on synthesis, we aimed to directly test multiple module boundaries (Fig. 1 and Table S1, ESI<sup>†</sup>). The first (Fig. 1A) was to use classical modules, replacing DEBS 1-TE module 2 by KS–AT–KR–ACP tetradomain regions from heterologous PKSs. The upstream fusion point was chosen within a highly-conserved region at the N-terminus of the KS domains (PIAI motif), such that the native linker joining module 1 to module 2 was preserved (Fig. S2, ESI<sup>†</sup>).<sup>22</sup> In the second approach (Fig. 1B), only the AT–KR–ACP tridomain of module 2 was replaced using a site at the conserved C-terminal end of the KS domain, thereby keeping intact the native DEBS ACP<sub>1</sub>/KS<sub>2</sub> junction (Fig. 1). The chosen fusion site (Fig. S2, ESI<sup>†</sup>) was used previously in successful module swaps within DEBS 1-TE,<sup>25</sup> and is located in close vicinity to the recently exploited GTNAH motif (Fig. S2, ESI<sup>†</sup>) that permitted efficient AT swapping.<sup>31</sup> A subset of these experiments were, in addition, carried out with two versions of the ACP domains (Fig. 1), as the ACP in DEBS 1-TE is a hybrid between DEBS ACPs 2 and 6.<sup>12</sup> We thus opted to retain the hybrid character of the ACPs in certain experiments (Fig. 1C and D) by carrying out exchanges up to the *SacI* site at the ACP<sub>2/6</sub> junction (Fig. S2, ESI<sup>†</sup>), but in addition, using the complete heterologous ACP domains. In this second case, we used as the point of fusion the conserved C-terminal end of the ACP (Fig. S2, ESI<sup>†</sup>), thereby preserving the native ACP–TE linker.<sup>32</sup>

The starting point for engineering was the gene encoding for DEBS 1-TE within an *Escherichia coli*/*Streptomyces* shuttle vector pMU3-DEBS 1-TE,<sup>10</sup> which enables heterologous expression in a previously-described strain of *Saccharopolyspora erythraea*, JC2 (Table S1, ESI<sup>†</sup>).<sup>33</sup> To allow for exchange of module 2 of DEBS 1-TE with heterologous modules by transformation associated recombination (TAR) cloning in *Saccharomyces cerevisiae* strain 23.344c, pMU3-DEBS 1-TE was first modified by insertion of the URA3-CEN6-ARS4 sequence,<sup>34,35</sup> to furnish the *E. coli*/*Streptomyces*/*Saccharomyces cerevisiae* shuttle vector pMU3-DEBS 1-TE-URA3 (Table S1, ESI<sup>†</sup>). The target modules were then amplified by PCR (Table S2 and S3, ESI<sup>†</sup>) to incorporate suitable regions of homology with pMU3-DEBS 1-TE-URA3,<sup>36</sup> and co-transformed together with linearised pMU3-DEBS 1-TE-URA3 into strain 23.344c. Successful, scarless module exchange was then verified by diagnostic digestion and DNA sequencing. The module modified DEBS 1-TE constructs were transformed into *Sac. erythraea* JC2 by transconjugation with *E. coli* ET12567 (pUZ8002). The resulting recombinants were grown as described previously,<sup>10</sup> and the triketide lactone products were extracted and analysed by GC-MS.

In *Sac. erythraea*, the mini-PKS use propionate (a) and acetate (b) as the starter unit. Thus overall, eight reduced lactone products (**1a/1b–4a/4b**, Table S4, ESI<sup>†</sup>) were possible in which the stereochemistry at carbons 2 and 3 varies (positions 4 and 5 are set by module 1, which was not altered in this experiment), as well as two 3-ketolactones (**5a/5b**). Assignment of product identity was based on comparison to synthetic standards (a 2.5:1 mixture of **1a** and **3a**, as well as **2a**) and biosynthetic standards (**1a/1b–3a/3b** and **5a/5b** from previously

**Table 1** Yields of lactones resulting from introduction of modules containing A2 and B2 KR into DEBS 1-TE

Module introduced	Strategy	Yields (mg/L) <sup>a</sup>				
						
Amph M11 (A2)	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)	A		nd		nd	14.6 ± 0.6
	B		nd		nd	6 ± 2
	C		nd		nd	nd
	D		nd		nd	nd
Bfm M4 (A2)	A		nd		nd	nd
	B		nd		nd	nd
Bfm M9 (A2)	A		nd		nd	nd
	B		nd		nd	nd
Hba M4 (A2)	A		nd		nd	nd
	B		nd		nd	nd
Hba M9 (A2)	A		nd		nd	nd
	B		nd		nd	nd
Lkm M1 (B2)	A		≈ 3.5		nd	≈ 2.7
	B		14 ± 1		traces	≈ 4.7
	C		nd		nd	nd
	D		nd		nd	nd
Las M7 (B2)	A		nd		nd	nd
	B		nd		nd	nd
	C		nd		nd	nd
	D		nd		nd	nd
Med M7a (B2)	A		nd		nd	nd
	B		nd		nd	nd

<sup>a</sup> Average yields and standard errors of experiments performed in triplicate are reported. nd indicates that the lactone was not detected.

described strains JC2/pJLK25, JC2/pJLK30 and JC2/pJLK35<sup>9</sup> (Fig. S3, ESI<sup>†</sup>; all retention times are provided in Table S4, ESI<sup>†</sup>). Lactone **4** was assigned by elimination, as only four reduced products are possible. Product yields were quantified based on a standard curve generated using the synthetic lactones **1a/3a**, for which there was a linear correlation between peak area and product concentrations in the range of 5–200 mg L<sup>-1</sup> (Fig. S4, ESI<sup>†</sup>). Three clones of each construct were analysed (the average yields are summarised in Table 1, while a representative analysis of each construct along with a sample yield calculation are provided in the ESI<sup>†</sup>).

The first important result is that none of the constructs incorporating hybrid ACPs was functional (Table 1). This result makes the reported strong activity<sup>9,12</sup> of parental DEBS 1-TE and certain of its derivatives surprising, but clearly underlines the need to keep ACP domains intact in module swap experiments. In our previous work, among all A2 KR tested, Amph KR<sub>11</sub> gave both the highest relative and absolute yields of the desired methyl-epimerised and reduced product **4** (17 mg L<sup>-1</sup> total lactone, 82.4% **4**, 17.6% ketolactone **5**). As hoped, exchanging a larger portion of the module resulted in substantially improved amounts of lactone **4**, albeit at the expense of relative yield (Table 1). Using the classical module boundaries (tetradomain KS-AT-KR-ACP; Amph M11-A), the overall yield of lactone increased almost 3-fold to 48.1 mg L<sup>-1</sup>, of which 75.0% was lactone **4** (36 mg L<sup>-1</sup>). The remaining product was evenly split between the ketolactone **5** observed previously (12.5%), and reduced, but non-epimerised lactone **2** (12.5%). Thus, the total amount of reduced lactone of either stereochemistry was also

boosted relative to the KR swap experiment (87.5% vs. 82.3%). As in our previous work,<sup>10</sup> no products (**1** or **3**) corresponding to incorrect B-type reduction were observed. Using the alternative module boundaries (tridomain AT-KR-ACP; Amph M11-B) produced an approximately 2-fold increase in total lactone yield compared to the KR exchange (35.7 mg L<sup>-1</sup>), but resulted in a greater proportion of **2** (19.0%), as well as ketolactone **5** (24.4%).

In keeping with the close homology between the Amph KR<sub>11</sub> and Nys KR<sub>1</sub> and **11** (84.1% and 69.3%, respectively), exchanging partial or whole Nys KR-containing modules resulted in the desired product **4**, and of the two modules, the Nys M11 gave the higher yields (Table 1). These results are reminiscent of swapping Amph KR<sub>1</sub> and **11** into DEBS 1-TE, wherein Amph KR<sub>11</sub> was superior,<sup>10</sup> showing that KR/modules sourced from later in the assembly lines operate better within a module 2 context. Again, the full module swap in each case produced both higher overall and relative yields of **4**, compared to the alternative module boundaries (Table 1; Nys M1-A, Nys M11-A). Stereochemical fidelity of both Nys modules was also moderately stronger than for Amph KR<sub>11</sub> (traces and 8.5% **2**, respectively). Taken together, these results demonstrate that heterologous KR function is improved by maintaining a native modular context, but that substrate specificity is nonetheless compromised, as the non-epimerised chains were also reduced. We attribute this loss of fidelity to the fact that the KR<sub>s</sub> were acting on chains which differ in both length and functionality to their native substrates. By contrast, the downstream TE has been shown to favour the α,β-stereochemistry present in **4** for both hydrolysis and ethanolysis,<sup>37</sup> and thus is unlikely to have biased the product distribution towards lactone **2**. Of the remaining modules containing A2 KR<sub>s</sub>, only Olm M5 was functional, yielding ketolactone **5** (Table 1; Olm M5-A, B). In addition, the **5** yields of 14.6 mg L<sup>-1</sup> and 6 mg L<sup>-1</sup>, respectively, for the full and partial swaps, were substantially lower than the amount of **5** produced by the corresponding KR exchange (52 mg L<sup>-1</sup> (ref. 10)). Thus, not only did Olm KR<sub>5</sub> fail to act, as observed previously, but the module swapping strategy further reduced module 2 chain extension activity. It is nonetheless notable that exploiting the classical boundaries to exchange the module again produced superior results.

In the case of the B2 KR-containing modules, only exchange of Lkm M1 produced the desired reduced and epimerised lactone **1**. Nonetheless, the yields (Table 1) were substantially improved relative to the corresponding KR swap, for which **5** predominated (6 mg L<sup>-1</sup>) and only traces of **1** were observed. In contrast to the A2 KR-containing modules, higher relative and overall yields of **1** were obtained by swapping the AT-KR-ACP tridomain instead of the full module (18.7 mg L<sup>-1</sup> total lactone, 75% **1** vs. ca. 6.2 mg L<sup>-1</sup>, 56% **1**; Table 1 (Lkm M1-B, A respectively)), although traces of the reduced, but non-epimerised lactone **3** were also observed. **3** was not present in extracts of the full module swap (Lkm M1-A), suggesting that Lkm KR<sub>1</sub> boosted the overall stereochemical fidelity of the module. However, its presence also led to reduced product yield, possibly due to a lower efficiency of intermodular chain transfer.<sup>28,38</sup>

None of the other module exchanges (A2-type Bfm M4/9 and HbaA M4/M9 and B2-type Las M7 and Med M7a) in either configuration yielded product (Table 1), despite the fact

that in our previous work, swapping in Las KR<sub>7</sub> produced a hybrid module 2 functional for chain extension (8 mg L<sup>-1</sup> of 5 (ref. 10)). In the case of the full module replacements, this failure is presumably due to lack of chain translocation between DEBS module 1 and the inserted module, while in the case of the AT–KR–ACP tridomain swaps which preserved the native DEBS ACP<sub>1</sub>/KS<sub>2</sub> interface, the problem likely arose at the stage of the condensation reaction catalysed between DEBS KS<sub>2</sub> and a heterologous ACP domain.

In conclusion, the work reported here establishes module exchange as a superior strategy to KR swapping to access polyketides incorporating unusual A2 and B2 stereochemistry. In this context, we have notably identified three A2-type and one B2-type modules which give both high overall and relative yields of the desired products, making them valuable additions to a PKS engineering toolbox. The data are, on the other hand, mixed in terms of which group of domains defines a functional model.<sup>26,27</sup> Although using classical module boundaries (KS–AT–KR–ACP) worked best for the A2-type modules, the alternative tridomain (AT–KR–ACP) gave better results for the B2-type module. Evidently, the relative negative effects of non-native chain transfer and chain extension interactions on chimeric PKSs can differ depending on the identity of the source modules. It is also notable that all of the functional modules are covalently joined to upstream modules within their native subunits. In such contexts, they have no ‘choice’ of donor acyl-ACP domain, and thus apparently have not evolved to discriminate strongly against alternative acyl-ACP substrates. In contrast, all of the non-functional swapped modules are located at the N-termini of their respective polypeptides. In these cases, the intermolecular nature of substrate delivery evidently imposes stricter specificity constraints on the module, both in terms of protein–protein (ACP/KS) and domain–substrate interactions.<sup>38</sup> Taken together, these findings argue that modules which are internal to subunits are intrinsically more adapted to relocation by genetic engineering into novel covalent contexts than their N-terminal counterparts.

We acknowledge support from the Agence Nationale de la Recherche (grant number ANR-16-CE92-0006-01, PKS STRUCTURE to K. J. W.), the Université de Lorraine, the Centre National de la Recherche Scientifique (CNRS), and the IMPACT Biomolecules project of the Lorraine Université d'Excellence (Investissements d'avenir – grant number ANR 15-004). Synthetic standards **1a** and **3a** were a kind gift of S. Heinrich and F. Schulz (Ruhr University Bochum, Germany), and **2a** was provided by A. T. Keatinge-Clay (University of Texas at Austin, USA).

## Conflicts of interest

There are no conflicts to declare.

## Notes and references

- 1 A. L. Demain, *J. Ind. Microbiol. Biotechnol.*, 2014, **41**, 185–201.
- 2 E. K. Davison and M. A. Brimble, *Curr. Opin. Chem. Biol.*, 2019, **52**, 1–8.
- 3 C. T. Walsh and M. A. Fischbach, *J. Am. Chem. Soc.*, 2010, **132**, 2469–2493.

- 4 E. M. Larsen, M. R. Wilson and R. E. Taylor, *Nat. Prod. Rep.*, 2015, **32**, 1183–1206.
- 5 J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380–416.
- 6 P. Caffrey, *ChemBioChem*, 2003, **4**, 654–657.
- 7 A. Garg, C. Khosla and D. E. Cane, *J. Am. Chem. Soc.*, 2013, **135**, 16324–16327.
- 8 A. T. Keatinge-Clay, *Chem. Biol.*, 2007, **14**, 898–908.
- 9 L. Kellenberger, I. S. Galloway, G. Sauter, G. Böhm, U. Hanefeld, J. Cortés, J. Staunton and P. F. Leadlay, *ChemBioChem*, 2008, **9**, 2740–2749.
- 10 T. Annaval, C. Paris, P. F. Leadlay, C. Jacob and K. J. Weissman, *ChemBioChem*, 2015, **16**, 1357–1364.
- 11 C. H. Eng, S. Yuzawa, G. Wang, E. E. K. Baidoo, L. Katz and J. D. Keasling, *Biochemistry*, 2016, **55**, 1677–1680.
- 12 J. Cortés, K. E. Wiesmann, G. A. Roberts, M. J. Brown, J. Staunton and P. F. Leadlay, *Science*, 1995, **268**, 1487–1489.
- 13 P. Caffrey, S. Lynch, E. Flood, S. Finnan and M. Oliyinyk, *Chem. Biol.*, 2001, **8**, 713–723.
- 14 S. Omura, H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki and M. Hattori, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 12215–12220.
- 15 K. Arakawa, F. Sugino, K. Kodama, T. Ishii and H. Kinashi, *Chem. Biol.*, 2005, **12**, 249–256.
- 16 L. Smith, H. Hong, J. B. Spencer and P. F. Leadlay, *ChemBioChem*, 2008, **9**, 2967–2975.
- 17 A. Nara, T. Hashimoto, M. Komatsu, M. Nishiyama, T. Kuzuyama and H. Ikeda, *J. Antibiot.*, 2017, **70**, 616–624.
- 18 E. M. Molloy, J. I. Tietz, P. M. Blair and D. A. Mitchell, *Bioorg. Med. Chem.*, 2016, **24**, 6276–6290.
- 19 T. Brautaset, O. N. Sekurova, H. Sletta, T. E. Ellingsen, A. R. Strøm, S. Valla and S. B. Zotchev, *Chem. Biol.*, 2000, **7**, 395–403.
- 20 L. Zhang, T. Hashimoto, B. Qin, J. Hashimoto, I. Kozono, T. Kawahara, M. Okada, T. Awakawa, T. Ito, Y. Asakawa, M. Ueki, S. Takahashi, H. Osada, T. Wakimoto, H. Ikeda, K. Shin-Ya and I. Abe, *Angew. Chem., Int. Ed.*, 2017, **56**, 1740–1745.
- 21 R. W. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay and K. J. Weissman, *Chem. Biol.*, 2003, **10**, 723–731.
- 22 R. S. Gokhale, S. Y. Tsuji, D. E. Cane and C. Khosla, *Science*, 1999, **284**, 482–485.
- 23 N. Wu, S. Y. Tsuji, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2001, **123**, 6465–6474.
- 24 R. McDaniel, C. M. Kao, S. J. Hwang and C. Khosla, *Chem. Biol.*, 1997, **4**, 667–674.
- 25 A. Ranganathan, M. Timoney, M. Bycroft, J. Cortés, I. P. Thomas, B. Wilkinson, L. Kellenberger, U. Hanefeld, I. S. Galloway, J. Staunton and P. F. Leadlay, *Chem. Biol.*, 1999, **6**, 731–741.
- 26 A. Nivina, K. P. Yuet, J. Hsu and C. Khosla, *Chem. Rev.*, 2019, **119**, 12524–12547.
- 27 A. T. Keatinge-Clay, *Angew. Chem., Int. Ed.*, 2017, **56**, 4658–4660.
- 28 M. Klaus, M. P. Ostrowski, J. Austerjost, T. Robbins, B. Lowry, D. E. Cane and C. Khosla, *J. Biol. Chem.*, 2016, **291**, 16404–16415.
- 29 H. Peng, K. Ishida, Y. Sugimoto, H. Jenke-Kodama and C. Hertweck, *Nat. Commun.*, 2019, **10**, 3918.
- 30 T. Miyazawa, M. Hirsch, Z. Zhang and A. T. Keatinge-Clay, *Nat. Commun.*, 2020, **11**, 80.
- 31 S. Yuzawa, K. Deng, G. Wang, E. E. K. Baidoo, T. R. Northen, P. D. Adams, L. Katz and J. D. Keasling, *ACS Synth. Biol.*, 2017, **6**, 139–147.
- 32 A. K. Joshi, A. Witkowski, H. A. Berman, L. Zhang and S. Smith, *Biochemistry*, 2005, **44**, 4100–4107.
- 33 C. J. Rowe, J. Cortés, S. Gaisser, J. Staunton and P. F. Leadlay, *Gene*, 1998, **216**, 215–223.
- 34 M. Grenson, *Eur. J. Biochem.*, 1969, **11**, 249–260.
- 35 D. Gietz, A. St. Jean, R. A. Woods and R. H. Schiestl, *Nucleic Acids Res.*, 1992, **20**, 1425.
- 36 N. Bonneaud, O. Ozier-Kalogeropoulos, G. Y. Li, M. Labouesse, L. Minvielle-Sebastia and F. Lacroute, *Yeast*, 1991, **7**, 609–615.
- 37 K. J. Weissman, C. J. Smith, U. Hanefeld, R. Aggarwal, M. Bycroft, J. Staunton and P. F. Leadlay, *Angew. Chem., Int. Ed.*, 1998, **37**, 1437–1440.
- 38 M. Klaus, L. Buyachuihan and M. Griningger, *ACS Chem. Biol.*, DOI: 10.1021/acscchembio.0c00405.

Electronic Supplementary Material (ESI) for Chemical Communications.

This journal is © The Royal Society of Chemistry 2020

## **Supporting Information**

### **Manipulating polyketide stereochemistry via exchange of polyketide synthase modules**

Jean-Malo Massicard,<sup>a</sup> Claire Soligot,<sup>b,c</sup> Kira J. Weissman\*<sup>a</sup> and Christophe Jacob\*<sup>a</sup>

<sup>a</sup> *Université de Lorraine, CNRS, IMoPA, F-54000 Nancy, France.*

<sup>b</sup> *Université de Lorraine, UR AFPA, USC 340 INRAE, F-54000 Nancy, France.*

<sup>c</sup> *Université de Lorraine, PASM, F-54000 Nancy, France.*

#### **Table of Contents**

1. Experimental methods
2. Additional tables and figures
3. Triketide lactone GC-MS data
4. References

## 1. Experimental methods

### *Strains, Plasmids, Media and Growth condition.*

The strains and plasmids used in this study are listed in **Table S1**. *E. coli* strains were grown in LB medium (agar and liquid). Appropriate antibiotics were added to the media when needed at the following concentrations: apramycin, 50 µg/mL; nalidixic acid, 50 µg/mL; kanamycin, 50 µg/mL; chloramphenicol, 25 µg/mL. *Saccharomyces cerevisiae* 23.344c was grown in rich medium (agar and liquid). Transformants were selected on uracil-deficient YNB (Difco) containing 20 mM ammonium and 3% (w/v) glucose. *Sac. erythraea* mutant strain JC2 NCIMB 40802 has been described previously.<sup>1</sup> Medium ABB13 was used for conjugation, and screening medium 3 (SM3) was used for triketide lactone production.<sup>2</sup>

### *Materials, DNA isolation and DNA manipulation.*

Primers used in this study are listed in **Table S2**. Standard procedures were used for DNA purification, PCR and molecular analysis. PCR was performed using Phusion Hot Start II DNA polymerase (Thermo Scientific, US) in a Mastercycler Pro (Eppendorf, Weseling-Berzdorf, Germany). Isolation of DNA fragments from agarose gel and purification of PCR products were carried out with a Nucleo-Spin Extract II kit (Marcherey-Nagel, Hoerd, France). Other enzymes used in this study were from Thermo Scientific. PCR primers were purchased from Sigma-Aldrich, and DNA sequencing was carried out by GATC Biotech (Mulhouse, France). All kits and enzymes were used according to the manufacturers' recommendations.

### *Plasmid construction.*

PCR was used to amplify the fragment containing *URA3* and *CEN6/ARS4* sites from plasmid pFL38.<sup>3</sup> The fragment contains a 15 bp homology region to the 5' end of the *Pcil* site from plasmid pMU3-DEBS 1-TE. After gel purification of the *URA3-CEN6/ARS4* fragment and the *Pcil* digested pMU3-DEBS 1-TE plasmid, they were assembled using the In-Fusion HD Cloning kit (Clontech, US) to yield plasmid pMU3-DEBS 1-TE-*URA3*. The identity of pMU3-DEBS 1-TE-*URA3* was confirmed by diagnostic digestion and sequencing. PCR was used to amplify each module from cosmid or genomic DNA template (**Table S3**), using appropriate pairs of primers (**Table S2**) containing 50 bp homologous to the appropriate region in the pMU3-DEBS 1-TE-*URA3* plasmid. Yeast homologous recombination was employed to recombine the module fragments with the pMU3-DEBS 1-TE-*URA3* plasmid. Each individual module fragment and *AvrII*-digested pMU3-DEBS 1-TE-*URA3* were co-transformed into *Saccharomyces cerevisiae* 23.344c as described elsewhere.<sup>4</sup> Correct plasmid assembly was confirmed by diagnostic digestion and sequencing.

