



**HAL**  
open science

# Contribution à l'étude des dernières étapes de la biosynthèse de l'anatoxine-a, une neurotoxine produite par les cyanobactéries

Guillaume Paci

► **To cite this version:**

Guillaume Paci. Contribution à l'étude des dernières étapes de la biosynthèse de l'anatoxine-a, une neurotoxine produite par les cyanobactéries. Biochimie, Biologie Moléculaire. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066478 . tel-01294533

**HAL Id: tel-01294533**

**<https://theses.hal.science/tel-01294533>**

Submitted on 29 Mar 2016

**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.

# Université Pierre et Marie Curie

Ecole doctorale 406 Chimie Moléculaire

*UMR 7223 Chimie Paristech / LIED UMR 8236 Université Paris Diderot / Equipe de  
Biochimie des microorganismes*

## **Contribution à l'étude des dernières étapes de la biosynthèse de l'anatoxine-a, une neurotoxine produite par les cyanobactéries**

Par Guillaume Paci

Thèse de doctorat de Chimie Moléculaire

Dirigée par le Pr Olivier Ploux

Présentée et soutenue publiquement le 10/11/2015

Devant un jury composé de :

M Lequin Olivier, Professeur, Président du jury

M Camadro Jean-Michel, Directeur de recherche, Rapporteur

Mme Gondry Muriel, Directeur de recherche, Rapporteur

M Ploux Olivier, Professeur, Directeur de Thèse

# Table des matières

|  |    |
|--|----|
| Résumé : .....   | 4  |
| Remerciements : .....  | 5  |
| Abréviations : .....   | 6  |
| Chapitre I : Introduction .....  | 9  |
| I.1) Les cyanobactéries.....   | 9  |
| I.2) Les métabolites secondaires de cyanobactéries .....   | 10 |
| I.3) Les polyketide synthases .....  | 11 |
| I.3.1) Les PKSs de type I.....   | 11 |
| I.3.2) Les différents domaines des PKSs de type I et leur fonctionnement .....                                     | 13 |
| I.3.3) Les PKSs de type II et de type III.....   | 20 |
| I.3.4) Les PKSs chez les cyanobactéries .....  | 20 |
| I.4) Les nonribosomal peptide-synthetases .....  | 21 |
| I.5) Les toxines de cyanobactéries .....   | 22 |
| I.6) L'anatoxine-a.....  | 24 |
| I.7) Biosynthèse de l'anatoxine-a.....   | 26 |
| I.7.1) Présentation de la biosynthèse .....  | 26 |
| I.7.2) Les dernières étapes de la biosynthèse de l'anatoxine-a .....   | 30 |
| I.8) Objectifs du projet de recherche.....   | 32 |
| Chapitre II : Analyse des clusters <i>ana</i> de différentes souches de cyanobactéries.....                        | 33 |
| II.1) Séquençage du cluster <i>ana</i> de la souche <i>Oscillatoria</i> sp. PCC 6506.....                          | 33 |
| II.1.1) Comparaison du cluster <i>ana</i> des <i>Oscillatoria</i> sp. PCC 6506 et 6407 .....                       | 33 |
| II.1.2) Réassemblage des contigs chez la souche <i>Oscillatoria</i> sp. PCC 6506 en 5' du cluster <i>ana</i> ..... | 35 |
| II.1.3) Réassemblage des contigs chez la souche <i>Oscillatoria</i> sp. PCC 6506 en 3' du cluster <i>ana</i> ..... | 36 |
| II.1.4) Bilan du réassemblage.....   | 38 |
| II.2) Comparaison des clusters <i>ana</i> de différentes souches.....  | 38 |
| II.2.1) Synténie du cluster <i>ana</i> .....   | 40 |
| II.2.2) Comparaison des séquences peptidiques de la protéine AnaG .....  | 42 |
| II.3) Discussion .....   | 45 |
| Chapitre III : Etude des dernières étapes de la biosynthèse .....  | 48 |
| III.1) Approche chimique : synthèse d'un analogue de substrat pour AnaG .....                                      | 48 |
| III.1.1) Voie de synthèse chimique de l'anatoxine-a.....   | 48 |
| III.1.2) Un analogue du substrat de l'AnaG .....   | 49 |
| III.1.3) Discussion .....  | 63 |
| III.2) Approche biochimique : obtention des protéines AnaG et AnaA.....  | 65 |
| III.2.1) Stratégies d'obtention d'AnaG.....  | 66 |
| III.2.2) Obtention d'AnaA.....   | 80 |
| III.2.3) Discussion .....  | 81 |
| III.3) Conclusion .....  | 82 |
| Chapitre IV : Etude de la biosynthèse de la dihydroanatoxine-a .....   | 83 |
| IV.1) La dihydroanatoxine-a .....  | 83 |
| IV.1.1) Synthèse de la <i>cis</i> et <i>trans</i> dihydroanatoxine-a.....  | 83 |
| IV.1.2) Toxicité de la dihydroanatoxine-a .....  | 83 |
| IV.1.3) Détection dans l'environnement .....   | 84 |
| IV.1.4) Dégradation de l'anatoxine-a.....  | 84 |

|  |     |
|--|-----|
| IV.2) Synthèse des standards dihydroanatoxine-a et dihydrohomoanatoxine-a .....  | 84  |
| IV.3) Souche productrice de dihydroanatoxine-a.....  | 85  |
| IV.4) Etude de la dihydroanatoxine-a en spectrométrie de masse .....   | 86  |
| IV.4.1) But de l'étude.....  | 86  |
| IV.4.2) Analyse des standards en LC-MS .....   | 86  |
| IV.4.3) Analyse des standards en GC-MS .....   | 89  |
| IV.4.4) Extrait de la <i>Cylindrospermum stagnale</i> PCC7417 .....  | 93  |
| IV.4.5) Expérience d'incorporation isotopique par feeding.....   | 94  |
| IV.4) Discussions .....  | 97  |
| IV.5) Présence de réductase dans le cluster <i>ana</i> de la <i>Cylindrospermum stagnale</i> PCC 7417 .....  | 98  |
| IV.6) Conclusion .....   | 99  |
| Chapitre V : Conclusion .....  | 100 |
| V.1) Bilan général .....   | 100 |
| V.2) Perspectives.....   | 101 |
| Partie Expérimentale de Biologie.....  | 103 |
| VI.1) Souches de cyanobactéries.....   | 103 |
| VI.2) Culture d' <i>Escherichia coli</i> .....   | 103 |
| VI.3) PCR.....   | 103 |
| VI.4) Préparation de bactéries compétentes .....   | 106 |
| VI.4.1) Préparation d' <i>E. coli</i> JM109 électrocompétentes .....   | 106 |
| VI.4.2) Préparation d' <i>E. coli</i> BL21(+)(DE3)-RIL(Cm <sup>R</sup> ) chimiocompétentes .....   | 106 |
| VI.5) Clonage .....  | 107 |
| VI.5.1) Digestion enzymatique .....  | 107 |
| VI.5.2) Déphosphorylation.....   | 107 |
| VI.5.3) Ligation.....  | 107 |
| VI.5.4) Transformation bactérienne .....   | 108 |
| VI.5.5) Criblage des clones .....  | 108 |
| VI.6) Expression et purification de protéines chez <i>E. coli</i> .....  | 109 |
| VI.6.1) Préculture .....   | 109 |
| VI.6.2) Culture et induction .....   | 110 |
| VI.6.3) Sonication de la culture.....  | 110 |
| VI.6.4) Purification des protéines sur colonne d'affinité .....  | 110 |
| VI.6.5) Gel d'électrophorèse SDS PAGE .....  | 111 |
| VI.6.6) Dessalage et concentration.....  | 112 |
| VI.6.7) Quantification des protéines .....   | 112 |
| VI.6.8) Récapitulatif des purifications effectuées .....   | 112 |
| VI.7) Etude de la <i>Cylindrospermum stagnale</i> PCC 7417.....  | 115 |
| VI.7.1) Expérience d'incorporation isotopique par feeding.....   | 115 |
| VI.7.2) Préparation d'extrait de milieu de culture de cyanobactéries pour analyse GC-MS .....  | 116 |
| Partie Expérimentale de Chimie.....  | 117 |
| VII.1) Matériels.....  | 117 |
| VII.2) Synthèse d'analogue de substrat d'AnaG.....   | 119 |
| VII.2.1) <i>rac</i> -(2 <i>S</i> ,3 <i>R</i> )-3-Azabicyclo[6.2.0]dec-7-ene-2-one (1).....   | 119 |
| VII.2.2) <i>rac</i> -(2 <i>S</i> ,3 <i>R</i> )-3-(tert-Butoxycarbonyl)-3-azabicyclo[6.2.0]dec-7-ene-2-one (2).....   | 120 |
| VII.2.3) <i>rac</i> -(2 <i>S</i> ,3 <i>R</i> )-3-((tert-Butoxycarbonyl)amino)cyclooct-7-ene-2-oate de méthyle (3) .....                                    | 121 |
| VII.2.4) <i>rac</i> -(2 <i>S</i> ,3 <i>S</i> ,6 <i>R</i> )-10-(tert-butoxycarbonyl)-6-(phenylselanyl)-3-azabicyclo[4.2.1]nonane-2-oate de méthyle (4)..... | 122 |

|   |     |
|---|-----|
| VII.2.5) <i>rac</i> -(2 <i>S</i> ,3 <i>S</i> ,6 <i>S</i> )-10-(tert-butoxycarbonyl)-3-azabicyclo[4.2.1]nonane-2-oate de methyle (5) .....       | 123 |
| VII.2.6) <i>rac</i> -(3 <i>S</i> ,6 <i>S</i> )-10-(tert-butoxycarbonyl)-2-(phenylselanyl)-3-azabicyclo[4.2.1]nonane-2-oate de methyle (6) ..... | 124 |
| VII.2.7) (3 <i>S</i> ,6 <i>S</i> )-10-(Butoxycarbonyl-tert) - tert-3-azabicyclo[4.2.1]non-9-ene-2-oate de methyle (7) .....                     | 125 |
| VII.2.8) (3 <i>S</i> ,6 <i>S</i> )-Acide 10-(tert-butoxycarbonyl)-9-azabicyclo[4.2.1]non-9-ene -2-oiïque(8) .....                               | 126 |
| VII.2.9) (3 <i>S</i> ,6 <i>S</i> )-10-(tert-butoxycarbonyl)-9-azabicyclo[4.2.1]thionon-9-ene-2-oate de N-ethyl-éthanamide (9).....              | 127 |
| VII.2.10) (3 <i>S</i> ,6 <i>S</i> )-9-azabicyclo[4.2.1]thionon-9-ene-2-oate de N-ethyl-éthanamide (10) .....                                    | 128 |
| VII.2.11) Analyse de (10) en LC-MS.....   | 129 |
| VII.2.12) Purification de (10) par HPLC analytique.....   | 129 |
| VII.2.13) Purification de (10) par HPLC semi-preparative.....   | 130 |
| VII.2.14) Analyse de (10) purifié par HPLC en LC-MS .....   | 130 |
| VII.3) Etude de la biosynthèse de la dihydroanatoxine-a .....   | 131 |
| VII.3.1) Déprotection des analogues dihydro d'anatoxine-a et d'homoanatoxine-a .....  | 131 |
| VII.3.2) Analyse des standards par GC-MS.....   | 132 |
| VII.3.3) Purification de la <i>cis</i> -dihydroanatoxine-a.....   | 132 |
| VII.3.4) Analyse de la <i>cis</i> -dihydroanatoxine-a .....   | 132 |
| VII.3.5) Analyse d'extrait de <i>Cylindrospermum stagnale</i> PCC 7417.....   | 133 |
| Annexes .....   | 134 |
| Annotation des gènes en 5' et 3' du cluster <i>ana</i> .....  | 134 |
| Annotation des gènes chez la souche <i>Anabaena</i> sp. strain 37.....  | 134 |
| Liste des Figures.....  | 134 |
| Liste des Tableaux.....   | 137 |
| Bibliographie .....   | 138 |

## Résumé :

Les cyanobactéries sont des procaryotes photosynthétiques ubiquitaires qui produisent un grand nombre de métabolites secondaires, dont des toxines. Parmi ces cyanotoxines, l'anatoxine-a et l'homoanatoxine-a sont deux neurotoxines puissantes qui provoquent une mort rapide, chez certains animaux, après ingestion. La mort est causée par asphyxie car ces alcaloïdes sont de puissants agonistes du récepteur nicotinique de l'acétylcholine. Ces toxines représentent donc un risque pour la santé animale et humaine.

L'équipe, au sein de laquelle j'ai effectué ma thèse, étudie la biosynthèse de l'anatoxine-a et de ses dérivés, chez les cyanobactéries. Des travaux précédents de l'équipe ont permis d'identifier le cluster de gènes responsable de la biosynthèse de l'anatoxine-a et de l'homoanatoxine-a, dans le génome de la cyanobactérie *Oscillatoria sp.* PCC 6506, une souche productrice d'homoanatoxine-a. Une voie de biosynthèse, à partir de la proline a été proposée par l'équipe, puis en partie reconstituée, in vitro.

J'ai travaillé sur l'étude des dernières étapes de cette voie de biosynthèse, qui met probablement en jeu une polyketide synthase (PKS) AnaG et une thioestérase AnaA. Lors de ces étapes le précurseur de l'homoanatoxine-a est condensé à une unité acétate, puis subirait une méthylation, une hydrolyse et une décarboxylation, pour donner l'homoanatoxine-a. Néanmoins, la PKS AnaG ne possède ni domaine thioestérase ni domaine décarboxylase, et les dernières étapes de la biosynthèse sont donc mal définies. Nous avons décidé d'exprimer différents domaines d'AnaG chez *Escherichia coli* pour obtenir plus d'informations sur ces étapes. Nous avons également tenté de préparer un analogue du substrat putatif d'AnaG par synthèse chimique. Avec ces outils et en utilisant la spectrométrie de masse, nous espérons pouvoir confirmer la voie de biosynthèse que nous avons proposée.

Par ailleurs, nous avons étudié la biosynthèse de la dihydroanatoxine-a chez *Cylindrospermum stagnale* PCC 7417. Nous avons réalisé des expériences d'incorporation isotopiques, couplées avec des études par spectrométrie de masse pour confirmer la production de cette toxine chez *C. stagnale* PCC 7417, et confirmer sa voie de biosynthèse.

## Remerciements :

Je tiens d'abord à remercier le professeur Olivier Ploux de m'avoir accepté comme étudiant de doctorat. J'ai apprécié ses conseils, son encadrement et son soutien tout au long de mon doctorat.

Je veux aussi remercier le professeur Annick Méjean pour son accueil dans son équipe et ses conseils tout au long du doctorat.

Je remercie les rapporteurs de cette thèse, pour avoir accepté de relire et de commenter ce mémoire. Mes remerciements vont aussi au Pr. Lequin pour présider mon jury.

Je remercie aussi les anciens membres de l'équipe (Stéphane Mann, Rabia Mazmouz, Perrine Zeller et Valérie Gautier) pour leur aide, leur soutien et leur amitié au cours de mon doctorat.

Je tiens à remercier monsieur Raphaël Cases de la Fondation Pierre Gilles de Gennes pour l'opportunité de travailler avec lui et les échanges que nous avons eu.

Je veux remercier les membres des équipes de Chimie ParisTech et du LIED que j'ai côtoyés pour les échanges que nous avons eu. Je pense particulièrement à l'ATCP et à mes co-bureaux (Julien et Simon en particulier).

Je pense aussi aux stagiaires qui sont passés par l'équipe (Juliette et Didier) que j'ai eu l'occasion de côtoyer et d'aider. Je leur souhaite de continuer à poursuivre leur passion.

Je remercie aussi mes amis pour leur présence et leur soutien tout au long de ces trois années.

Je veux remercier ma compagne et ma famille pour leur amour et leur soutien.

## **Abréviations :**

2-ME : 2-mercaptoéthanol

A : Domaine d'adénylation

ACP : Acyl carrier protein

ADN : Acide désoxyribonucléique

AIBN : Azobisisobutyronitrile

ARN : Acide ribonucléique

AT : Acyltransférase

ATX : Anatoxine-a

Bp : paire de bases

C : Domaine de condensation

Cf : Concentration finale

CI : Ionisation chimique

Ci : Concentration initiale

CLC : Claisen-like cyclase

Cm : Chloramphénicol

Cm<sup>f</sup> : Chloramphénicol résistant

CoASH : Coenzyme A

DCC : Dicyclohexylcarbodiimide

DEBS : Deoxyerythronolide B synthase

DH : Dehydratase

dhATX : Dihydroanatoxine-a

dhHATX : Dihydrohomoanatoxine-a

DMAP : 4-diméthylaminopyridine

DMF : Diméthylformamide

DO : Densité Optique

DL<sub>50</sub> : dose létale 50%

*E. coli* : *Escherichia coli*

EDCI : 1-Ethyl-3-(3-diméthylaminopropyl)carbodiimide

EDTA : Acide éthylène diamine tétra acétique

ER : Enoylreductase

ESI : Ionisation à électrospray

FAS : Fatty acid synthase  
GC-MS : Gas chromatography–mass spectrometry  
HATX : Homoanatoxine-a  
His-tag : Étiquette poly-histidine  
HOBt : Hydroxybenzotriazole  
HPLC : High Performance Liquid Chromatography  
HRP : Horseradish peroxidase  
IPTG : Isopropyl-beta-D-thiogalactopyranoside  
ITS : Internal transcribed sequence  
Kan : Kanamycine  
Kan<sup>r</sup> : Kanamycine résistant  
KR : Ketoreductase  
Ki : Constante d'inhibition  
KS : Keto-synthase  
LB : Lysogeny broth  
LC-MS : Liquid chromatography-mass spectrometry  
LDA : Diisopropylamidure de lithium  
mCPBA : Acide métachloroperbenzoïque  
MT : Methyl transférase  
NAC : N-Acetylcystéamine  
NRPS : Non-ribosomal peptides synthetase  
PAGE : PolyAcrylamide gel electrophoresis  
PCC : Pasteur culture collection of cyanobacteria  
PCP : Domaine peptidyl carrier protein  
PCR : Polymerase chain reaction  
PKS : Polyketide synthases  
PPTase : Phosphopantéthéine transférase  
Rdt : Rendement  
Rf : Rapport frontal  
RMN : Résonance Magnétique Nucléaire  
Rpm : Rotation par minute  
SDS : Dodécylsulfate de sodium  
SOB : Super Optimal Broth  
TA : Température ambiante

Taq Pol : Thermophilus aquaticus polymérase

TCEP : Tris (2-carboxyethyl)phosphine

TE : Thioestérase

TFA : Acide trifluoroacétique

TIC : Total Ion Current

THF : Tétrahydrofurane

# Chapitre I : Introduction

## I.1) Les cyanobactéries

Les cyanobactéries sont des procaryotes photosynthétiques parmi les plus anciens connus. Leur apparition sur Terre date en effet de 3,5 milliards d'années. Elles appartiennent au groupe des bactéries et elles ont laissé une empreinte sur Terre sous forme de fossiles, les stromatolithes, que l'on retrouve en Afrique et en Australie. Ce sont des bactéries à Gram négatif. Les cyanobactéries peuplent tous les milieux terrestres et aquatiques mais la majorité des espèces connues se trouve en eaux douces et dans les eaux marines. Certaines espèces peuvent même survivre dans des milieux extrêmes comme l'Antarctique et les sources chaudes<sup>1</sup>.



Figure 1. *Lyngbya* sp. (gauche) et *Anabaena azollae* (droite) observées au microscope au grossissement  $\times 400$ . <http://www-cyanosite.bio.purdue.edu/images/images.html>

Les cyanobactéries sont capables de photosynthèse oxygénique mais certaines cyanobactéries sont aussi capables de fixer le diazote de l'atmosphère. Ces différentes caractéristiques leur donnent donc une place essentielle dans la biosphère terrestre. Les cyanobactéries sont également connues pour leur capacité à former des efflorescences dans les eaux douces comme les eaux marines<sup>2</sup>. Ces efflorescences sont dues à une surabondance de cyanobactéries et elles sont responsables de problèmes environnementaux. La capacité de certaines cyanobactéries à produire des toxines rend ces efflorescences dangereuses pour la santé des hommes et des animaux.

Les cyanobactéries sont aussi très étudiées pour leur capacité à produire une grande diversité de métabolites secondaires. Ce sont ces métabolites secondaires produits par les

cyanobactéries et plus particulièrement les toxines qui sont le sujet principal de mon travail de recherche.

## **I.2) Les métabolites secondaires de cyanobactéries**

Les métabolites secondaires sont des molécules, généralement de petites tailles, produites majoritairement par les bactéries, les champignons et les plantes<sup>3</sup>. L'ensemble des rôles des métabolites secondaires n'est pas encore bien connu même s'il est généralement admis que ces composés doivent donner un avantage à l'organisme producteur dans son environnement. Parmi les exemples connus d'avantages on peut citer les molécules appelées suncreens comprenant la scytonémine et la mycosporine qui permettent aux cyanobactéries de se protéger des rayonnements UVs<sup>4</sup>. Cependant pour la plupart des métabolites secondaires, leur rôle reste inconnu.

Les métabolites secondaires des cyanobactéries présentent une extraordinaire diversité structurale<sup>5-7</sup>. Parmi eux, nous retrouvons plusieurs classes de composés : alcaloïdes, terpénoïdes, glycosides, phénols, polycétides, acides gras, alcanes, alcènes, peptides non-ribosomiaux et ribosomiaux. Ces classes de composés peuvent bien sûr s'interpénétrer. Les métabolites secondaires de cyanobactéries sont étudiés pour des applications de ces composés. Leur étude permet de découvrir de nombreuses pistes pour l'identification de nouvelles molécules thérapeutiques<sup>7</sup>. Certains des métabolites secondaires peuvent aussi être utilisés comme biocarburants<sup>8-10</sup>. Il y a finalement de nombreux métabolites secondaires qui sont toxiques pour les animaux supérieurs et sont appelés cyanotoxines<sup>11</sup>. Le risque qu'ils posent pour l'environnement et pour l'homme en a fait un des axes de recherche principaux concernant l'étude des métabolites secondaires de cyanobactéries d'eaux douces. Nous aborderons ce point plus en détail dans un sous-chapitre suivant.

La très grande diversité chimique des métabolites secondaires produits par les cyanobactéries se reflète dans les nombreuses voies de biosynthèse qui sont responsables de la synthèse de ces composés. L'identification des gènes impliqués dans ces voies de biosynthèses a été réalisée très récemment (depuis 2000) grâce à des efforts importants<sup>5</sup>. Ces voies de biosynthèses font intervenir de nombreuses enzymes différentes et nous allons nous intéresser à deux types d'enzymes bien particulières : les polyketide synthases (PKSs) et les non-ribosomal peptide synthetases (NRPSs).

## **I.3) Les polyketide synthases**

Les PKSs sont des enzymes présentes chez les bactéries, les plantes, les champignons et chez certains animaux. Elles sont apparentées aux Fatty Acid Synthases (FASs). Les FASs sont responsables de la biosynthèse des acides gras<sup>12</sup>. Elles sont classées en deux types selon leur organisation. Nous retrouvons le type I chez les animaux et les champignons et le type II chez les bactéries. La FAS de type I est une macrosynthase multifonctionnelle homodimérique de grande taille alors que la FAS de type II est constituée de plusieurs enzymes distinctes et monofonctionnelles.

Les PKSs sont responsables de la synthèse de nombreux macrolides possédant des propriétés thérapeutiques comme l'érythromycine (antibiotique), la rapamycine (immunosuppresseur) et l'épothilone (anticancéreux). Il existe trois groupes de PKSs connus : les PKSs de type I, les PKSs de type II et les PKSs de type III. Les différences observées entre les trois types de PKSs portent sur l'organisation de l'enzyme. Nous allons particulièrement nous intéresser aux PKSs de type I.

Ces différents types de PKSs partagent toutes des mêmes types de réactions. Elles vont ajouter deux ou trois carbones à une chaîne en élongation à partir de différents monomères et d'un composé de départ. L'ajout de ces carbones se fait par le biais d'une réaction de condensation de Claisen. La liaison entre l'enzyme et la chaîne en élongation est covalente sous forme d'un thioester.

### **I.3.1) Les PKSs de type I**

Les PKS de type I sont des enzymes modulaires<sup>13</sup>. Chaque module assure l'ajout de deux ou trois carbones venant d'un monomère comme le malonyl-CoA ou le méthylmalonyl-CoA, à la chaîne carbonée en cours d'élongation. Ces protéines sont de véritables chaînes de montage moléculaire, c'est-à-dire que la chaîne carbonée va passer d'un module à un autre dans un ordre bien défini lors de l'élongation comme le châssis d'une voiture sur une chaîne de montage industriel. De plus ces PKSs obéissent à une règle de colinéarité. En effet, l'organisation des gènes codant pour les PKSs suit l'organisation de la voie de biosynthèse. Par exemple, le premier module d'une PKS correspondra au premier gène présent dans le génome. Dans le cas de modules répartis sur plusieurs protéines, la liaison et la

reconnaissance entre les différentes protéines porteuses de modules se fait par des « docking domains »<sup>14-16</sup>.

Un module est composé de plusieurs domaines et chaque domaine possède une fonction catalytique bien précise lors du processus d'élongation. Les PKSs de type I présentent toujours trois domaines essentiels à la réaction d'élongation. Ces domaines sont : Keto-synthase (KS), Acyltransferase (AT) et Acyl Carrier Protein (ACP). Ensuite il existe d'autres domaines « optionnels » qui se chargent de catalyser d'autres réactions. S'il n'y qu'un seul module répétant à chaque fois une certaine réactions d'élongation, alors c'est une PKS de type I itérative<sup>17-19</sup>. Si par contre, la composition en domaine varie d'un module à l'autre, la PKS est de type modulaire.

Les PKSs de type I sont donc des protéines de grande taille, allant de plusieurs centaines de kDa à plus d'un million. La plus grande PKS de type I connue est la PKS appelée MlsA1 responsable de la biosynthèse des mycolactones. Cette protéine possède 62 domaines repartis sur 8 modules pour un poids moléculaire de 1,8 MDa<sup>20,21</sup>. Les PKSs de type I ont été abondamment étudiées, en particulier la voie de biosynthèse de l'érythromycine et la voie de biosynthèse de la pikromycine<sup>22-25</sup>. L'étude de ces deux PKSs a beaucoup apporté sur la connaissance du fonctionnement des PKSs. Nous prendrons comme exemple la voie de biosynthèse de l'érythromycine présentée en Figure 2. La structure des domaines des PKSs de type I a beaucoup été étudiée ces dernières années, de façon à donner des informations sur leur mécanisme<sup>26-28</sup>. La structure tridimensionnelle du troisième module de la pikromycine synthase a, d'ailleurs, été résolue récemment<sup>29</sup>. Ceci a permis de confirmer que ces protéines s'assemblent sous forme d'homodimère. Avant cela, seules les structures tridimensionnelles des domaines avaient été obtenues.