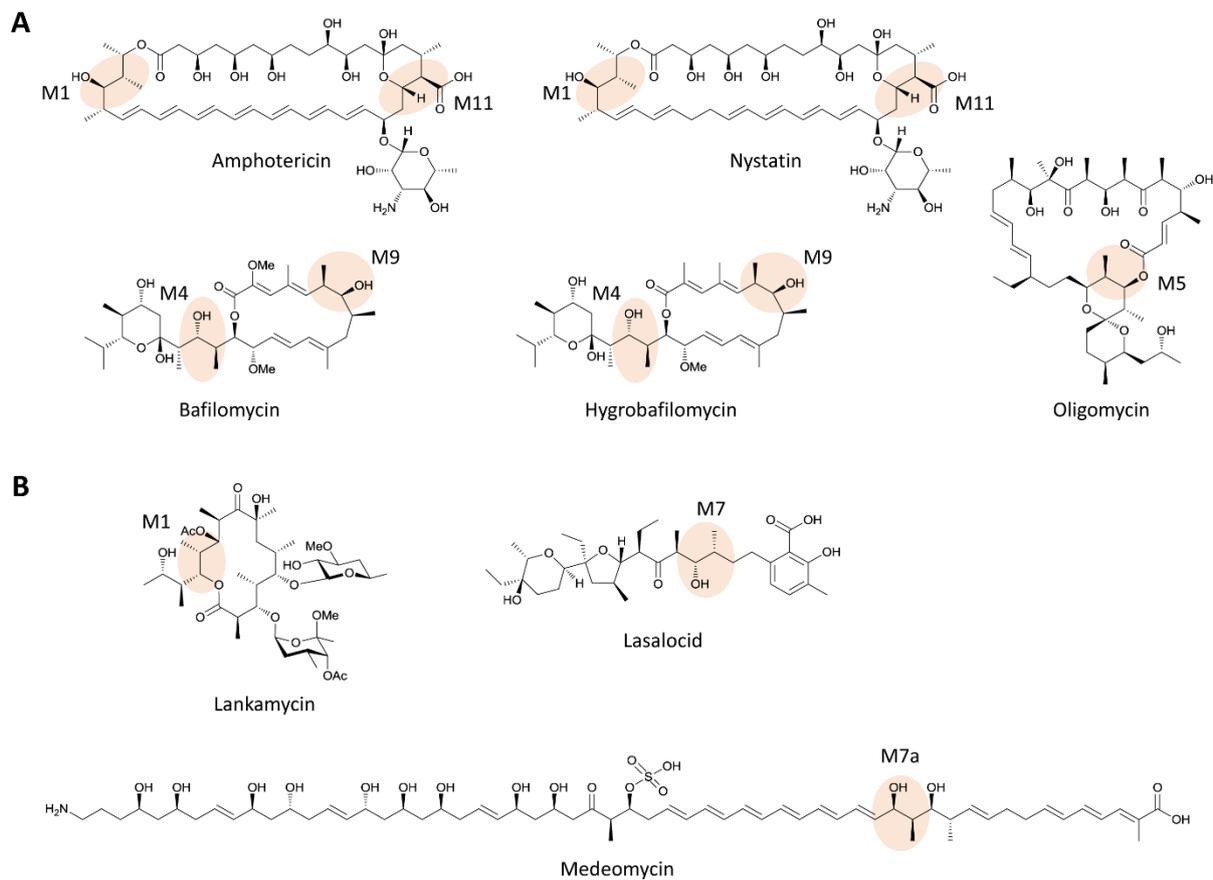
*Transformation of Sac. erythraea JC2 by transconjugation, and growth for triketide lactone production.* The pMU3-DEBS 1-TE-URA3 derivatives were transferred into *E. coli* 12567/pUZ8002 cells by heat shock transformation. Transconjugation between *E. coli* ET12567/pUZ8002 and *Sac. erythraea* JC2 was carried out as described previously.<sup>1</sup> The plates were incubated for 14 days at 30 °C, or until exconjugates became visible. At least 4 clones per plasmid were then grown in liquid TSB for two days at 30 °C. Equal portions of the cultures were then used to inoculate SM3 production plates, which were incubated for 14 days at 30 °C. Equally-sized portions of each SM3 plate (2 cm<sup>2</sup>) were removed and frozen so that all samples could be analysed together.

*Analysis of triketide fermentation products.*

Extraction of triketides was carried out using the “lactonex” procedure, as previously described:<sup>5</sup> the excised 2 cm<sup>2</sup> pieces of mycelium and agar were chopped into pieces, and transferred into 2 mL Eppendorf tubes. Ethyl acetate (1.2 mL) and formic acid (20 µL) were added to each tube with mixing. The tubes were incubated at 50 °C for 15 min, and then vortexed for 30 min. The mixtures were centrifuged (20 000g) for 1 min, and the supernatants carefully removed and placed in fresh 1.5 mL Eppendorf tubes. After evaporation to dryness under reduced pressure, the residue in each tube was dissolved in ethyl acetate (200 µL), centrifuged for 1 min and the supernatant was transferred to a GC-MS vial. GC-MS was performed on a model 7890A-5975C system (Agilent Technologies) equipped with a DB 5 MS column (30 m × 0.25 mm × 0.25 µm; Agilent technologies): 40 °C for 2 min; 10 °C/min to 250 °C; 25 °C/min to 300 °C, hold for 7 min. The products were analysed by EI-MS in positive-ion MS mode, in both the full-scan and extracted ion modes ( $m/z = 58$  and  $56$ ).

The identities of all lactones present in each sample were determined by comparison of the observed retention times and mass spectra with those of synthetic (a 2.5:1 mixture of **1a** and **3a**, as well as **2a**) and/or biosynthetic standards (**1ab–3ab** and **5ab** from previously-described strains JC2/pJLK25, JC2/pJLK30 and JC2/pJLK35).<sup>5</sup> In order to quantify lactone yields, we constructed a calibration curve with known quantities of the 2.5:1 mixture of lactones **1a** and **3a**. This yielded a linear correlation between lactone quantity and peak area in the total ion chromatogram (TIC). We could further confirm that this correlation held for key extracted ions (58 for reduced lactones **1–4**, and 56 for ketolactone **5**), allowing us to also quantify lactone yields based on the EICs (this was critical for lactones **1b** and **5a** whose peaks overlap in the TIC).

## 2. Additional tables and figures



**Fig. S1** Structures of polyketides whose modular polyketide synthases (PKSs) were used as sources of heterologous modules. (A) Sources of A2 modules. (B) Sources of B2 modules. In (A) and (B), the relevant stereocentres and the modules which establish them, are indicated (shading).

**Table S1.** List of strains and plasmids used in this study.

Name	Description	Source and reference
DH5 $\alpha$	<i>Escherichia coli</i> for routine plasmid maintenance, <i>supE44</i> , $\Delta$ <i>lacU169</i> ( $\odot$ 80 <i>lacZ</i> $\Delta$ M15), <i>hsdR17</i> , <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>relA1</i>	Life Technologies
23.344c	<i>Saccharomyces cerevisiae</i> for yeast homologous recombination, <i>Mat<math>\alpha</math></i> , <i>ura3</i>	6
ET12567/pUZ8002	<i>Escherichia coli</i> for conjugation, <i>F<sup>-</sup></i> , <i>dam<sup>-</sup></i> 13::TN9, <i>dcm-6</i> , <i>hsdM</i> , <i>Cml<sup>R</sup></i> , carrying helper plasmid pUZ8002	7
JC2	<i>Saccharopolyspora erythraea</i> for lactone production, $\Delta$ <i>eryA</i>	1
pMU3-DEBS 1-TE-URA3	pMU3-DEBS 1-TE modified to incorporate the URA3-CEN6/ARS4 DNA sequence	This study
pMU3-DEBS Amph <sub>M11</sub> -A	1-TE-URA3- pMU3-DEBS 1-TE-URA3 modified to incorporate Amph M11 (KS $\rightarrow$ ACP)	This study
pMU3-DEBS Amph <sub>M11</sub> -B	1-TE-URA3- pMU3-DEBS 1-TE-URA3 modified to incorporate Amph M11 (linker KS-AT $\rightarrow$ ACP)	This study
pMU3-DEBS Amph <sub>M11</sub> -C	1-TE-URA3- pMU3-DEBS 1-TE-URA3 modified to incorporate Amph M11 (KS $\rightarrow$ ACP <sub>SacI</sub> )	This study
pMU3-DEBS Amph <sub>M11</sub> -D	1-TE-URA3- pMU3-DEBS 1-TE-URA3 modified to incorporate Amph M11 (linker KS-AT $\rightarrow$ ACP <sub>SacI</sub> )	This study
pMU3-DEBS 1-TE-URA3-Bfm <sub>M4</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Bfm M4 (KS $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Bfm <sub>M4</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Bfm M4 (linker KS-AT $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3-Bfm <sub>M9</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Bfm M9 (KS $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Bfm <sub>M9</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Bfm M9 (linker KS-AT $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3-Hba <sub>M4</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Hba M4 (KS $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Hba <sub>M4</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Hba M4 (linker KS-AT $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Hba <sub>M9</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Hba M9 (KS $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Hba <sub>M9</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Hba M9 (linker KS-AT $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3-Las <sub>M7</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Las M7 (KS $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Las <sub>M7</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Las M7 (linker KS-AT $\rightarrow$ ACP)	This study
pMU3-DEBS 1-TE-URA3- Las <sub>M7</sub> -C	pMU3-DEBS 1-TE-URA3 modified to incorporate Las M7 (KS $\rightarrow$ ACP <sub>SacI</sub> )	This study
pMU3-DEBS 1-TE-URA3- Las <sub>M7</sub> -D	pMU3-DEBS 1-TE-URA3 modified to incorporate Las M7 (linker KS-AT $\rightarrow$ ACP <sub>SacI</sub> )	This study
pMU3-DEBS 1-TE-URA3-Lkm <sub>M1</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Lkm M1 (KS $\rightarrow$ ACP)	This study

*Annexes*

pMU3-DEBS 1-TE-URA3- Lkm <sub>M1</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Lkm M1 (linker KS-AT → ACP)	This study
pMU3-DEBS 1-TE-URA3- Lkm <sub>M1</sub> -C	pMU3-DEBS 1-TE-URA3 modified to incorporate Lkm M1 (KS → ACP <sub>Sacl</sub> )	This study
pMU3-DEBS 1-TE-URA3- Lkm <sub>M1</sub> -D	pMU3-DEBS 1-TE-URA3 modified to incorporate Lkm M1 (linker KS-AT → ACP <sub>Sacl</sub> )	This study
pMU3-DEBS 1-TE-URA3- Nys <sub>M1</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Nys M1 (KS → ACP)	This study
pMU3-DEBS 1-TE-URA3- Nys <sub>M1</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate NysM1 (linker KS-AT → ACP)	This study
pMU3-DEBS 1-TE-URA3-Nys <sub>M11</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Nys M11 (KS → ACP)	This study
pMU3-DEBS 1-TE-URA3- Nys <sub>M11</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate NysM11 (linker KS-AT → ACP)	This study
pMU3-DEBS 1-TE-URA3-Med <sub>M7a</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Med M7a (KS → ACP)	This study
pMU3-DEBS 1-TE-URA3- Med <sub>M7a</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Med M7a (linker KS-AT → ACP)	This study
pMU3-DEBS 1-TE-URA3-Olm <sub>M5</sub> -A	pMU3-DEBS 1-TE-URA3 modified to incorporate Olm M5 (KS → ACP)	This study
pMU3-DEBS 1-TE-URA3- Olm <sub>M5</sub> -B	pMU3-DEBS 1-TE-URA3 modified to incorporate Olm M5 (linker KS-AT → ACP)	This study
pMU3-DEBS 1-TE-URA3- Olm <sub>M5</sub> -C	pMU3-DEBS 1-TE-URA3 modified to incorporate Olm M5 (KS → ACP <sub>Sacl</sub> )	This study
pMU3-DEBS 1-TE-URA3- Olm <sub>M5</sub> -D	pMU3-DEBS 1-TE-URA3 modified to incorporate Olm M5 (linker KS-AT → ACP <sub>Sacl</sub> )	This study

Table S2. List of forward and reverse primers.

Plasmid	Primers	Sequence (5' → 3') <sup>a</sup>
pMU3-DEBS 1- TE-URA3- Amph <sub>M11</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCcatcggcattgagctgccgctaccccggcgga
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGtccttcggcatcagcgggtaccaacgcccac
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCcagttcggagtgagggaagccggccagggc
	Reverse ACP <sub>SacI</sub>	GCCTGCTGGAGCTGGTTGCGCAGCCCCGACCGGTCAGCGAGTC GAAGCCGAcgtcgcggaaggcgcgctggtcggtgag
pMU3-DEBS 1- TE-URA3-Bfm <sub>M4</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCgtctcagatggcctgccgctaccccggcggc
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGgccttcggcgtcagcggcaccaatgcccacctg
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCcaggccggtctccaggcgcggcgagcgc
pMU3-DEBS 1- TE-URA3-Bfm <sub>M9</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCgtggcgatgagctgccgcttcccggcggg
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGtcgcttcggcatgagcggcaccaacgcgcacgtc
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCcagttcggcgcggatctcccgggcgagcgc
pMU3-DEBS 1- TE-URA3-Hba <sub>M4</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCgtctccatgggctgccgctaccccgggtggc
	Forward AT	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGgcgcttcggcgtcagcggcaccaacgcccac
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCgaggagggcgcggcaggtggcgggcgagggc
pMU3-DEBS 1- TE-URA3-Hba <sub>M9</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCgcggtcgtctcagatggcctgccgcttcccc
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGtccttcggcatgagcggcaccaacgcccac
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCgagttccgtgccgatgtgccgggacagggc
pMU3-DEBS 1- TE-URA3- Med <sub>M7a</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCgtcgcgatgagctgccgctaccccgggggc
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGtccttcggcttcagcggcaccaacgcccac
	Reverse ACP <sub>full</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCgagctcggcgcgcaggtgcctggcgagcac
pMU3-DEBS 1- TE-URA3-NyS <sub>M1</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCgtcggaatgggctgccgacttcccggcggc
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGtcgcttcggcatcagcggcaccaacgcgcac
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCcgccgctcgtcttcaggaaacgcggcgagcgc
pMU3-DEBS 1- TE-URA3-NyS <sub>M11</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCcatcggcattgacctgccgcttcccggcggc
	Forward KS-AT linker	GGGAGTGGTCGCCC GCCCGGACGGGGTGC GCCGGGCAGGTGTG TCGtcgcttcggcatcagcgggtaccaacgcccac
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCcagctcggcgtgcagataggcggcgagcgc
	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCcatcgtcggcatggcctgccgctaccccggcggc

Annexes

pMU3-DEBS 1- TE-URA3-Olm <sub>M5</sub>	Forward KS-AT linker	GGGAGTGGTCGCCCCGCCGCGGACGGGGTTCGCCGGGCAGGTGTG TCGgcggttcggagtcagcggcaccaacgcacat
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCgagggccgcgggcccaggaaatccgccagttg
	Reverse ACP <sub>SacI</sub>	GCCTGCTGGAGCTGGTTGCGCAGCCCCGACCGCGGTTCAGCGAGTC GAAGCCGAgctccttgagcgcgcgggcccgggtcgac
pMU3-DEBS 1- TE-URA3-La <sub>M7</sub>	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCggcatggcctgcccgttgcccgggtggc
	Forward KS-AT linker	GGGAGTGGTCGCCCCGCCGCGGACGGGGTTCGCCGGGCAGGTGTG TCGgccttcggcatcagcggcaccaacgcccac
	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCgagttcggcccgcaggtggcgtaccaccgc
pMU3-DEBS 1- TE-URA3-Lkm <sub>M1</sub>	Reverse ACP <sub>SacI</sub>	GCCTGCTGGAGCTGGTTGCGCAGCCCCGACCGCGGTTCAGCGAGTC GAAGCCGAggtcccggaaggcacgcgtcgggtcgac
	Forward KS	AGCAGGCGGCACCGGCGACCACGGCCCCCGTCGACGAGCCGATC GCGATCcgtegcctatggccgtacggtgcccggcggg
	Forward AT	GGGAGTGGTCGCCCCGCCGCGGACGGGGTTCGCCGGGCAGGTGTG TCGtcgaagaggcaccacgcgacgtccgggcccgc
Sequencing of pMU3-DEBS 1- TE-URA3	Reverse ACP <sub>complete</sub>	TAGCCGTCGCGAAGAGCGCTGCTCGCTTCCCGGGCGGGAGTCCC GCTGTCgagttcggcctccaggcgcagggcgacggc
	Reverse ACP <sub>SacI</sub>	GCCTGCTGGAGCTGGTTGCGCAGCCCCGACCGCGGTTCAGCGAGTC GAAGCCGAgttccgtgaacgcccgggtccgcggggac
	Forward complete module	cggagcgcgagaaagcgctg
	Forward KS-AT linker	tcggtgaagtogaacatcgg
	Reverse complete and SacI-fused module	gccgggctggggcacggcccg

<sup>a</sup>Sequence homologous to pMU3-DEBS 1-TE-URA3 is in all caps, while that in lower case represents sequence homologous to the genomic/cosmid DNA from the various actinomycete strains used during this study.

**Table S3.** Sources of DNA used in this work.

<b>Module</b>	<b>Source of DNA</b>
Amph M1 and M11	<i>Streptomyces nodosus</i> strain ATCC 14899 (Prof. P. Caffrey, University College Dublin, Ireland)
Bfm M4 and M9	<i>Kitasatospora setae</i> DSM 43861 strain from the DSMZ (Germany)
Hba M4 and M9	<i>Streptomyces varsoviensis</i> strain DSM 40346 from the DSMZ (Germany)
Las M7	<i>Streptomyces lasaliensis</i> strain NRRL 3382 (Prof. P.F. Leadlay, University of Cambridge, United Kingdom)
Lkm M1	Cosmid 14F1 (Prof. K. Arakawa, Hiroshima University, Japan)
Nys M1 and M11	<i>Streptomyces noursei</i> strain DSM 40635 from the DSMZ (Germany)
Med M7a	<i>Streptomyces blastmyceticus</i> strain DSM 40029 from the DSMZ (Germany)
Olm M5	<i>Streptomyces avermitilis</i> strain ATCC 31267 from the DSMZ (Germany)

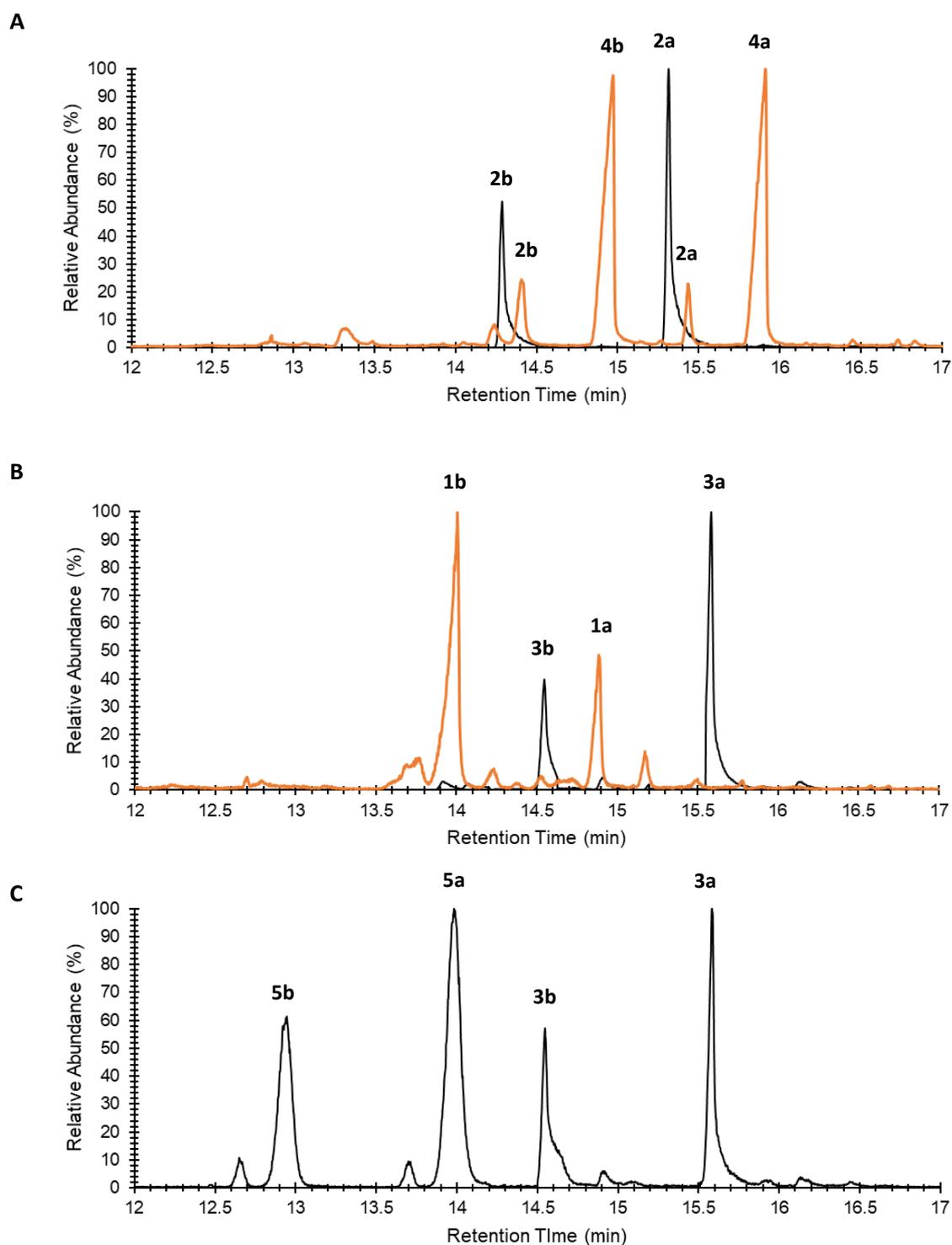
**Table S4.** Possible products resulting from exchanging module 2 of DEBS 1-TE for modules containing A2 and B2 KR<sub>s</sub>.

Catalytic events (Introduced module)	Product	Retention time (min) <sup>a</sup>	
		R = Et (a)	R = Me (b)
B2: Epimerisation + B-type reduction		14.9	13.9
A1: No epimerisation, A-type reduction		15.3	14.3
B1: No epimerisation, B-type reduction		15.6	14.6
A2: Epimerisation + A-type reduction		15.8	14.8
No reduction		14.1	13.1

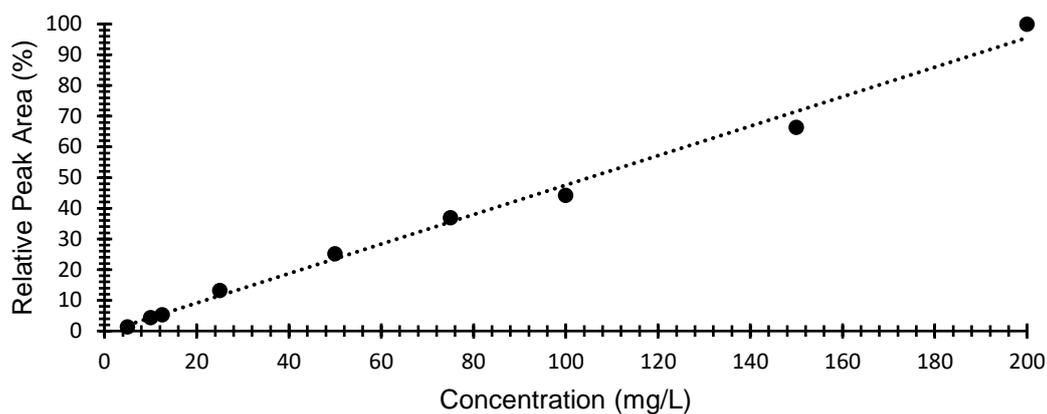
<sup>a</sup>We observed a dependence of the retention time on product yields, with low quantities of lactone leading to variation of as much as 0.1 min. However, these discrepancies were never large enough to preclude conclusive assignment of stereochemistry.

<b>A</b>	DEBS_M2	TGAEQAAPAT	TAPVDE	<b>PIAI</b>	VGMACRLPGE	VDSPELWEL
	Amph_M11	VLEVAGPVAT	GGTDEE	<b>PIAI</b>	IGMSCRYPGG	VSSPEQLWDL
	Bfm_M4	GDARRQLREA	EQARHE	<b>PIAI</b>	VSMACRYPGG	ADTPELLWDL
	Bfm_M9	RDTRQRLREA	EARSSE	<b>PIAI</b>	VAMSCRFPGG	VRDPEELWDL
	Hba_M4	GEAKRRLRDA	ENARRE	<b>PIAI</b>	VSMGCRYPGG	ADTPERLWDL
	Hba_M9	QNTRRRLLADA	ETRAHE	<b>PIAV</b>	VSMACRFPGG	VRTPDDLWEL
	Las_M7	SRVRRQLQET	EAASRE	<b>PIAI</b>	IGMACRLPGG	VDSPEGLWEL
	Lkm_M1	PDPRRREEAA	GHTPDE	<b>PIAV</b>	VMAVRLPGG	VRTPEQFWEL
	Nys_M1	RRARRRIGEL	ESKDNE	<b>PIAI</b>	VGMGCRPFGG	VNSPESLWDL
	Nys_M11	VLEVAGPVAT	GGADDE	<b>PIAI</b>	IGMACRFPGG	VSSPEQLWDL
	Med_M7a	RQARRRLREV	EDRHQE	<b>PIAI</b>	VAMSCRYPGG	VRTPEDLWRL
	Olm_M5	VRQHEDEVAP	TADHDD	<b>PIAI</b>	VGMACRYPGG	VRGPEDLWDL
	<b>B</b>	DEBS_M2	--GEIELADG	VREWSPAADG	VRRAGVSAFG	VS
Amph_M11		-PGTVRLLGE	NTDW-PQTGR	PRRAAVSSFG	IS	<b>GTNAH</b> VVIL
Bfm_M4		--GAVELLTE	ARAW-DDTGR	PRRAAVSAFG	VS	<b>GTNAH</b> LIV
Bfm_M9		--GAVELLTE	ARDW-PAVDR	PRRAAVSSFG	MS	<b>GTNAH</b> VVVL
Hba_M4		-AGAVRVLSR	SEPW-PETGR	PRRAGVSAFG	VS	<b>GTNAH</b> LILL
Hba_M9		-AGAVELLTD	ARFW-PETGR	PRRAGVSSFG	MS	<b>GTNAH</b> LVL
Las_M7		-SGAVTLLTE	PVDW-PDSR	PRRAGVSAFG	IS	<b>GTNAH</b> VVIL
Lkm_M1		-RGEVRLLE	PVPW-AGAAR	PRRAGVSSFG	IS	<b>GTNAH</b> VVVL
Nys_M1		-AGSVRLLTE	GQQW-PETGR	PRRAAVSSFG	IS	<b>GTNAH</b> HALL
Nys_M11		-AGTVRLLTQ	ARAW-PETGR	PRRAAVSSFG	IS	<b>GTNAH</b> VVLL
Med_M7a		SVGAVSLLTE	TPPW-PETGR	PRRAAVSSFG	FS	<b>GTNAH</b> TIL
Olm_M5		-AGTVGLLTE	NREW-PHTGK	PRRAGVSAFG	VS	<b>GTNAH</b> VVVL
<b>C</b>		DEBS_M2	TLVFDHPNA	SAVAGFLDAE	<b>L</b> GTVEVRGEAP	SALAGLDALE
	Amph_M11	TMVFDYPNP	AALAGFLHSE	LADVHS--AG	AVAVTAGAPV	
	Bfm_M4	TVVFDYSSA	TALAGRLETG	LGAADQPAA	GRPAAVRRPP	
	Bfm_M9	TLIFDHPSA	TALAREIRAE	LLG--RPAAD	PAPAGPDTAA	
	Hba_M4	TVVFDYSSA	TALARHLGAL	LLGDAST-GE	GAPLAAGALP	
	Hba_M9	SLIFDYPSA	AALSRHIGTE	LMGGDAPPTA	PAPSAEPAAA	
	Las_M7	TLVFDHPTP	EAVVRHLRAE	LGLEGDGAPD	PVFDELDGLE	
	Lkm_M1	TLVYDHPHA	RAVALRLEAE	LFGGARPEPA	AGVAGSGSG-	
	Nys_M1	TMVFDHPNC	AALAAFLKTT	ALGVPG-AAP	QQHAATGTPA	
	Nys_M11	TMVFDYPNP	AALAAYLHGE	LAGARSAAG	AAAVPTGAP-	
	Med_M7a	TLIFDYPTP	AVLARHLRAE	LGGQLATAA	PLPTAAALAD	
	Olm_M5	TMVFDHPTI	AELADFLARG	LTPEAAVPAE	PATVVRVDQD	
	<b>D</b>	DEBS_M2	PKAVRATTPF	<u>KEL</u> <b>GFDSL</b> LAA	VRLRNLLNAA	TGLRRLPSTLV
Amph_M11		AEDLTDQRAF	RDV <b>GFDSL</b> TA	VGLRNRLASV	TGLTLPSTMV	
Las_M7		LEAVEPTRAF	RDL <b>GFDSL</b> MA	VELRNRIKAA	TGLRRLPSTLV	
Lkm_M1		AERVPADRAF	TEL <b>GFDSL</b> AS	VELRNRLTAA	TGLRRLPSTLV	
Olm_M5		ARSVDPARAL	KEL <b>GFDSL</b> TA	VELRNRLSTA	TGLRRLPATMV	

**Fig. S2** Location of splice sites for introducing the heterologous modules into the DEBS 1-TE (indicated with arrows). (A) Junction retained to exchange the module while preserving intact the intermodular linker region. In purple, the highly conserved **PIAI** motif at the beginning of the KS.<sup>8</sup> (B) Junction retained to exchange the module while preserving the ACP/KS intermodular interface (underlined *HindIII* restriction site).<sup>9</sup> The conserved **GTNAH** motif used by others to swap AT domains<sup>10</sup> is also indicated. (C) Junction retained to exchange the module while keeping the entire ACP domain intact. In green, the conserved **L** corresponds to the end of the last  $\alpha$ -helix of the ACP. (D) Junction retained to exchange the module by generating a chimeric protein ACP<sub>x</sub>/DEBS ACP<sub>6</sub>, adjacent to a well-conserved **GFDSL** motif (blue). The underlined *SacI* restriction site was used for the construction of the original DEBS 1-TE system.<sup>11</sup>



**Fig. S3** GC-MS identification of the lactones. (A) Analysis of previously-characterised strain JC2/pJLK25<sup>2</sup> (black) producing **2a** and **2b** superimposed with JC2/Amph M11-A (orange). As discussed in the main text, we observed some variation of retention time with lactone yield (as much as 0.1 min), explaining why peaks **2a** and **2b** in the two samples do not strictly superimpose. (B) Analysis of previously-characterised strain JC2/pJLK30<sup>2</sup> (black) giving rise to **3a** and **3b** and traces of **1a**, **1b** (black), superimposed with JC2/Lkm M1-A (orange). (C) Analysis of previously-characterised strain JC2/pJLK35<sup>2</sup> giving rise to lactones **3a**, **3b**, **5a** and **5b**. The identities of lactones **1a**, **2a** and **3a** were confirmed by comparison to the chromatographic behaviour of synthetic standards, while those of **4a** and **4b** (shown here in (A), JC2/Amph<sub>M11</sub>-A) were assigned by elimination.

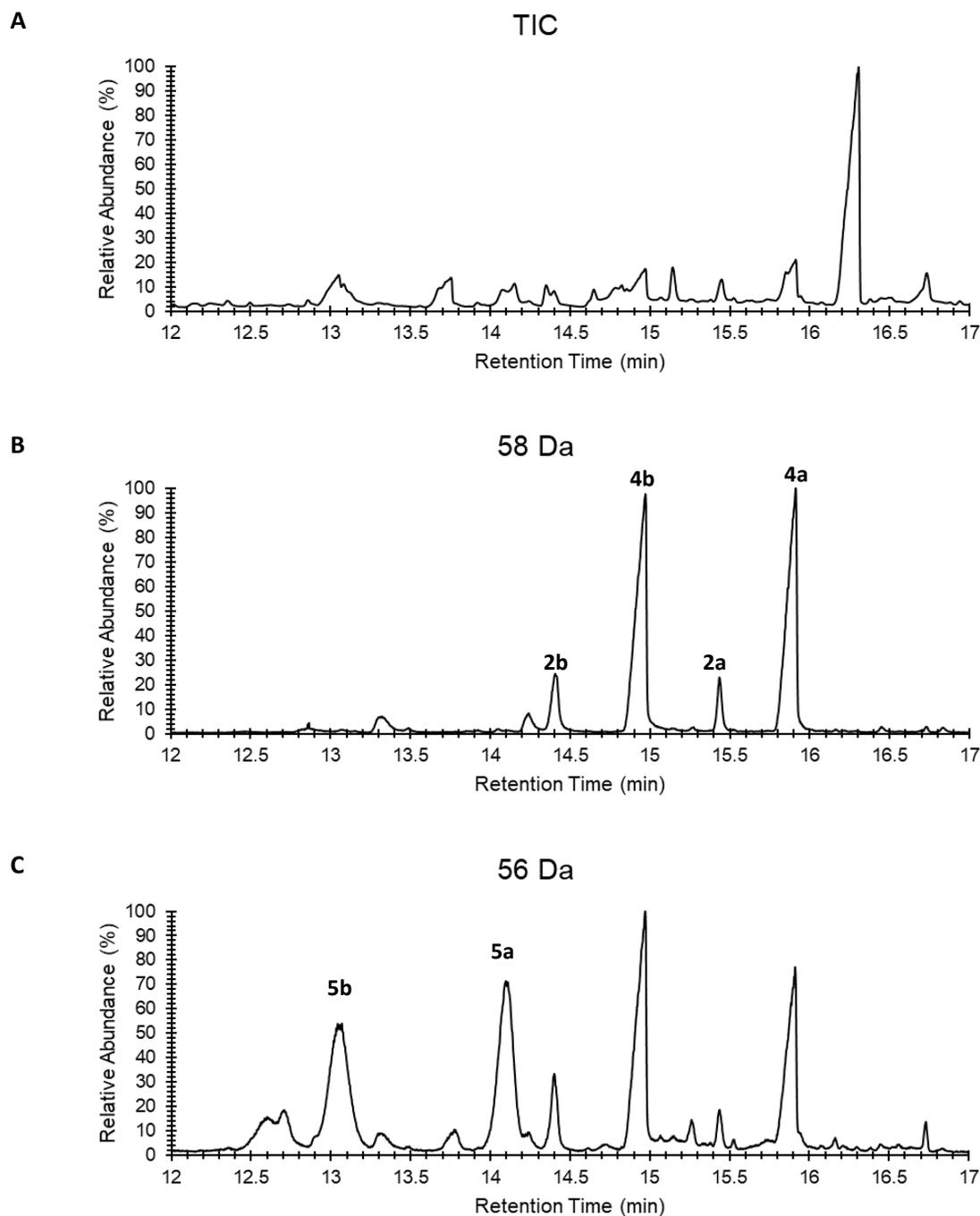


Data used to generate the calibration curve:

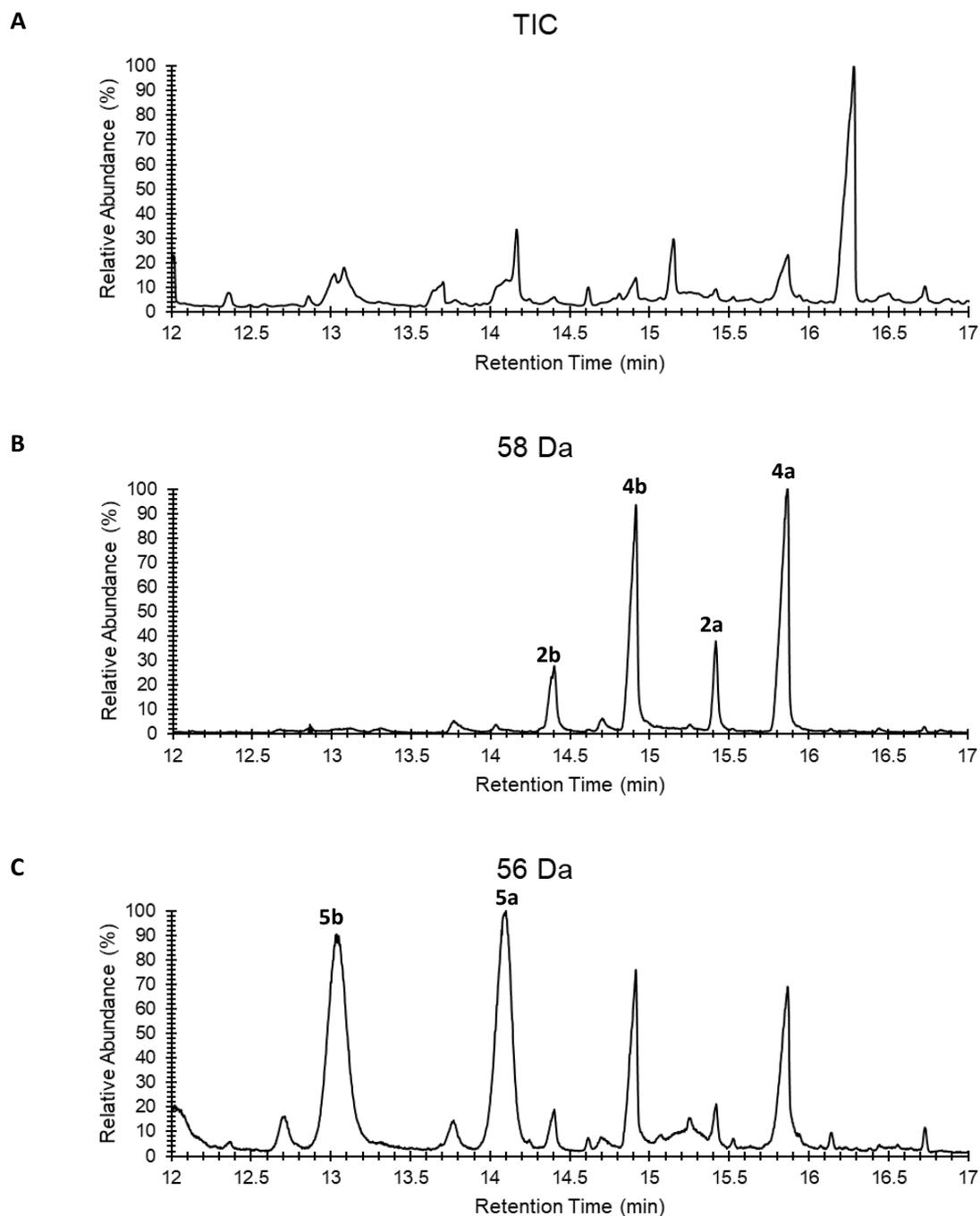
Concentration (mg/L)	Calculated peak area	Relative peak area (%)
200	56653020	100
150	37583022	66.3
100	25055348	44.2
75	20915773	36.9
50	14281548	25.2
25	7514826	13.2
12.5	3026725	5.3
10	2505535	4.4
5	751482	1.3

**Fig. S4** Calibration of lactone yields using known quantities of 2.5:1 mixture of synthetic standards of reduced lactones **1a** and **3a**. Peak areas were calculated from extracted ion chromatograms ( $m/z = 58$ ). This analysis yielded the following equation which was used to determine lactone yields in the biological samples:  $y = 0.4802x - 0.4732$  ( $R^2 = 0.9925$ ).

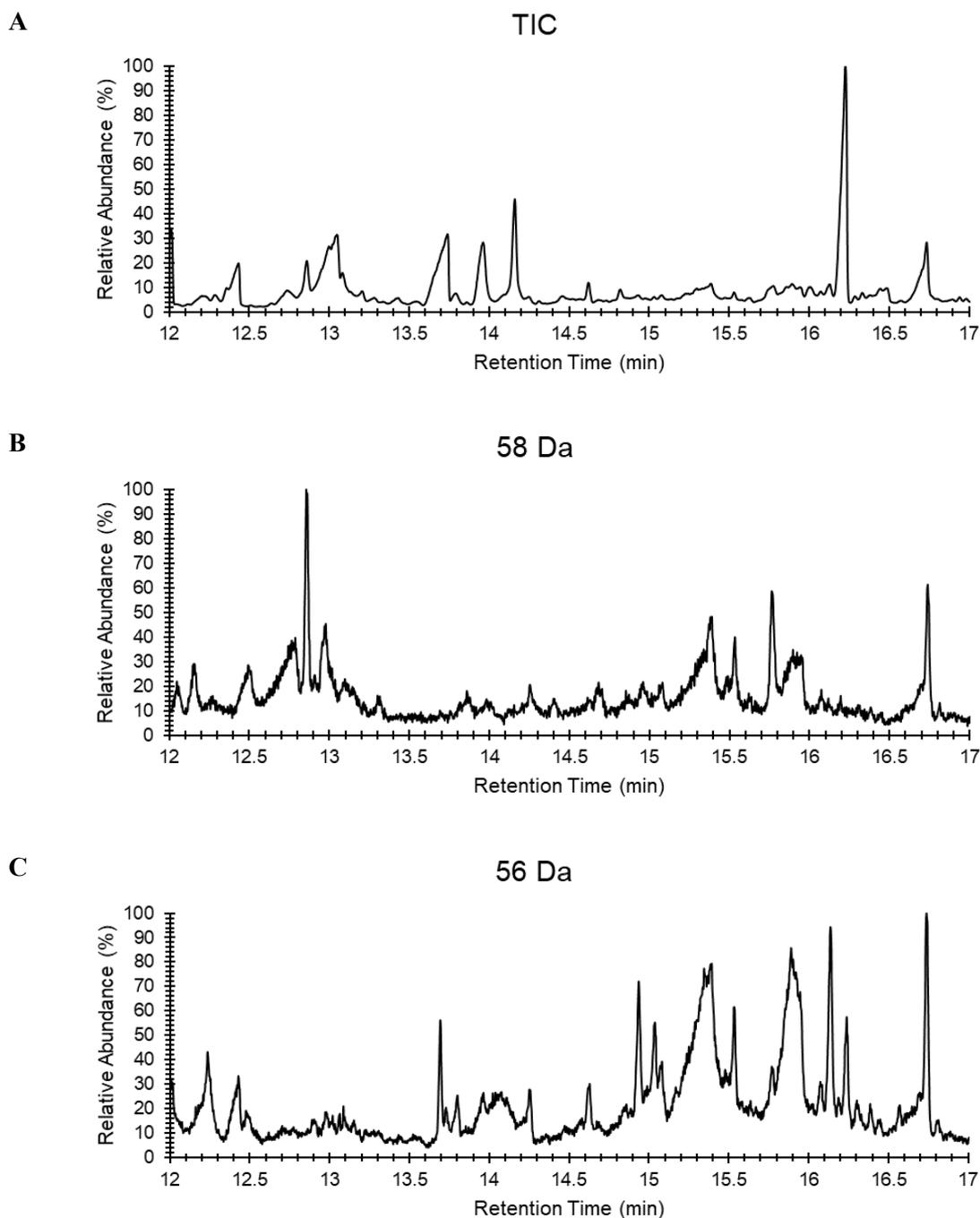
## 3. Triketide lactone GC-MS data

Amph<sub>M11</sub>-A:

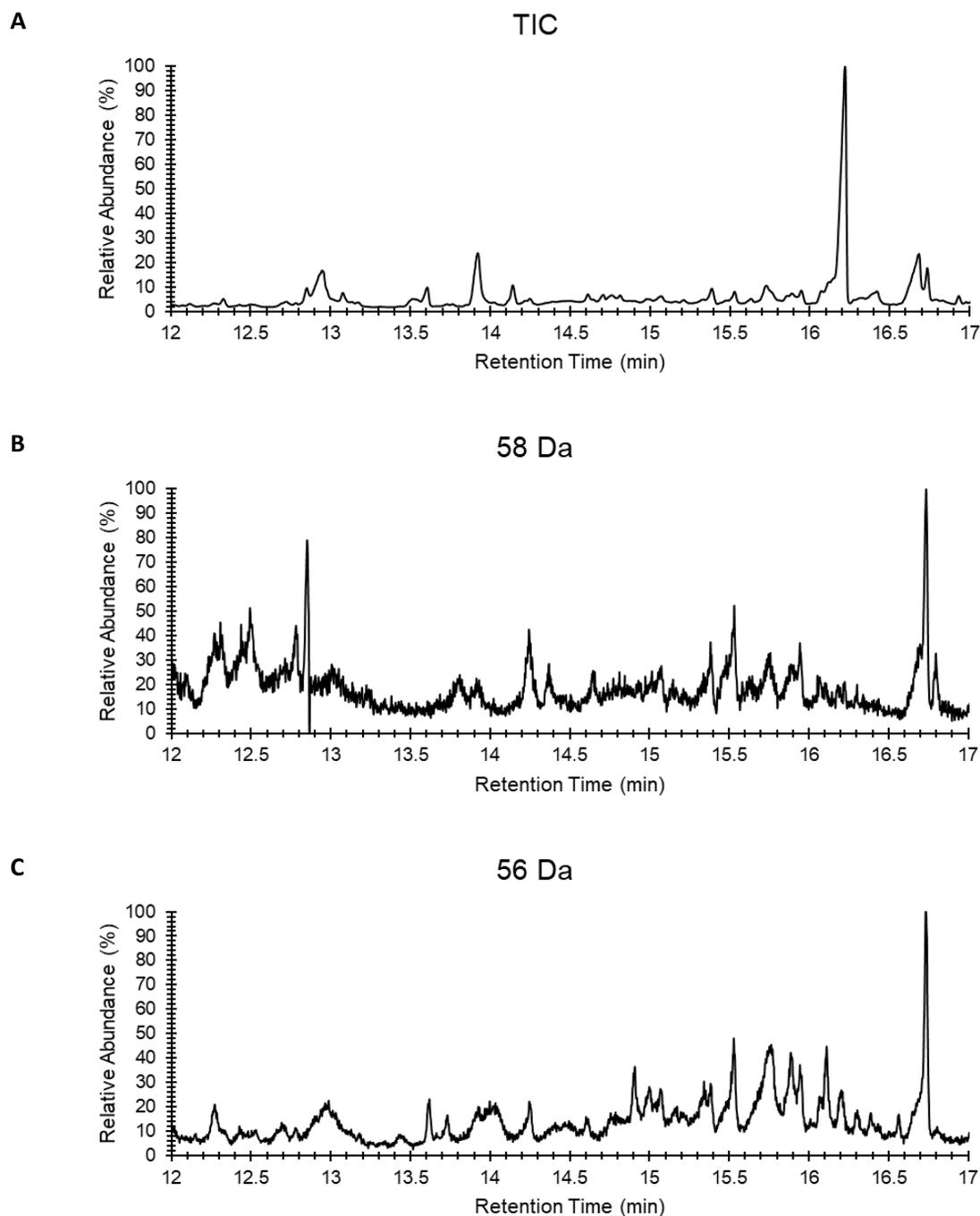
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Amph<sub>M11</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). The 58 Da fragment was used systematically for identification/quantification of the reduced lactones, as it is not generated by the ketolactones **5a/5b** and is therefore diagnostic. (C) Extracted ion chromatogram ( $m/z = 56$  Da). The 56 Da fragment allowed identification/quantification of the ketolactones **5a/5b**, although peaks corresponding to the reduced lactones were also observed.