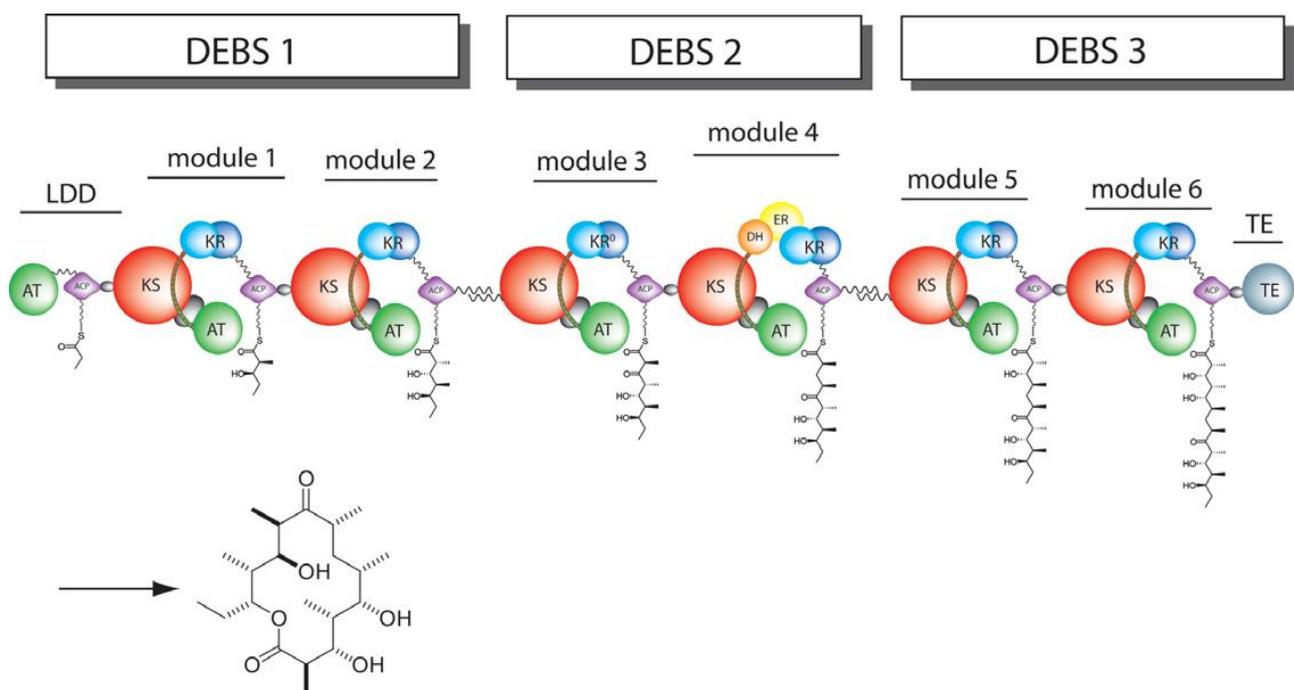


Figure 2. PKS responsable de la biosynthèse de l'érythromycine, LDD correspond au module d'initiation

[http://proteopedia.org/wiki/index.php/6-deoxyerythronolide\\_B\\_synthase\\_%28DEBS%29](http://proteopedia.org/wiki/index.php/6-deoxyerythronolide_B_synthase_%28DEBS%29)

### I.3.2) Les différents domaines des PKSs de type I et leur fonctionnement

#### I.3.2.1) Les domaines KS, AT et ACP

Le domaine ACP permet de porter la chaîne en cours d'élongation grâce à un bras phosphopantéthéinyl. Ce bras doit être assez flexible pour pouvoir avoir accès aux sites actifs des différents domaines présents sur le module.

Le domaine AT va catalyser le chargement du monomère sur le bras du domaine ACP. Le domaine AT assure la spécificité du chargement du monomère sur le domaine ACP et donc son incorporation dans le métabolite final. Si le monomère chargé est le méthylmalonyl-CoA, il se pose la question de la stéréochimie du composé chargé. Dans ce cas, le domaine AT charge seulement l'un des deux isomères. Par exemple, dans la voie de biosynthèse de l'érythromycine, seuls les isomères 2*S* du méthylmalonyl-CoA sont chargés sur les domaines ACP<sup>30</sup>.

Le domaine KS va catalyser le transfert de la chaîne en élongation du précédent module sur lui-même. C'est au niveau de son site catalytique que va avoir lieu la

condensation de Claisen entre la chaîne en élévation et le monomère. Concernant la stéréochimie de la réaction de condensation, dans le cas d'un méthyldmalonyl-CoA, l'inversion de la configuration au niveau de la position C-2 a toujours lieu comme c'est le cas chez les FAS de type I. Par contre, dans la voie de biosynthèse de l'érythromycine, bien que seuls des isomères 2S du méthyldmalonyl-CoA soient utilisés, nous retrouvons des groupes méthyles ayant une configuration L et d'autres ayant une configuration D<sup>31</sup>. Il peut aussi y avoir une réaction d'épimérisation chez certains modules après la condensation ce qui explique le changement de configuration. L'origine de cette épimérisation provient du domaine Ketoreductase<sup>32,33</sup>. Le mécanisme de ces domaines est présenté sur la Figure 3.

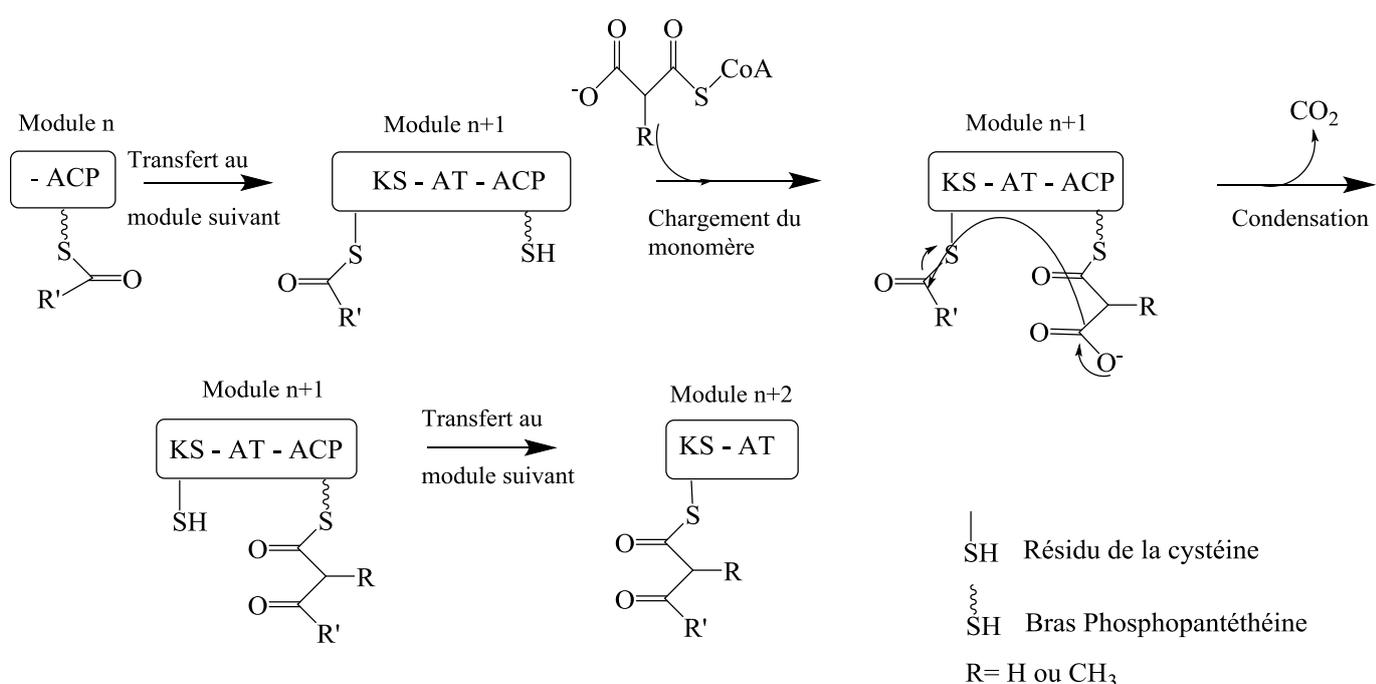


Figure 3. Fonctionnement des domaines KS, AT et ACP avec un monomère de malonyl-CoA.

### I.3.2.2) Les domaines « optionnels »

D'autres domaines, permettant de réduire plus ou moins la chaîne carbonée, sont présents de façon optionnelle sur certains modules de PKSs : Ketoreductase (KR), Déshydratase (DH), Enoylreductase (ER). Ces domaines interviennent après la réaction de condensation et avant le transfert sur un autre module.

Le domaine KR va, tout d'abord, réduire la fonction cétone en alcool. Les domaines KR des PKSs de type I peuvent produire des groupements hydroxyle de configuration D ou L

en fonction des modules. La stéréospécificité de la réaction de cétoreduction est déterminée par l'orientation du substrat dans le site actif. Cette orientation est influencée par la composition en acide aminés des boucles non structurées proches du site actif<sup>33,34</sup>. Il existe donc deux types de domaines KR, le type A et B<sup>35,34</sup>. De plus lorsque le monomère était le méthylmalonyl-CoA, il y a une deuxième stéréospécificité à prendre en compte. Il existe dans ce cas des types I et des types II en fonction de la réduction par le module KR de l'isomère 2*S* ou 2*R*. Ceci nous donne finalement 4 types de domaines KR : A1, A2, B1 et B2.

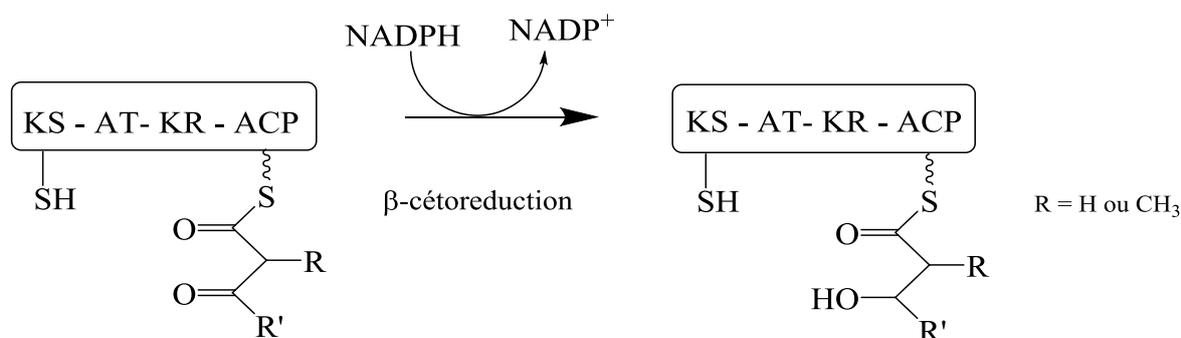


Figure 4. Réaction catalysée par le domaine KR.

Le domaine DH va ensuite déshydrater la fonction alcool en fonction alcène. La plupart des insaturations présentes chez les polycétides sont de configuration *E*. Il semble que la configuration des insaturations présentes chez ces composés soit généralement liée à la configuration de la fonction alcool avant la réaction de déshydratation<sup>30</sup>. Le domaine DH serait capable d'accueillir les deux types de substrats pour former une insaturation *Z* ou *E*<sup>36</sup>. Les alcools de configuration *D* subissent une *syn*-déshydratation qui donne une insaturation de configuration *E* et les alcools de configuration *L* subissent une *syn*-déshydratation qui donne une insaturation de configuration *Z*. Il arrive aussi que les insaturations de configuration *Z* soient produites par des enzymes n'appartenant pas à la PKS<sup>37</sup>.

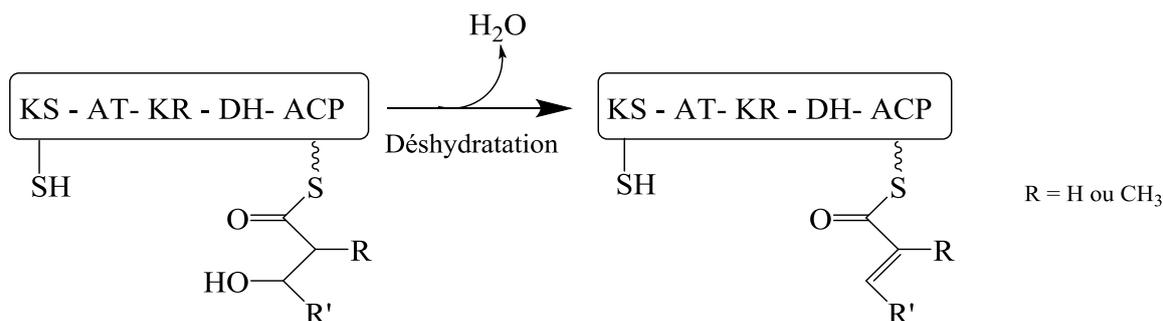


Figure 5. Réaction catalysée par le domaine DH

Enfin, le domaine ER va réduire la fonction alcène en fonction alcane. Il appartient à la famille des déhydrogénases/réductases de chaînes moyennes (MDR)<sup>38,39</sup>. Dans le cas de composés portant une fonction sur le C-2, il existe de nombreux exemples de composés ayant une configuration 2*S* ou 2*R*. Il existe donc deux types de domaine ER chacun étant stéréospécifique pour un composé d'une configuration donnée<sup>40,41</sup>.

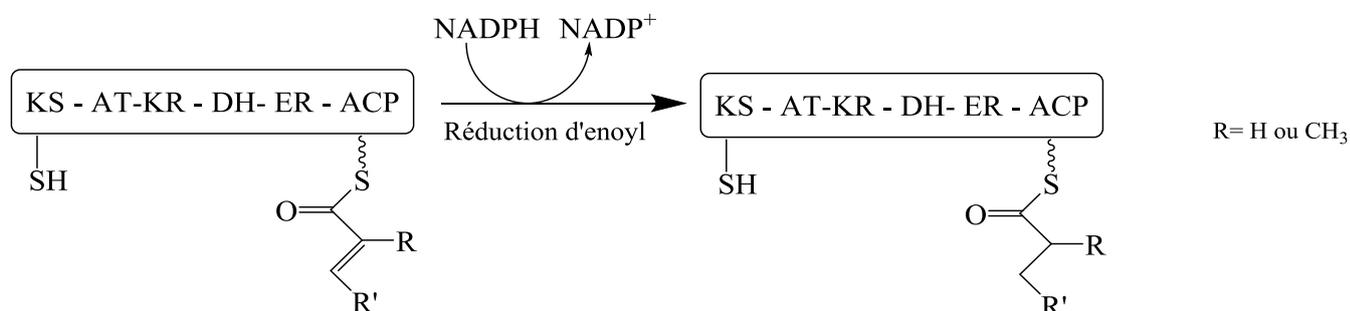


Figure 6. Réaction catalysée par le domaine ER

Il arrive aussi que dans une voie de biosynthèse naturelle, un domaine soit présent sur un module mais inactif comme le domaine KR du module 3 chez l'érythromycine. La structure modulaire des PKSs rend la modification des PKSs très attractive. Il est possible de modifier le produit de synthèse en inactivant certains domaines dans la voie de biosynthèse. Ceci permet d'utiliser la machinerie des PKSs pour produire des composés qui ne sont pas produits naturellement dans ces voies de biosynthèse. De nombreux travaux impliquant la modification de PKSs à différents niveaux, pour exploiter pleinement leur potentiel, sont actuellement en cours<sup>42,43</sup>.

### I.3.2.3) Initiation et terminaison

Des domaines bien particuliers chez les PKSs servent à l'initiation et à la terminaison de la biosynthèse médiée par une PKS.

On trouve d'abord un module d'initiation qui charge le produit de départ de la voie de biosynthèse sur un domaine ACP. Généralement, ce module comporte un domaine AT et un domaine ACP comme dans la voie de biosynthèse de l'érythromycine. Le domaine AT va catalyser le transfert du produit initial, comme un propionyl-CoA dans le cas de l'érythromycine, sur le bras du domaine ACP. Le composé pourra ensuite être transféré au module suivant pour commencer l'élongation. Il arrive aussi qu'un domaine KR soit présent

sur le module d'initiation mais il est alors inactif. Il peut aussi arriver que le module d'initiation soit particulier de façon à permettre au composé de départ d'être différent des habituels acyl-CoA<sup>44</sup>.

A la fin d'une suite de réactions médiées par des PKSs, lorsque l'élongation de la chaîne carbonée est terminée, le composé final doit être libéré. Souvent, dans le dernier module d'une PKS, il existe un domaine Thioestérase de type I (TE I) qui a pour fonction de libérer la molécule<sup>45-47</sup>. Les domaines TE I sont des variations des sérine hydrolases où une sérine, présente au site actif, va accepter le transfert de la chaîne en élongation depuis le dernier domaine ACP de la PKS. Une histidine, présente au site actif, jouera le rôle de catalyseur basique. Le composé en cours d'élongation sera transféré sur la fonction alcool de la sérine.

Le composé est ensuite libéré par un mécanisme de cyclisation intramoléculaire<sup>48</sup>. Cette réaction va donner des produits cycliques sous forme de lactones ou de lactames. C'est ce qui se passe dans la voie de biosynthèse de l'érythromycine, mais les produits libérés ne sont pas toujours cycliques. En effet, un mécanisme d'hydrolyse, donnant un carboxylate, peut avoir lieu. Cette hydrolyse rentre alors en compétition cinétique avec la cyclisation. La différence de vitesse entre les deux réactions va déterminer le produit majoritaire obtenu à la fin de la voie de biosynthèse. Ces deux réactions sont présentées sur la Figure 7.

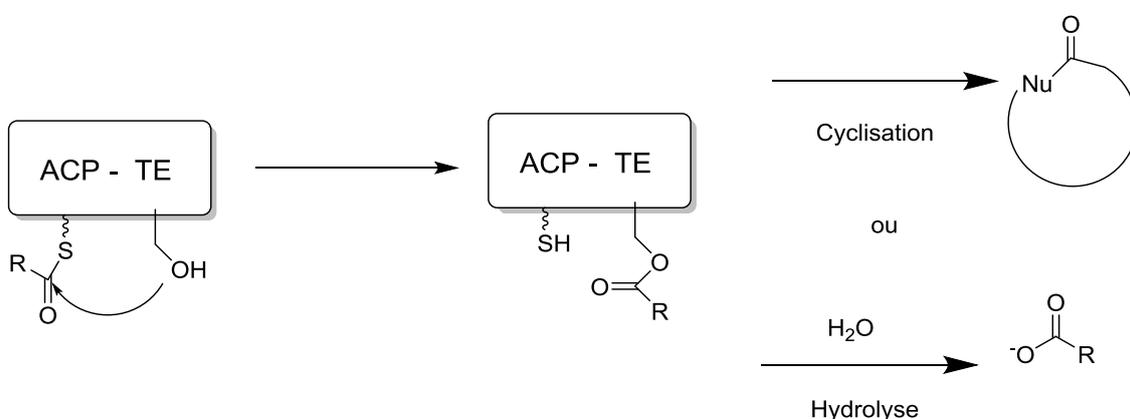


Figure 7. Mécanisme d'un domaine TE

Il existe aussi d'autres mécanismes de terminaison ne faisant pas intervenir de domaine TE mais d'autres domaines présents sur une PKS. Il peut y avoir un domaine

réductase qui va permettre la libération du composé à la fin de l'élongation. C'est le cas dans la biosynthèse du xenovulène<sup>49</sup>. Il existe aussi un domaine particulier de TE appelé TE/CLC (Claisen-like cyclase). Ce domaine va catalyser une condensation de Claisen qui va pouvoir former un nouveau cycle et libérer le composé. C'est le cas dans la biosynthèse de la naphthopyrone<sup>50</sup>. Tous ces mécanismes font intervenir un domaine en *cis* (présent sur la PKS) au niveau C-terminal de la protéine.

Il arrive que de tels domaines de terminaison ne soient pas présents dans une voie de biosynthèse. C'est alors une enzyme en *trans*, indépendante qui réalisera la terminaison de l'élongation. Il existe, cependant, une très grande variété d'enzymes pouvant être impliquées dans cette étape de terminaison. Parmi ces enzymes, on trouve les lactamases, comme dans la voie de biosynthèse de l'actrochryson<sup>51</sup> ou les monooxygénases, comme dans la voie de biosynthèse de l'aurafuron<sup>52</sup>. Il arrive aussi que l'enzyme responsable de la terminaison ne soit pas encore identifiée pour une voie de biosynthèse donnée<sup>53</sup>.

Il existe aussi un autre type de domaine thioestérase appelé Thioestérase de type II (TE II). Les domaines TE II ont un rôle de correcteur dans les PKSs de type I<sup>54,55</sup>. Les gènes codant pour ces protéines sont souvent présents dans des clusters de gènes faisant intervenir des PKSs. Ces domaines TE II doivent « nettoyer » les domaines ACP d'une PKS de type I en hydrolysant les composés bloquant la voie de biosynthèse qui peuvent s'y trouver, permettant ainsi à la PKS de retrouver sa fonction. En effet, il arrive que la réaction de condensation ne s'effectue pas correctement. A la place, une simple réaction de décarboxylation a lieu et elle produit un résidu ne portant plus qu'un seul groupement thioester fixé sur le domaine ACP comme présenté sur la Figure 8. Ce résidu ne peut pas servir pour l'élongation d'un polycétide et il bloque donc l'élongation pour des réactions suivantes. Les domaines TE II vont ainsi permettre à la PKS de retrouver sa fonction et l'enzyme est ainsi plus processive. Des études récentes montrent que les domaines TE II peuvent aussi servir de domaine de terminaison en *trans* comme dans la voie de biosynthèse de la Nanchangmycine<sup>56</sup>.

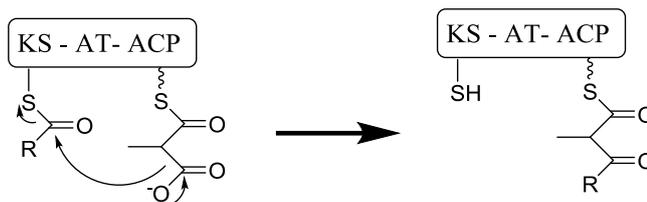
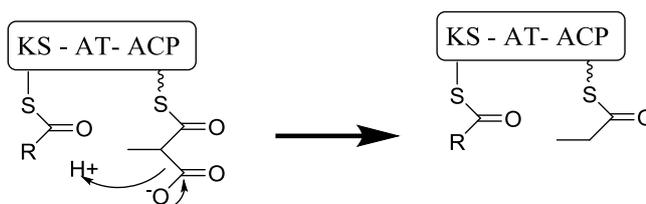
**A****B**

Figure 8. A: mécanisme classique de condensation , B: Condensation erronée

#### I.3.2.4) Les « tailoring » enzymes

La diversité des structures produites par les PKSs ne peut pas s'expliquer seulement par les domaines que nous avons déjà présentés. Il existe une grande variété d'enzymes qui agissent sur les produits de PKSs aussi bien au cours de l'élongation qu'après la terminaison. Ces domaines peuvent se trouver sur la PKS ou être des protéines indépendantes même si les gènes codant pour ces enzymes se retrouvent la plupart du temps dans le cluster de la voie de biosynthèse. Ces enzymes sont appelées « tailoring » enzymes<sup>57</sup>. Une grande variété d'enzymes appartient à la catégorie des « tailoring » enzymes. Nous pouvons trouver des enzymes catalysant des glycosylations, des cyclisations, des halogénations, des méthylations, des oxydations (déshydrogénations, époxydations, hydroxylations)<sup>57,58</sup>.

Par exemple, dans la voie de biosynthèse de l'érythromycine, le passage de la 6-déoxyérythronolide B à l'érythromycine est dû à l'intervention de ce type d'enzymes. Il y a d'abord deux hydroxylases régiospécifiques qui interviennent. Elles vont réaliser

l'hydroxylation de la 6-déoxyérythronolide B au niveau du carbone 6 puis au niveau du carbone 12<sup>59,60</sup>. Différentes glycosylations vont ensuite avoir lieu pour donner l'érythromycine<sup>61</sup>.

### **I.3.3) Les PKSs de type II et de type III**

Contrairement au type I, les domaines des PKSs de type II ne sont pas organisés en module. Ils sont apparentés au FAS de type II. Ici, chaque domaine est sur une protéine à part, monofonctionnelle<sup>62,63</sup>. Les PKSs de type II sont toutes itératives et chaque domaine catalyse une réaction. Par exemple, pour une voie donnée, il n'y aura qu'un seul domaine KS.

La fonction et la structure des PKSs de type III diffèrent considérablement de celles des PKSs de type I et de type II<sup>64</sup>. Il n'y a pas de domaine ACP chez les PKSs de type III. C'est une seule enzyme homodimérique qui va se charger des réactions de condensation de façon itérative. Les PKSs de type III sont donc de taille modeste (40-45 kDa). Ce dernier type est de plus en plus étudié à cause de leur relative simplicité.

### **I.3.4) Les PKSs chez les cyanobactéries**

Chez les cyanobactéries, de nombreuses voies de biosynthèse des métabolites secondaires sont médiées par des PKSs. Une étude récente des différents génomes connus de cyanobactéries a montré que sur 89 cyanobactéries, 262 clusters de gènes de voies de biosynthèse présentent des gènes de PKSs et 290 clusters de gènes de voies de biosynthèse présentent des gènes de NRPSs<sup>65</sup>.

Plusieurs de ces clusters possèdent des gènes de PKSs de type I et de PKSs de type III. Il est intéressant de noter que chez certains clusters, différents types de PKSs interviennent dans la même voie de biosynthèse. Chez 100 de ces clusters, il existe des PKSs reliées à des modules de NRPSs ce qui donne des hybrides de PKSs et de NRPSs. Les voies de biosynthèse de ces hybrides font intervenir des modules de PKSs et des modules de NRPSs. Lorsque les deux protéines sont distinctes, les composés synthétisés sont transférés d'une enzyme à l'autre<sup>66</sup>. Il arrive aussi que les modules PKSs et NRPSs se trouvent sur la même protéine. Ces hybrides permettent d'ajouter des acides aminés sur des polycétides ou à l'inverse, d'ajouter des monomères acyl-CoA sur des peptides. Ceci permet d'accroître la diversité structurale des métabolites secondaires qui peuvent être produits par les cyanobactéries.