Amph<sub>M11</sub>-B:

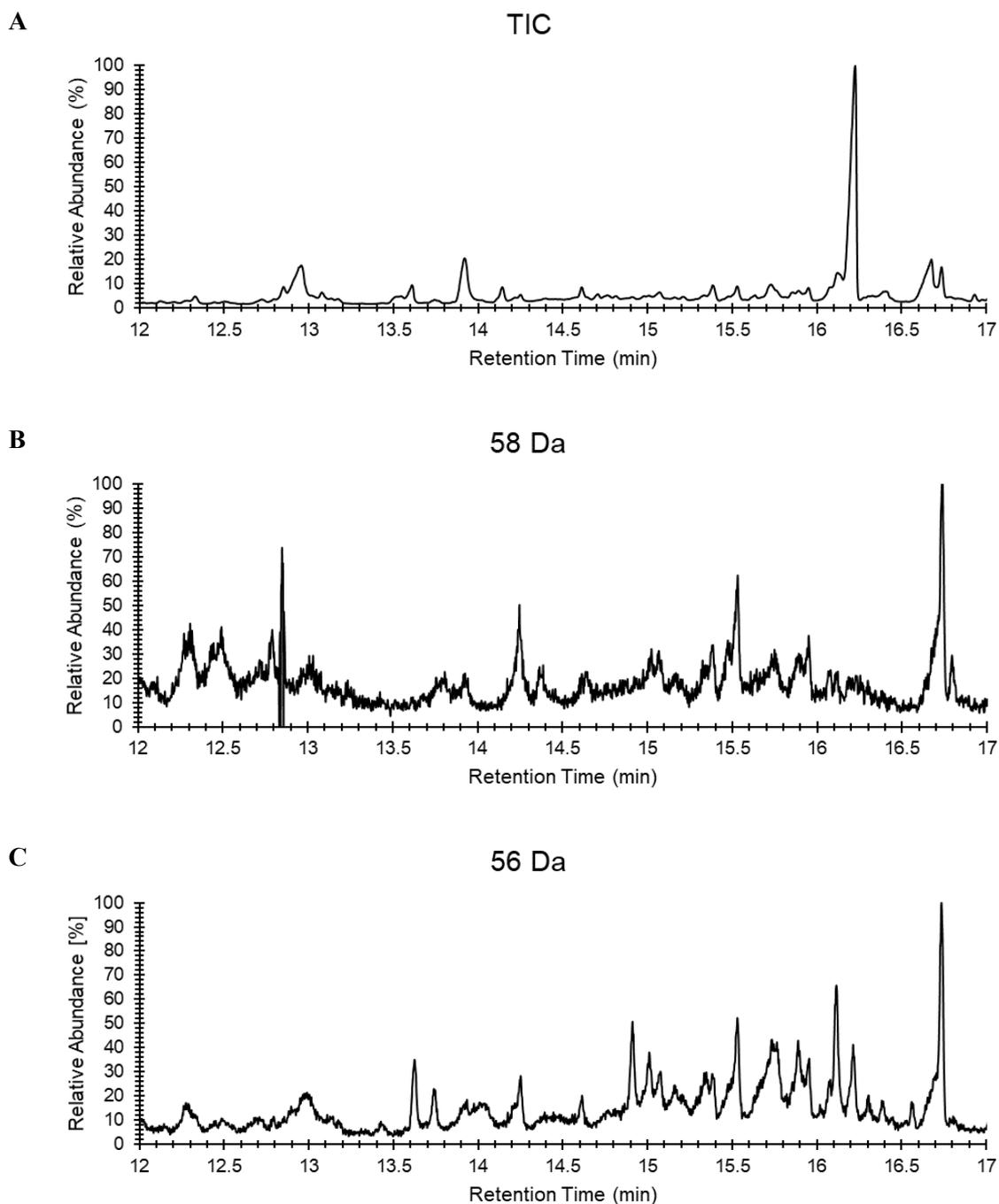
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Amph<sub>M11</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da)]. (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Amph<sub>M11</sub>-C:

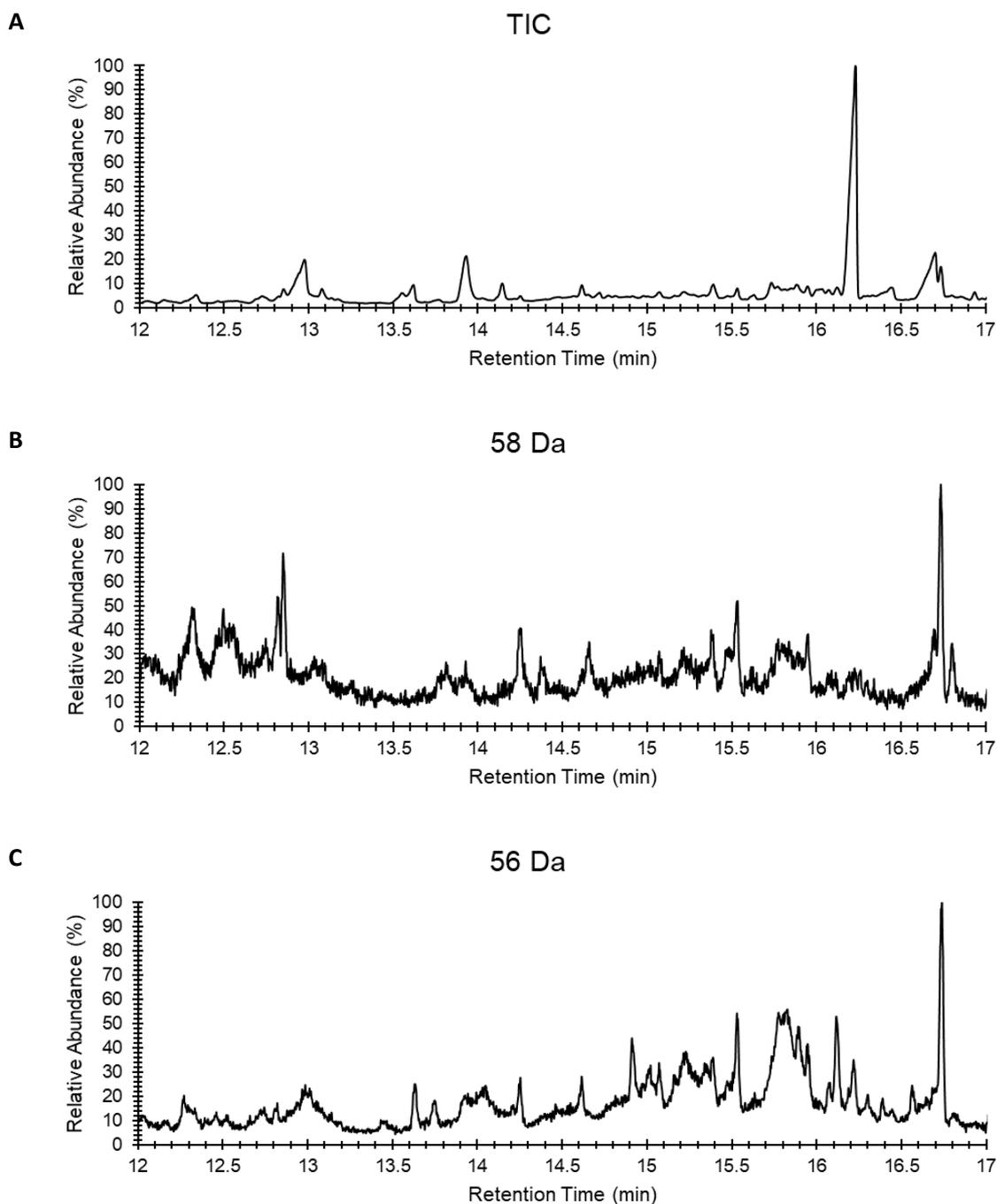
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Amph<sub>M11</sub>-C. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Amph<sub>M11</sub>-D:

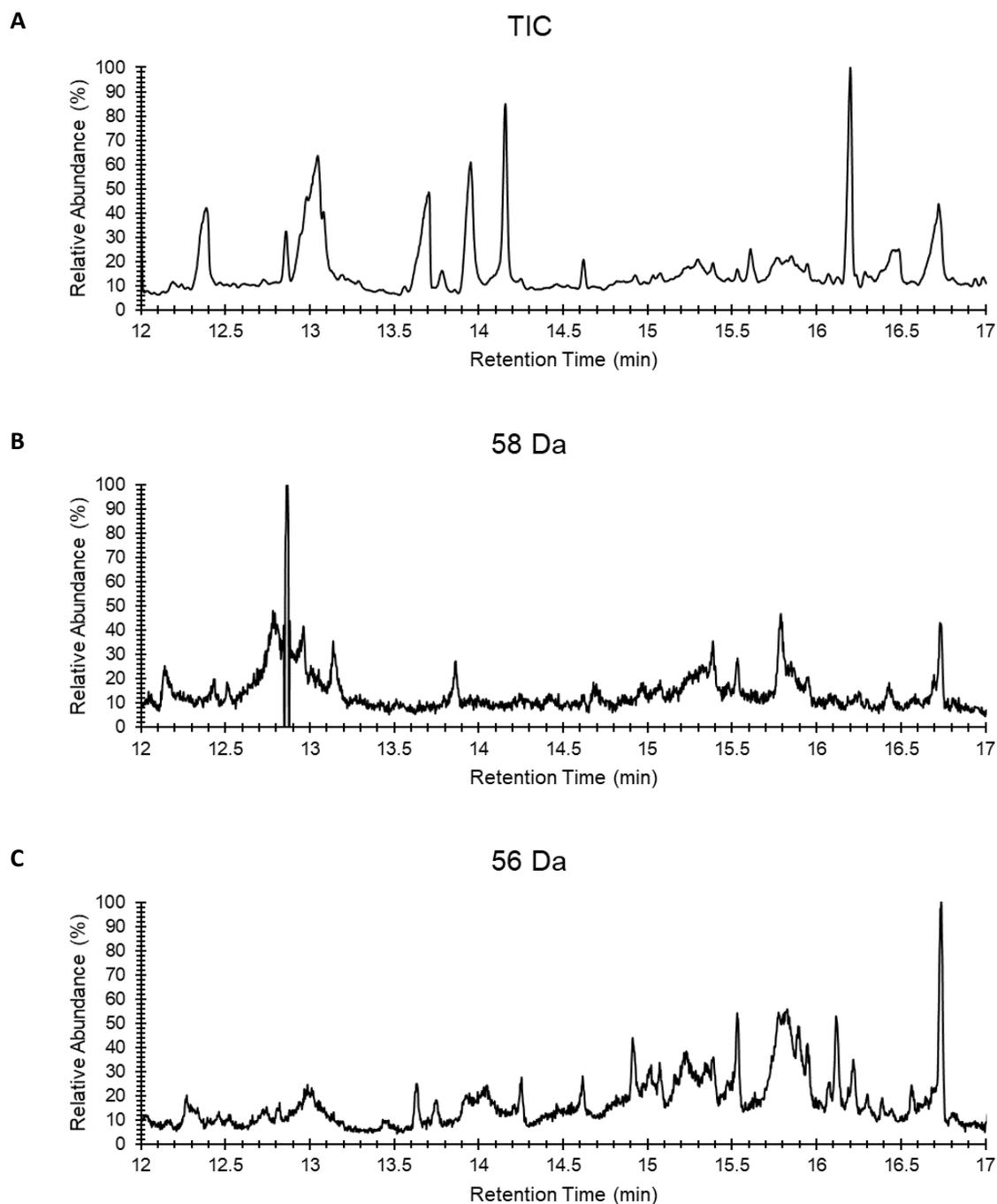
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Amph<sub>M11</sub>-D. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Bfm<sub>M4</sub>-A:

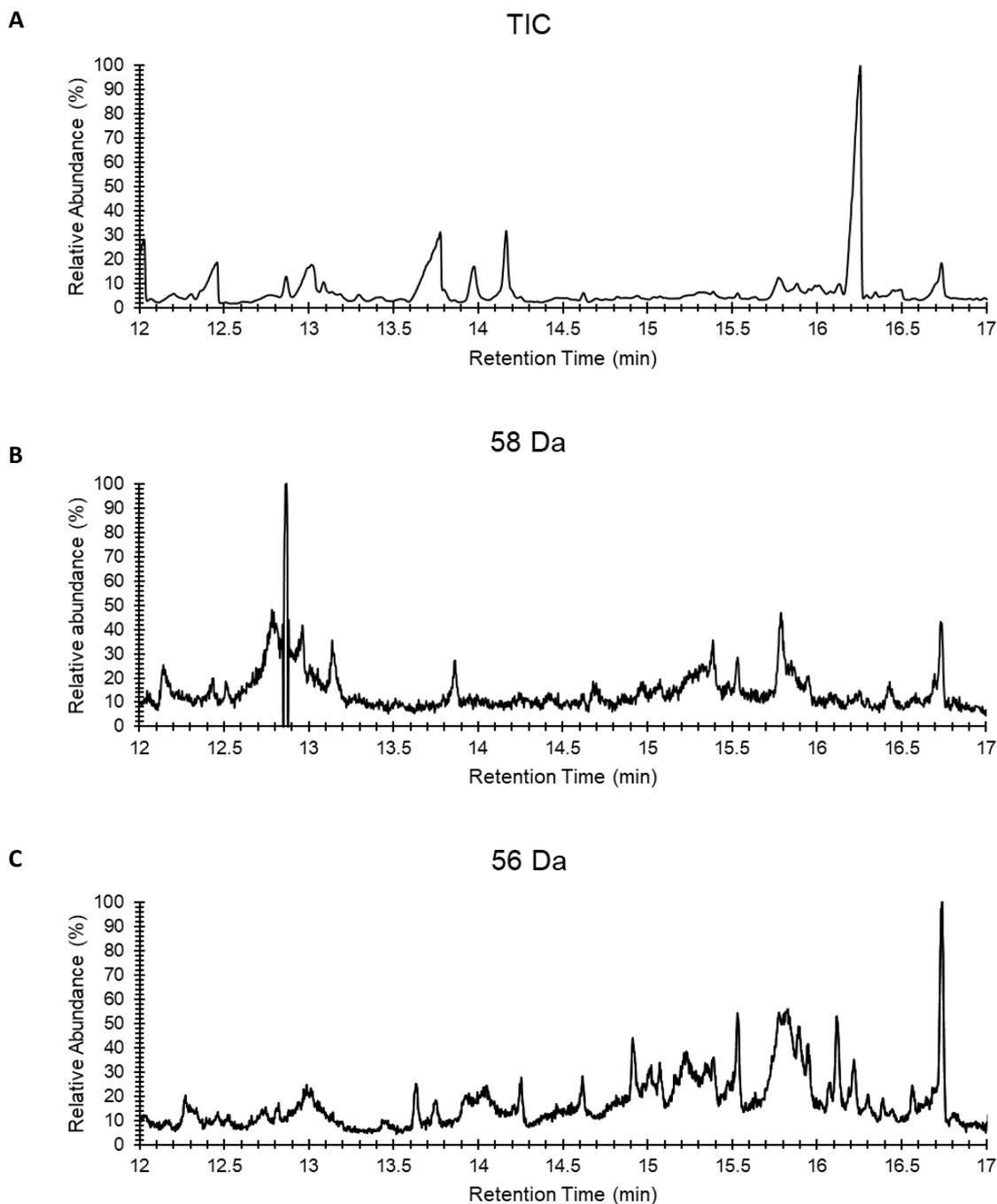
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Bfm<sub>M4</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Bfm<sub>M4</sub>-B:

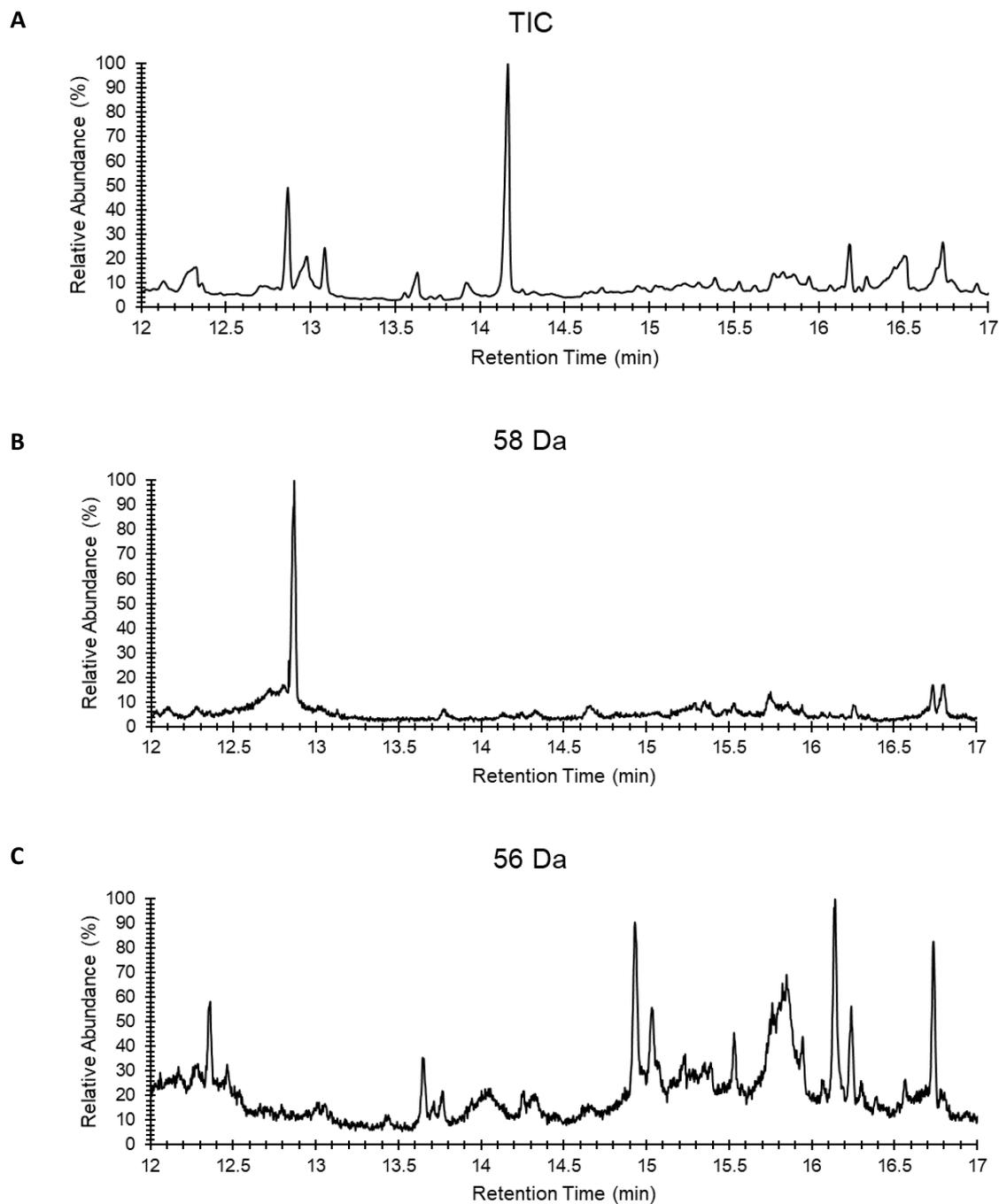
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Bfm<sub>M4</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Bfm<sub>M9</sub>-A:

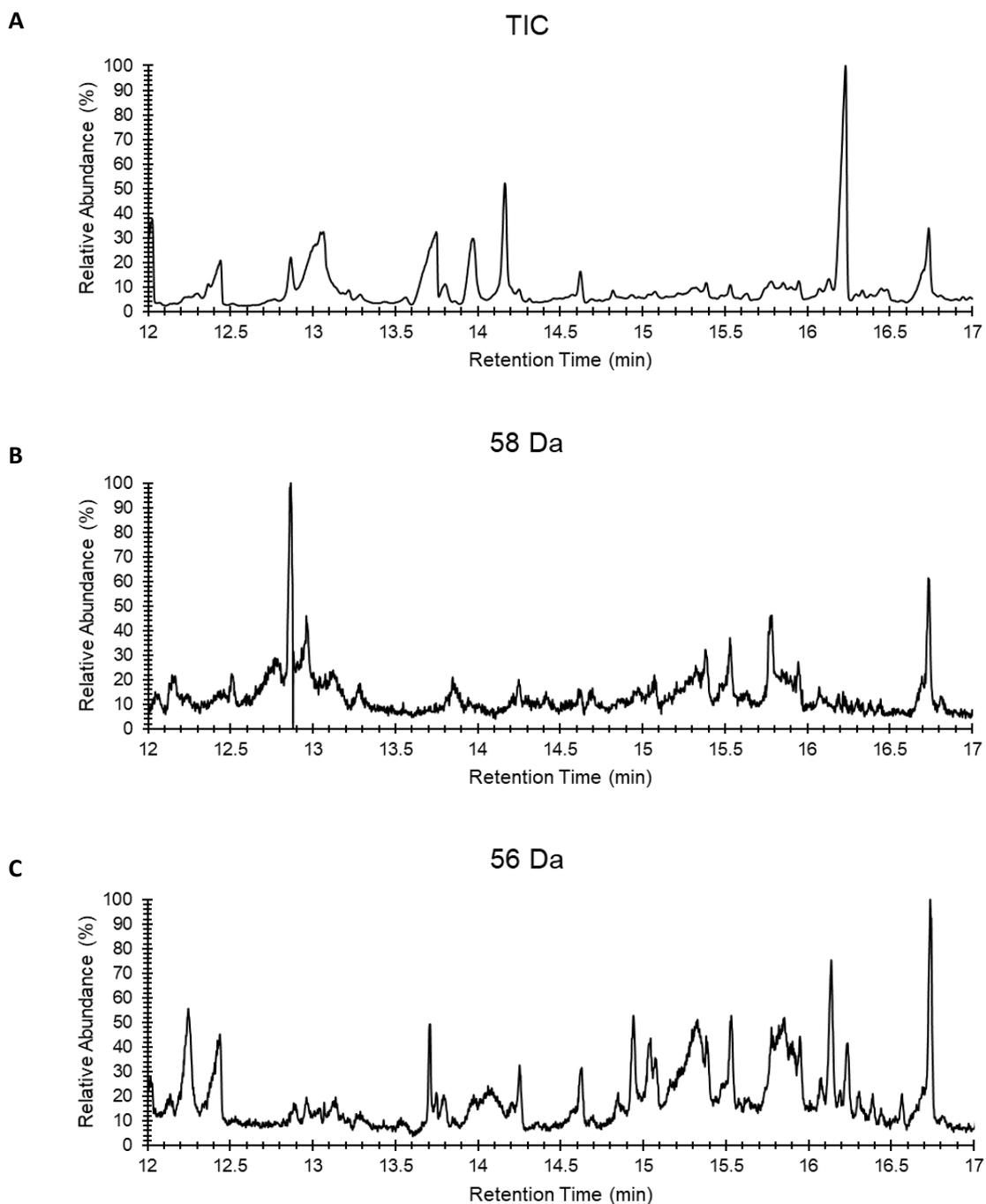
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Bfm<sub>M9</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Bfm<sub>M9</sub>-B:

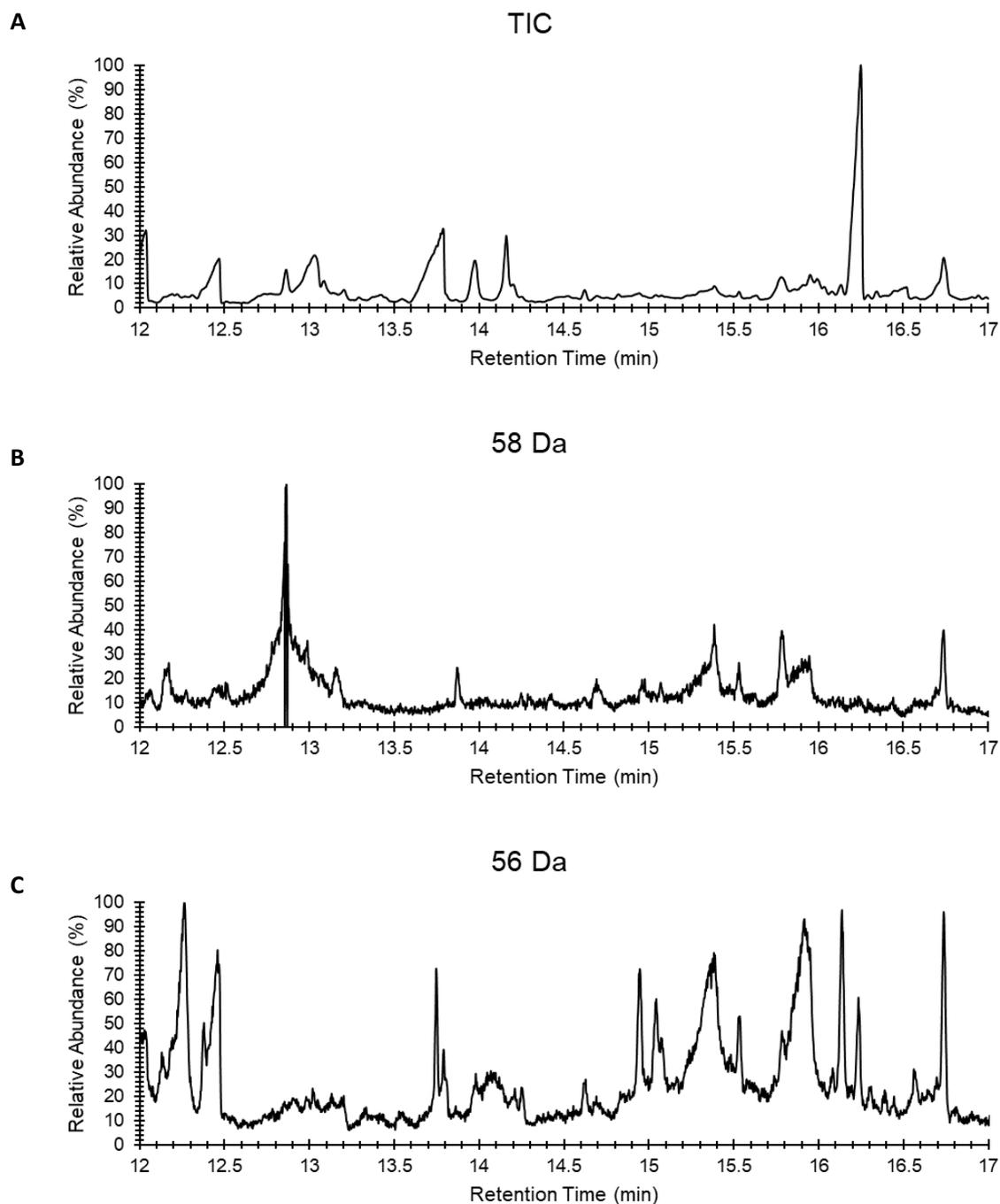
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Bfm<sub>M9</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Hba<sub>M4</sub>-A:

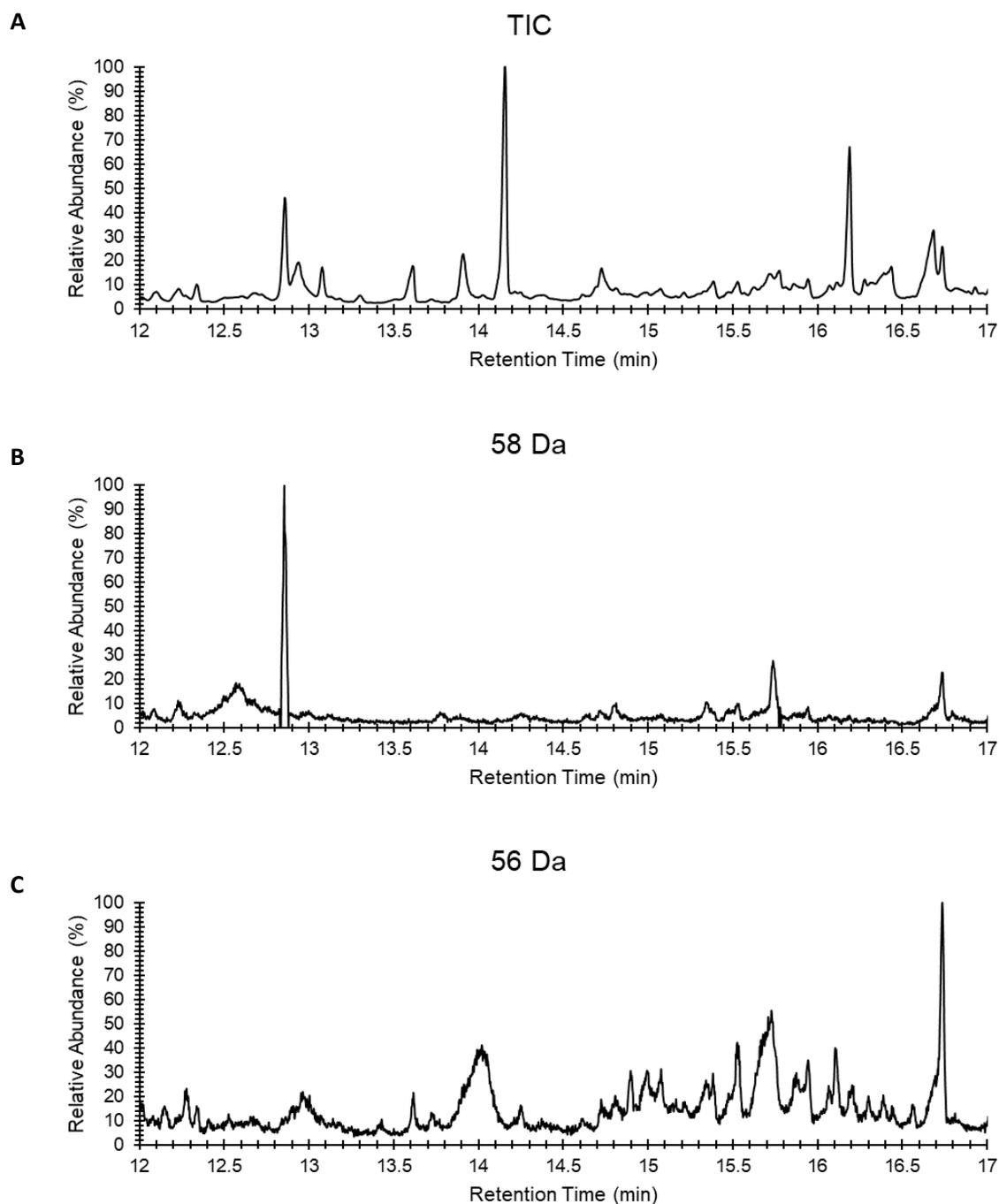
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Hba<sub>M4</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Hba<sub>M4</sub>-B:

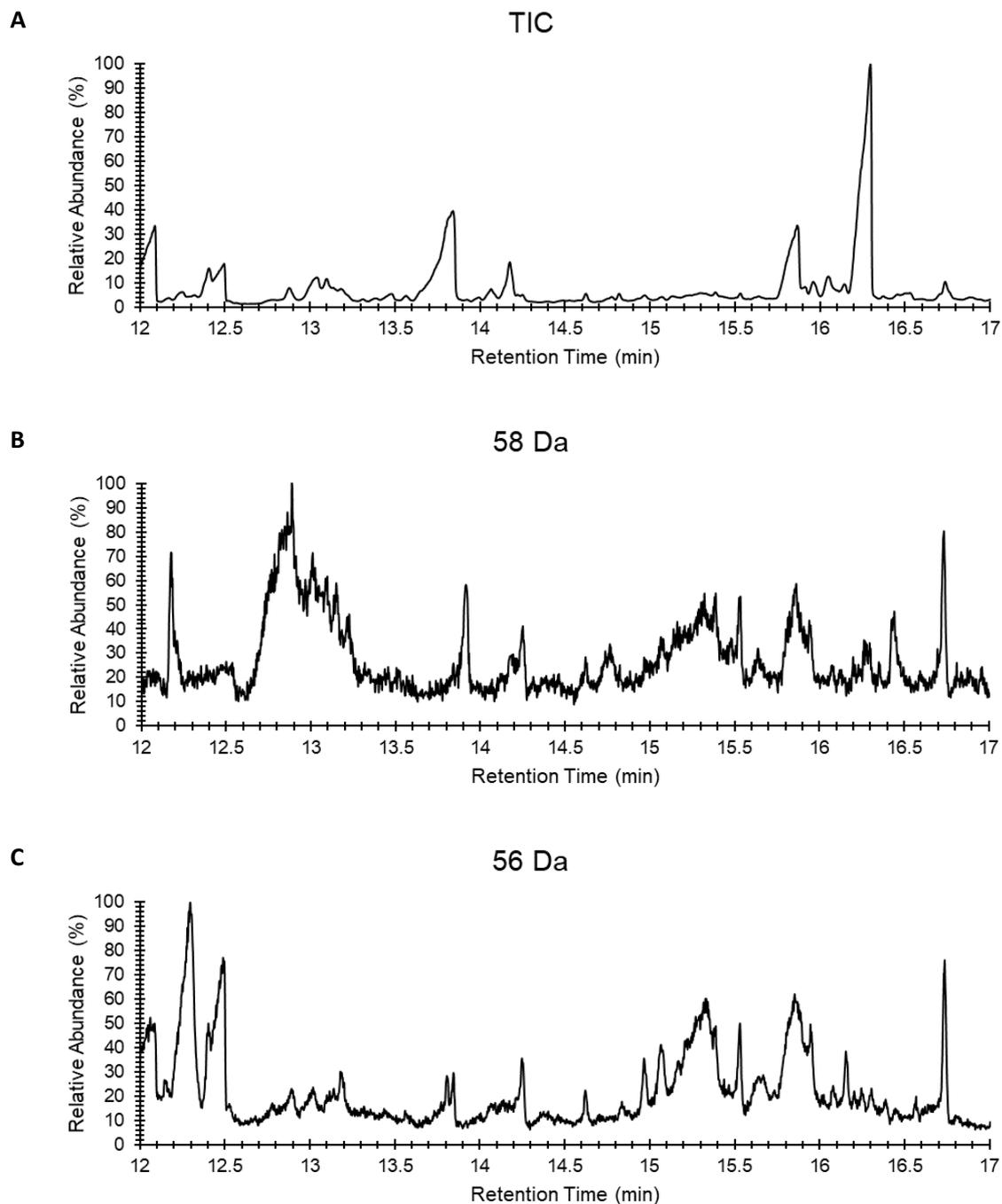
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Hba<sub>M4</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Hba<sub>M9</sub>-A:

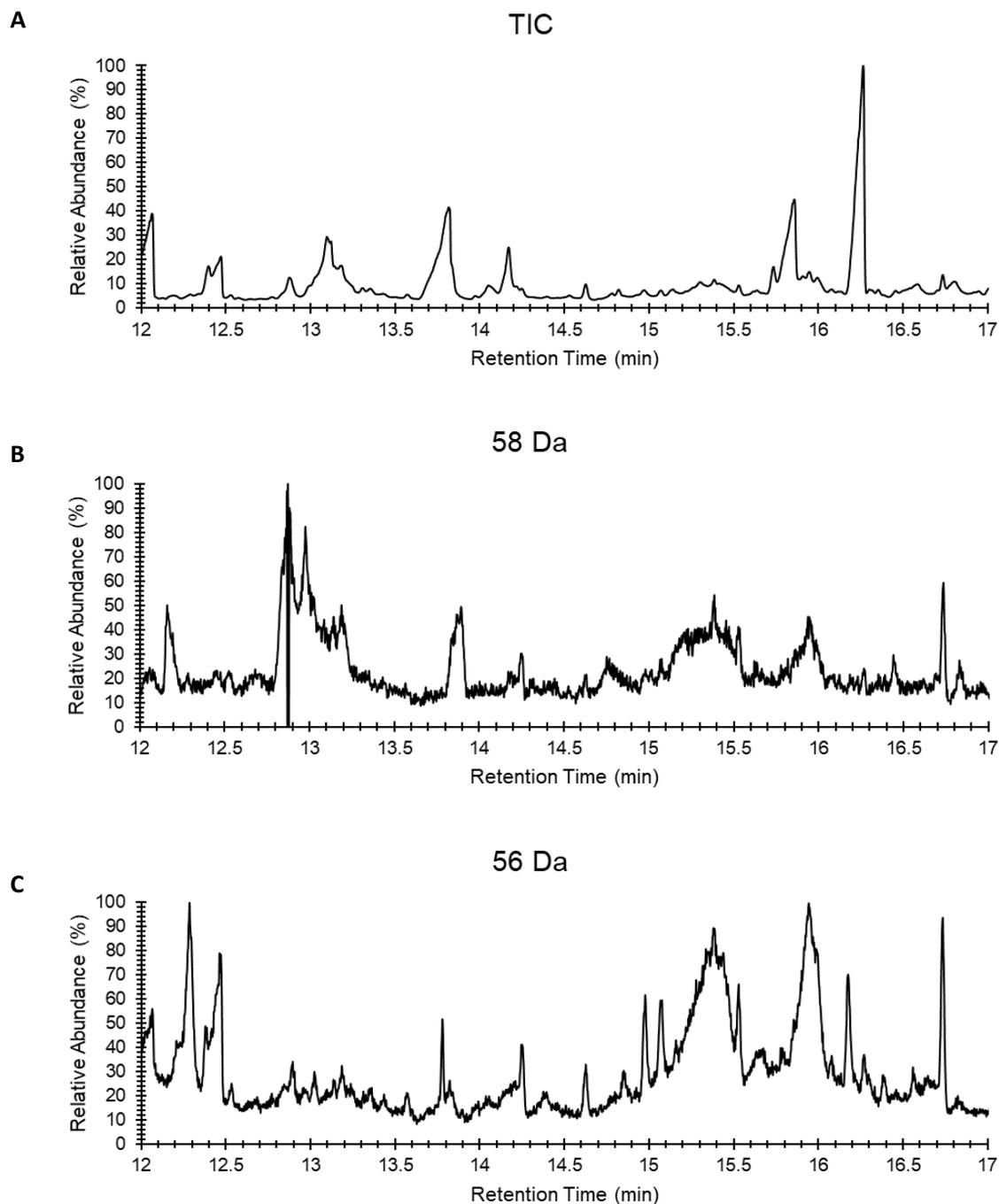
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Hba<sub>M9</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Hba<sub>M9</sub>-B:

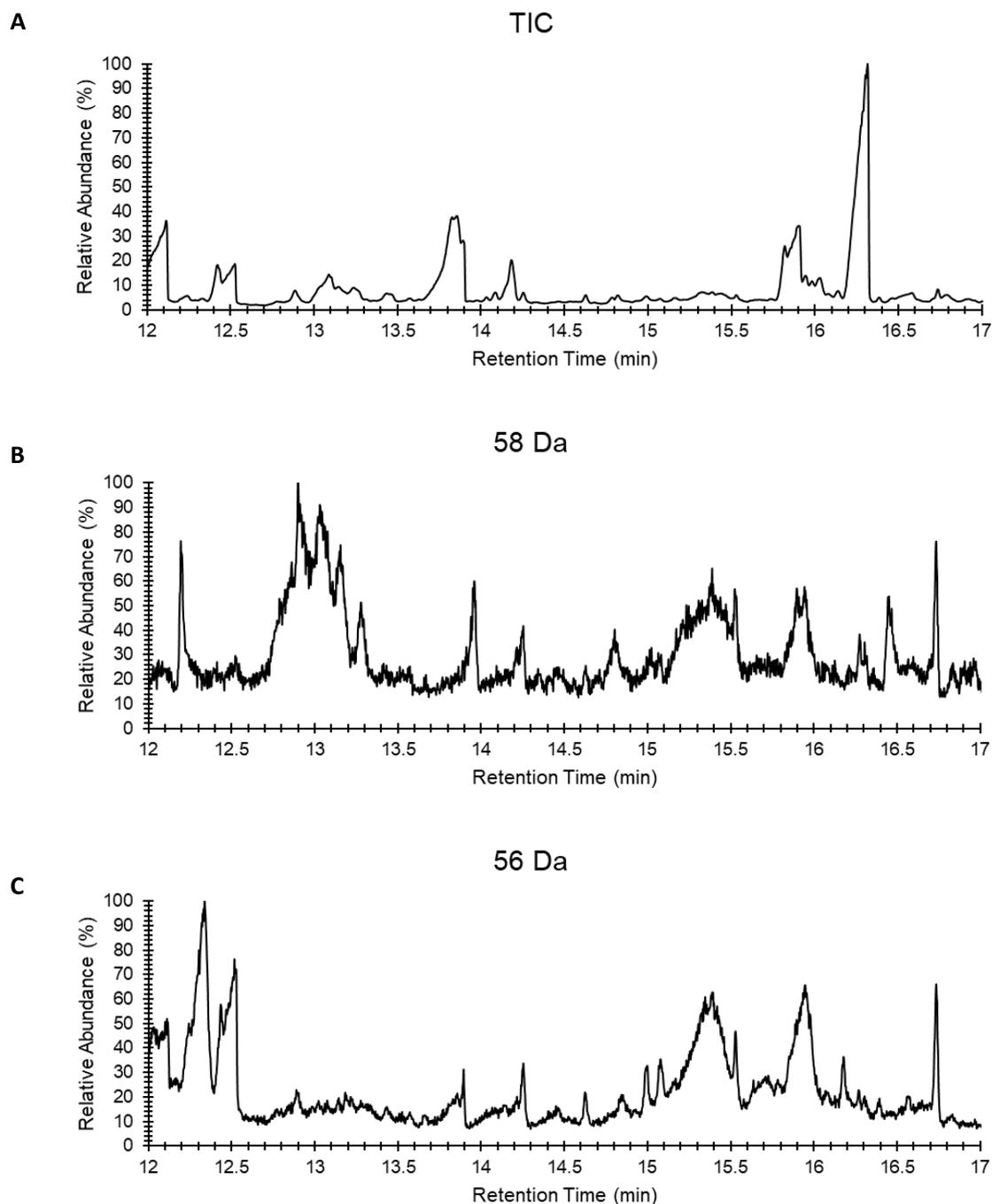
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Hba<sub>M9</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Las<sub>M7</sub>-A:

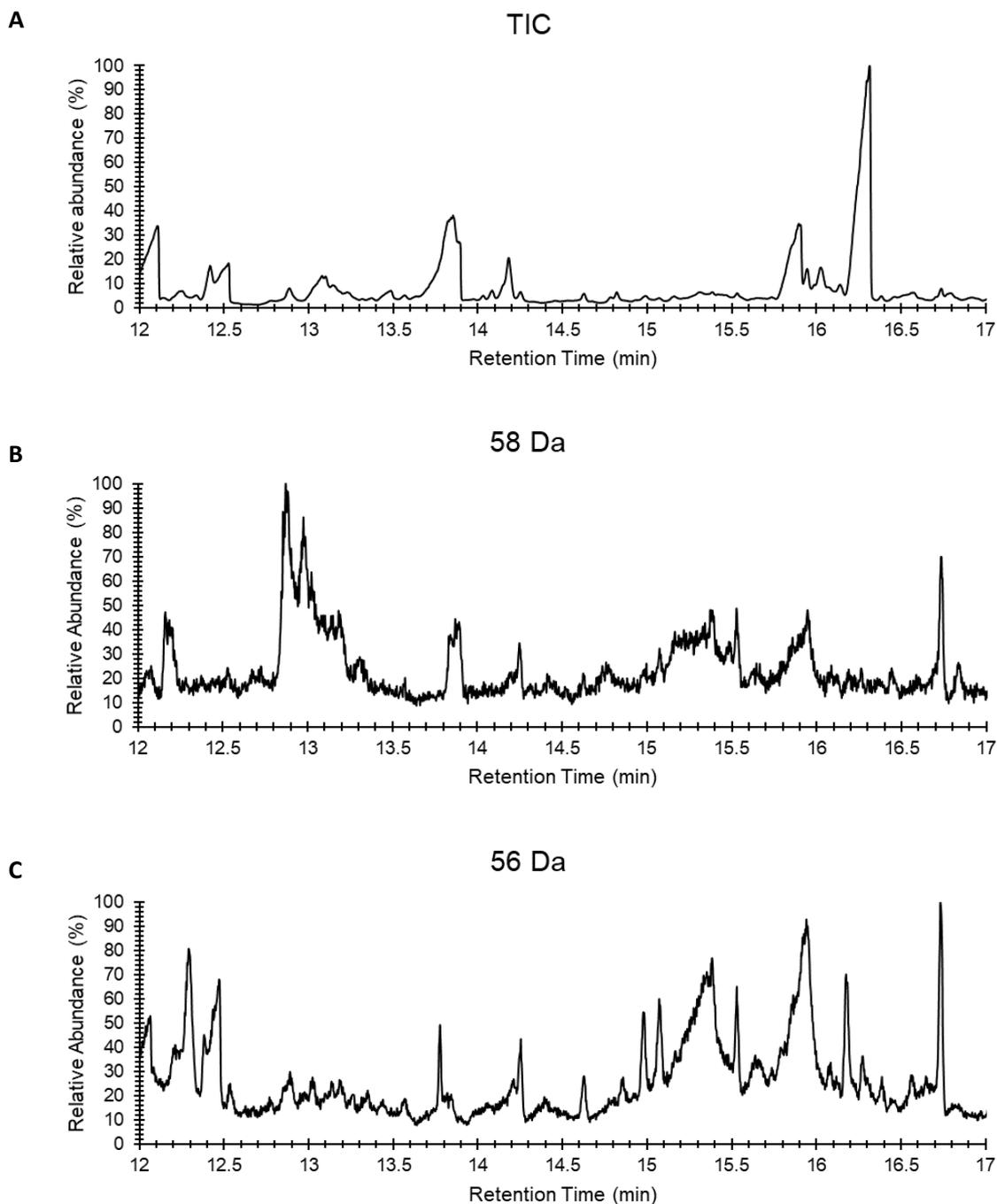
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Las<sub>M7</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Las<sub>M7</sub>-B:

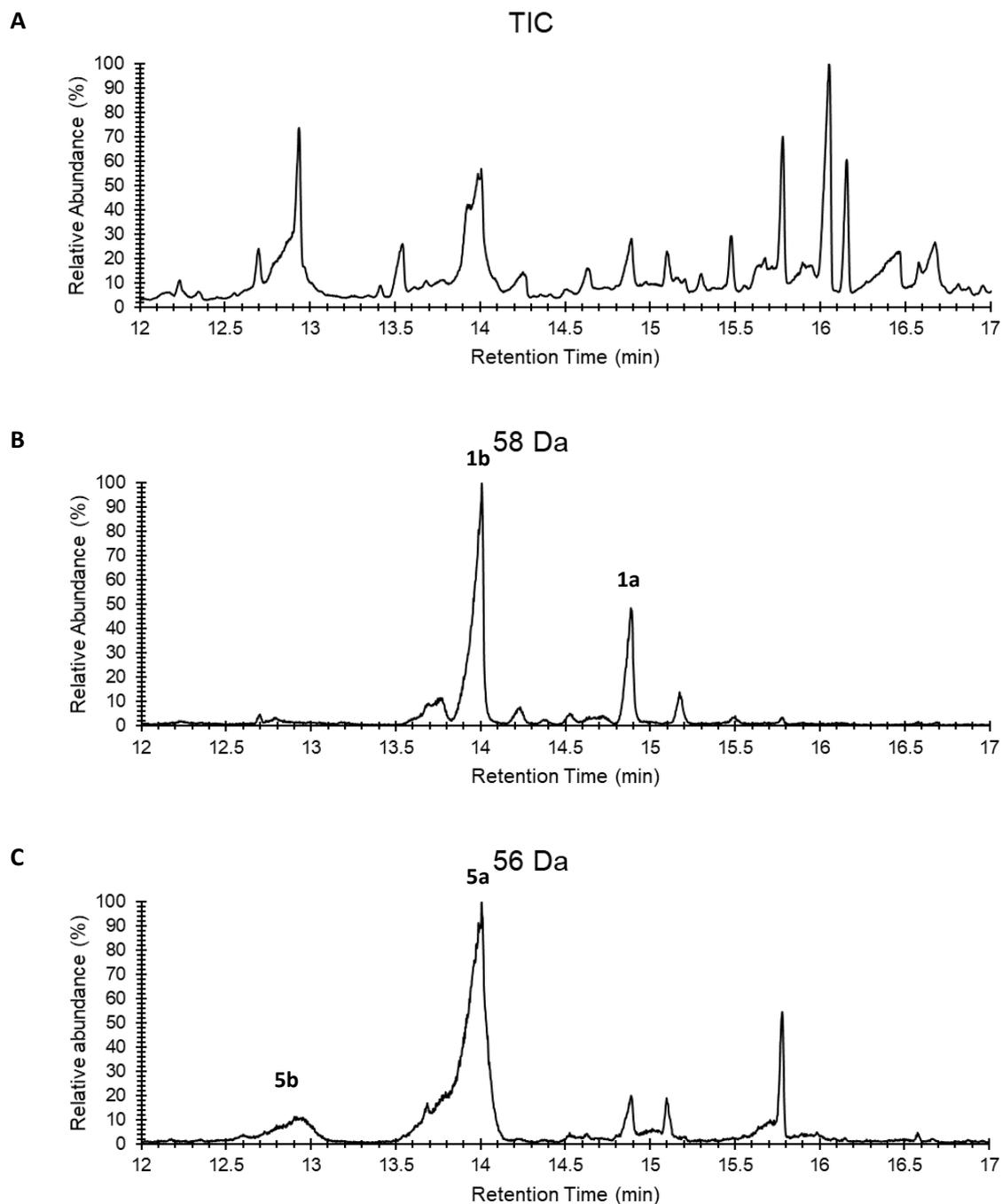
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Las<sub>M7</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Las<sub>M7</sub>-C:

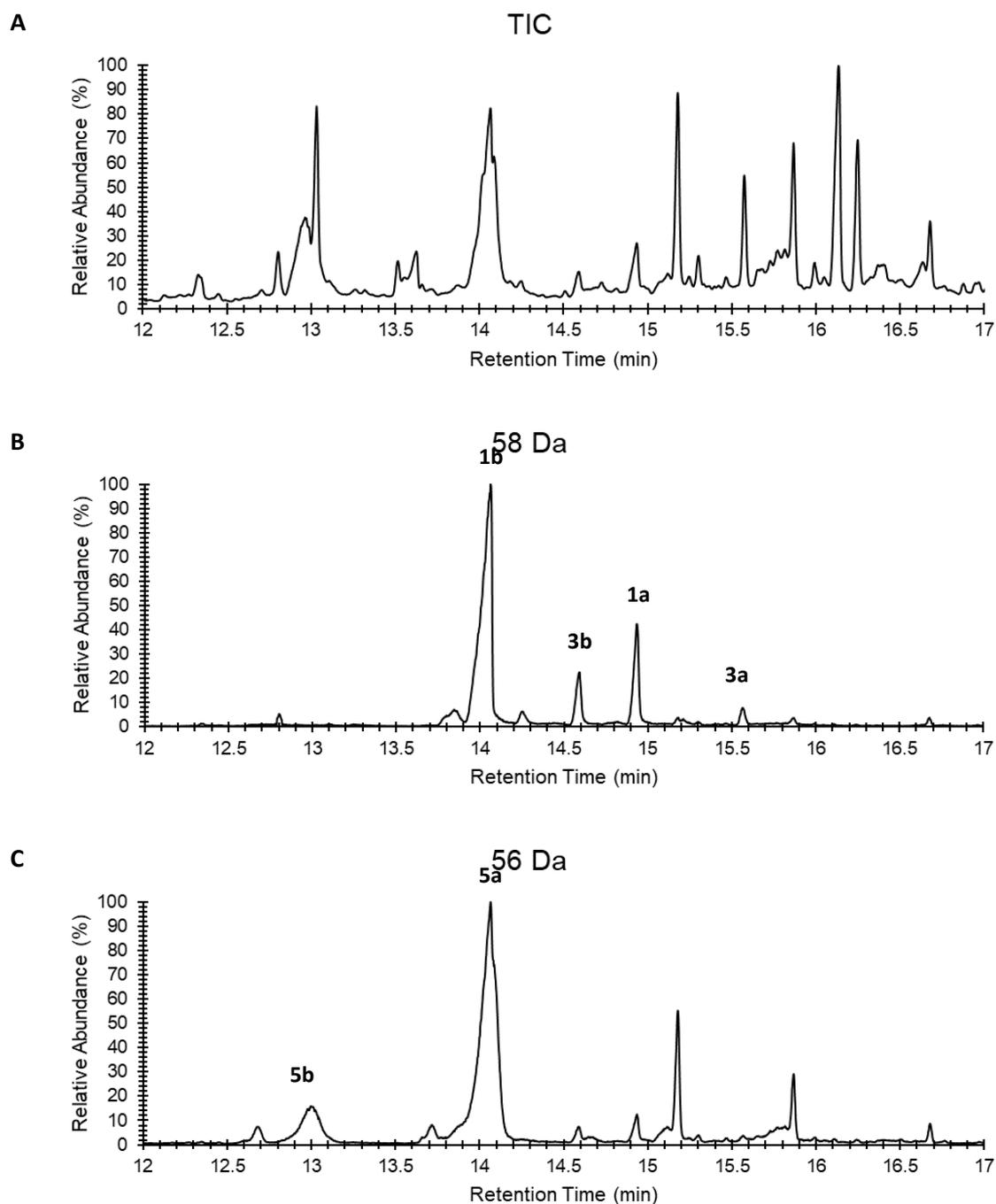
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Las<sub>M7</sub>-C. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Las<sub>M7</sub>-D:

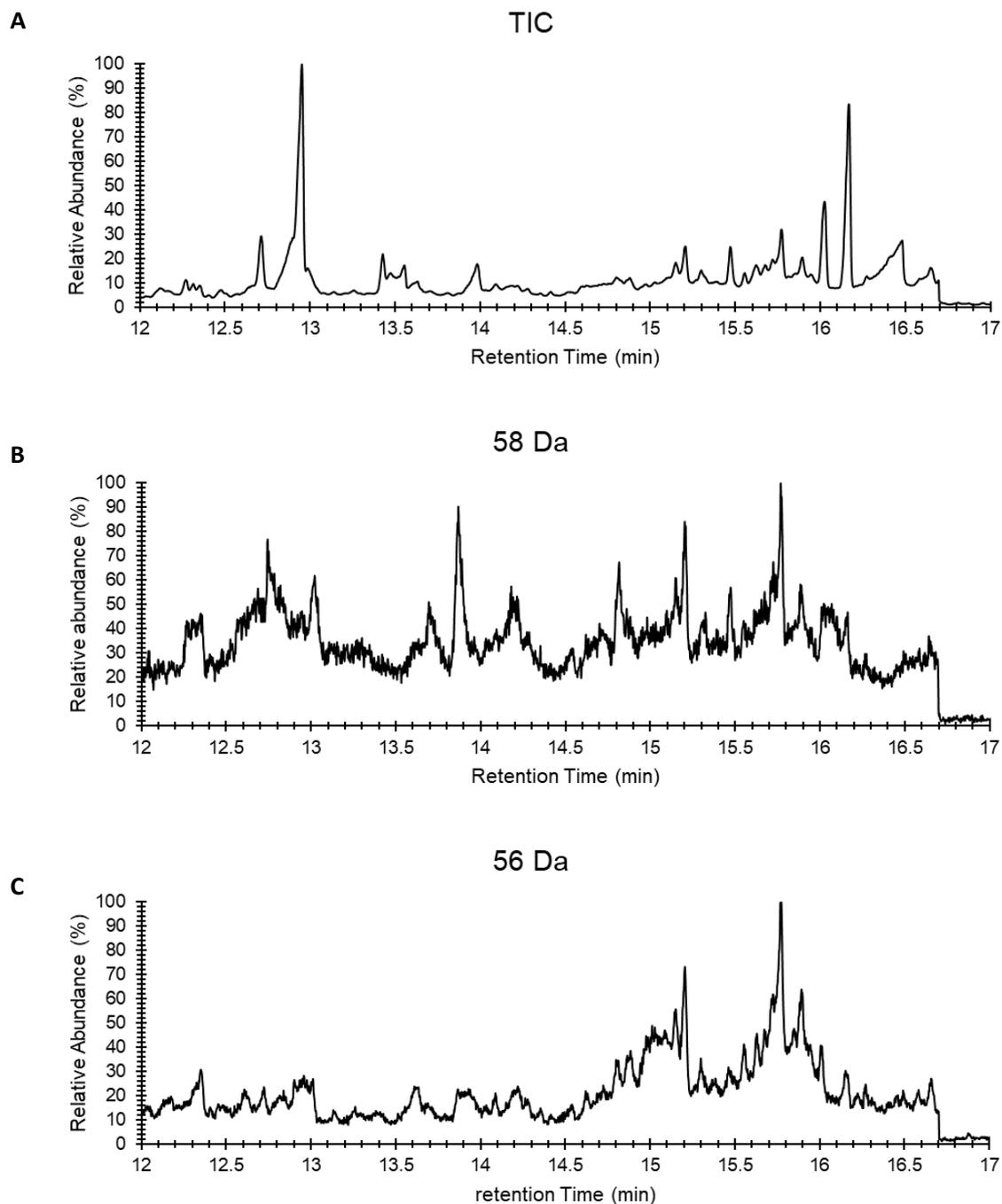
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Las<sub>M7</sub>-D. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Lkm<sub>M1</sub>-A:

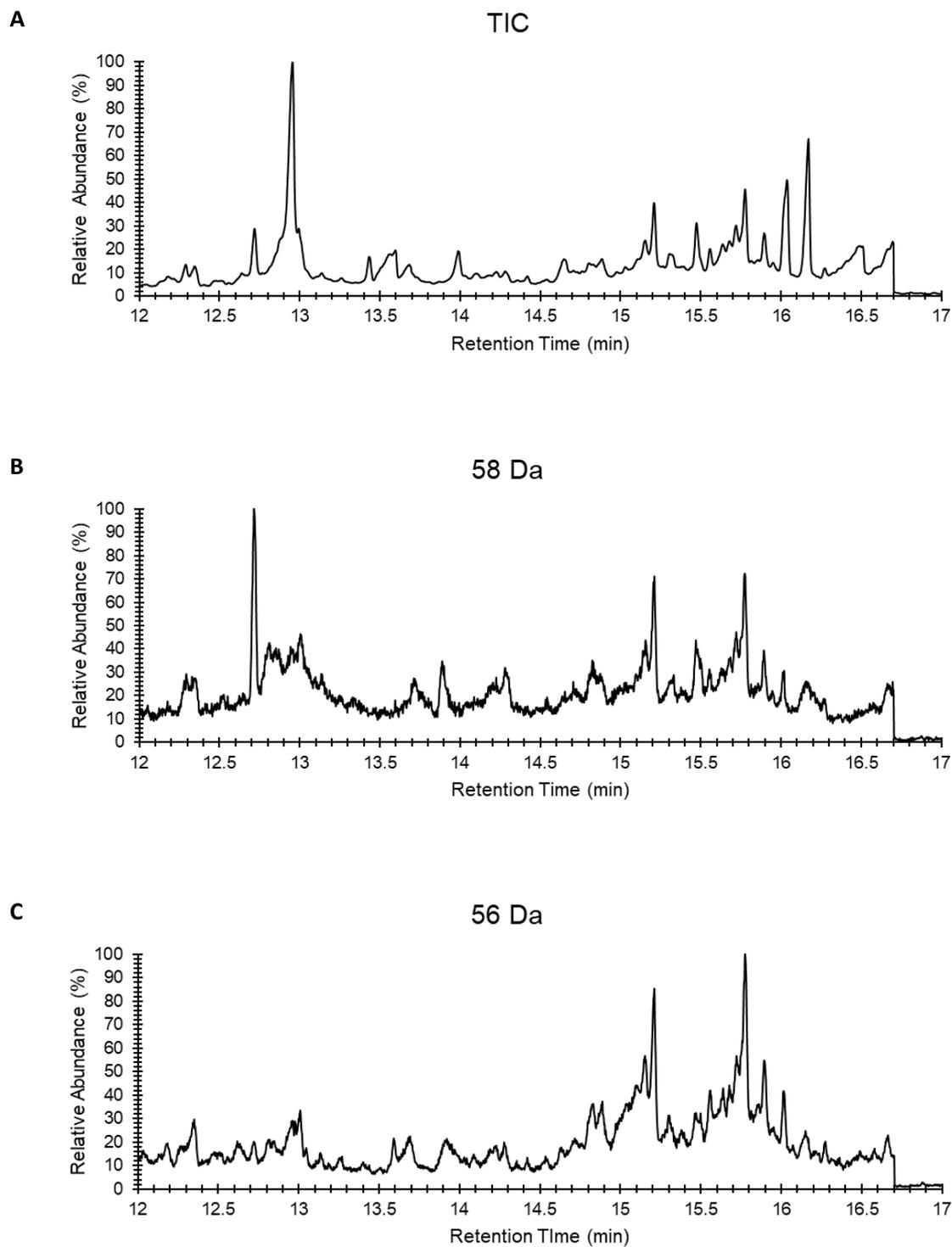
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Lkm<sub>M1</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Lkm<sub>M1</sub>-B:

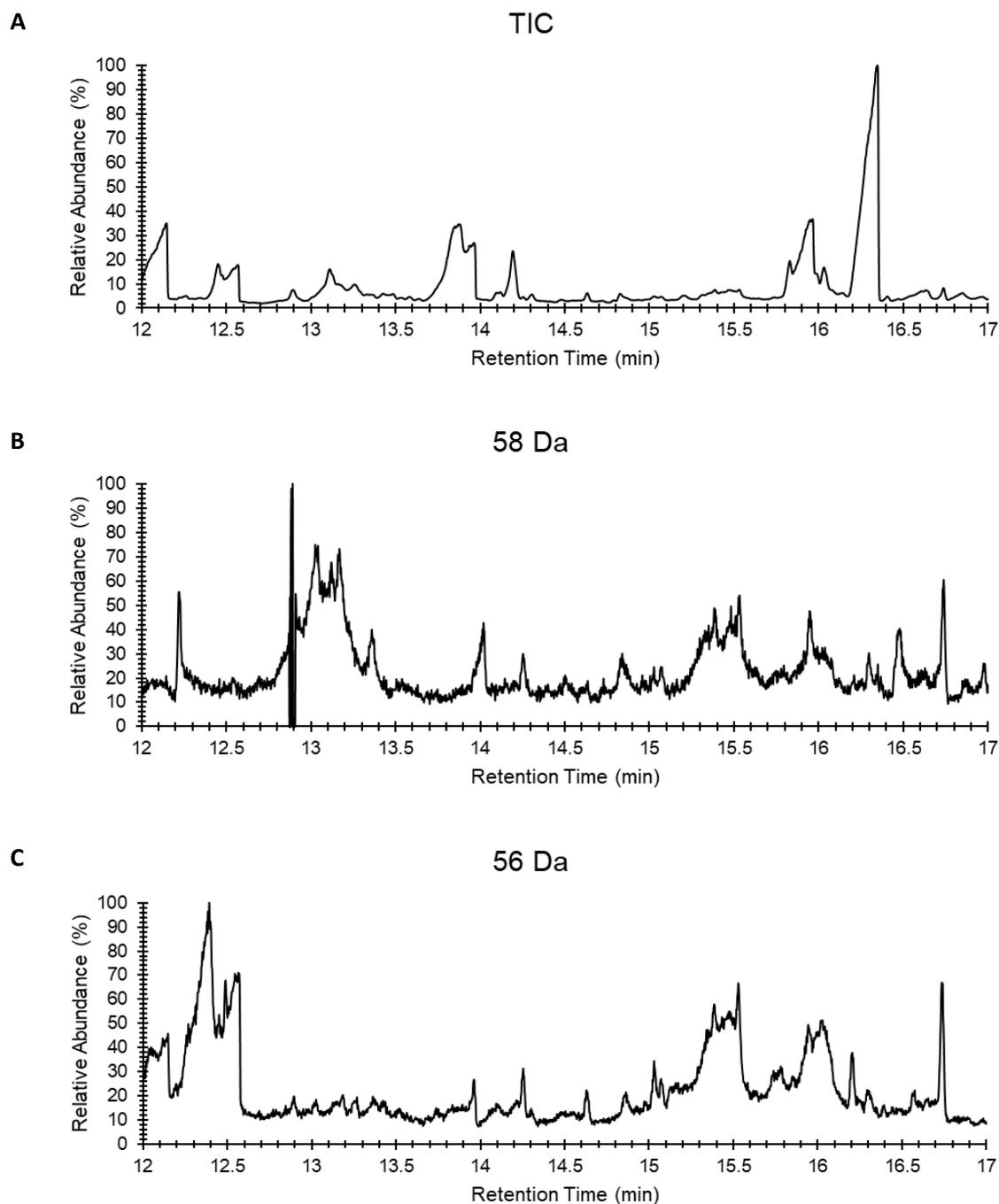
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Lkm<sub>M1</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Lkm<sub>M1</sub>-C:

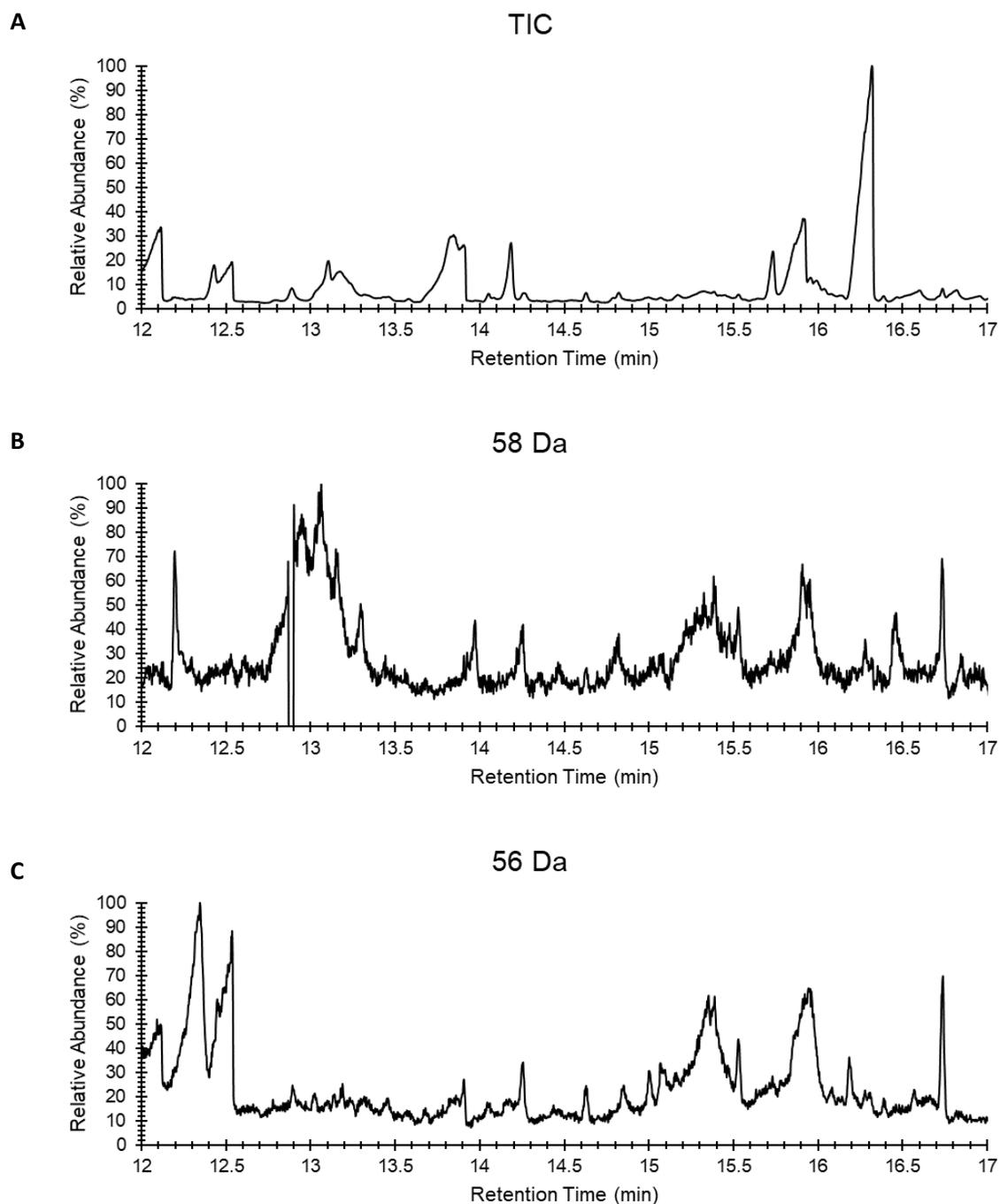
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Lkm<sub>M1</sub>-C. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Lkm<sub>M1</sub>-D:

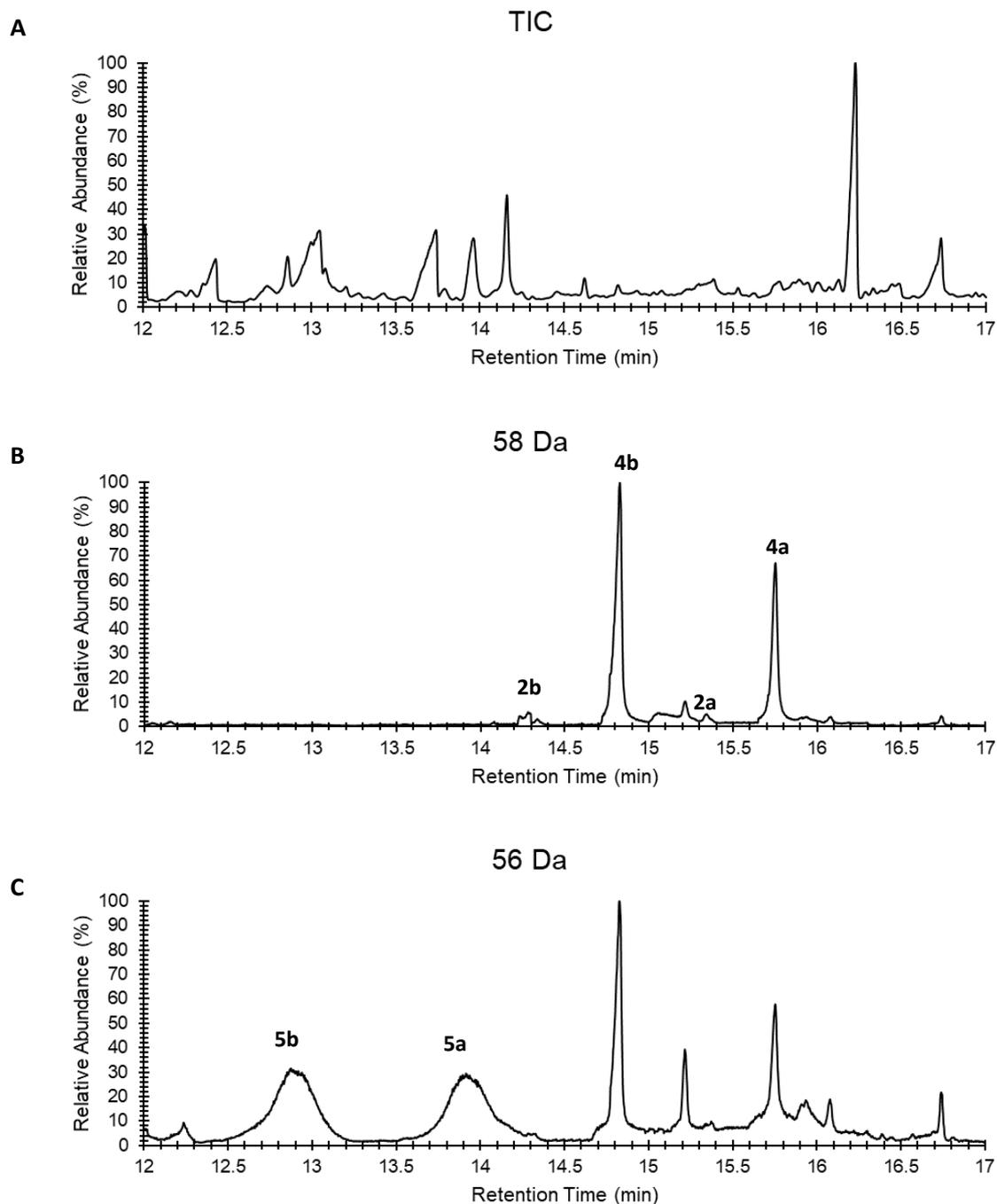
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Lkm<sub>M1</sub>-D. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Med<sub>M7a</sub>-A:

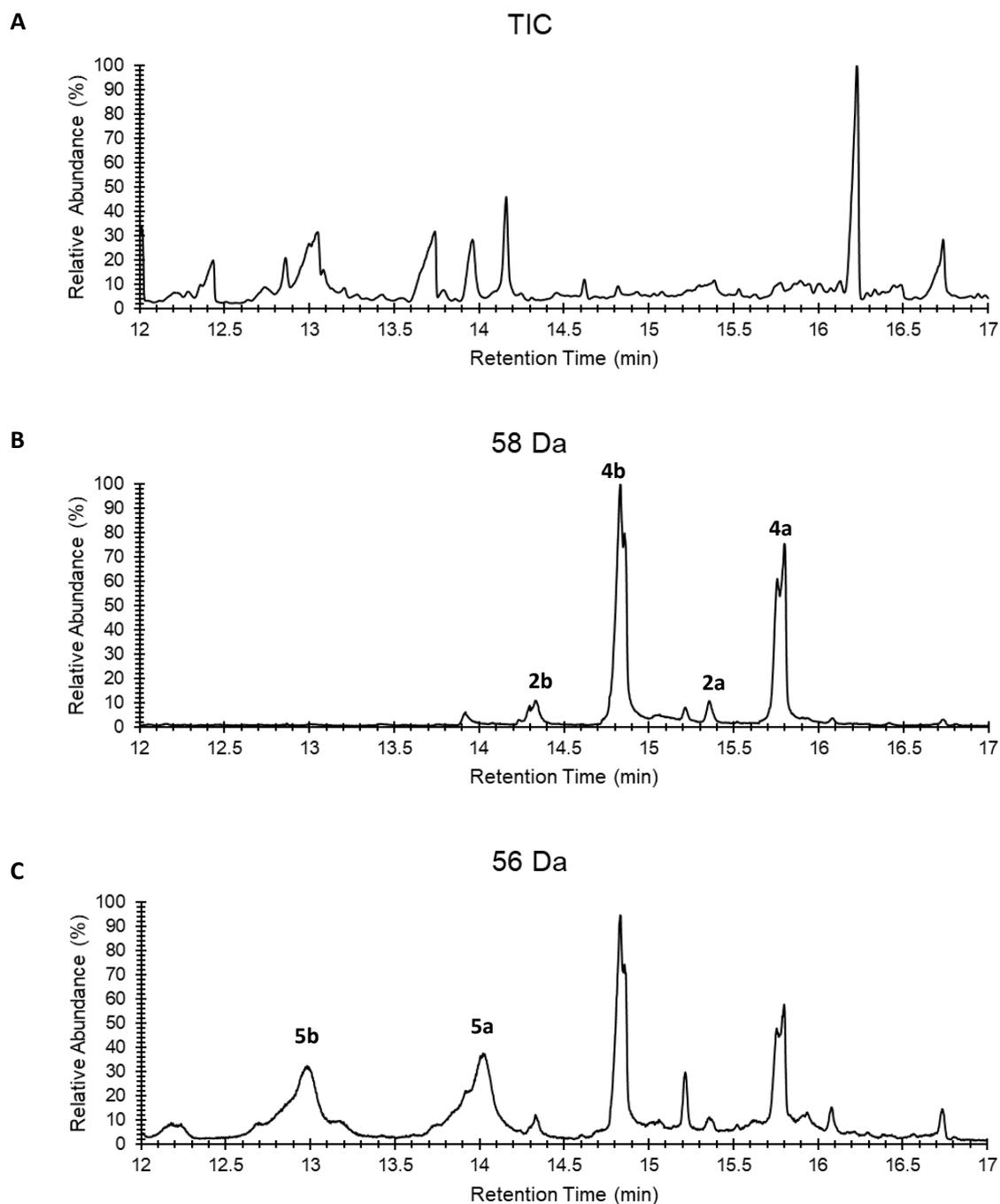
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3-Med<sub>M7a</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Med<sub>M7a</sub>-B:

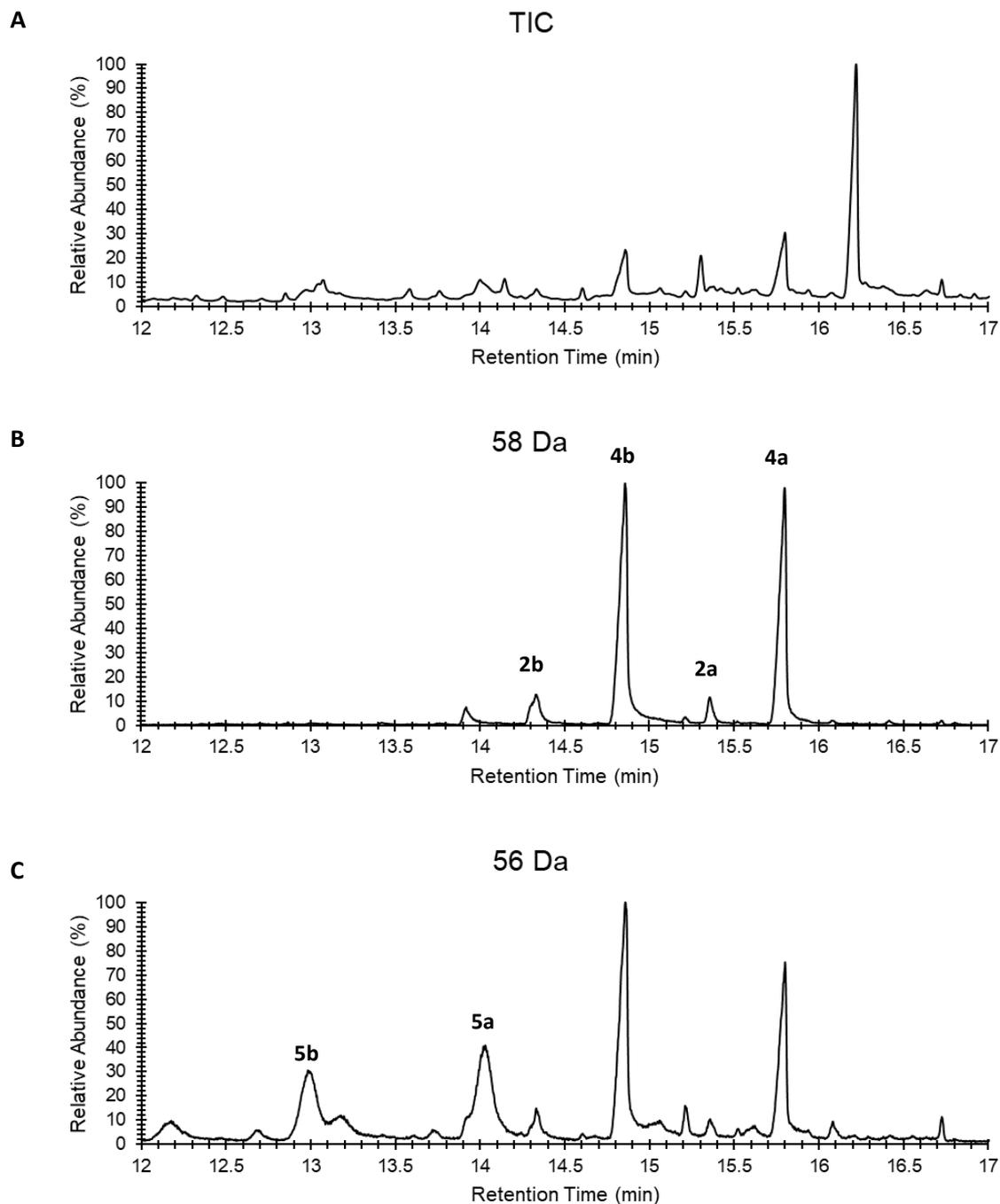
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Med<sub>M7a</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Nys<sub>M1</sub>-A:

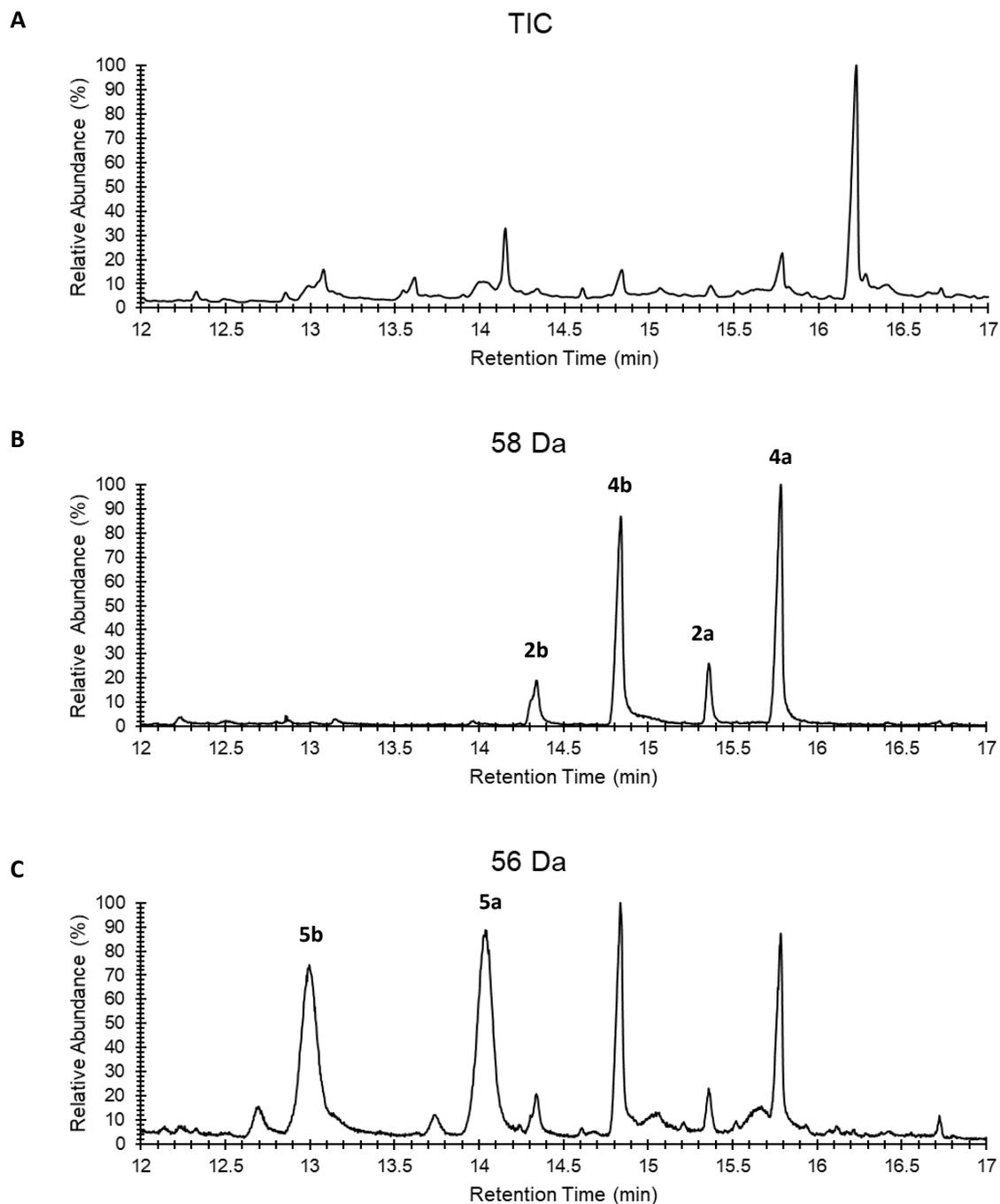
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Nys<sub>M1</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Nys<sub>M1</sub>-B:

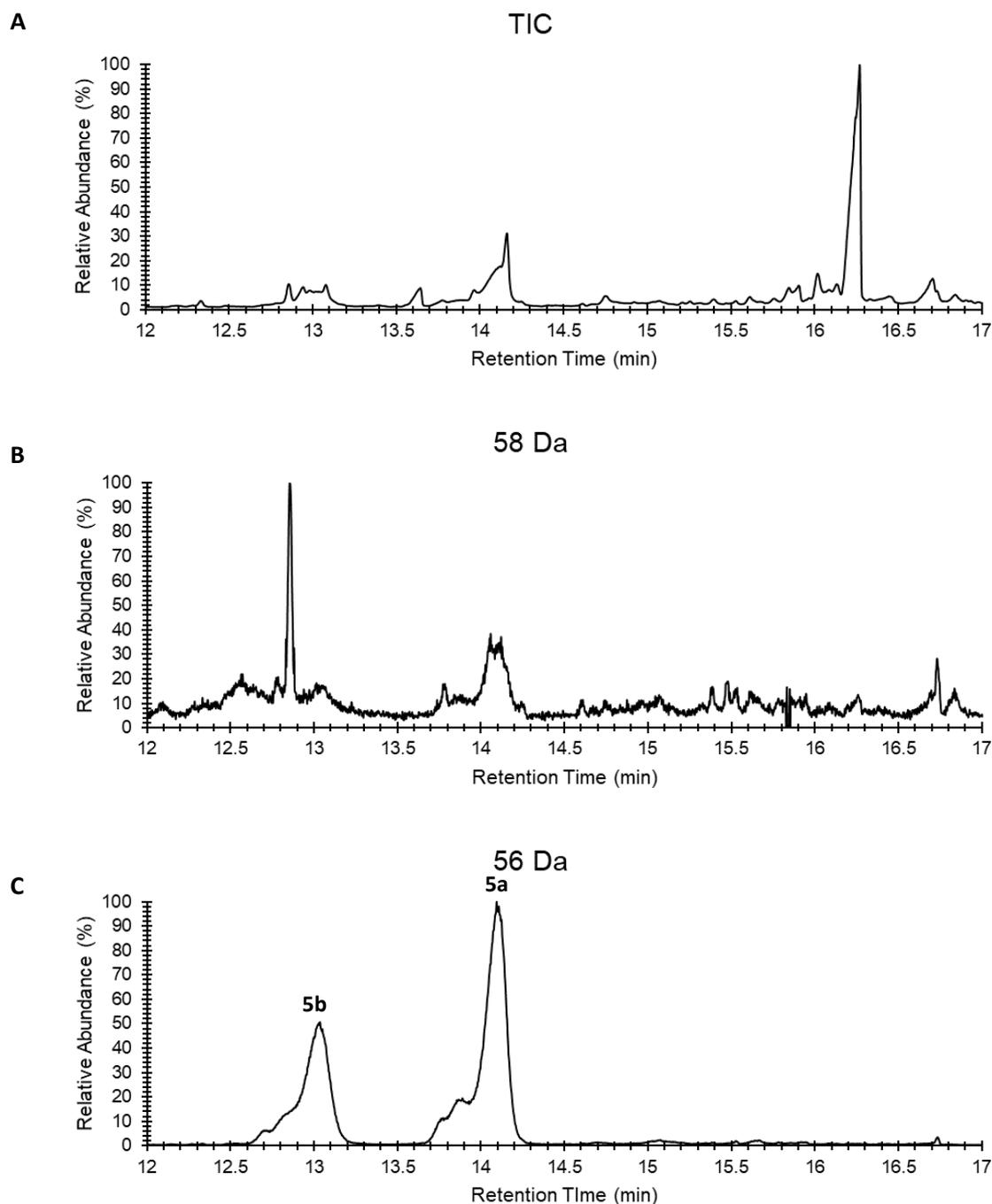
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Nys<sub>M1</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Nys<sub>M11</sub>-A:

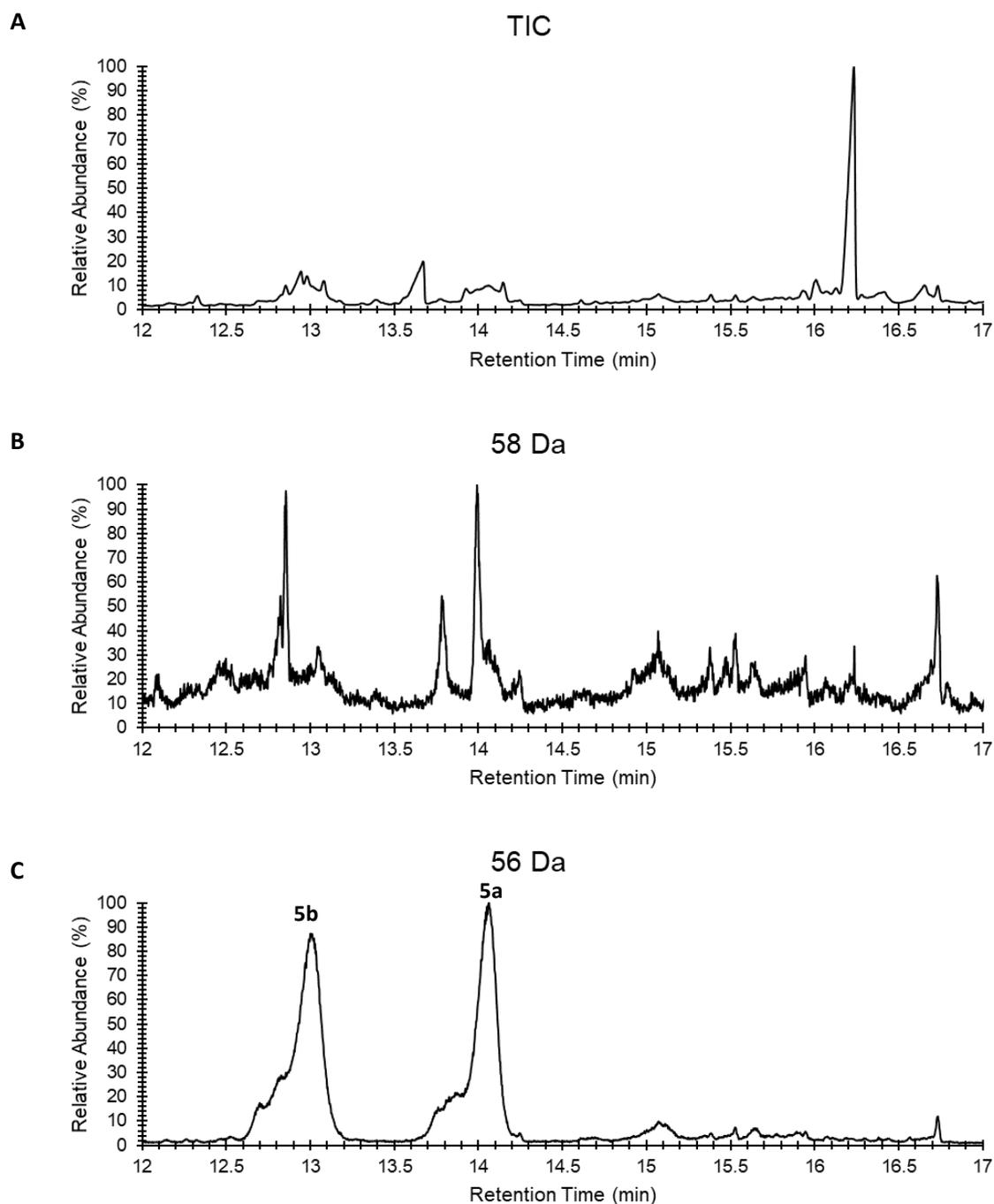
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Nys<sub>M11</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Nys<sub>M11</sub>-B:

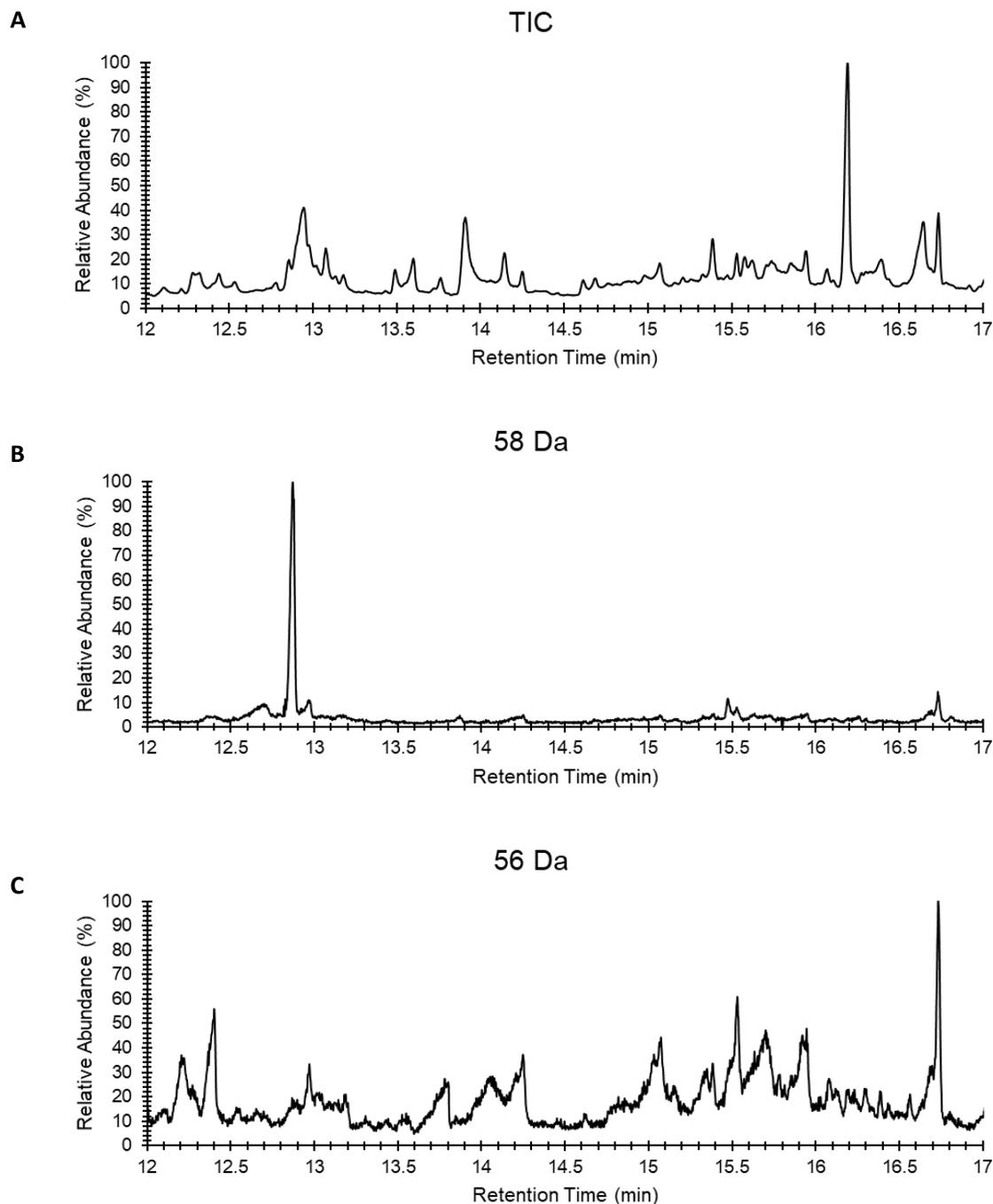
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Nys<sub>M11</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Olm<sub>M5</sub>-A:

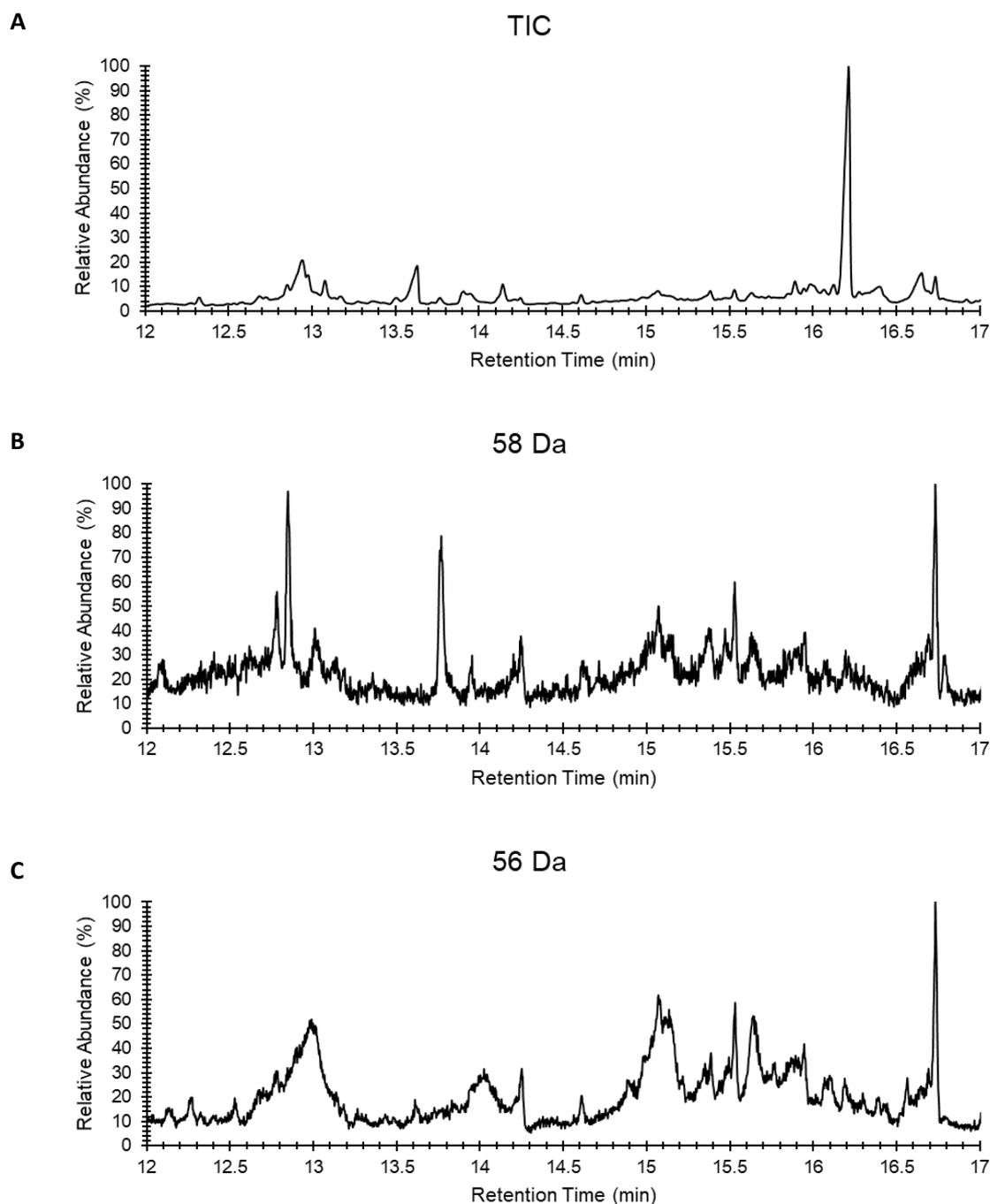
GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Olm<sub>M5</sub>-A. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Olm<sub>M5</sub>-B:

GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Olm<sub>M5</sub>-B. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Olm<sub>M5</sub>-C:

GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Olm<sub>M5</sub>-C. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Olm<sub>M5</sub>-D:

GC-MS separation and identification of triketide lactones from fermentation of *Sac. erythraea* JC2, containing pMU3-DEBS 1-TE-URA3- Olm<sub>M5</sub>-D. (A) Total ion chromatogram (TIC). (B) Extracted ion chromatogram ( $m/z = 58$  Da). (C) Extracted ion chromatogram ( $m/z = 56$  Da).

Peak areas for the chromatograms shown:

Construct	Lactone product									
	1a	1b	2a	2b	3a	3b	4a	4b	5a	5b
Amph <sub>M11</sub> -A			3870717	5116032			26191115	25083510	17839138	16905214
Amph <sub>M11</sub> -B			6952497	3369644			14877348	12647122	25663243	22060377
Ny <sub>S</sub> M <sub>11</sub> -A			nq	nq			4334385	7632632	8608377	8261861
Ny <sub>S</sub> M <sub>11</sub> -B			nq	nq			933964	1660804	4625551	4252341
Ny <sub>S</sub> M <sub>11</sub> -A			1027298	1898647			13487590	10867466	6544640	3347253
Ny <sub>S</sub> M <sub>11</sub> -B			193686	1557642			7408472	6557642	8232459	7926235
Olm <sub>M5</sub> -A			nd	nd			nd	nd	36330627	51743140
Olm <sub>M5</sub> -B			nd	nd			nd	nd	23801927	22205200
Lkm <sub>M1</sub> -A	1228182	3684001			nd	nd			368401	nq
Lkm <sub>M1</sub> -B	5194822	15863262			nq	nq			25863262	nq

nq: no quantification possible; nd: not detected

#### Example yield calculation for Amph<sub>M11</sub>-A, 4a/4b products:

In each case, 1 mL of growth medium was extracted and analysed.

Equation (from Fig. S4): relative peak area = 0.4802(conc. (mg/L)) – 0.4732

Total peak area corresponding to 4a + 4b = 26191115 + 25083510 = 51274625

Conversion to relative peak area: (51274625/56653020)(100%) = 90.5% (where 56653020 represents 200 mg/L, 100% relative peak area)

Solving for concentration:

Conc. (mg/L) = (relative peak area + 0.4732)/0.4802

Conc. (mg/L Amph<sub>M11</sub>-A) = (90.5 + 0.4732)/0.4802 = 189.4

As the extracts analysed were 5× more concentrated than the growth samples, the final concentration was obtained by dividing by 5.

This gave 37.9 mg/L for this particular sample.

**Note: the values given in Table 2 in the main text represent the average yield from three biological replicates.**

#### 4. References

- 1 C. J. Rowe, J. Cortés, S. Gaisser, J. Staunton and P. F. Leadlay, *Gene*, 1998, **216**, 215–223.
- 2 T. Annaval, C. Paris, P. F. Leadlay, C. Jacob and K. J. Weissman, *ChemBioChem*, 2015, **16**, 1357–1364.
- 3 N. Bonneaud, O. Ozier-Kalogeropoulos, G. Y. Li, M. Labouesse, L. Minvielle-Sebastia and F. Lacroute, *Yeast*, 1991, **7**, 609–615.
- 4 D. Gietz, A. St Jean, R. A. Woods and R. H. Schiestl, *Nucleic Acids Res.*, 1992, **20**, 1425.
- 5 L. Kellenberger, I. S. Galloway, G. Sauter, G. Böhm, U. Hanefeld, J. Cortés, J. Staunton and P. F. Leadlay, *ChemBioChem*, 2008, **9**, 2740–2749.
- 6 M. Grenson, *Eur. J. Biochem.*, 1969, **11**, 249–260.
- 7 P. Mazodier, R. Petter and C. Thompson, *J. Bacteriol.*, 1989, **171**, 3583–3585.
- 8 R. S. Gokhale, S. Y. Tsuji, D. E. Cane and C. Khosla, *Science*, 1999, **284**, 482–485.
- 9 A. Ranganathan, M. Timoney, M. Bycroft, J. Cortés, I. P. Thomas, B. Wilkinson, L. Kellenberger, U. Hanefeld, I. S. Galloway, J. Staunton and P. F. Leadlay, *Chem. Biol.*, 1999, **6**, 731–741.
- 10 S. Yuzawa, K. Deng, G. Wang, E. E. K. Baidoo, T. R. Northen, P. D. Adams, L. Katz and J. D. Keasling, *ACS Synth. Biol.*, 2017, **6**, 139–147.
- 11 J. Cortés, K. E. Wiesmann, G. A. Roberts, M. J. Brown, J. Staunton and P. F. Leadlay, *Science*, 1995, **268**, 1487–1489.



Aucun micro-organisme n'a été maltraité au cours de cette thèse de doctorat.





Les polycétides réduits sont des produits naturels à la base de plusieurs médicaments à fortes valeurs thérapeutiques et économiques, tels que certains antibiotiques, immunosuppresseurs et agents anti-cancéreux. Ces biomolécules sont synthétisés par les polycétide synthases (PKS) qui figurent parmi les complexes enzymatiques les plus grands, en taille et en complexité, dans le domaine du vivant. Par conséquent, il est d'un grand intérêt de trouver de nouveaux polycétides ou de modifier les structures existantes afin de pouvoir créer des analogues. Dans ce contexte, une approche prometteuse consiste à modifier les PKS par la biologie synthétique. Pour que cette stratégie soit efficace, elle nécessite la compréhension en détail des relations structure/fonction de ces systèmes protéiques. Une des questions majeures qui reste posée concernant les PKS est le rôle du contexte macromoléculaire sur le fonctionnement de ces systèmes. Cet aspect a été quasiment négligé pendant les tentatives précédentes d'ingénierie, ce qui peut expliquer, au moins en partie, le faible taux de réussite.

Afin d'évaluer l'influence du contexte modulaire sur le stéréocontrôle exercé par certains domaines, les échanges de modules ont été effectués au sein de notre système PKS modèle bimodulaire. Ces modules ont été introduits par un système de recombinaison homologue chez la levure en suivant plusieurs stratégies d'échange de modules. Puis l'expression de ces protéines hybrides a été effectuée chez *Saccharopolyspora erythraea*, qui est la souche contenant le système PKS parental, suivie par l'analyse des produits par chromatographie en phase gazeuse couplée à la spectrométrie de masse. Nos résultats suggèrent que l'échange de module entier serait plus efficace car il permet de préserver les contacts entre les domaines KS et ACP d'un même module, afin de maintenir une réaction de condensation décarboxylative efficace. De plus, l'échange de module interne à une sous-unité est plus efficace, car le domaine KS de ce type de module possède une spécificité de substrat plus grande. Les modules ayant montré les résultats les plus significatifs ont ensuite été utilisés pour modifier une PKS entière pour tenter pour la première fois de modifier la stéréochimie de l'intermédiaire érythromycine.

Concernant l'organisation macromoléculaire des PKS, il a été décrit que la PKS de type *trans*-AT responsable de la synthèse de la bacillaene chez *Bacillus subtilis* s'assemblent pour former un mégacomplexe, et que ce mégacomplexe se localise proche de la membrane interne. Ces découvertes doivent être généralisées à la deuxième grande classe de PKS, celle des PKS de type *cis*-AT. Pour pouvoir localiser la PKS de type *cis*-AT responsable de la synthèse de la coelimycine (Cpk) au sein de son hôte natif *Streptomyces coelicolor*, nous avons fusionné la première et/ou la dernière sous-unité de Cpk à une protéine fluorescente. La création des fusions de protéine a été effectuée avec l'outil d'ingénierie génétique CRISPR/Cas9. Par la suite, nous avons obtenu des images de la localisation des protéines de fusion par microscopie confocale. Nous montrons pour la première fois que les sous-unités d'une PKS de type *cis*-AT se localisent également en macrocomplexes, mais que la PKS de type *cis*-AT posséderait une organisation macromoléculaire différente. Nous montrons également la colocalisation d'une des sous-unités de Cpk avec le transporteur CpkF associé, afin d'évaluer la formation d'une véritable usine biosynthétique.

Mots clés : polycétide synthase, ingénierie génétique, usine biosynthétique, contexte modulaire, stéréochimie

Reduced polyketides are natural products that form the basis of several drugs with high therapeutic and economic value, such as certain antibiotics, immunosuppressants and anti-cancer agents. These biomolecules are synthesized by polyketide synthases (PKS) which are among the largest enzymatic complexes, in size and complexity, in the living world. Therefore, it is of great interest to find new polyketides or modify existing structures to be able to create analogs. In this context, a promising approach is to modify PKS by synthetic biology. For this strategy to be effective, it requires a detailed understanding of the structure / function relationships of these protein systems. One of the major questions that remain about PKS is the role of the macromolecular context on the functioning of these systems. This aspect has been largely overlooked during previous engineering attempts, which may explain, at least in part, the low success rate.

In order to assess the influence of the modular context on the stereocontrol exerted by certain domains, the module exchanges were carried out within our bimodular model PKS system. These modules were introduced by a homologous recombination system in yeast following several module exchange strategies. The expression of these hybrid proteins was then performed in *Saccharopolyspora erythraea*, which is the strain containing the parental PKS system, followed by analysis of the products by gas chromatography coupled with mass spectrometry. Our results suggest that whole-module-to-whole-module exchange would be more efficient because it preserves contacts between the KS and ACP domains of the same module, to maintain an efficient decarboxylative condensation reaction. In addition, the module exchange internal to a subunit is more efficient, because the KS domain of this type of module has a greater substrate specificity. The modules which showed the most significant results were then used to modify an entire PKS to attempt for the first time to modify the stereochemistry of the erythromycin intermediate.

Regarding the macromolecular organization of PKS, it has been described that the *trans*-AT-type PKS responsible for the synthesis of bacillaene in *Bacillus subtilis* assemble to form a mega-complex, and that this mega-complex is localized close to the inner membrane. These findings should be generalized to the second major class of PKS, that of *cis*-AT-type PKS. To be able to locate the *cis*-AT-type PKS responsible for the synthesis of coelimycin (Cpk) within its native host *Streptomyces coelicolor*, we fused the first and / or the last subunit of Cpk to a fluorescent protein. Creation of the protein fusions was performed with the CRISPR / Cas9 genetic engineering tool. Subsequently, we obtained images of the localization of the fusion proteins by confocal microscopy. We show for the first time that the subunits of a *cis*-AT-like PKS also localize in macrocomplexes, but that the *cis*-AT-like PKS has a different macromolecular organization. We also show the colocalization of one of the Cpk subunits with the associated CpkF transporter, to assess the formation of a true biosynthetic factory.

Key words : polyketide synthase, genetic engineering, biosynthetic factory, modular context, stereochemistry.