Plusieurs composés, dont des cyanotoxines, sont les produits de ces voies de biosynthèse hybrides chez les cyanobactéries comme la microcystine<sup>67</sup>, l'anatoxine-a<sup>68</sup>, la cylindrospermopsine<sup>69</sup> et l'apratoxine<sup>70</sup>.

#### **I.4) Les nonribosomal peptide-synthetases**

Il existe de fortes analogies entre les PKSs et les NRPSs. Elles se comportent de façon similaire et ont la même logique de fonctionnement. Les NRPSs possèdent trois domaines essentiels, i) un domaine Peptidyl Carrier Protein (PCP) qui a un rôle analogue au domaine ACP des PKSs ; ii) un domaine de condensation (C) analogue du domaine KS qui permet l'élongation de la chaîne peptidique ; iii) et un domaine d'adénylation (A) analogue au domaine AT qui sélectionne l'acide aminé à incorporer et l'active sous forme d'acide aminé-AMP<sup>13</sup>. Le fonctionnement d'une NRPS portant ces domaines est présenté sur la Figure 9.

Les NRPSs présentent également des domaines optionnels tels que le domaine d'épimérisation qui permet de transformer un acide aminé L en son isomère, le domaine de cyclisation permet la cyclisation des acides aminés sérine, cystéine (pour former des thiazolines) et la thréonine (pour donner des oxazolines), les domaines de méthylation et de formylation sont responsables, respectivement, de l'addition d'un groupement méthyle et d'un groupement formyle.

La terminaison des voies de biosynthèse des NRPs fait, elle aussi, intervenir des domaines TE I, avec un mécanisme similaire à celui rencontré chez les PKSs.

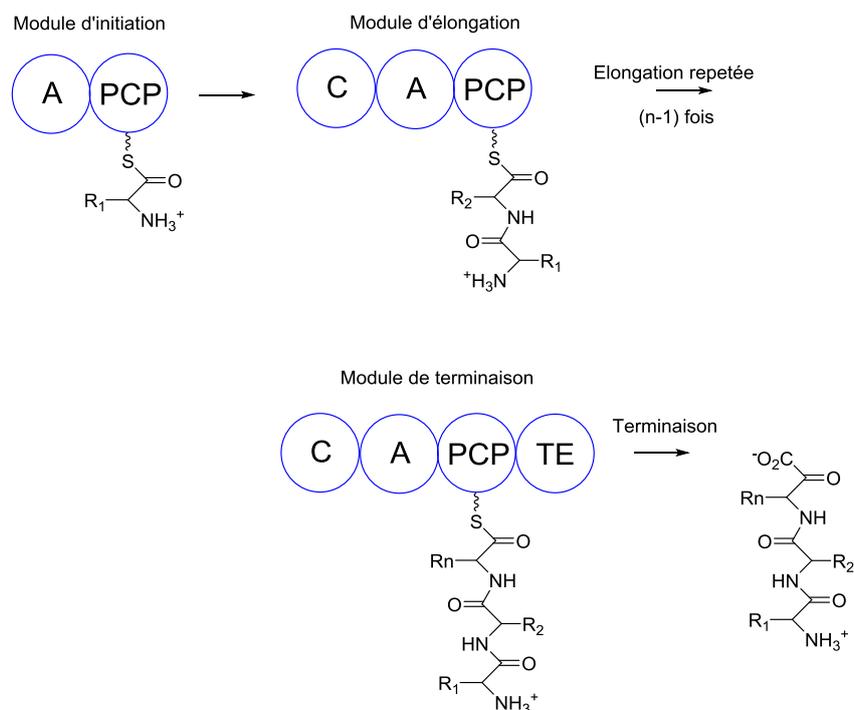


Figure 9. Fonctionnement d'une NRPS

## I.5) Les toxines de cyanobactéries

De nombreux métabolites secondaires produits par les cyanobactéries sont des toxines pour les animaux supérieurs<sup>11</sup>. La présence de ces toxines dans les eaux douces représente un risque pour la vie animale et humaine<sup>71,72</sup>, surtout lors de l'apparition d'efflorescence. Une cyanobactérie peut produire plusieurs toxines. Ces toxines peuvent être classées en quatre catégories : dermatotoxines, hépatotoxines, cytotoxines et finalement neurotoxines.

Les cyanotoxines les plus connues sont : les cytotoxines avec la microcystine et la cylindrospermopsine, les neurotoxines avec l'anatoxine-a, l'homoanatoxine-a et la saxitoxine. Il est intéressant de noter que ces toxines présentent toutes des analogues. Cela est dû à la grande variété de structures possibles provenant des différents mécanismes enzymatiques que nous avons présentés auparavant.

La microcystine est la toxine de cyanobactérie la plus connue et la plus étudiée. Elle cible principalement le foie. Cet hydride PK/NRP inhibe les protéines phosphatases 1 et 2A<sup>73</sup>. Il existe un grand nombre d'analogues de la microcystine (plus de 90 répertoriés). Cette toxine est produite par des cyanobactéries appartenant aux genres *Microcystis*, *Anabaena*, *Planktothrix* et *Nostoc*<sup>74</sup>.

La cylindrospermopsine, qui a elle un effet cytotoxique, est une cyanotoxine de type alcaloïdes. Il existe deux analogues connus pour cette toxine, la 7-déoxycylindrospermopsine<sup>75</sup> et la 7-épi-cylindrospermopsine<sup>76</sup>. La cylindrospermopsine est produite par des cyanobactéries appartenant aux genres : *Umezakia*<sup>77</sup>, *Aphanizomenon*<sup>78,79</sup>, *Raphidiopsis*<sup>80</sup>, *Anabaena*<sup>81</sup>, et *Lyngbya*<sup>82</sup>. La biosynthèse de cette toxine a été étudiée par l'équipe<sup>83,84,69</sup>.

La saxitoxine, initialement nommée PSP (pour « Paralytic Shellfish Poisons ») est un alcaloïde. C'est la cyanotoxine la plus létale avec une DL<sub>50</sub> (dose létale 50%) de 10 µg/kg en injection chez la souris<sup>85</sup>. Elle bloque l'activité des canaux sodiques à la jonction neuro-musculaire provoquant la paralysie musculaire conduisant à une mort par asphyxie<sup>86</sup>. La saxitoxine et ses analogues sont produits par des cyanobactéries appartenant aux genres *Aphanizomenon*, *Anabaena*, *Cylindrospermopsis* et *Lyngbya*<sup>87</sup>.

L'anatoxine-a étant le sujet de cette étude, nous allons nous y intéresser plus en détail.

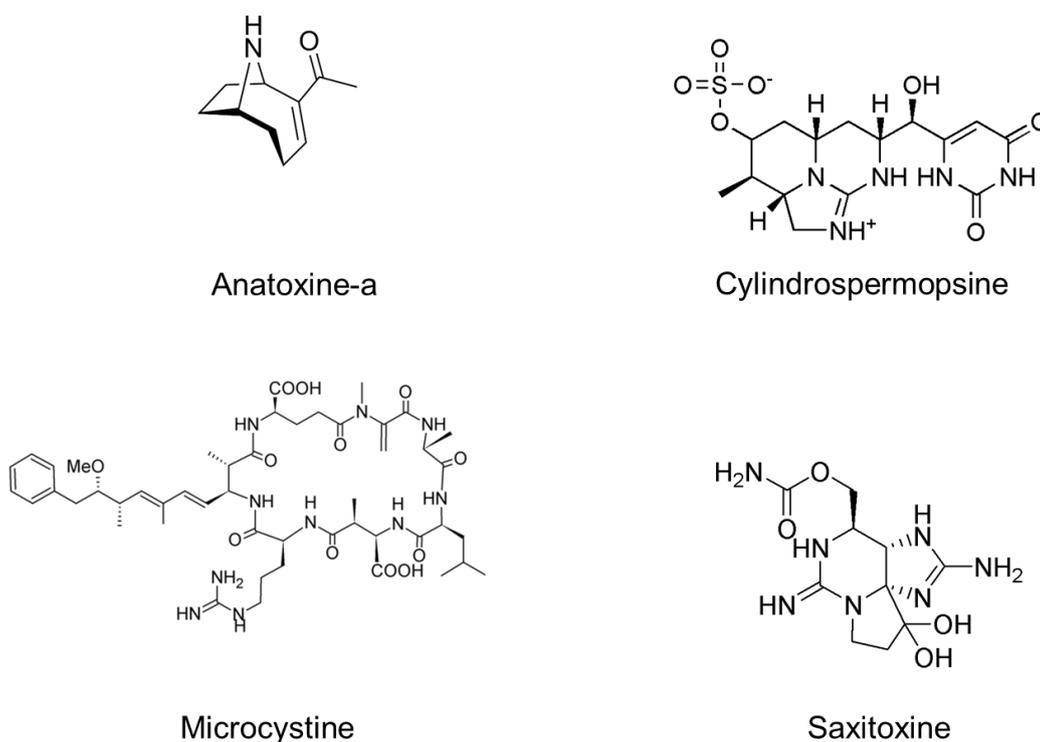


Figure 10. Différentes structures des cyanotoxines les plus connues

## I.6) L'anatoxine-a

L'anatoxine-a est un puissant agoniste du récepteur nicotinique de l'acétylcholine<sup>88</sup> qui entraîne une paralysie des muscles dont les muscles respiratoires et, donc, une mort rapide par asphyxie. L'anatoxine-a est d'ailleurs aussi connue sous le nom de Very Fast Death Factor en anglais ce qui atteste de sa létalité. La  $DL_{50}$  de cette toxine est de 250  $\mu\text{g}/\text{kg}$  en injection chez la souris<sup>89,90</sup>. L'anatoxine-a a été isolée dans les années soixantes par Gorham et ses collaborateurs à partir d'*Anabaena flos-aquae*, une souche provenant d'une efflorescence ayant empoisonné du bétail au Canada<sup>91</sup>. Sa structure a ensuite été résolue par des méthodes spectroscopiques et par synthèse chimique<sup>92</sup>. Son effet, en tant que neurotoxine, a été étudié par Carmichael et al.<sup>93</sup> ainsi que par Spivak et al.<sup>94</sup>. Il existe des analogues de cette toxine comme l'homoanatoxine-a, qui est la plus connue<sup>95</sup>, la dihydroanatoxine-a<sup>96,97</sup> ainsi que l'époxyanatoxine-a dont la stéréochimie n'a pas été déterminée<sup>97</sup>. La structure de ces composés est présentée sur la Figure 11.

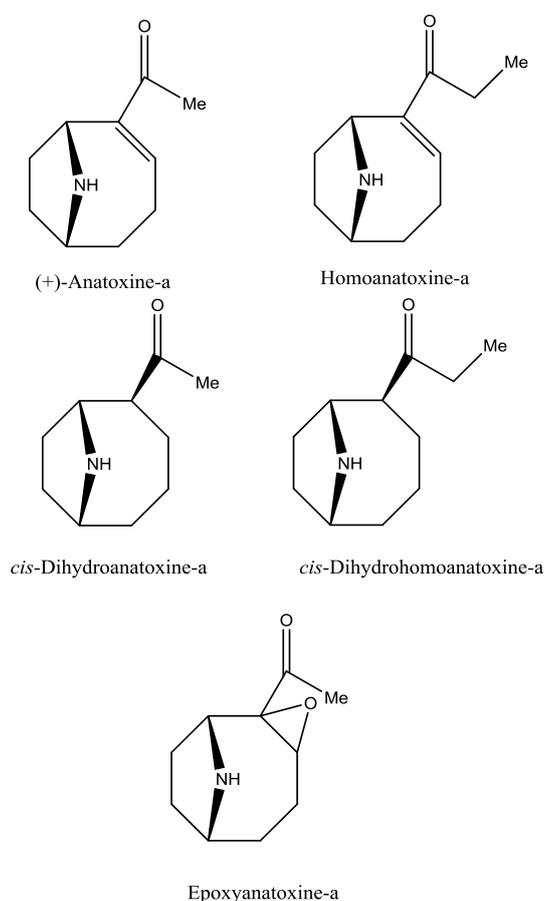


Figure 11. Structure de l'anatoxine-a et de ses analogues

Notre équipe de recherche s'est intéressée à la biosynthèse de ces alcaloïdes chez les cyanobactéries après des épisodes d'empoisonnement canins, provoqué par ces toxines, dans la rivière du Tarn en France<sup>72</sup>. L'équipe a décidé de séquencer le génome d'une cyanobactérie productrice de cette toxine de façon à pouvoir déterminer les gènes responsables de la biosynthèse de l'anatoxine-a. Le choix de l'équipe s'est porté sur *Oscillatoria* sp. PCC 6506 car elle produit de l'anatoxine-a et de l'homoanatoxine-a dans un rapport 2% d'anatoxine-a et 98% d'homoanatoxine-a<sup>98</sup>. C'est aussi une cyanobactérie facilement cultivable en laboratoire<sup>99</sup>.



Figure 12. Micrographie de bactéries de la souche *Oscillatoria* sp. PC 6506 cultivée au laboratoire (grossissement  $\times 400$ )

L'équipe a identifié un cluster de gènes, appelé *ana* par l'équipe, responsable de la biosynthèse de l'anatoxine-a et de l'homoanatoxine-a<sup>100</sup>. Ceci a permis de définir de façon plus précise les gènes du cluster et leurs fonctions. Ces fonctions ont été attribuées par comparaison de séquence avec des gènes existants, comme montré dans la Figure 13. On notera la présence de trois gènes codant pour des PKSs de type I parmi les gènes présents dans ce cluster<sup>99</sup>. L'équipe a aussi pu, par ce travail, identifier un cluster *cyn*, responsable de la biosynthèse de la cylindrospermopsine, présent chez cette souche<sup>84</sup>.



| Gène        | Fonction putative                    | Gène        | Fonction putative       |
|-------------|--------------------------------------|-------------|-------------------------|
| <i>anaI</i> | Protéine d'efflux                    | <i>anaD</i> | Domaine ACP             |
| <i>anaJ</i> | Cyclase ?                            | <i>anaE</i> | PKS de type I modulaire |
| <i>anaA</i> | Thioestérase de type II              | <i>anaF</i> | PKS de type I modulaire |
| <i>anaB</i> | Oxidase de prolyl-ACP                | <i>anaG</i> | PKS de type I modulaire |
| <i>anaC</i> | Protéine d'adénylation de la proline | <i>anaH</i> | Transposase             |

Figure 13. Le cluster *ana* responsable de la biosynthèse de l'anatoxine-a

## I.7) Biosynthèse de l'anatoxine-a

### I.7.1) Présentation de la biosynthèse

L'analyse bioinformatique du cluster *ana* chez *Oscillatoria* sp. PCC 6506, des expériences d'incorporations isotopiques à partir de différents acides aminés marqués et en se basant sur la règle de colinéarité des gènes de PKSs ont permis d'établir un schéma réactionnel de la biosynthèse de l'anatoxine-a<sup>101</sup>. Cette voie de biosynthèse proposée par l'équipe est montrée sur la Figure 14.

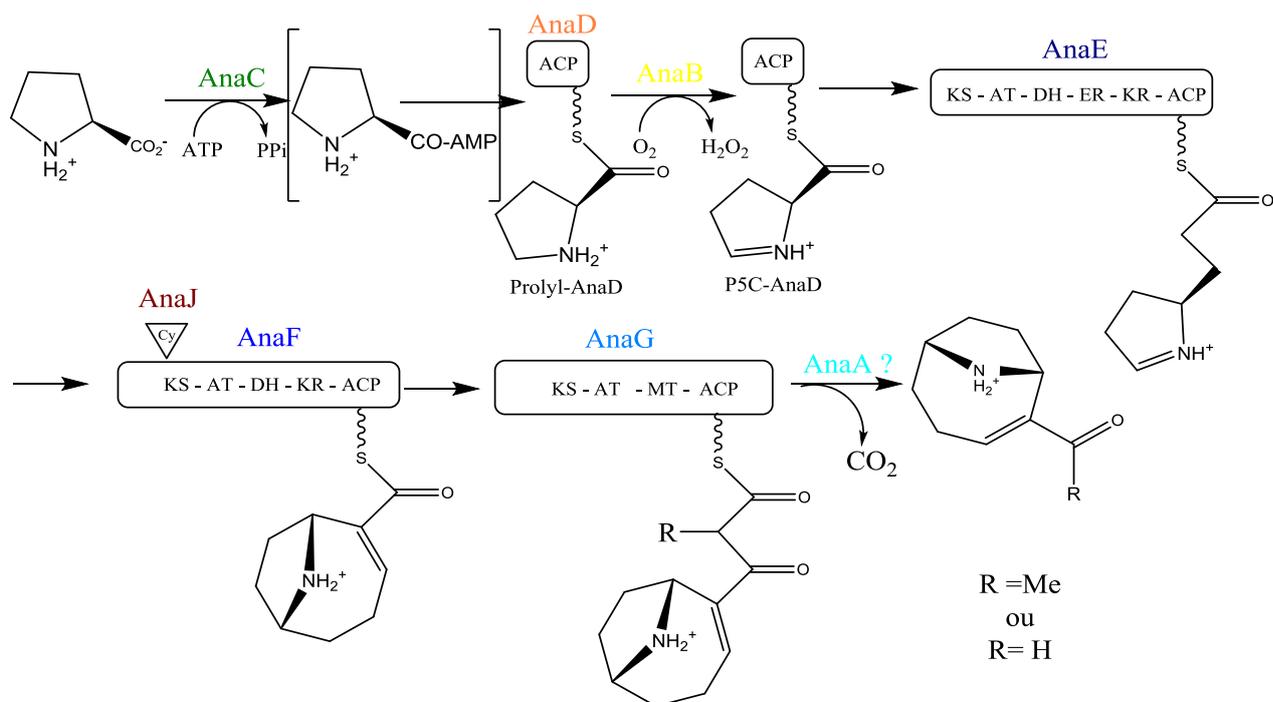


Figure 14. Schéma de biosynthèse de l'anatoxine-a (R=H) ou de l'homoanatoxine-a (R=Me)

Le composé de départ de cette voie de biosynthèse est la proline. Les trois premières étapes de la voie de biosynthèse vont servir à former l'intermédiaire qui pourra être chargé sur AnaE. La première étape est l'activation de la L-proline par AnaC avec un mécanisme similaire à celui utilisé par les domaines d'adénylation chez les NRPSs. La proline activée est ensuite chargée sur l'holo-AnaD pour former le prolyl-AnaD. AnaB va, ensuite, oxyder le cycle pyrrolidine pour former une dehydropyrrolidine selon le mécanisme présenté sur la Figure 15.

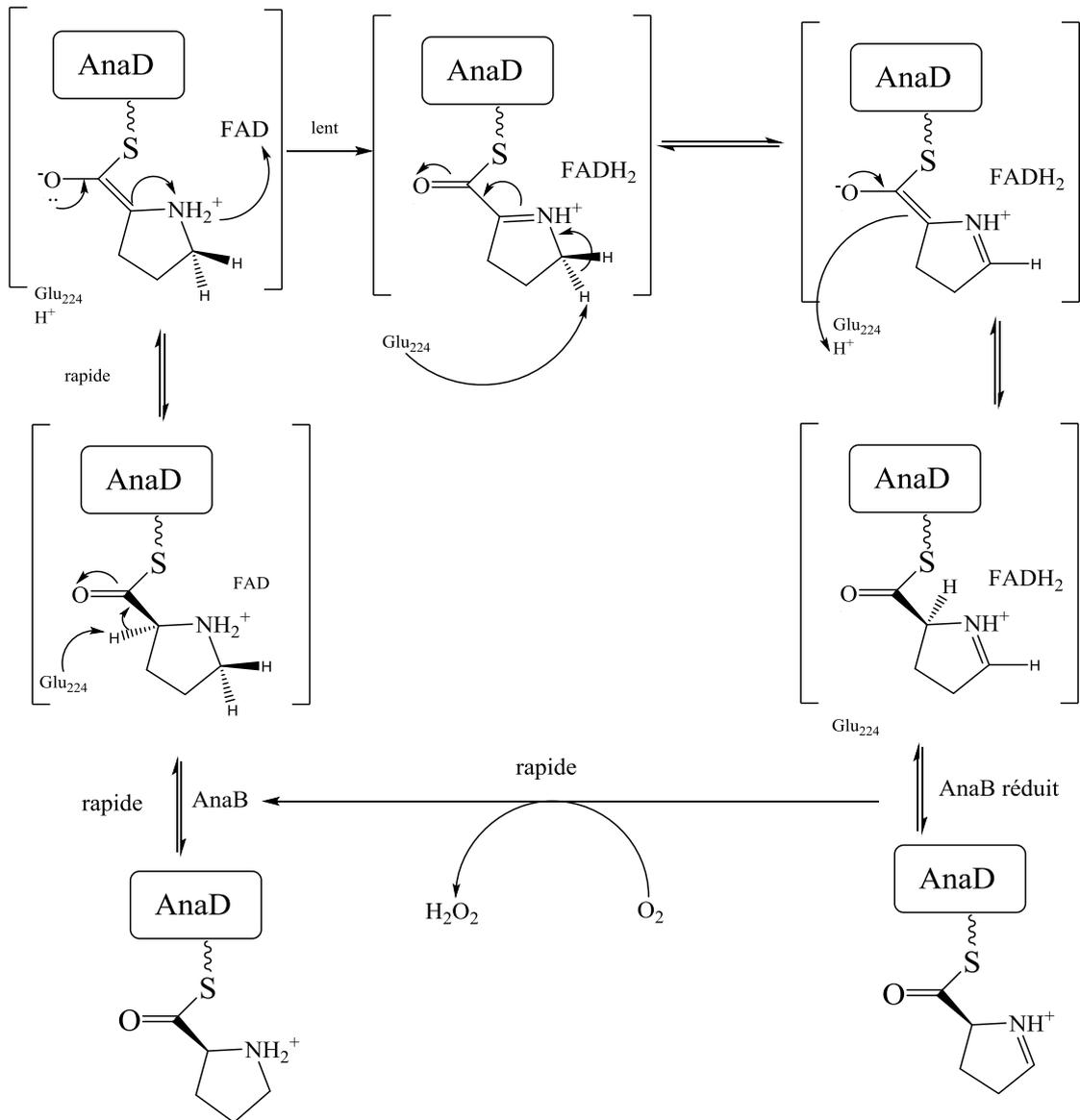


Figure 15. Mécanisme de la réaction catalysée par AnaB

Les trois premières étapes de ce schéma ont été reconstituées *in vitro* par l'équipe, avec confirmation des fonctions prédites et suivi des carbones marqués de la proline<sup>101</sup>. L'équipe a aussi pu obtenir la protéine AnaB sous forme de cristaux en collaboration avec Karine Moncoq à l'IBPC. Ceci a permis de déterminer la structure 3D de la protéine et de confirmer sa fonction<sup>102</sup>.

L'intermédiaire va, ensuite, être transféré sur AnaE pour que deux carbones soient ajoutés à partir du malonyl-CoA comme montré sur la Figure 16.

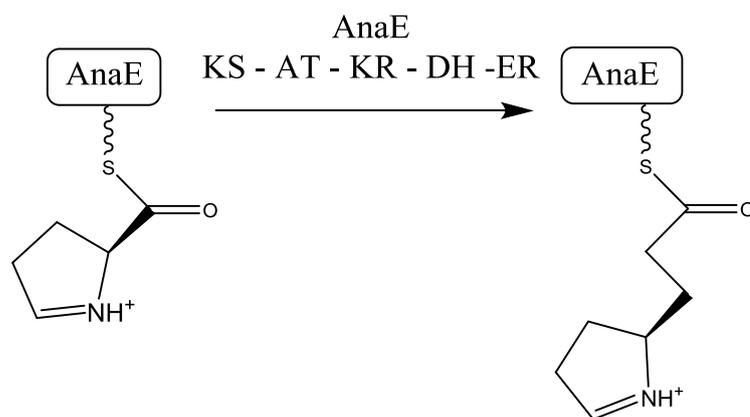


Figure 16. Elongation de la chaîne sur AnaE

L'intermédiaire sera ensuite transféré à AnaF pour l'élongation et la cyclisation. Nous pensons que c'est AnaJ qui se charge de la cyclisation selon le mécanisme proposé sur la Figure 17. En effet, cette protéine présente des similarités avec le domaine Cyclase de StiJ chez *Stigmatella aurantiaca*. Cette enzyme aurait un rôle similaire dans la biosynthèse de la stigmatelline<sup>103</sup>. L'intermédiaire est ensuite transféré sur AnaG et nous allons nous intéresser aux dernières étapes de cette biosynthèse.

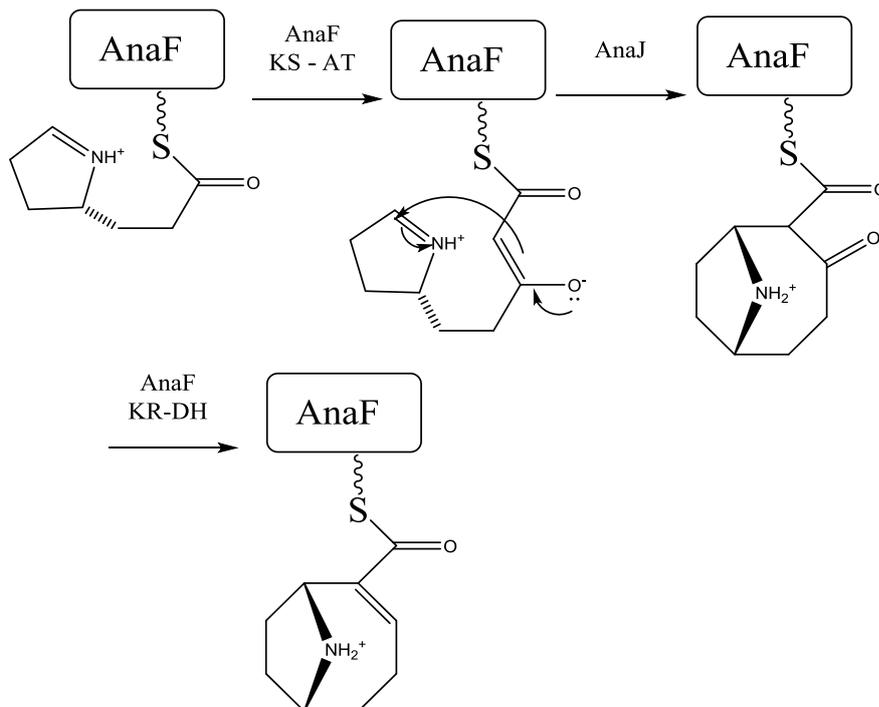


Figure 17. Mécanisme proposé pour la cyclisation sur AnaF

### I.7.2) Les dernières étapes de la biosynthèse de l'anatoxine-a

L'équipe s'intéresse maintenant aux étapes catalysées par les PKSs et en particulier à celle catalysée par AnaG comme présenté dans la Figure 18. En effet, AnaG présente certaines différences par rapport au fonctionnement classique des PKSs de type I. 1) Il est possible que deux domaines de fonctions inconnues se trouvent entre le domaine AT et MT ainsi qu'entre le domaine MT et ACP (voir chapitre II), 2) AnaG présente des différences importantes dans son organisation chez certaines souches (voir chapitre II), 3) AnaG est la dernière PKS de la voie de biosynthèse de l'anatoxine-a. Elle devrait donc contenir un domaine TE I, responsable de l'hydrolyse du produit final, mais ce domaine est absent chez AnaG. Une autre enzyme doit donc effectuer cette hydrolyse. L'équipe a ainsi trouvé dans le cluster *ana* un gène codant pour une TE II appelée AnaA.

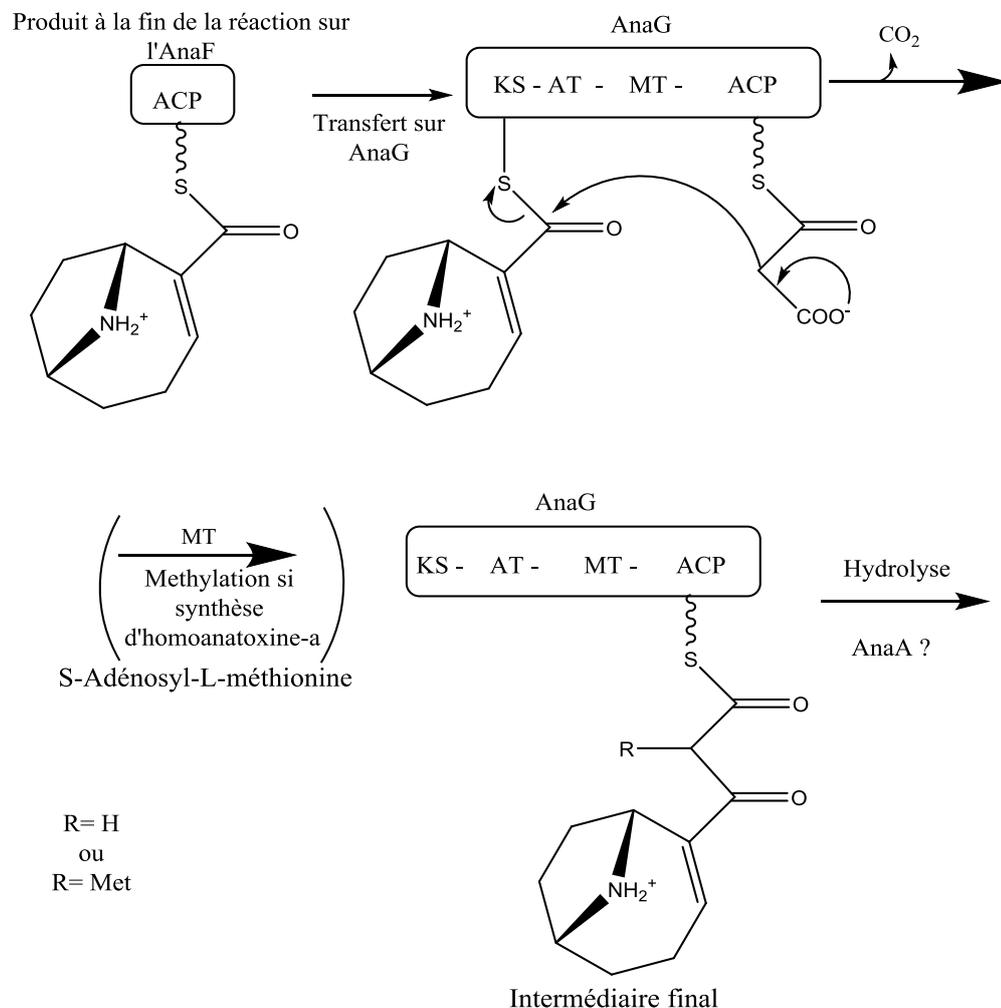


Figure 18. Etapes catalysées sur AnaG

Pour la voie de biosynthèse de l'anatoxine-a, nous nous posons donc la question de savoir si AnaA ne peut pas hydrolyser l'intermédiaire final de la voie de biosynthèse après l'élongation. Si ce n'est pas le cas, nous ne voyons pas d'autres enzymes dans le cluster *ana* qui puisse effectuer cette réaction.

Une fois que le composé chargé sur AnaG est hydrolysé, nous nous trouvons en présence du carboxy-anatoxine-a. Pour obtenir l'anatoxine-a, il faut que le composé subisse une décarboxylation comme montré sur la Figure 19. Nous ne savons pas si cette décarboxylation est catalysée par une enzyme. Si la réaction se déroule spontanément et que la réaction de décarboxylation est assez lente, nous pouvons penser qu'il y a accumulation du carboxy-anatoxine-a. Cet intermédiaire peut donc être détecté par des techniques analytiques. Cependant, si la catalyse enzymatique est assez lente, il se peut aussi qu'il y ait aussi une accumulation du carboxy-anatoxine-a et donc que sa détection soit possible. Une équipe néo-zélandaise a utilisé la spectrométrie de masse pour suivre la présence de ce précurseur dans une culture de cyanobactéries<sup>104</sup>. Il faut cependant rester prudent car seuls les spectres de masse en ionisation à électrospray (ESI) ont été obtenus. Sur ces spectres, seul l'ion MH<sup>+</sup> a été détecté et l'équipe ne possède pas de standards pour la carboxy-anatoxine-a. Il n'y a pas eu d'étude mécanistique de cette réaction. Le composé n'a pas, non plus, été isolé et sa stabilité ne peut donc pas être étudiée. Aucune enzyme connue appartenant au cluster *ana* ne peut catalyser cette réaction à notre connaissance.

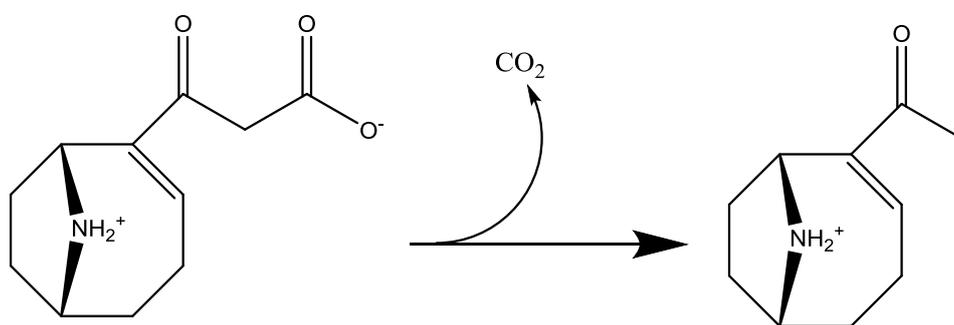


Figure 19. Mécanisme proposé pour la décarboxylation de la carboxy-anatoxine-a

## I.8) Objectifs du projet de recherche

Lors de mon doctorat, notre objectif était d'obtenir un maximum d'informations sur les étapes finales de la voie de biosynthèse de l'anatoxine-a et de l'homoanatoxine-a. Pour cela, il nous fallait avoir tous les outils en main pour mettre en œuvre cette étude. Cela implique l'obtention et la purification des protéines participant à ces étapes ainsi que le développement d'un analogue de substrat pour les dites enzymes. Nous savions qu'obtenir une PKS est un processus complexe et que nous aurions besoin de mettre au point la méthode de purification. Pour le développement de l'analogue de substrat, nous pouvions nous appuyer sur l'expertise de l'équipe concernant la synthèse chimique de dérivés de l'anatoxine-a.

Une fois ces outils en main, nous nous proposons de déterminer grâce à des méthodes de spectrométrie de masse la nature des produits d'AnaG et s'ils valident le mécanisme proposé pour les dernières étapes de la voie de biosynthèse. Nous pourrions ainsi clarifier le rôle d'AnaA dans la voie de biosynthèse mais aussi vérifier si la décarboxylation de la carboxy-anatoxine-a est spontanée ou médiée par une enzyme.

Nous voulons aussi obtenir plus d'informations sur le cluster *ana* en comparant les différentes séquences de clusters *ana* obtenues chez les souches productrices d'anatoxine-a et de ses dérivés. Ces informations peuvent nous permettre de mieux délimiter le cluster *ana* et de découvrir de nouveaux gènes impliqués dans la voie de biosynthèse pour étoffer notre modèle. Nous pourrions aussi observer la conservation de ces gènes entre les différentes souches productrices d'anatoxine-a et de ses analogues ainsi que leurs positions dans le cluster *ana*.

Nous avons aussi décidé de nous intéresser à un dérivé de l'anatoxine-a, la dihydroanatoxine-a. La publication du génome d'une souche<sup>105</sup>, *Cylindrospermum stagnale* PCC 7417, pouvant produire cette toxine, nous a permis d'avoir une base de travail pour commencer notre étude de cette toxine. Nous allons tenter maintenant de confirmer que cette souche produit bien de la dihydroanatoxine-a mais aussi caractériser la voie de biosynthèse utilisée par la cyanobactérie pour produire cette toxine en utilisant des techniques de spectrométrie de masse et en nous appuyant sur les standards de dihydroanatoxine-a et de dihydrohomoanatoxine-a qui ont été synthétisés par l'équipe.

# Chapitre II : Analyse des clusters *ana* de différentes souches de cyanobactéries

## II.1) Séquençage du cluster *ana* de la souche *Oscillatoria* sp. PCC 6506

Le cluster *ana* a été découvert par l'équipe, pour la première fois, lors du séquençage du génome d'*Oscillatoria* sp. PCC 6506 en 2010<sup>84</sup>. Ce séquençage a été réalisé en utilisant la méthode de séquençage 454. Cette méthode est une technique de séquençage à haut débit qui a été développée dans les années 2000 par Roche Diagnostics<sup>106-108</sup>. Le génome est obtenu, par cette méthode, sous forme de milliers de séquences nucléotidiques de taille et de nombre variables appelées reads. Nous avons obtenu, ici, 677 304 reads pour une taille totale de 164 mégabases ce qui représente une couverture estimée du génome d'*Oscillatoria* sp. PCC 6506 à 24. Ces reads sont ensuite assemblés par des moyens bioinformatiques en fonction de leur identité de séquences pour donner des séquences de nucléotides plus grandes appelées contigs. L'assemblage a été assuré par la société GATC dans le cas d'*Oscillatoria* sp. PCC 6506. Mais l'assemblage en contigs à partir de reads obtenus par le séquençage 454 ne garantit pas une couverture complète du génome<sup>109</sup>. La présence de séquences répétées dans le génome peut notamment entraîner des erreurs d'assemblage.

Grâce au séquençage 454, l'équipe a obtenu la séquence du génome d'*Oscillatoria* sp. PCC 6506 sur 419 contigs. Le génome a ensuite été soumis au logiciel MaGe<sup>110</sup> du LABGeM pour l'annoter automatiquement. Le logiciel MaGe a détecté des gènes liés à des PKSs sur le contig 407 qui est un contig de 28909 bp et représenté sur la Figure 20. Des expériences ont ensuite été réalisées permettant d'identifier ces gènes comme étant lié à la biosynthèse de l'anatoxine-a. Ces expériences ont permis l'identification d'un cluster de gènes appelé *ana* par l'équipe, il a ensuite été réalisé une annotation experte manuelle des gènes du cluster *ana*<sup>84</sup>.

### II.1.1) Comparaison du cluster *ana* des *Oscillatoria* sp. PCC 6506 et 6407

L'équipe s'est intéressée, en parallèle, à la souche *Oscillatoria* sp. PCC 6407. Cette souche produit aussi de l'homoanatoxine-a. Il avait été montré par amplification PCR que le génome de cette souche contenait la séquence *ks2* qui correspond à une partie du domaine KS

d'*anaF*, une séquence 16S et une séquence internal transcribed sequence (ITS) identique à 100% à celles d'*Oscillatoria* sp. PCC 6506<sup>99</sup>. Le fait que les séquences 16S et ITS soient identiques indique que les deux souches sont très proches phylogénétiquement. En 2013, la séquence du génome d'*Oscillatoria* sp. PCC 6407 a été publiée avec, en tout, 112 contigs<sup>105</sup>. Le cluster *ana* chez *Oscillatoria* sp. PCC 6407 se trouve sur le contig 69 qui est pratiquement deux fois plus étendu que le contig 407 de la souche *Oscillatoria* sp. PCC 6506. Le séquençage étant plus récent, il est probable que les reads obtenus soient de plus grande taille et qu'il soit possible de les assembler en des contigs plus grands. Ceci nous a permis de comparer le cluster *ana* de cette souche avec celui d'*Oscillatoria* sp. PCC 6506. Les génomes des souches *Oscillatoria* sp. PCC 6506 et *Oscillatoria* sp. PCC 6407 sont presque identiques. Nous avons comparé les deux génomes sur MaGe et nous avons trouvé quelques différences entre les deux génomes, mais ils sont en effet très proches.

Les séquences en nucléotides des clusters, du gène *anaJ* jusqu'au nucléotide 400 du gène *anaH*, sont identiques chez les deux souches comme montré dans la Figure 20. Le contig 69 de la souche *Oscillatoria* sp. PCC 6407 est, cependant, plus étendu. Nous avons donc décidé d'examiner les séquences bordant le cluster *ana* chez *Oscillatoria* sp. PCC 6407. En 5' du cluster *ana* d'*Oscillatoria* sp. PCC 6407, nous avons découvert un gène que nous avons appelé *anaI* codant pour une protéine de transport qui pourrait être responsable du transport de l'anatoxine-a. En réalisant un BLASTP de sa séquence peptidique, nous avons découvert que c'est une protéine de transport de la famille MATE<sup>111-113</sup>. Nous avons aussi retrouvé ce gène chez *Oscillatoria* sp. PCC 6506 mais il était présent sur le contig 13 du génome d'*Oscillatoria* sp. PCC 6506. Nous nous sommes alors demandé si les clusters *ana* des souches *Oscillatoria* sp. PCC 6506 et *Oscillatoria* sp. PCC 6407 n'étaient pas entièrement identiques.

En 3' du cluster, nous avons constaté que la séquence différait entre les souches *Oscillatoria* sp. PCC 6506 et 6407 à partir d'*anaH*. Nous voulions savoir d'où provenait cette différence et nous voulions analyser cette partie du génome chez *Oscillatoria* sp. PCC 6506. En effet, les génomes des deux souches étant très similaires, nous pensions que des erreurs d'assemblage étaient probables chez la souche *Oscillatoria* sp. PCC 6506.

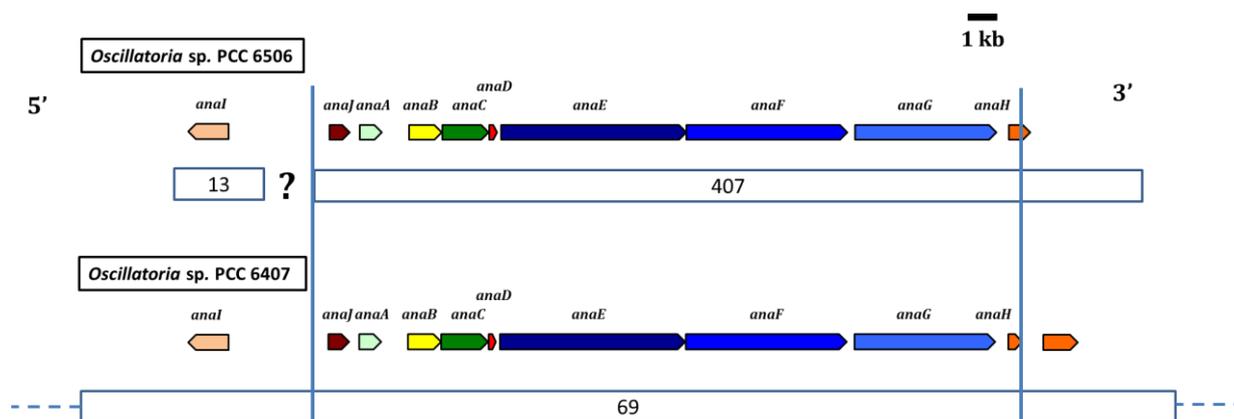


Figure 20. Comparaison des clusters *ana* d'*Oscillatoria* sp. PCC 6506 et PCC 6407 avant le travail de reséquençage par l'équipe. Les nombres en dessous des gènes désignent les numéros des contigs, les deux barres verticales bleues délimitent la zone à l'intérieur de laquelle les séquences nucléotidiques sont identiques

### II.1.2) Réassemblage des contigs chez la souche *Oscillatoria* sp. PCC 6506 en 5' du cluster *ana*

Nous avons d'abord réalisé une PCR sur de l'ADN génomique provenant de la souche *Oscillatoria* sp. PCC 6506 avec une amorce dans le contig 13 (amorce GP12, voir chapitre VI.3) et une amorce dans le contig 407 (amorce GP13, voir chapitre VI.3) d'*Oscillatoria* sp. PCC 6506. Nous voulions vérifier que le contig 13 contenant *anaI* était bien en 5' du contig 407 et en séquençant le produit d'amplification de cette PCR, nous avons bien confirmé la présence d'*anaI* en 5' d'*anaA* chez *Oscillatoria* sp. PCC 6506. Nous avons voulu savoir à ce moment si nous pouvions assembler d'autres contigs en 5' du contig 13. Nous pouvions nous appuyer sur la séquence nucléotidique du contig 69 de la souche *Oscillatoria* sp. PCC 6407 pour réaliser cet assemblage.

Nous avons d'abord réalisé un BLASTN de la séquence du contig 69 de la souche *Oscillatoria* sp. PCC 6407 contre le génome d'*Oscillatoria* sp. PCC 6506. Nous avons ainsi obtenu un enchaînement de contigs provenant du génome d'*Oscillatoria* sp. PCC 6506 ce qui nous a permis de reconstruire a priori la partie en 5' du contig 13 d'*Oscillatoria* sp. PCC 6506. Nous avons donc dessiné des amorces PCR de 25 à 30 paires de bases de façon à vérifier l'enchaînement de ces contigs chez *Oscillatoria* sp. PCC 6506. Nous avons atteint l'extrémité 5' du contig 69 d'*Oscillatoria* sp. PCC 6407 qui correspond à l'extrémité 3' du contig 7 chez *Oscillatoria* sp. PCC 6506. Les PCRs ont fonctionné et nous avons donc pu réassembler du

contig 7 jusqu'au contig 13 comme montré sur la Figure 21. Les séquences obtenues sont identiques à 100% avec celles provenant du génome de la souche *Oscillatoria* sp. PCC 6407. Pour les contigs plus importants en taille comme le contig 13 ou 215, nous avons seulement séquencé les extrémités en supposant qu'il n'y avait pas d'erreurs à l'intérieur des contigs.

### **II.1.3) Réassemblage des contigs chez la souche *Oscillatoria* sp. PCC 6506 en 3' du cluster *ana***

Pour l'exploration de la partie en 3' du cluster *ana* de la souche *Oscillatoria* sp. PCC 6506, nous avons rencontré des difficultés. En effet, nous avons d'abord voulu confirmer l'extrémité 3' du contig 407 de la souche *Oscillatoria* sp. PCC 6506 en amplifiant cette partie par PCR. Nous avons ainsi pu confirmer la partie 3' d'*anaG* et le début d'*anaH*. Néanmoins les PCRs ne fonctionnaient pas au-delà d'*anaH*. Nous avons donc suspecté une erreur dans cette partie du contig 407 de la souche *Oscillatoria* sp. PCC 6506.

Nous avons donc dû dessiner de nouvelles amorces en nous basant sur la séquence du contig 69 de la souche *Oscillatoria* sp. PCC 6407 comme nous l'avions fait pour la partie 5'. Les PCRs ont fonctionné et nous avons observé la présence de deux transposases. La première correspond à la séquence d'*anaH* chez *Oscillatoria* sp. PCC 6407 qui est similaire à celle d'*Oscillatoria* sp. PCC 6506 jusqu'au nucléotide 400. La deuxième correspond à une transposase se trouvant sur le contig 228 d'*Oscillatoria* sp. PCC 6506. Nous avons retrouvé les séquences de ces transposases à deux autres endroits du génome d'*Oscillatoria* sp. PCC 6506. Pour *anaH*, nous l'observons au niveau du contig 176 et 419 avec une identité aux alentours de 90%. La deuxième transposase se retrouve au niveau du contig 418 et 176 avec une identité aux alentours de 89 et 87%.

Nous pensons que c'est la présence de ces deux transposases qui se retrouvent à différents endroits du génome qui est à l'origine de l'erreur d'assemblage chez la souche *Oscillatoria* sp. PCC 6506. L'assemblage des reads chez la souche *Oscillatoria* sp. PCC 6407 ne présente pas ce défaut d'assemblage à ce niveau. Ce séquençage étant plus récent que celui d'*Oscillatoria* sp. PCC 6506, l'assemblage des reads est de meilleure qualité. La version de 2010 du contig 407 est donc erronée et il faudra la remplacer par une nouvelle version.

Nous avons donc pu réassembler plusieurs contigs jusqu'au contig 400 chez la souche *Oscillatoria* sp. PCC 6506 comme nous le montrons dans la Figure 21.

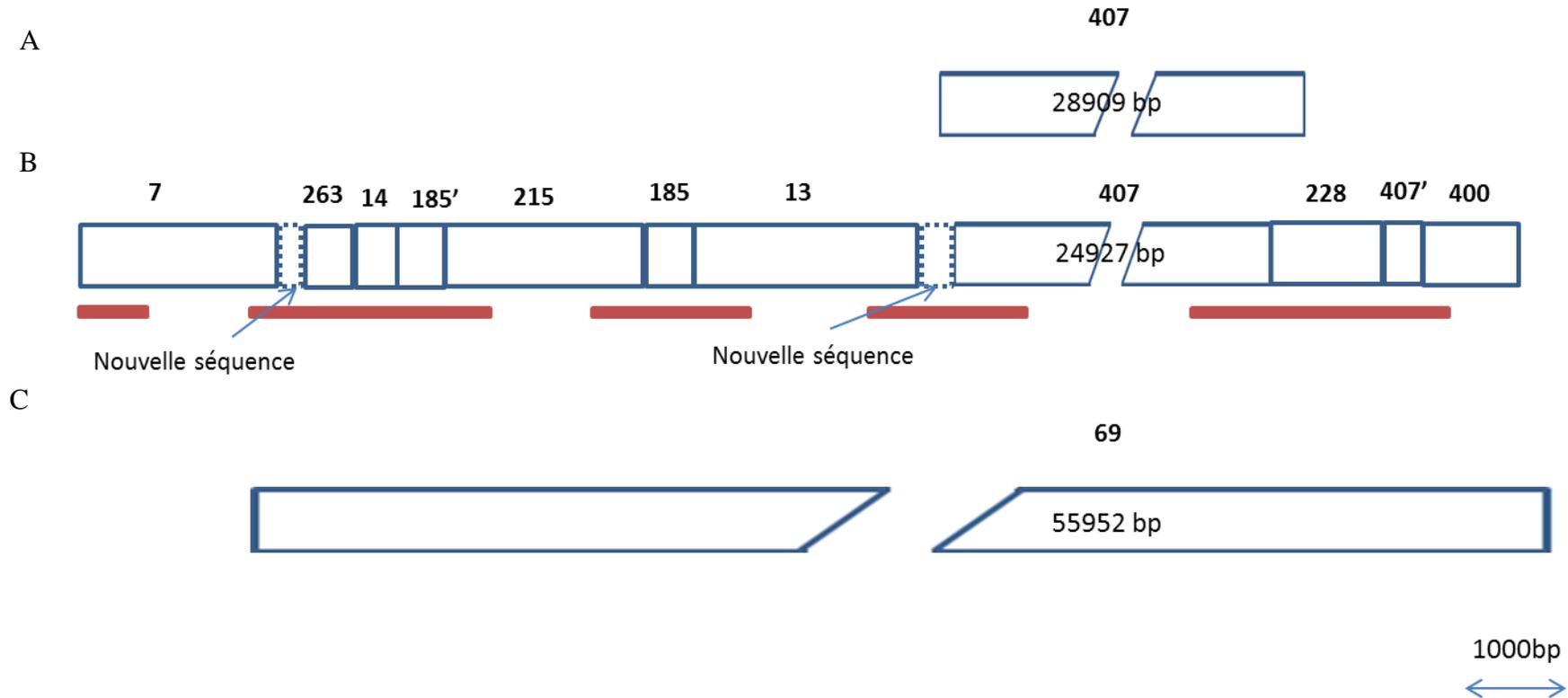


Figure 21. A : Contig 407 chez la souche *Oscillatoria* sp. PCC 6506 (version 2010). B : Assemblage des différents contigs à cet endroit du génome chez la souche *Oscillatoria* sp. PCC 6506 (version 2015). Les barres en rouge représentent les parties qui ont été reséquencées à partir d’amplification PCR. Les nombres en gras correspondent au numéro des contigs dans sa version 2010. Les caractères ‘ indiquent des différences par rapport à la séquence du contig disponible en 2010. Les quadrilatères en pointillés représentent des séquences qui étaient absents du génome dans sa version 2010. C : Contig 69 chez la souche *Oscillatoria* sp. PCC 6407.

#### **II.1.4) Bilan du réassemblage**

Au cours de ce travail, nous avons réussi à assembler plusieurs contigs en un seul contig auquel nous avons ajouté deux séquences qui n'étaient pas présentes dans le génome d'*Oscillatoria* sp. PCC 6506. Nous avons pu définir que les clusters *ana* pour les deux souches *Oscillatoria* sp. PCC 6506 et *Oscillatoria* sp. PCC 6407 s'étendaient d'*anaI* à *anaG* et qu'ils étaient identiques à 100%. De plus, toutes les séquences que nous avons obtenues chez la souche *Oscillatoria* sp. PCC 6506 se sont révélées identiques à 100% avec celles de la souche *Oscillatoria* sp. PCC 6407.

A part *anaI*, aucun autre gène pouvant appartenir au cluster *ana* n'a été découvert en 5' et en 3' du contig 407 chez la souche *Oscillatoria* sp. PCC 6506. Nous avons bien trouvé d'autres gènes mais il ne nous semble pas qu'ils puissent participer à la voie de biosynthèse de l'anatoxine-a (voir liste des gènes annotés en annexe).

## **II.2) Comparaison des clusters *ana* de différentes souches**

Il existe plusieurs souches de cyanobactéries productrices d'anatoxine-a et d'homoanatoxine-a, mais nous ne disposons de la séquence nucléotidique du cluster *ana* que chez 7 souches qui sont listées dans le Tableau 1. La majorité des séquences nucléotidiques de ces clusters ont été obtenues au cours des cinq dernières années<sup>105,114</sup>. Concernant *Raphidiopsis mediterranea* LBRI 48, elle est citée dans deux publications de 2003 et 2004<sup>115,116</sup>, mais une publication récente lui donne le nom de *Cuspidothrix issatschenkoi* LBRI48<sup>117</sup>. Les deux noms de souche désignent la même cyanobactérie.

| Souche  | Origine  | Toxine produite majoritairement | Référence |
|---|----------|---------------------------------|-----------|
| <i>Anabaena</i> sp. strain 37   | Finlande | ATX                             | 114,118   |
| <i>Oscillatoria</i> sp. PCC 6506  | Inconnue | HATX                            | 98,100    |
| <i>Oscillatoria</i> sp. PCC 6407  | USA      | HATX                            | 98,105    |
| <i>Cylindrospermum stagnale</i> PCC 7417  | Suède    | dhATX                           | 119       |
| <i>Cuspidothrix issatschenkoi</i> LBRI48 anciennement<br><i>Raphidiopsis mediterranea</i> LBRI 48 | Chine    | HATX                            | 116,117   |
| <i>Cuspidothrix issatschenkoi</i> RM-6  | Japon    | HATX                            | 117       |
| <i>Cuspidothrix issatschenkoi</i> CHABD3  | Japon    | HATX et ATX non détectées       | 117       |

Tableau 1 : Souches de cyanobactéries productrices d'anatoxine-a et de ses dérivés connues et dont le cluster *ana* a été séquencé.

Actuellement, nous avons accès à la séquence nucléotidique de 7 clusters *ana*. Le premier cluster à avoir été séquencé, en 2009, appartient à *Oscillatoria* sp. PCC 6506<sup>84</sup>. Ont suivi, les clusters d'*Anabaena* sp. strain 37, d'*Oscillatoria* sp. PCC 6407 et de *Cylindrospermum stagnale* PCC 7417. Il faut noter que le génome d'*Anabaena* sp. strain 37 a été séquencé dans son entièreté mais seul le cluster *ana* a été publié. Le génome de *Cylindrospermum stagnale* PCC 7417 a été complètement séquencé et circularisé. Dernièrement, trois séquences nucléotidiques de cluster *ana* ont été publiées sur le NCBI. Ce sont ceux de *Cuspidothrix issatschenkoi* LBRI48, *Cuspidothrix issatschenkoi* CHABD3 et *Cuspidothrix issatschenkoi* RM-6<sup>117</sup>.

Les clusters *ana* de la souche *Oscillatoria* sp. PCC 6506 et de la souche *Oscillatoria* sp. PCC 6407 étant identiques, nous allons seulement comparer les séquences des gènes du cluster *ana* de l'*Oscillatoria* sp. PCC 6506 avec les génomes provenant des souches *Anabaena* sp. strain 37, *Cylindrospermum stagnale* PCC 7417, *Cuspidothrix issatschenkoi* LBRI48, *Cuspidothrix issatschenkoi* CHABD3 et *Cuspidothrix issatschenkoi* RM-6. Pour cela nous avons effectué un alignement de séquence à l'aide du programme BLASTN et BLASTP avec les différentes séquences nucléotidiques et peptidiques obtenues à partir du cluster *ana*. Les résultats sont résumés dans le Tableau 2.

| Gène du cluster <i>ana</i> chez <i>Oscillatoria</i> sp. PCC 6506 | % d'identité chez <i>Anabaena</i> sp. strain 37 | % d'identité chez <i>Cylindrospermum stagnale</i> PCC 7417 | % d'identité chez <i>Cuspidothrix issatschenkoi</i> LBRI48 | % d'identité chez <i>Cuspidothrix issatschenkoi</i> CHABD3 | % d'identité chez <i>Cuspidothrix issatschenkoi</i> RM-6 |
|--|---|--|--|--|--|
| <i>anaA</i>  | 89%/85%   | 92%/86%  | 86%/78%  | 85%/80%  | 85%/78%  |
| <i>anaB</i>  | 85%/86%   | 91%/89%  | 83%/84%  | 83%/84%  | 83%/84%  |
| <i>anaC</i>  | 82%/79%   | 90%/89%  | 83%/82%  | 83%/82%  | 83%/82%  |
| <i>anaD</i>  | 86%/87%   | 90%/90%  | 84%/87%  | 84%/87%  | 84%/87%  |
| <i>anaE</i>  | 86%/85%   | 90%/89%  | 86%/85%  | 86%/85%  | 86%/85%  |
| <i>anaF</i>  | 86%/86%   | 85%/83%  | 86%/85%  | 86%/85%  | 86%/85%  |
| <i>anaG</i>  | 88%/87%   | 85%#/84%   | 88%/87%  | 87%#/87%   | 88%/87%  |
| <i>anaH</i>  | absent  | absent   | absent   | absent   | absent   |
| <i>anaI</i>  | 82%   | 89%/83%  | absent   | 81%*   | absent   |
| <i>anaJ</i>  | 88%/85%   | 85%/74%  | 78%/78%  | 71%/60%  | 78%/78%  |

Tableau 2 : Comparaison des différents gènes *ana* entre les différents clusters. La première valeur correspond à celle obtenus avec un BLASTN et la deuxième avec un BLASTP. \* : identité obtenue avec une couverture de seulement 29% de la séquence. # : identité obtenue sur les parties de séquence présentes dans le cluster

### II.2.1) Synténie du cluster *ana*

Nous avons ensuite établi la synténie des gènes du cluster *ana* avec les informations que nous avons obtenues par la comparaison des différents clusters. Les résultats sont présentés sur la Figure 22.

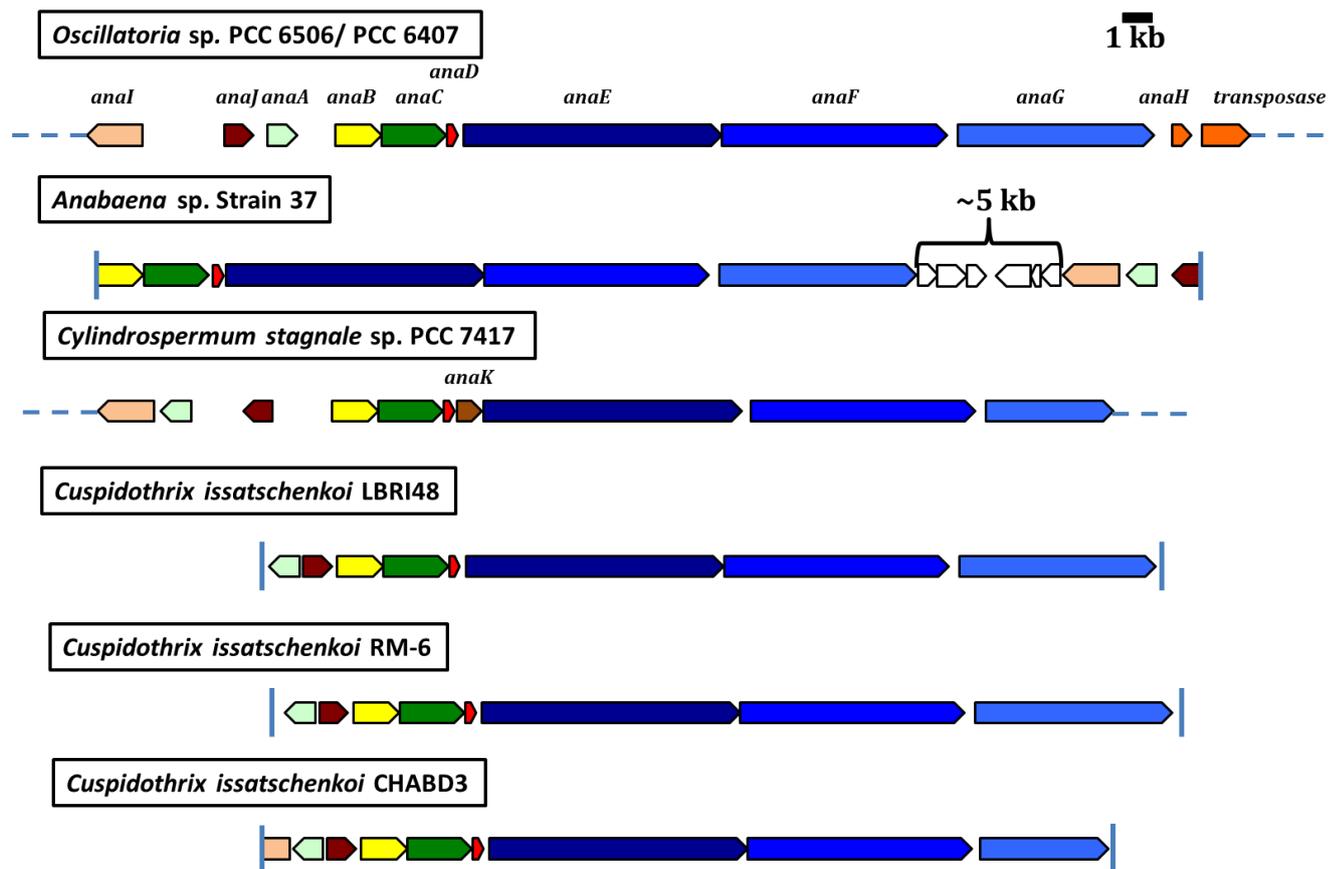


Figure 22. Comparaison de la synténie des clusters *ana* de souches productrices d’anatoxine-a et d’homoanatoxine-a. Les traits en pointillés montrent que nous avons accès aux séquences nucléotidiques encadrant le cluster. Les barres verticales représentent des clusters dont le génome n’est pas disponible et les séquences nucléotidiques disponibles sont donc limitées en taille. L’alignement a été réalisé à partir du gène *anaB*.

Globalement, les gènes sont conservés d’une souche à l’autre et sont très similaires. Les gènes *anaB*, *anaC*, *anaD*, *anaE*, *anaF* et *anaG* sont les plus conservés. Leur synténie est conservée dans toutes les souches analysées, sauf chez la souche *Cylindrospermum stagnale* PCC 7417 qui présente un gène *anaK* entre *anaD* et *anaE*. Nous reviendrons sur ce point plus loin dans ce chapitre. C’est la souche *Anabaena* sp. strain 37 qui présente la disposition de gènes la plus différente des autres souches. En effet, chez elle, *anaJ*, *anaI* et *anaA* sont séparés des autres gènes du cluster *ana* par environ 5 kilobases mais les gènes présents dans cet intervalle ne sont pas liés à la biosynthèse de l’anatoxine-a (voir liste des gènes annotés en annexe)<sup>114</sup>. Cette disposition peut provenir d’un réarrangement des gènes dû à une transposition ou à une possible erreur d’assemblage. En effet, il nous semble étrange qu’une seule souche présente un enchaînement des gènes ainsi modifié. C’est pour cette raison que nous avons pensé à une erreur d’assemblage.

*anaA* et *anaJ* sont présents dans toutes les souches mais leur disposition l'un par rapport à l'autre peut varier d'une souche à une autre. Dans la majorité des cas, ils se trouvent en 5' d'*anaB*. On ne trouve *anaH*, codant pour une transposase, que dans deux souches : *Oscillatoria* sp. PCC 6506 et PCC 6407. Ce gène avait été intégré au cluster *ana* dans le passé par l'équipe à cause de sa proximité avec *anaG* mais nous pouvons nous poser la question de sa place dans le cluster *ana* car une transposase n'a pas de fonction dans la voie de biosynthèse de l'anatoxine-a.

Le cluster *ana* de la souche *Cylindrospermum stagnale* PCC 7417 est particulièrement intéressant car il présente un gène supplémentaire par rapport aux autres clusters. Lors de la comparaison avec le cluster d'*Oscillatoria* sp. PCC 6506, nous avons remarqué qu'il existe un gène supplémentaire, placé entre les gènes *anaD* et *anaE* qui a été nommé par l'équipe *anaK*<sup>68</sup>. L'annotation du génome de la souche *Cylindrospermum stagnale* PCC 7417 indique que le gène *anaK* code pour une réductase putative nécessitant un cofacteur F420<sup>68,120</sup>. Ceci est cohérent avec le fait que la souche *Cylindrospermum stagnale* PCC 7417 produise de la dihydroanatoxine-a (voir chapitre IV).

### II.2.2) Comparaison des séquences peptidiques de la protéine AnaG

Ce n'est pas la seule différence que nous pouvons observer au niveau du cluster. Le gène *anaG* est plus court chez la souche *Cylindrospermum stagnale* PCC 7417 que chez la souche *Oscillatoria* sp. PCC 6506. Nous retrouvons cette séquence raccourcie chez la souche *Cuspidothrix issatschenkoi* CHABD3. En alignant les deux gènes avec BLASTP, nous trouvons une identité de 85% avec une couverture de séquence de 100%. La protéine AnaG obtenue est donc plus courte chez ces deux souches. *Oscillatoria* sp. PCC 6506 est une souche productrice d'homoanatoxine-a et *Cylindrospermum stagnale* PCC 7417 produit de la dihydroanatoxine-a.

Nous regardons la structure de la protéine AnaG plus en détail en utilisant un logiciel de prédiction de domaines (antiSMASH 3.0<sup>121</sup>) pour les clusters de gènes de biosynthèse de métabolites secondaires. Ce logiciel nous permet d'obtenir une prédiction des différents domaines chez une PKS et une estimation de sa position sur la séquence peptidique. Nous avons comparé les séquences peptidiques de la protéine AnaG chez les 7 souches productrices d'anatoxine-a et ses dérivés. Les résultats sont présentés sur la Figure 23.

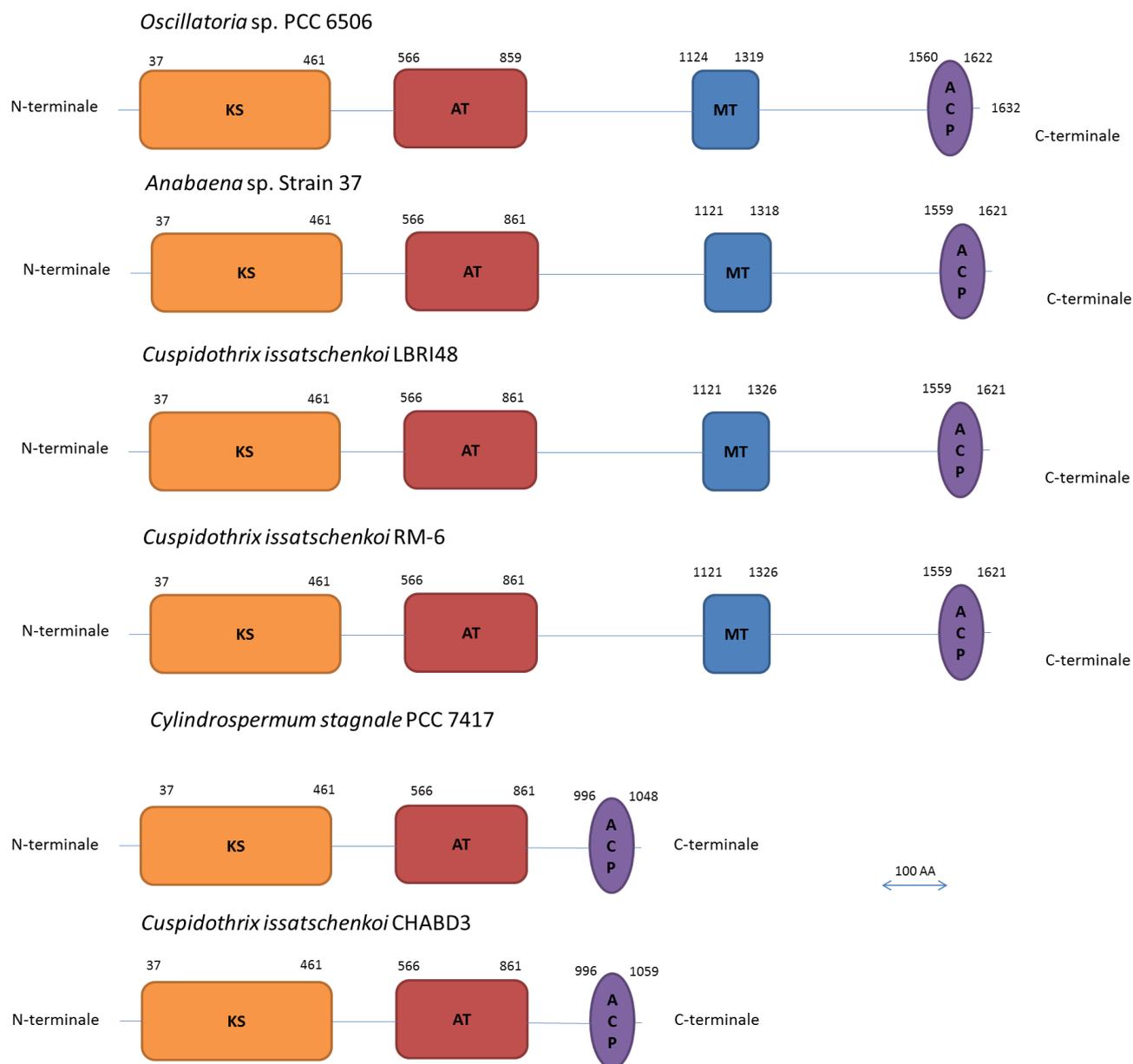


Figure 23. Comparaison des domaines présents dans la protéine AnaG dans les souches productrices d'anatoxine-a et ses dérivés. Les limites des domaines sont issues du logiciel de prédiction de domaine antiSMASH 3.0 et sont donc à envisager avec prudence

Chez les souches *Cylindrospermum stagnale* PCC 7417 et *Cuspidothrix issatschenkoi* CHABD3, nous observons l'absence du domaine MT, se trouvant normalement entre les domaines AT et ACP d'AnaG chez la souche *Oscillatoria* sp. PCC 6506. Ce domaine est probablement en charge de la méthylation du précurseur de l'homoanatoxine-a qui permet d'obtenir l'homoanatoxine-a dans la voie de biosynthèse. La présence de ce domaine ne peut, cependant, pas expliquer à lui seul la production d'homoanatoxine-a par cette souche. L'équipe a prouvé, dans ses travaux précédents, que deux souches ayant une protéine AnaG

identique peuvent produire quasiment exclusivement de l'anatoxine-a pour l'une et de l'homoanatoxine-a pour l'autre comme c'est le cas pour *Oscillatoria* sp. PCC 6506 et PCC 9127<sup>98</sup>. Il est intéressant de noter qu'une souche productrice d'anatoxine-a produira toujours une petite quantité d'homoanatoxine-a. L'inverse est, aussi, vrai pour les souches productrices d'homoanatoxine-a. La régulation de l'activité de méthylation du domaine MT d'AnaG n'est donc pas liée à sa seule présence dans la séquence protéique, nous pouvons imaginer qu'elle est due à d'autres facteurs. Nous pourrions imaginer une compétition cinétique entre le domaine MT d'AnaG et l'hydrolyse du composé chargé sur l'AnaG. Si la méthylation est la plus rapide, nous aurions en majorité de l'homoanatoxine-a et si c'est l'hydrolyse qui est la plus rapide, le composé majoritaire sera l'anatoxine-a.

Nous avons aussi observé une particularité au niveau des séquences peptidiques entre les domaines AT et MT ainsi qu'entre les domaines MT et ACP chez *Oscillatoria* sp. PCC 6506. En effet, le logiciel antiSMASH ne détecte pas de domaines connus de PKSs à ces niveaux de la séquence peptidique de la protéine mais antiSMASH nous permet de délimiter ces domaines. Nous avons donc utilisé un logiciel de prédiction de structure secondaire PSIPRED pour obtenir une prédiction des structures secondaires à ces niveaux<sup>122</sup>, nous avons remarqué la présence d'un grand nombre de brins  $\beta$  et d'hélices  $\alpha$  dans ces séquences. Ceci nous a amené à penser que nous pourrions être face à deux domaines de fonctions inconnues ou en face de vestiges de domaines PKSs. Nous avons donc nommé D1, le domaine potentiel se trouvant entre les domaines AT et MT et D2, le domaine potentiel se trouvant entre les domaines MT et ACP. Nous montrons sur la Figure 24 la prédiction de structures secondaires obtenue pour le domaine D2. Nous avons ensuite réalisé une analyse avec BLASTP sur ces deux séquences peptidiques. Nous avons remarqué que le domaine D2 présentait des identités de séquence avec des domaines KR de PKSs. Ces identités étant assez faibles, nous pouvons penser que ce domaine n'est qu'un vestige.

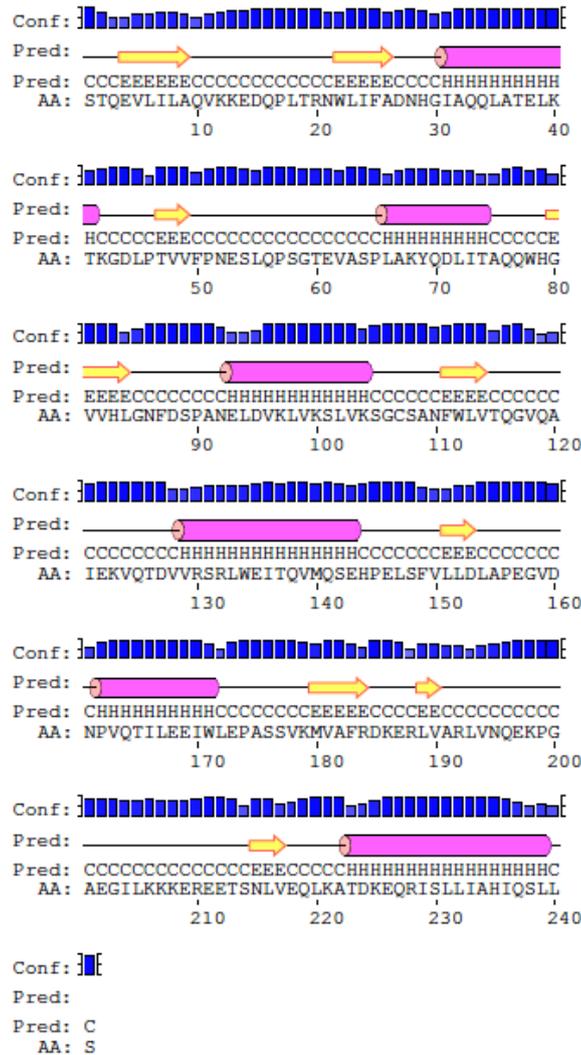


Figure 24. Séquence peptidique du domaine D2 chez la souche *Oscillatoria* sp. PCC 6506 allant de l'acide aminé 1320 à 1559 avec prédiction de structure secondaire obtenue avec le logiciel PSIPRED, les cylindres roses représentent les hélices  $\alpha$  prédites, les flèches jaunes les brins  $\beta$ . Les barres bleues correspondent à la probabilité de la prédiction.

### II.3) Discussion

Nous avons maintenant un contig d'une plus grande taille en termes de séquence nucléotidique contenant le cluster *ana* de la souche *Oscillatoria* sp. PCC 6506. Ce contig sera publié dans la base de données de MaGe. L'élément essentiel que nous a apporté ce travail de reséquençage est la présence d'*anaI* chez la souche *Oscillatoria* sp. PCC 6506 et le fait que nous avons retrouvé ce gène dans la majorité des autres souches. Par ailleurs, l'équipe

envisage actuellement de faire séquencer avec des nouvelles technologies de séquençage le génome entier de la souche *Oscillatoria* sp. PCC 6506 de façon à pouvoir le circulariser.

Nous avons aussi pu confirmer avec nos travaux que les deux clusters *ana* de la souche *Oscillatoria* sp. PCC 6506 et de la souche *Oscillatoria* sp. PCC 6407 sont 100% identiques.

Nous avons comparé plusieurs clusters *ana* provenant de différentes souches de cyanobactéries productrices de toxines en nous basant sur leurs séquences et avons ainsi observé des similarités dans l'organisation de certains gènes. La position des gènes *anaB*, *anaC*, *anaD*, *anaE* et *anaF* est conservée dans tous les clusters observés. Ces gènes pourraient donc être utilisés comme marqueurs dans la recherche de cluster *ana* chez différentes souches de cyanobactéries et c'est déjà le cas pour *anaC*, *anaE* et *anaF*<sup>114,123,124</sup>. La position des autres gènes est plus variable. La séquence nucléotidique des différents gènes est bien conservée entre les différentes souches avec un pourcentage d'identité allant de 71 à 92%.

Concernant *anaI*, nous avons remarqué qu'il y a une partie tronquée de ce gène chez *Cuspidothrix issatschenkoi* CHABD3 mais cette séquence a été obtenue par PCR et gene walking. Il est très probable que le reste de la séquence nucléotidique d'*anaI* se trouve bien chez *Cuspidothrix issatschenkoi* CHABD3. Les séquences provenant de *Cuspidothrix issatschenkoi* RM-6 et de *Cuspidothrix issatschenkoi* LBRI48 ont aussi été obtenues par PCR et gene walking. Donc même si *anaI* en est absent, nous pensons qu'il est probable que ces souches possèdent bien le gène *anaI*. Les trois clusters *ana* des souches *Cuspidothrix issatschenkoi* CHABD3, *Cuspidothrix issatschenkoi* RM-6 et de *Cuspidothrix issatschenkoi* LBRI48 sont très similaires entre eux. Nous obtenons 95% d'identité sur un BLASTN et il serait donc étonnant que ce ne soit pas le cas.

La présence d'*anaK* dans le cluster *ana* de la souche *Cylindrospermum stagnale* PCC 7417 est particulièrement intéressante. En effet, c'est le seul gène que l'on retrouve dans une seule souche et qui a une fonction bien précise dans ce cluster. Il serait intéressant de voir si ce gène est aussi présent chez la souche *Cuspidothrix issatschenkoi* CHABD3.

Avec ces éléments, nous pouvons réfléchir à l'évolution du cluster *ana* chez les cyanobactéries. Il n'y a, cependant, pas assez de clusters *ana* séquencés pour pouvoir reconstituer l'évolution de ce cluster dans les différentes souches. De plus, les phénomènes de

transfert horizontaux rendent la tâche plus complexe. Nous pouvons nous demander si l'évolution du cluster *ana* et celle des différentes souches de cyanobactéries productrices d'anatoxine-a et de ses dérivés sont les mêmes ? Nous savons que c'est déjà le cas pour d'autres clusters de biosynthèses de cyanotoxines où la phylogénie de la séquence 16S n'est pas liée avec la toxicité d'une souche<sup>125</sup>. Pour obtenir une réponse définitive, il nous faut cependant attendre que de nouveaux clusters soient séquencés et que de nouvelles souches soient découvertes dans les prochaines années de façon à pouvoir réaliser un arbre phylogénétique du cluster *ana*.

# Chapitre III : Etude des dernières étapes de la biosynthèse

Plusieurs éléments sont nécessaires pour mener à bien l'étude des étapes finales de la biosynthèse de l'anatoxine-a. Nous avons d'abord besoin des protéines AnaG et AnaA sous forme de protéines pures pour déterminer leurs activités enzymatiques et leur mécanisme d'action. Nous avons aussi besoin d'un analogue du substrat putatif d'AnaG pour étudier la voie de biosynthèse proposée par l'équipe. Pour ce faire, nous avons développé deux approches. Une première approche fait intervenir de la chimie de synthèse qui va nous permettre d'obtenir l'analogue de substrat. Nous allons, ensuite, utiliser une approche biochimique de façon à exprimer les deux protéines chez *Escherichia coli* et finalement les purifier.

## III.1) Approche chimique : synthèse d'un analogue de substrat pour AnaG

### III.1.1) Voie de synthèse chimique de l'anatoxine-a

Après la découverte de l'anatoxine-a, des voies de synthèses chimiques de cette toxine ont rapidement été proposées. Les premiers essais de synthèses à partir de la (-) cocaïne ont été menés par Campbell et al. de façon à confirmer la structure et à déterminer la configuration de l'anatoxine-a<sup>92</sup>. Deux contributeurs majeurs en ce qui concerne la synthèse de l'anatoxine-a dans les années 80 et 90 sont l'équipe de Rapoport<sup>126-129</sup> et l'équipe de Gallagher<sup>95,130</sup>. Une revue, écrite en 1996, résume toutes les voies de synthèse et les différentes approches qui ont été tentées pour obtenir l'anatoxine-a par synthèse chimique<sup>131</sup>. Ensuite, au début des années 2000, Philip Parson propose une voie de synthèse de l'anatoxine-a racémique plus simple et plus rapide commençant par le 1,5-cyclooctadiène comme représenté sur la Figure 25<sup>132</sup>.

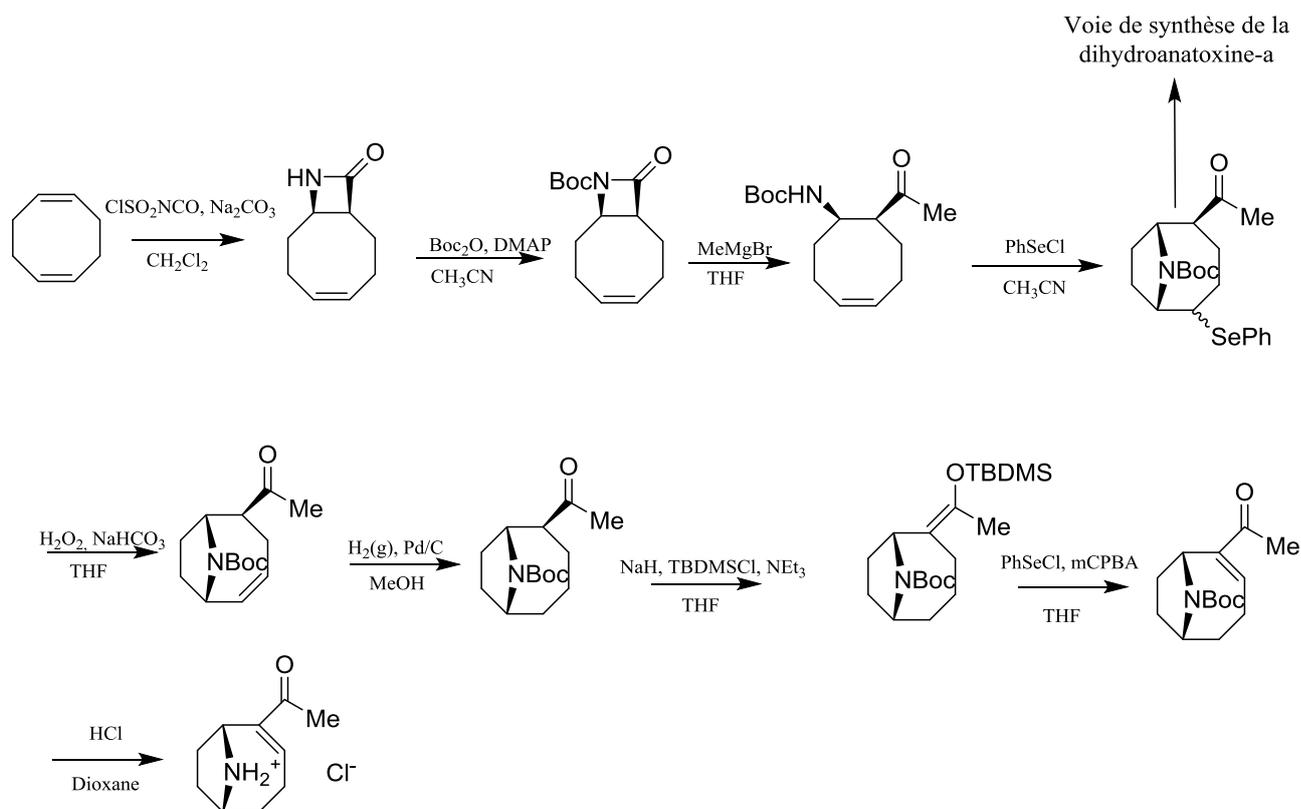


Figure 25. Voie de synthèse en série racémique de l'anatoxine-a développée par Parson<sup>132</sup>.

Cette méthode de synthèse a été reprise par plusieurs équipes pour former l'anatoxine-a mais aussi ses dérivés. Notre équipe s'est inspirée de cette voie de synthèse de Parson et l'a adaptée pour former la *cis*- et *trans*-dihydroanatoxine-a ainsi que la *cis*- et *trans*-dihydrohomoanatoxine-a (voir chapitre IV) mais aussi des dérivés de l'anatoxine-a en remplaçant le méthyle par le butyle, le propyle et l'isopropyle<sup>98</sup>. C'est sur cette voie de synthèse que nous nous sommes appuyés pour synthétiser l'analogue de substrat d'AnaG. Ces travaux doivent beaucoup au Dr. Stéphane Mann qui a réalisé la voie de synthèse de la dihydroanatoxine-a.

### III.1.2) Un analogue du substrat de l'AnaG

#### III.1.2.1) La structure de l'analogue

L'analogue de substrat désiré doit être similaire au substrat putatif d'AnaG selon la voie de biosynthèse que nous avons proposée. Ce substrat est composé de la chaîne en cours d'élongation relié au bras phosphopantéthéine du domaine ACP de l'AnaF. Le substrat doit

être ensuite transféré sur le domaine KS d'AnaG. Il est difficile pour nous d'obtenir un analogue d'une PKS entière, c'est pour cette raison que nous avons décidé de mimer seulement le bras phosphopantéthéine et son substrat chargé. Pour mimer le bras, nous avons choisi d'utiliser un thioester avec le composé N-acylé cystéamine ou NAC, un thioester qui a déjà servi à remplacer un bras phosphopantéthéine de PKSs dans de nombreuses expériences et dans ces exemples, le composé attaché au bras a bien été transféré à la PKS suivante<sup>133-138</sup>.

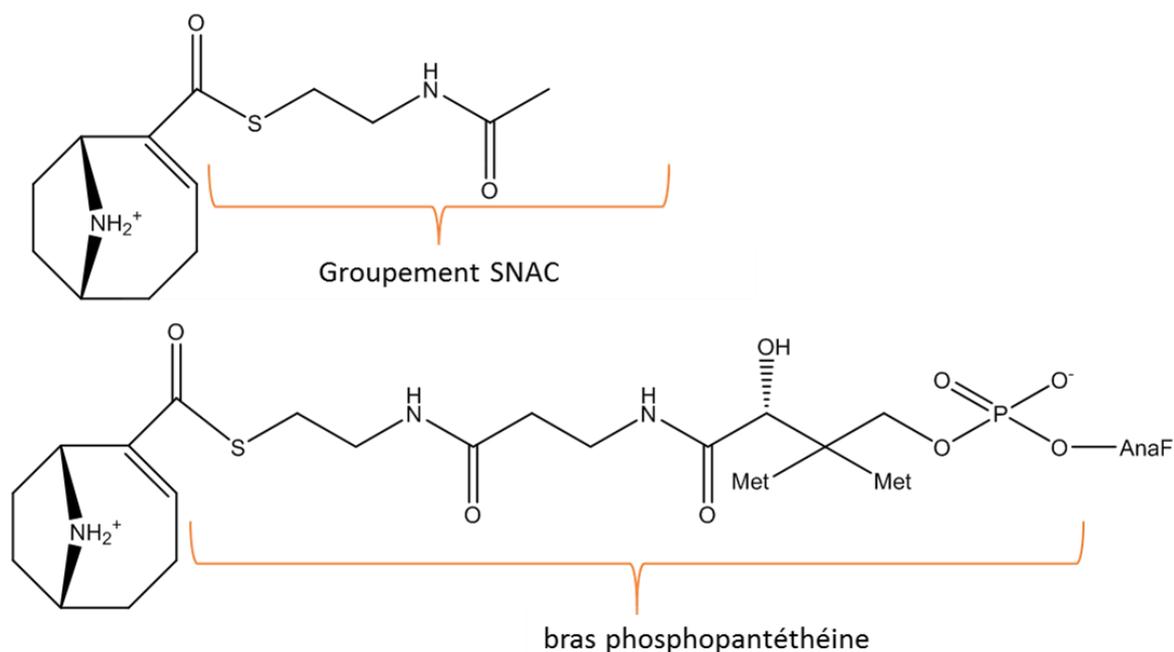


Figure 26. Structure de l'analogue et du substrat putatif d'AnaG.

### III.1.2.2) Voie de synthèse de l'analogue

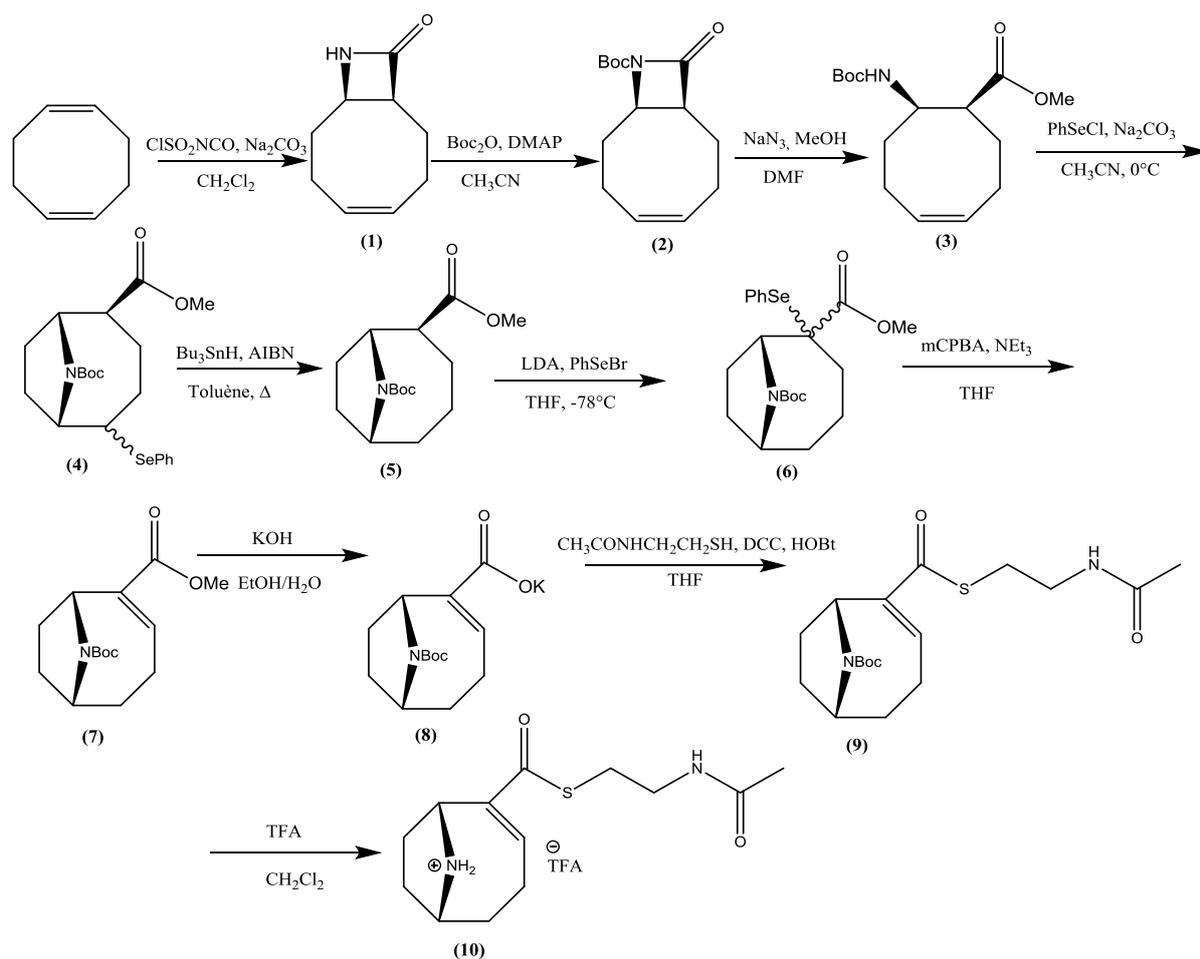


Figure 27. Voie de synthèse de l'analogue du substrat putatif d'AnaG

L'obtention des dérivés SNAC se fait le plus souvent dans la littérature par une réaction de thioestérification et donc nous voulions un composé portant une fonction acide carboxylique. Cette fonction sera protégée sous forme de fonction ester pendant la voie de synthèse. C'est pour cela que nous avons mis au point une voie de synthèse qui en s'inspirant de la voie Parson, nous permette d'arriver au composé désiré. Cette voie dont j'ai effectué les étapes vous est présentée sur la Figure 27. Nous n'aborderons pas en détails les deux premières réactions car elles ont déjà été décrites dans de nombreux articles<sup>98,132,139,140</sup>. J'ai obtenu en tout 5,66 g de composé (1) avec un rendement moyen de 20% et 28,3 g de composé (2) avec un rendement moyen de 78%.

### III.1.2.2.1) Ouverture du $\beta$ -lactame

La voie de synthèse diffère de celle de Parson à partir de l'ouverture du  $\beta$ -lactame pour donner le composé (3). Nous avons utilisé de l'azoture de sodium pour réaliser l'ouverture. Cet azoture est ensuite substitué avec du méthanolate selon le mécanisme présenté en Figure 28. L'azoture de sodium sert ici à effectuer une catalyse nucléophile. Il est plus nucléophile que le méthanolate et aussi un bon groupe partant. Lorsque j'ai effectué cette réaction au laboratoire, j'ai obtenu 12,9 g de produit (3) avec un rendement de 85%.

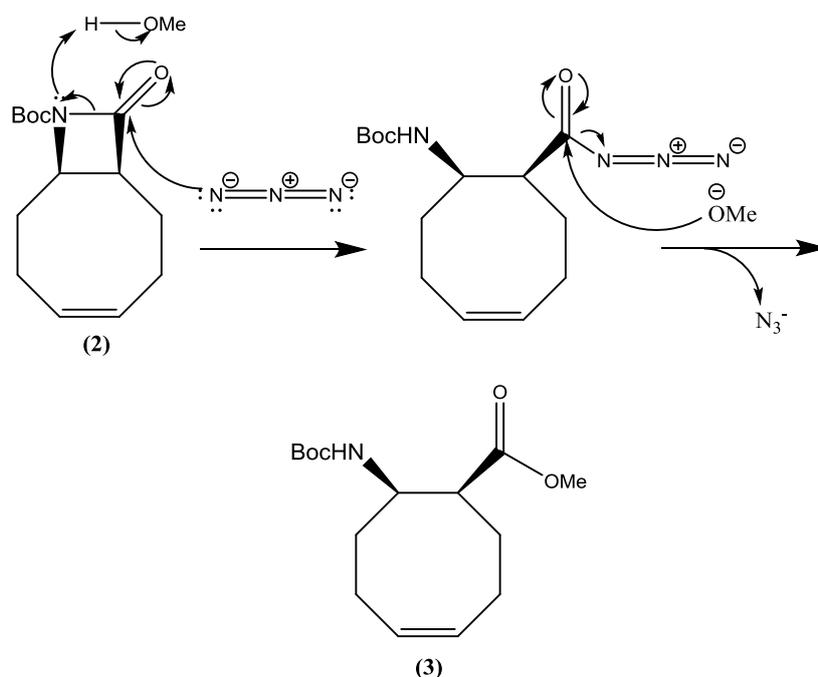


Figure 28. Mécanisme proposé pour l'ouverture du  $\beta$ -lactame

### III.1.2.2.2) Formation du cycle à 5 chaînons

Il nous faut ensuite former le cycle à 5 chaînons dans le cycle à 8 pour donner le composé (4). Pour cela nous utilisons du chlorure de phényl sélénium. Cette réaction a été problématique car nous obtenions de nombreux sous-produits ce qui diminuait le rendement de la réaction. Nous avons caractérisé ces produits comme étant le disélénium de diphenyle et un analogue du composé (4) portant une fonction chlore à la place du phényl sélénium. Le disélénium de diphenyle doit diminuer le rendement de la réaction car sa formation fait disparaître du chlorure de phényl sélénium qui n'est plus disponible pour la réaction. Le produit chloré doit être formé à partir du composé (4) et il a tendance à être co-élué sur colonne avec le produit ce qui le rend difficile à séparer.

Si nous observons le mécanisme proposé de cette réaction qui est représentée sur la Figure 29, nous voyons la formation d'un pont entre le sélénium et la double liaison<sup>141</sup>. Il y a ensuite une attaque de l'azote qui a été déprotoné par le carbonate de sodium pour former un cycle à 5 chaînons. Même si l'attaque peut se faire au niveau des deux carbones du pont, nous n'observons pas la formation du cycle à 6 chaînons. Ceci indique que la réaction est irréversible et donc probablement sous contrôle cinétique. Concernant la stéréochimie de l'attaque, elle se fait en *anti* du sélénium et donc sur le sélénium en arrière du plan du cycle. Ceci nous amène à penser que le groupement phényl séléniure est en arrière du cycle mais nous n'avons pas pu le démontrer.

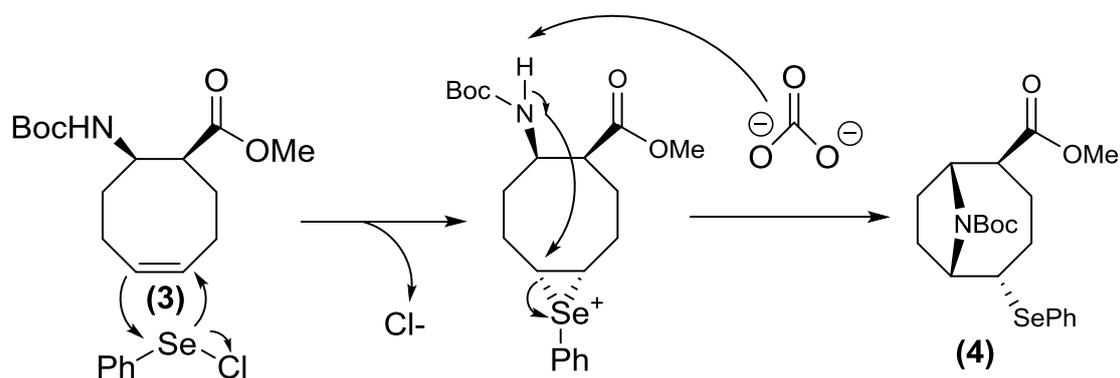


Figure 29. Mécanisme proposé de formation du produit sélénié

A ce moment, pour former le produit chloré, nous pouvons imaginer une attaque de l'ion chlorure sur le carbone portant le groupement sélénié sur le composé (4). Le mécanisme réactionnel doit être celui d'une substitution nucléophile. Sachant qu'il n'y a pas de facteur pour stabiliser un carbocation potentiel et que l'on est en présence d'un anion chlorure, nous avons tendance à penser que le mécanisme doit être de type SN2 représenté sur la Figure 30. Nous avons ici représenté une stéréochimie compatible avec le mécanisme proposé dans la Figure 29 mais la configuration du composé (4) n'a pas été résolue. La formation de phényl sélénoate lors de la formation du composé chloré peut expliquer la formation du diséléniure de diphenyle. En effet nous pouvons imaginer l'attaque du sélénoate sur chlorure de phényl séléniure ce qui donnerait bien du diséléniure de diphenyle.

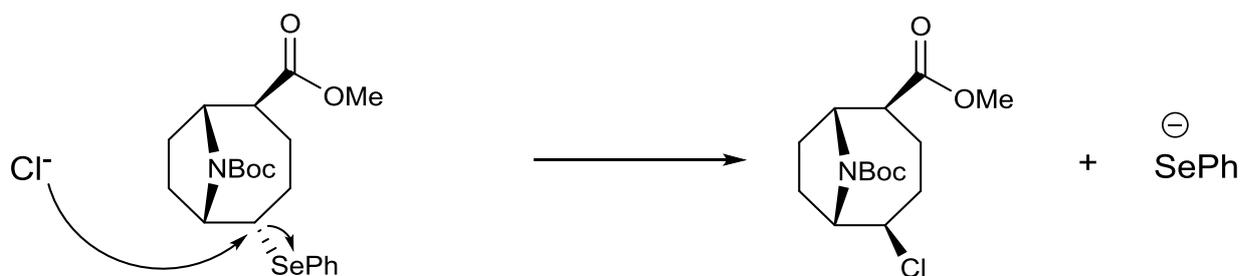


Figure 30. Mécanisme de la SN2

Pour diminuer la formation des deux sous-produits, nous avons effectué cette réaction en diminuant la température réactionnelle aux alentours de 0°C. Nous avons observé par la suite que la formation de sous-produits avait en effet diminué puisque nous n'observions plus de pics en RMN correspondant au composé chloré. Le rendement s'en est trouvé amélioré, passant de 62% à 86%. Lorsque j'ai effectué cette réaction au laboratoire, j'ai obtenu 8,14 g de produit **(4)**.

### III.1.2.2.3) Réduction du groupement phényl séléniure

Nous avons ensuite réduit le composé **(4)** de façon à substituer le groupement phényl séléniure par un hydrogène pour former le composé **(5)**. Nous avons utilisé l'hydrure de tributylétain ( $\text{Bu}_3\text{SnH}$ ) et l'azobisisobutyronitrile ( $\text{AIBN}$ )<sup>142</sup>. L'AIBN produit par chauffage deux radicaux isobutyronitriles. Le radical isobutyronitrile va pouvoir former des radicaux  $\text{Bu}_3\text{Sn}\cdot$ . Ce radical va ensuite réagir avec le composé **(4)** comme montré dans la Figure 31.

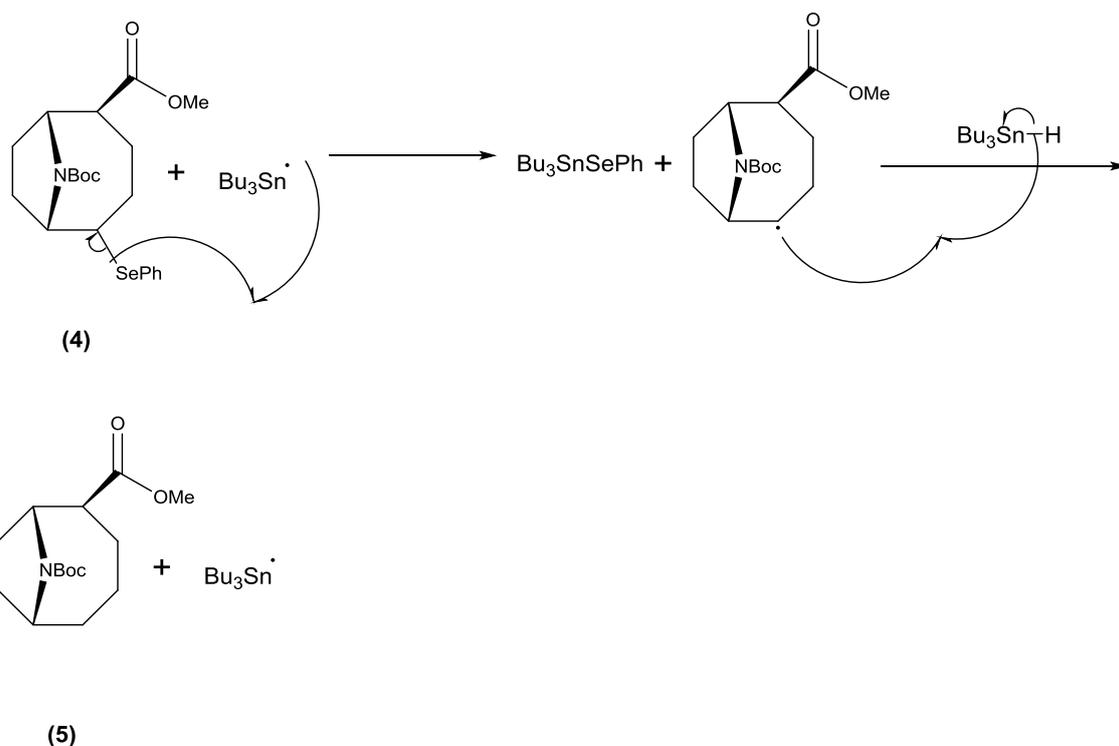


Figure 31. Mécanisme de la formation du composé (5)

Par rapport à la voie de synthèse développée par Parson, nous faisons cette réaction en une étape alors qu'il en prenait deux. Il s'agit d'une réaction en chaîne et le radical  $\text{Bu}_3\text{Sn}\cdot$  va de nouveau réagir avec du composé (4). Lorsque j'ai effectué cette réaction au laboratoire, j'ai obtenu 1,8 g de produit (5) avec un rendement de 89%.

#### III.1.2.2.4) Ajout du groupement phényl séléniure au pied du carbonyle

Pour former le composé (6), nous avons utilisé du bromure de phényl séléniure avec du diisopropylamidure de lithium de façon à former l'énolate. Il va ensuite réagir avec le bromure de phényl séléniure pour donner le composé (6) comme montré dans la Figure 31. Nous avons remarqué qu'avec cette réaction, nous avons aussi la formation de sous-produits dont le diséléniure de diphenyle. La réaction est aussi très sensible à la qualité du THF utilisé. En effet, si le THF n'est pas fraîchement distillé, la réaction n'a pas lieu. Nous avons réussi à limiter la formation de sous-produits en contrôlant l'ajout de  $\text{PhSeBr}$  pour qu'il soit le plus lent possible. Cet ajout contrôlé permet de limiter au maximum la formation de diséléniure de diphenyle. Nous avons aussi remarqué que la réaction est très sensible à la température et c'est pour cela que nous devons maintenir la température strictement à  $-78^\circ\text{C}$ . Même dans ces

conditions, nous n'avons pas réussi à empêcher totalement la formation de ces sous-produits et nous n'avons pas pu les isoler pour les caractériser.

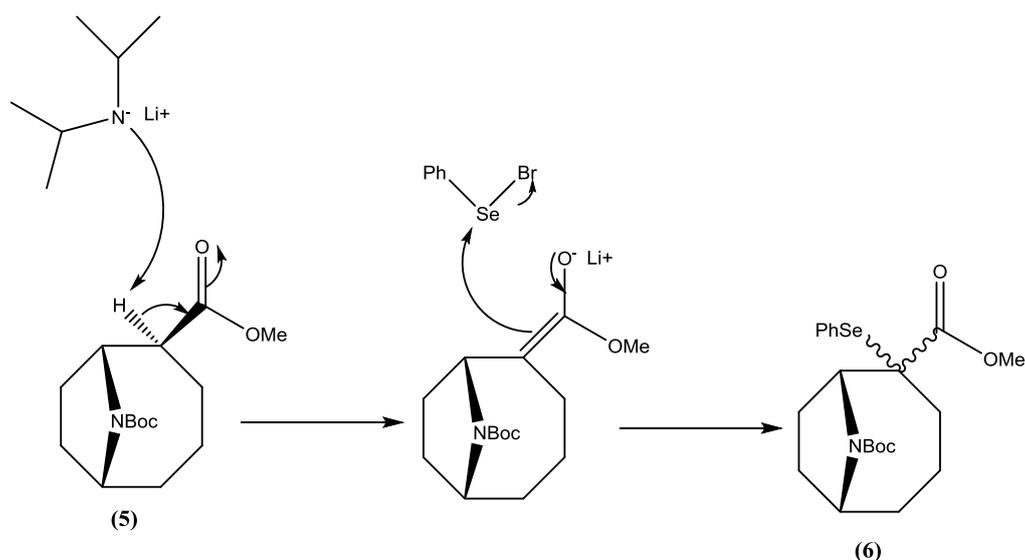


Figure 32. Réaction avec le LDA et le PhSeBr

Lorsque j'ai effectué cette réaction au laboratoire, j'ai obtenu 0,94 g de produit (6) avec un rendement de 74%.

### III.1.2.2.5) Réaction d'élimination du sélénoxyde

Nous allons ensuite former une double liaison pour former le composé (7) par une réaction d'élimination du sélénoxyde avec de l'acide métachloroperbenzoïque (mCPBA) et de la triéthylamine ( $\text{NEt}_3$ )<sup>143,144</sup>. Le mCPBA va permettre d'oxyder le groupement phényl séléniure en sélénoxyde. Ce composé va ensuite subir une élimination *syn* intramoléculaire pour donner la double liaison comme montré sur la Figure 33. La  $\text{NEt}_3$  va permettre de neutraliser l'acide carboxylique formé par le mCPBA après l'oxydation du groupement phényl séléniure.

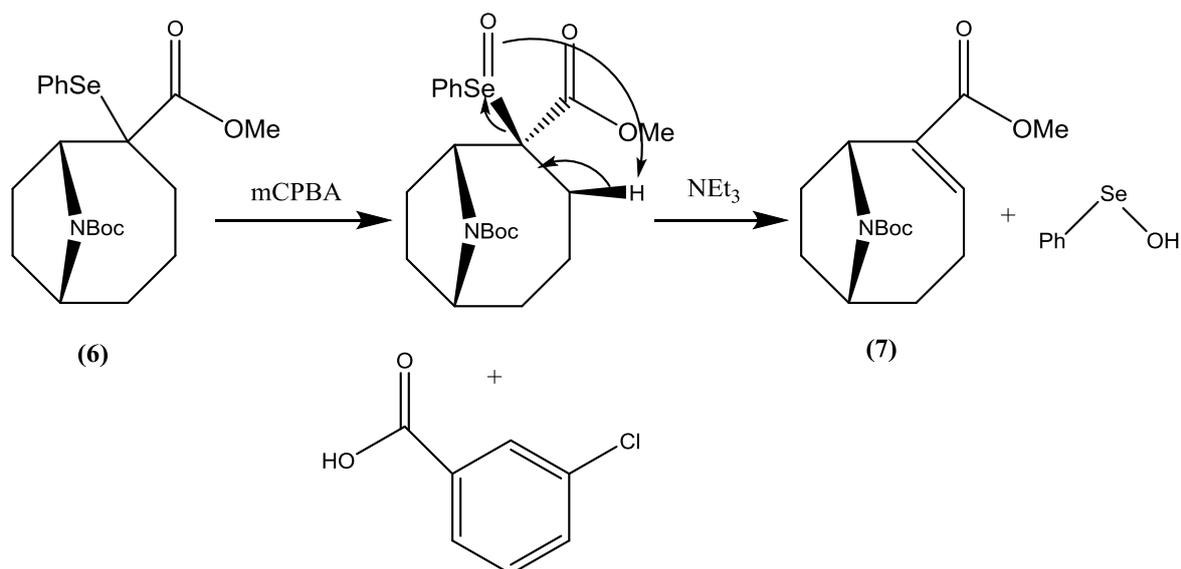


Figure 33. Mécanisme d'élimination du sélénoxyde

Lorsque j'ai effectué cette réaction au laboratoire, j'ai obtenu 0,39 g de produit (7) avec un rendement de 90%.

#### III.1.2.2.6) Thioestérification

Une fois que nous avons formé une double liaison au pied de la fonction ester, nous pouvons passer à l'étape de saponification pour former le composé (8). Lorsque j'ai effectué cette réaction au laboratoire, j'ai obtenu 0,276 g de produit (8) avec un rendement de 80%.

A partir de cette étape, nous avons réfléchi aux meilleures conditions réactionnelles pour la thioestérification permettant d'obtenir le composé (9). De nombreuses conditions sont proposées dans la littérature pour effectuer cette réaction. L'équipe avait auparavant utilisé les conditions utilisées par Walsh et son équipe<sup>145,146</sup>. Ces conditions font intervenir le DCC et le HOBT. Nous nous sommes cependant rendu compte qu'avec ces conditions, une addition 1,4 sur la double liaison pouvait avoir lieu. En effet, lorsque nous avons observé le spectre de masse des produits obtenus, nous observons un pic majoritaire à  $m/z$  369 qui peut correspondre au produit désiré. Nous avons, ensuite, deux pics significatifs à  $m/z$  488 et 386 comme montré sur la Figure 34. Les pics à  $m/z$  120 et 137 doivent correspondre à la N-acétylcysteamine n'ayant pas réagi sous forme  $[M+H]^+$  et  $[M+NH_4]^+$ .

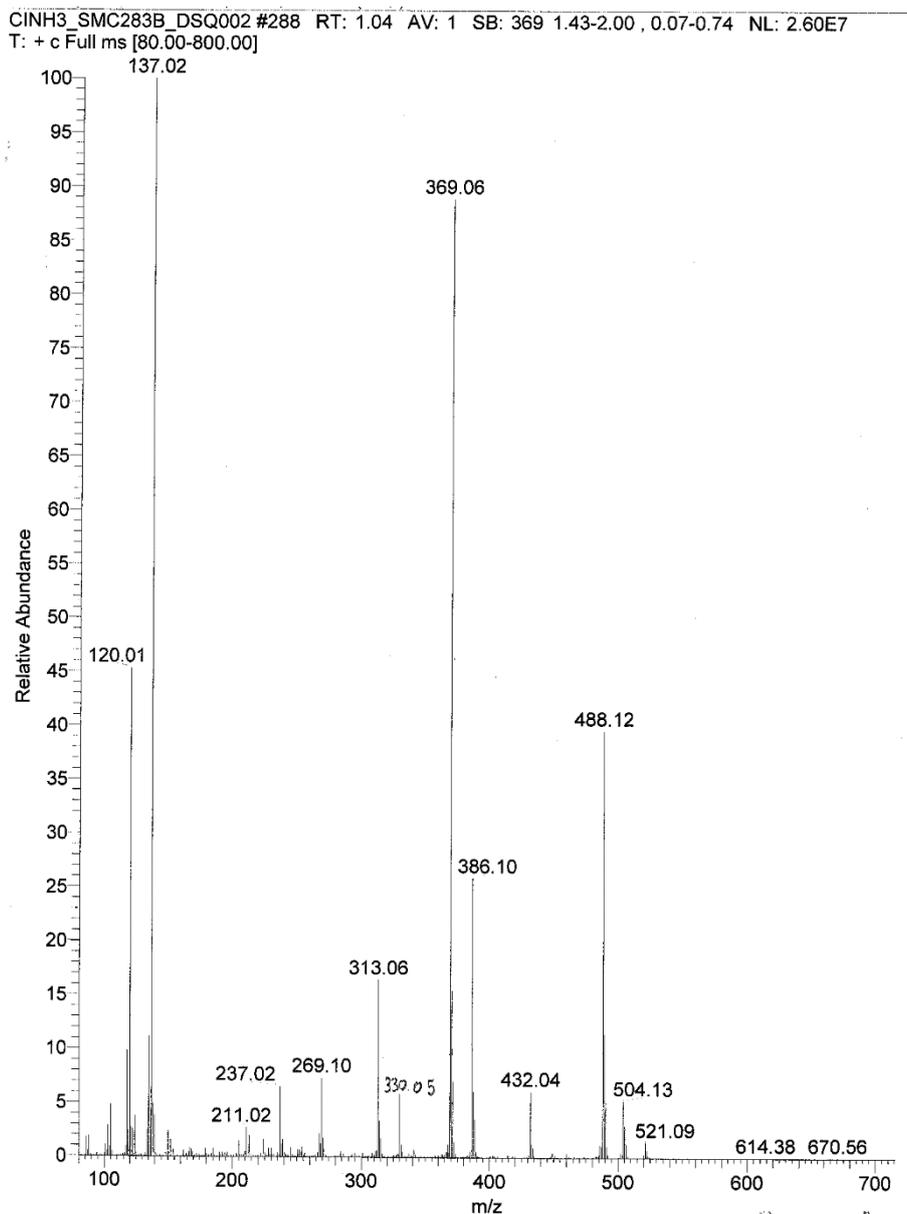


Figure 34. Spectre de masse du mélange après thioesterification en  $MH^+$  en CI (ionisation chimique) par le Dr Stéphane Mann au service de spectrométrie de masse de Chimie Paristech

La  $m/z$  488 correspond au produit **(9)** ayant subi une réaction d'addition 1,4 pour donner le produit **(9)'**. Le pic à une  $m/z$  386 peut correspondre au composé **(9)** sous forme  $[M+NH_4]^+$ . Même si la spectrométrie de masse n'est pas quantitative, la hauteur des pics pour le pic à une  $m/z$  de 369 et de 488 montre que les produits sont bien obtenus. La Figure 35 présente la structure du sous-produit proposé.

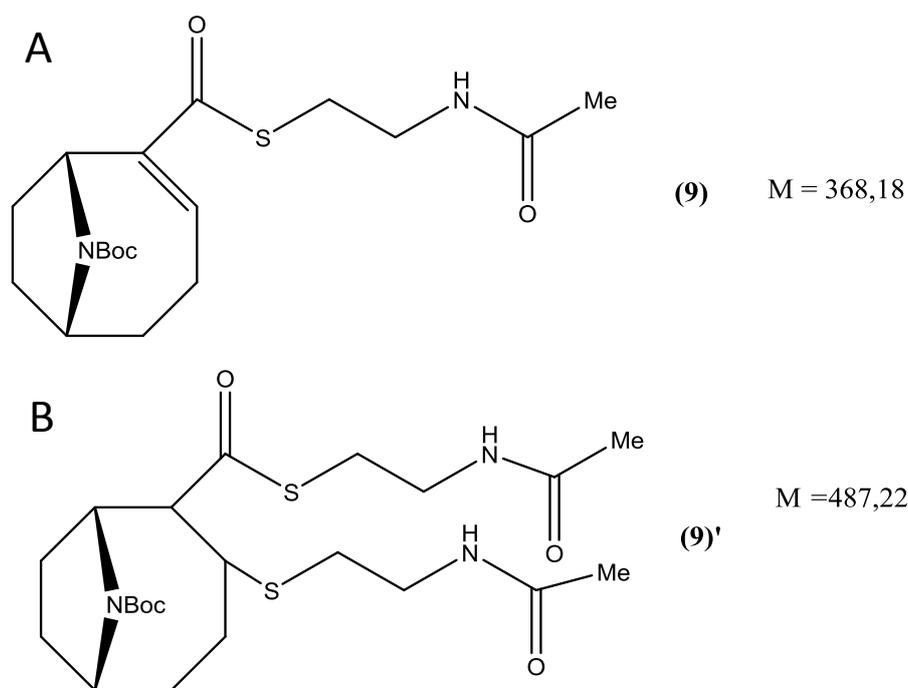


Figure 35. A : Produits attendus lors de la réaction de thioestérification ; B : Sous-produit obtenu lors de la réaction de thioesterification, les masses sont en  $\text{g}\cdot\text{mol}^{-1}$

Nous n'avons pas été capable de séparer le produit (9) du sous-produit (9)' et nous avons donc décidé d'employer d'autres conditions réactionnelles pour la thioestérification du produit (8). Nous avons regardé dans la littérature pour des conditions de thioestérification de NAC faisant intervenir des réactifs portant des doubles liaisons en alpha d'une fonction carbonyle. Nous avons alors trouvé des conditions adéquates faisant intervenir le DMAP et l'EDCI pour des composés cycliques portant des insaturations avec des rendements allant de 80 à 90%<sup>134,147,137</sup>. Nous avons pu effectuer la réaction au laboratoire et nous avons remarqué sur les spectres RMNs du composé obtenu que nous avons effectivement réussi à obtenir le thioester mais nous avons également observé des impuretés dues à la présence de pics non attribués en RMN <sup>13</sup>C. Ce sont les pics aux  $\delta$  ppm 124.50, 123.99 et 119.07. Ces pics peuvent correspondre à des carbones insaturés.

Nous avons alors essayé de purifier à nouveau par chromatographie flash sur silice le mélange de produits obtenus mais nous n'avons pas été capables d'éliminer les impuretés. Nous avons donc utilisé des méthodes analytiques plus précises pour nous assurer de la présence du composé (9) et le purifier.

### III.1.2.3) Purification et caractérisation du produit final

Pour nous assurer que nous avons bien obtenu l'analogue de substrat d'AnaG, nous avons décidé de caractériser le produit obtenu par LC-MS. Dans un premier temps, nous avons dû effectuer la déprotection du composé (9) afin de retirer le groupement Boc avec de l'acide trifluoroacétique (TFA) pour rendre le produit soluble dans l'eau comme montré dans la Figure 36.

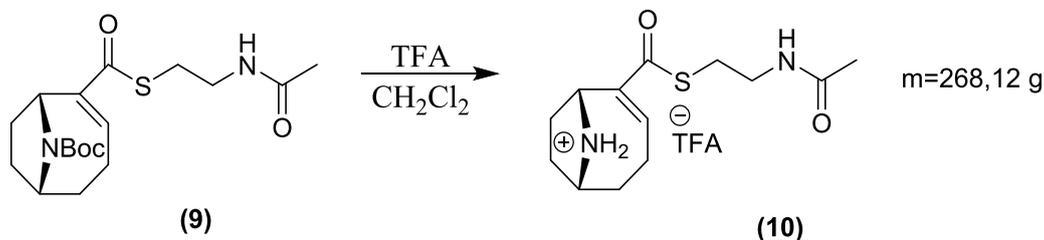


Figure 36. Déprotection du produit de thioestérification

Après analyse LC-MS du produit et une mise au point des conditions analytiques, nous avons pu confirmer qu'il y avait bien plusieurs composés dans le mélange obtenu par synthèse chimique. Nous obtenons un chromatogramme avec 3 pics comme montré sur la Figure 37 dont nous avons observé les spectres de masse.

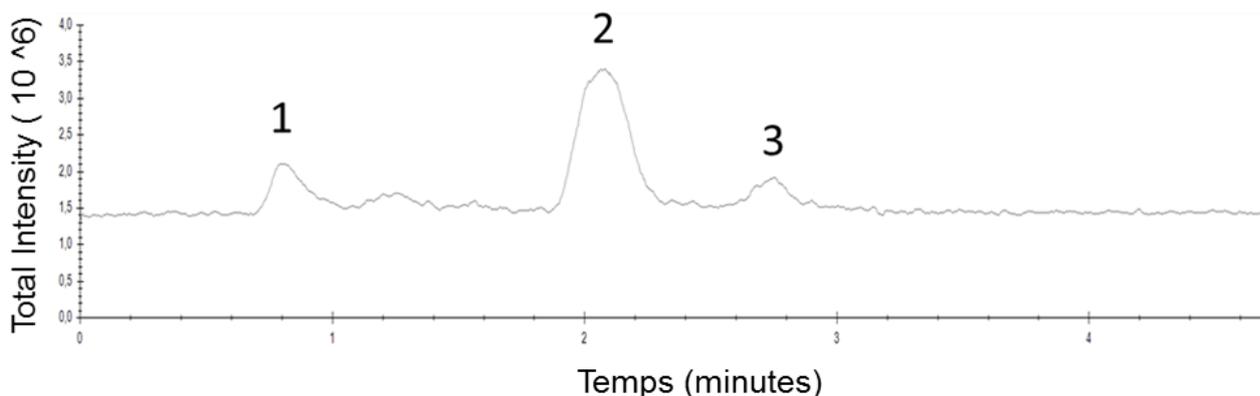


Figure 37. Chromatogramme de la LC-MS du produit, la colonne utilisé est une C18 analytique avec un volume d'injection de 8  $\mu\text{L}$  avec un gradient isocratique 20% méthanol/ 80% d'eau milliporée avec 0,1% en volume de TFA

Le premier pic du chromatogramme correspond à des impuretés. En effet, on observe de très nombreux pics avec des masses diverses comprises entre  $m/z$  100 et 200. Le second pic, qui est le plus important sur le chromatogramme, contient un pic majoritaire en masse à  $m/z$  269,5 ce qui peut correspondre à l'analogue de substrat. Le troisième pic du chromatogramme présente, lui aussi, un pic majoritaire en masse à  $m/z$  269,5.

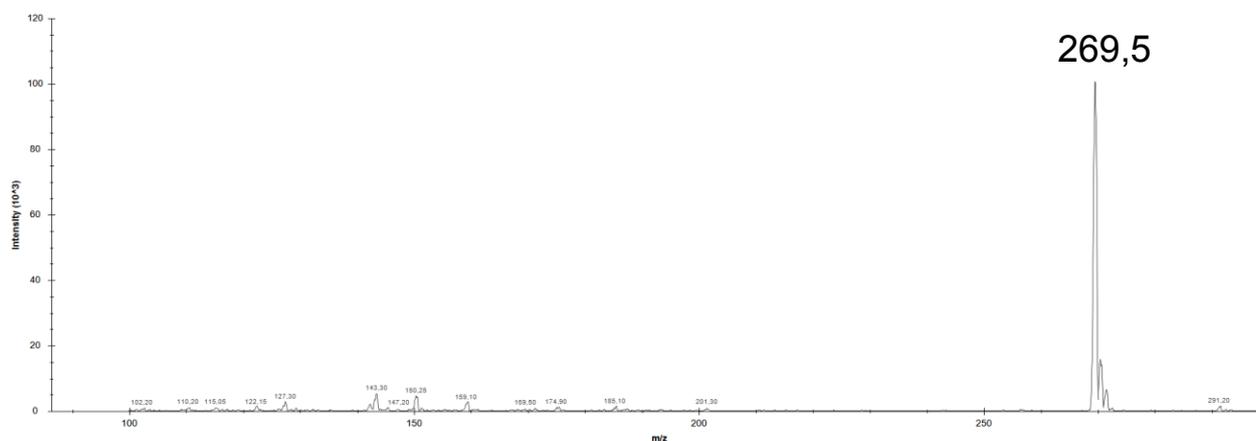


Figure 38. Spectre de masse en ESI du deuxième pic

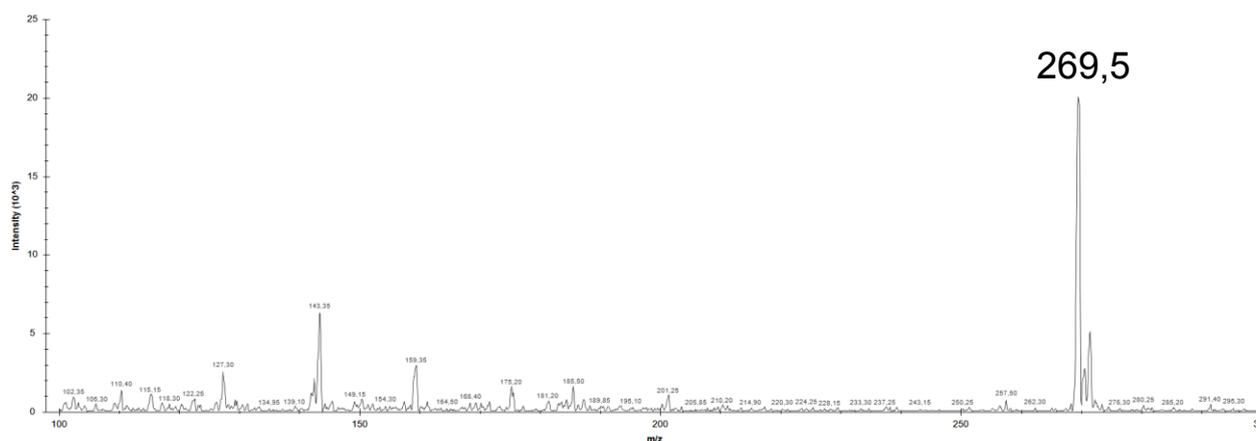


Figure 39. Spectre de masse en ESI du troisième pic

La présence de ce pic sur le spectre de masse aux alentours d'une  $m/z$  269,5 dans le deuxième et le troisième pic du chromatogramme est surprenante. En effet, si nous avons bien obtenu l'analogue de substrat, il devrait sortir dans un seul et même pic. Il est probable qu'il y ait deux composés avec une masse proche dans le mélange obtenu. Il est possible que nous soyons face à un isomère du composé **(10)**.

Nous avons donc décidé d'essayer de séparer les deux pics obtenus par une purification sur un système HPLC avec une colonne C18 qui correspond aux conditions utilisés en LC-MS. Nous ne sommes pas arrivés à séparer les pics présents en HPLC de façon satisfaisante malgré l'utilisation d'une HPLC analytique et d'une HPLC semi-préparative. Finalement après plusieurs tentatives en HPLC, nous avons collecté deux fractions. Nous les avons à nouveau injectées en LC-MS de façon à vérifier si nous n'avions pas une fraction contenant seulement le composé **(10)**. Dans la première fraction, nous avons deux pics majoritaires et un seul pic majoritaire dans la deuxième fraction comme présenté sur la Figure 40.

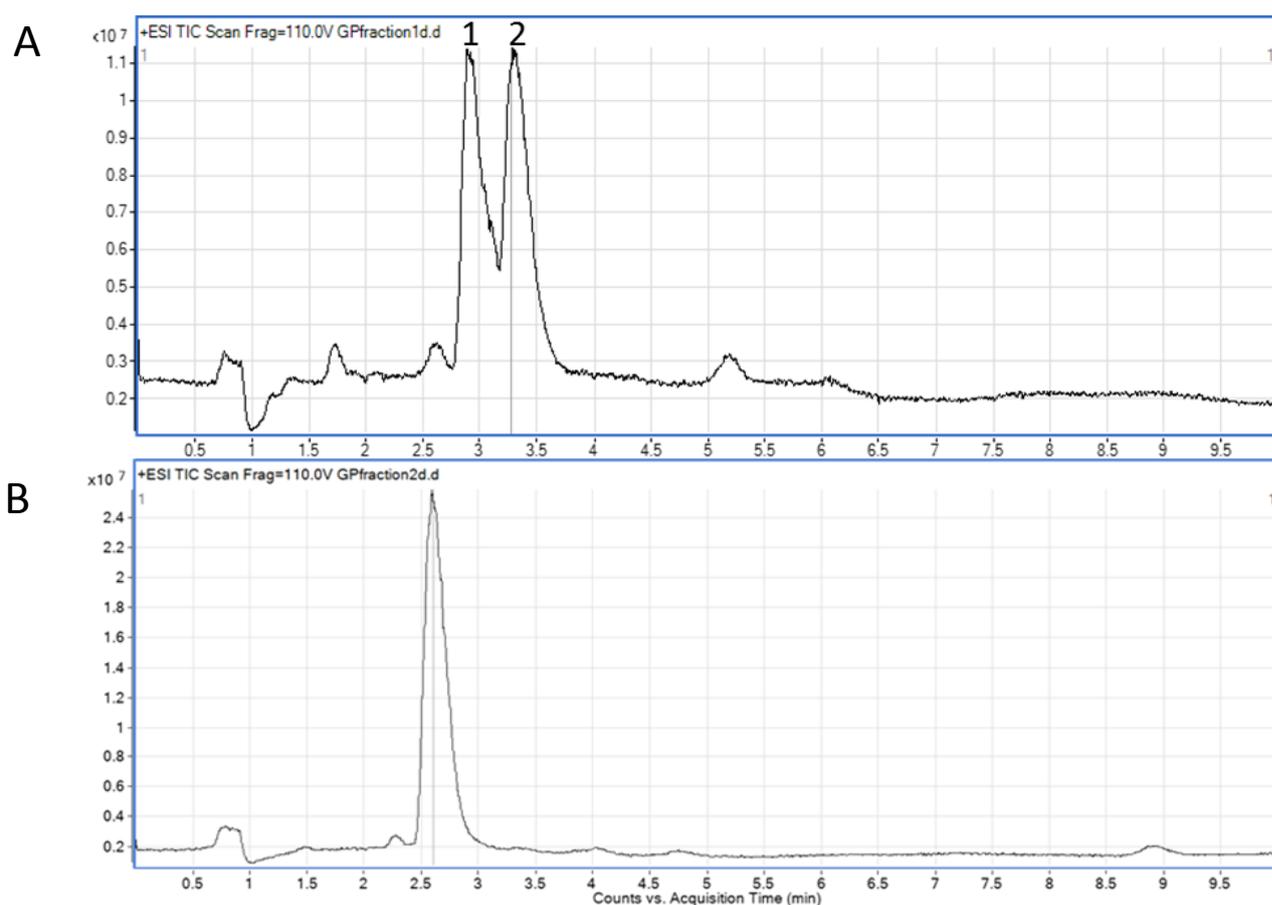


Figure 40. Chromatogrammes des injections LC-MS. A : Première fraction. B : Deuxième fraction. La colonne utilisée est une C18 analytique avec un volume d'injection de 8  $\mu$ L avec un gradient isocratique 20% Méthanol/ 80% d'eau milliporée avec 0,1% en volume de TFA.

Pour la première fraction, le spectre de masse du premier pic donne un pic majoritaire à  $m/z$  269,5. Pour le spectre de masse du deuxième pic, nous observons un pic majoritaire à  $m/z$  271,5. Pour la deuxième fraction, le spectre de masse du pic majoritaire est similaire au spectre de masse du deuxième pic de la première fraction.

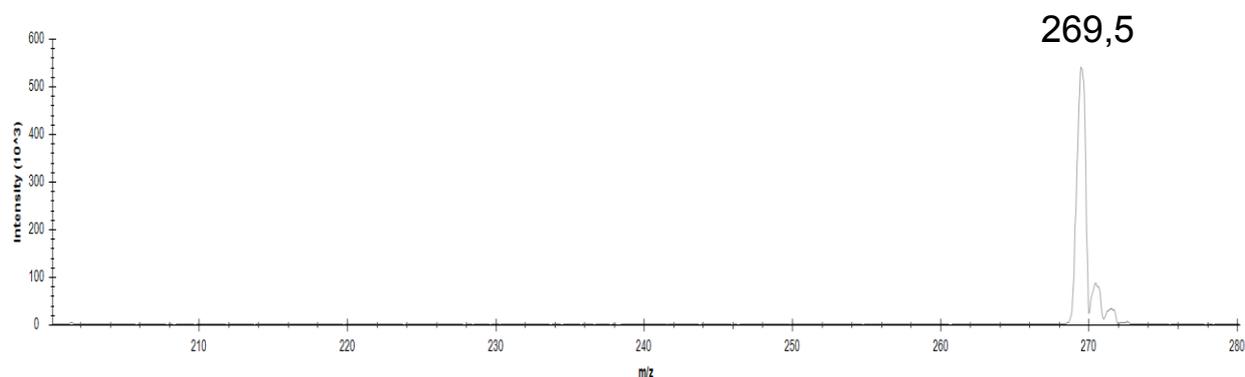


Figure 41. Spectre de masse en ESI du premier pic de la première fraction

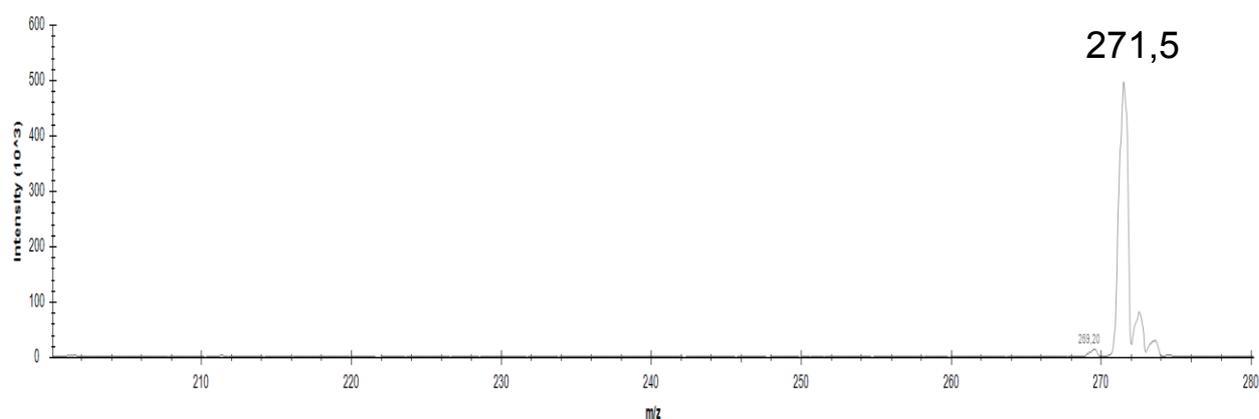


Figure 42. Spectre de masse en ESI du deuxième pic

### III.1.3) Discussion

Nous sommes parvenus avec la voie de synthèse développée au laboratoire à obtenir l'analogue de substrat d'AnaG d'après les diverses analyses effectuées (RMN + spectrométrie de masse). Nous ne sommes, cependant, pas parvenus à séparer ce composé d'une impureté co-élue. L'impureté doit donc avoir une structure très proche de celle du composé puisqu'elle est indissociable en chromatographie sur silice. Selon le spectre RMN et le spectre de masse que nous avons obtenu, nous pouvons penser que nous avons un dérivé dihydro de l'analogue du substrat d'AnaG. Cependant, nous ne pouvons pas le confirmer car nous n'avons pas réussi à l'isoler. Il nous faudrait arriver à déterminer à quelle réaction nous obtenons cette impureté. Cette impureté peut se former au moment de la réaction d'élimination ou lors d'une réaction ultérieure. Lors de précédentes expériences au laboratoire

avec ces réactions, nous avons rencontré un problème similaire, où une impureté de structure proche au produit désiré rendait difficile la séparation.

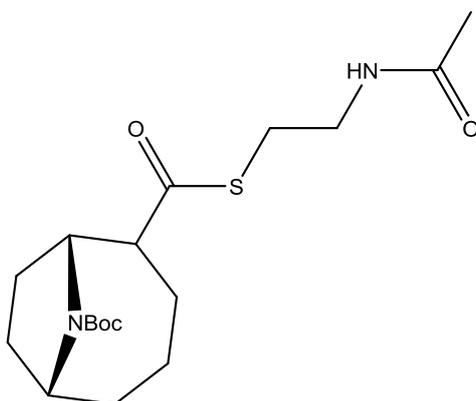


Figure 43. Structure proposée de l'impureté

Il nous faut donc réussir à identifier quelle est l'étape qui génère cette impureté. Nous avons pensé reprendre la voie de synthèse avec des quantités de réactifs plus importantes de façon à avoir assez de réactifs pour tester une mise au point de conditions réactionnelles des dernières étapes.

## III.2) Approche biochimique : obtention des protéines AnaG et AnaA

Pour l'obtention d'AnaG et AnaA pures, nous avons choisi d'exprimer ces protéines chez *Escherichia coli*. En effet, l'équipe a déjà effectué, par le passé, l'expression hétérologue chez cette bactérie de protéines de cyanobactéries avec des résultats probants<sup>83,148,69,101</sup>.

Pour cela, l'équipe utilise majoritairement un plasmide d'expression, le pET-28a(+). Ce plasmide permet la surexpression d'un gène sous le contrôle d'un promoteur T7 en présence d'isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Ce système d'expression permet d'obtenir en grande quantité une protéine dont le gène a été cloné dans le plasmide. Pour faciliter le clonage d'un gène d'intérêt dans le plasmide, de nombreux sites de restriction sont présents en 3' du promoteur T7.

Même si la protéine est surexprimée et est donc disponible en grande quantité, il nous faut la purifier de façon à éliminer les protéines naturellement exprimées par *E. coli*. Nous utilisons, dans l'équipe, une étiquette poly-histidine ou His-tag. Cette étiquette est constituée de six résidus histidine adjacents. Ces histidines sont de bons ligands pour complexer le nickel ou le cobalt. Cette complexation permet de retenir la protéine porteuse de l'His-tag sur une résine greffée portant des ions de nickel ou de cobalt. Des résines greffées portant du nickel ou du cobalt sont disponibles dans le commerce sous différentes formes notamment chargées dans une colonne. Ceci nous permet d'utiliser ces colonnes au laboratoire conjointement à un automate de purification pour récupérer rapidement et efficacement la protéine en une seule étape de purification dans la plupart des cas<sup>149-151</sup>.

Pour que la protéine soit bien retenue sur la colonne, il faut que l'étiquette His-tag soit aisément accessible pour pouvoir complexer l'ion métallique. En fonction du repliement de la protéine, une position de l'His-tag en N-terminal ou en C-terminal est plus accessible. Mais même avec la prédiction de structure secondaire de la protéine, il est souvent compliqué de savoir à l'avance quelle extrémité sera la plus accessible. C'est pour cette raison que nous préparons parfois deux constructions pour la même protéine avec un placement de l'His-tag différent. Ceci nous permet de voir la construction qui est la mieux retenue par une colonne d'affinité. Il est d'ailleurs possible de mettre deux His-tag sur une même protéine, l'un en C-terminal et l'autre en N-terminal, pour s'assurer une rétention maximale sur la colonne.

Deux séquences nucléotidiques codant pour des His-tags sont présentes en 3' du promoteur T7 dans le plasmide pET-28a(+). Ces séquences sont placées proches de sites de restriction de façon à pouvoir facilement insérer la séquence d'un gène d'intérêt à proximité. Nous allons utiliser à chaque fois des sites de restriction différents pour les clonages en fonction du placement de l'His-tag désiré. L'expression hétérologue chez *E. coli* et la purification sur une colonne d'affinité vont nous permettre de récupérer les protéines désirées pures.

### III.2.1) Stratégies d'obtention d'AnaG

Pour la purification d'AnaG, la doctorante précédente du laboratoire, Rabia Mazouz, avait essayé d'obtenir la protéine entière. AnaG est une protéine de taille importante (184 kDa). Il est connu que l'expression hétérologue chez *E. coli* de protéines de plus de 70 kDa est difficile<sup>152,153</sup> même si des protéines de taille importante ont déjà été obtenues par le passé<sup>154</sup>. Le problème vient du fait que des protéines de cette taille sont particulièrement sensibles à la protéolyse. De plus, la machinerie traductionnelle d'*E. coli* supporte mal la traduction de protéines d'une telle taille et le ribosome peut s'arrêter avant la fin de la séquence protéique. Lors de la construction du plasmide pET28a(+)(Kan<sup>r</sup>)-anaG, la doctorante avait décidé de mettre un His-Tag en N-terminal et en C-terminal pour mieux récupérer la protéine.

L'AnaG n'a jamais pu être obtenue entière et aucune des purifications tentées n'a donné le résultat escompté. Après purification, une seule bande était obtenue à un poids moléculaire de 108 kDa. Après analyse par spectrométrie de masse, il s'avère que c'est une partie tronquée de la protéine AnaG. Cette partie comprend le domaine KS et le domaine AT de l'AnaG et la séquence protéique continue jusqu'à l'acide aminé 963 comme montré sur la Figure 44.

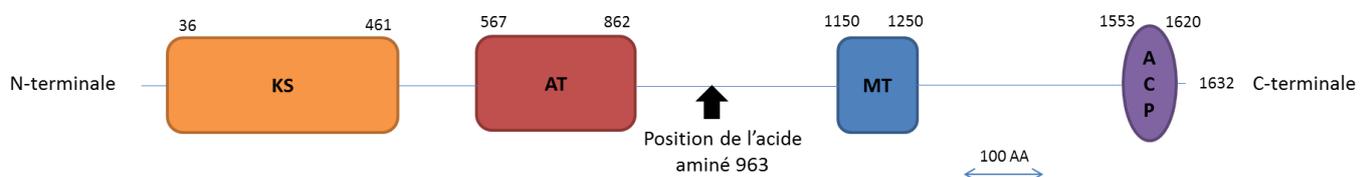


Figure 44. AnaG avec les domaines connus

Il n'a pas été détecté d'autres fragments d'AnaG. La doctorante a ensuite modifié les conditions de purification de la protéine en ajoutant un cocktail d'inhibiteurs de protéase commercial dans le tampon de lyse mais également dans les tampons d'équilibration et d'élution de la colonne d'affinité. Elle a ainsi vu apparaître sur gel deux bandes d'intensités égales, l'une aux environs de 120 kDa et l'autre à 60 kDa. Ceci tend à prouver que nous sommes bien face à un phénomène de protéolyse.

En parallèle, la doctorante a tenté une expression hétérologue chez un hôte différent. Son choix s'est arrêté sur les souches de cyanobactérie *Synechocystis* PCC 6803 et *Synechococcus elongatus* PCC 7942. Ce sont des cyanobactéries unicellulaires naturellement transformables et des organismes modèles<sup>155</sup>. Elle a donc entrepris de cloner *anaG* en collaboration avec l'équipe de Franck Chauvat au CEA dans un vecteur d'expression adapté à la transformation de ces deux cyanobactéries<sup>156</sup>. Elle a ensuite réalisé un premier test d'expression d'AnaG à partir de la souche *Synechocystis* PCC 6803 et a obtenu sur gel une bande entre 150 et 250 kDa qui peut correspondre à AnaG. La bande protéique a été analysée en spectrométrie de masse et la séquence de l'AnaG est bien retrouvée entière. Ce résultat confirme qu'AnaG a bien été exprimée chez *Synechocystis* PCC 6803. Mais la quantité obtenue est malheureusement trop faible pour que nous puissions la purifier facilement et nous ne pouvons pas lancer de culture de cyanobactéries sur d'assez grands volumes pour en récupérer une quantité suffisante.

Nous sommes donc revenus à une expression hétérologue chez *E. coli* avec comme objectif de vérifier qu'avec le bon cocktail d'inhibiteurs nous pouvions obtenir l'AnaG entière et cela même en faible quantité. En effet nous pouvons facilement lancer une culture sur un grand volume d'*E. coli* au laboratoire. Pour vérifier la présence d'AnaG, nous allons réaliser un Western blot.

Pour ce Western blot, nous avons décidé d'utiliser l'His-tag comme cible pour la détection d'AnaG. Plutôt qu'un anticorps anti His-tag, nous avons préféré utiliser un Hisprobe. L'Hisprobe est un ligand tridendate avec chaque ligand complexant un ion  $Ni^{2+}$ , ce ligand est lié de façon covalente à l'enzyme peroxydase de raifort (HRP). L'ion de nickel va être complexé par l'His-tag et la HRP permet ensuite la détection du complexe entre l'Hisprobe et l'His-tag. L'Hisprobe est aussi spécifique et moins onéreux qu'un anticorps anti His-tag, c'est pour cela que notre choix s'est arrêté sur ce composé.

Le Western blot avec l' Hisprobe n'a pas donné des résultats très concluants. En effet, nous n'avons pas détecté d'AnaG entière lors de la révélation de la membrane. Cela nous a, cependant, confortés dans notre idée que nous ne pouvions pas obtenir AnaG intacte chez *E. coli*.

Nous avons donc décidé d'obtenir AnaG en plusieurs fragments protéiques de façon à contourner ce problème. Une PKS de type I peut garder sa fonction si on sépare ses domaines en des protéines distinctes<sup>157</sup>. Nous avons donc décidé de fragmenter AnaG d'un côté en sa partie N-terminale contenant le domaine KS et le domaine AT et de l'autre côté, la partie C-terminale contenant le domaine D1, D2, ACP et MT. Nous avons aussi décidé d'obtenir le domaine ACP d'AnaG seul. En effet, nous voulions savoir si le domaine seul est fonctionnel et si les domaines D1 et D2 présents sur la partie N-terminale jouent un rôle dans la régulation de son activité. Nous avons aussi voulu réaliser une protéine AnaG synthétique contenant uniquement les domaines KS, AT et ACP.



Figure 45. Domaines présents chez AnaG

Notre objectif, une fois ces protéines obtenues, était de valider notre proposition de voie de biosynthèse. Pour cela, il nous faut réaliser le chargement du bras phosphoantéthéine sur le domaine ACP chez les protéines porteuses de ce domaine et charger le bras avec le monomère de malonyl-CoA. Nous pensions essayer différentes combinaisons de protéines, chacune comprenant les domaines nécessaires à la réaction d'élongation. La première combinaison serait de réunir la partie N-terminale et C-terminale d'AnaG. Nous pourrions ensuite mettre également en présence la partie N-terminale avec le domaine ACP seul. Finalement, nous voulions utiliser l'AnaG synthétique.

Nous comptons ensuite préparer un milieu réactionnel associant ces différentes combinaisons de protéines et l'analogue du substrat. Dans un premier temps, nous voulions vérifier que la réaction de condensation a bien eu lieu et si le composé obtenu a été hydrolysé. Si le composé n'a pas été hydrolysé, il serait alors possible d'ajouter AnaA au mélange pour voir si cette protéine est capable d'hydrolyser l'intermédiaire. Une fois le produit de la réaction

libéré, nous serions en mesure d'observer si nous sommes en présence d'anatoxine-a ou de son dérivé  $\beta$ -cétoacide. Chaque produit obtenu aurait été analysé par des méthodes de spectrométrie de masse.

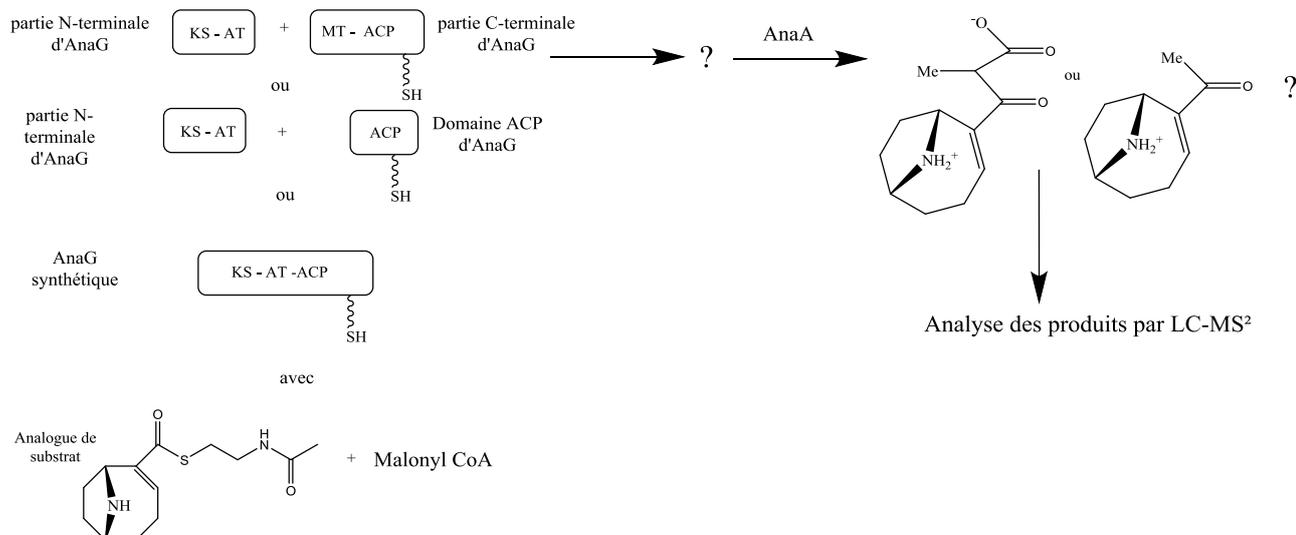


Figure 46. Récapitulatifs des étapes prévues pour l'étude des dernières étapes de la voie de biosynthèse

### III.2.1.1) Obtention de la partie N-terminale d'AnaG

Pour la partie N-terminale d'AnaG, nous avons cloné la séquence nucléotidique allant du premier codon d'AnaG jusqu'au codon de la Glycine 945 avec un His-tag en C-terminal. J'ai utilisé les sites de restriction NcoI (amorce GP34, voir chapitre VI.3) et HindIII (amorce GP35, voir chapitre VI.3) pour obtenir le plasmide pET28a(+)(Kan<sup>r</sup>)-anaG pGP02. La protéine doit avoir un poids moléculaire de 105 kDa.

Je n'ai pas eu de problèmes de protéolyse lors de l'expression de la protéine et nous n'avons donc pas eu besoin d'utiliser des inhibiteurs de protéases. J'ai rassemblé les fractions les plus pures, concentrées et évaluées par gel. Sachant que la protéine a bien été retenue par la colonne, nous pouvons penser que l'His-tag en C-terminal est intacte et que la protéine n'a donc pas subi de protéolyse en C-terminal. Nous pouvons voir que le poids moléculaire de la protéine est similaire à celui que nous attendions, nous pouvons donc penser que la protéine n'a pas subi de protéolyse en N-terminal.

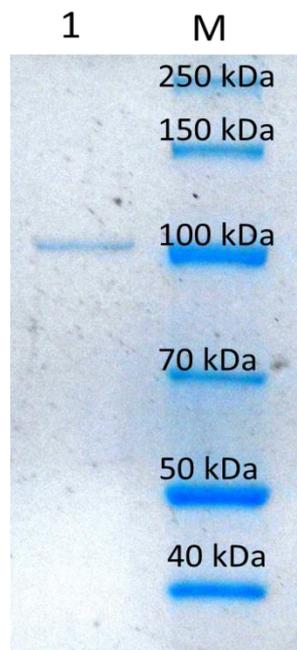


Figure 47. 1 : Partie N-terminale d'AnaG après purification, dépôt de 15  $\mu\text{L}$  à 0,126  $\text{mg}\cdot\text{mL}^{-1}$

### III.2.1.2) Obtention de la partie C-terminale d'AnaG

Une construction de la partie C-terminale d'AnaG avait été effectuée au laboratoire avant mon arrivée en utilisant les sites de restriction NcoI et XhoI pour donner le plasmide pET28a(+)(Kanr)-anaG V954-stop. La séquence nucléotidique choisie commence après le domaine AT au niveau du codon de la valine 954 et va jusqu'au codon de l'acide aspartique 1633 de la protéine avec un His-tag en C-terminal. La protéine doit avoir un poids moléculaire de 77 kDa.

Après les premières purifications, j'ai obtenu deux bandes majoritaires sur le gel d'électrophorèse. La première est à 77 kDa ce qui correspond à la taille attendue pour la partie C-terminale d'AnaG et une deuxième bande à 56 kDa. Cette deuxième bande n'est pas présente sur les gels en début de purification, mais elle apparaît progressivement avec le temps. Des analyses en spectrométrie de masse nous ont permis de confirmer que la première bande contient bien des séquences peptidiques correspondant à la séquence de la partie C-terminale d'AnaG. Nous avons pensé que nous étions face à une protéolyse étant donné les résultats obtenus précédemment avec AnaG. Sachant que la partie C-terminale d'AnaG contient les domaines D1, MT, D2 et ACP, nous avons pensé que l'un des domaines D1 ou D2 dont la fonction est inconnue pourrait avoir une fonction de décarboxylation ou d'hydrolyse pour permettre la terminaison de la voie de biosynthèse de l'anatoxine-a. Si c'est

le cas, ces domaines pourraient être responsables d'un phénomène d'autoprotéolyse une fois la protéine purifiée. Une autre hypothèse est que la protéolyse peut venir d'une protéase présente dans le milieu de culture ou libérée lors de la lyse des cellules.

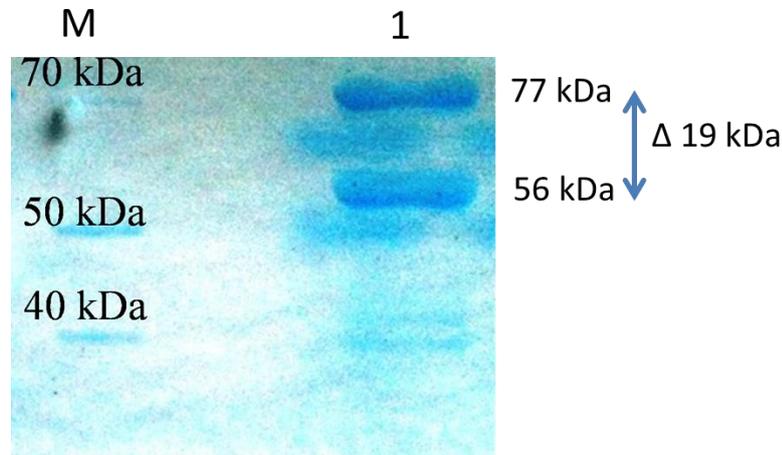


Figure 48. 1: Partie C-terminale d'AnaG après purification, dépôt de 15  $\mu\text{L}$  à 0,88  $\text{mg}\cdot\text{mL}^{-1}$

Nous avons donc modifié notre approche pour éviter la digestion par l'ajout d'un mélange d'inhibiteurs de protéases au tampon de lyse. Nous voulions aussi quantifier l'ampleur de la protéolyse et l'efficacité des inhibiteurs de protéases. Sachant que la composition du cocktail commercial d'inhibiteurs de protéases n'est pas connue, nous avons voulu déterminer lesquels produisaient un effet sur la protéolyse de la partie C-terminale d'AnaG.

Nous avons ensuite choisi différents inhibiteurs de protéase avec le plus large spectre d'inhibition pour préparer une solution de chaque inhibiteur à sa concentration effective<sup>158,159</sup>. Le pepabloc est inhibiteur irréversible des protéases à sérine<sup>160</sup>. La pepstatine est un puissant inhibiteur des protéases acides<sup>161</sup>. L'aprotinine est un inhibiteur des protéases à sérine<sup>162</sup>. La bestatine est inhibiteur de l'aminopeptidase B3<sup>163</sup>, de la leucotriène A4 hydrolase<sup>164</sup> et de l'alanine aminopeptidase<sup>165</sup>. L'Acide éthylène diamine tétra acétique (EDTA) est un inhibiteur de metalloprotéases<sup>166</sup>.

J'ai alors réalisé plusieurs solutions. La première est une solution témoin du domaine C-terminal d'AnaG purifié sans inhibiteur de protéases. J'ai ensuite réalisé 6 solutions contenant chacune une certaine quantité de domaine C-terminal d'AnaG purifié et l'un des inhibiteurs de protéases précédemment cité. J'ai aussi utilisé un domaine C-terminal d'AnaG déjà protéolysé comme marqueur de dégradation. J'ai incubé les solutions à 28°C pendant

trois heures puis pendant six heures et observé l'évolution par gel d'électrophorèse. A trois heures, l'évolution était peu marquée. A six heures, nous avons remarqué une digestion plus marquée dans la solution sans inhibiteur de protéases et plus faible avec le cocktail commercial ainsi qu'avec le Pefabloc mais la différence est minime comme montré sur la Figure 49.

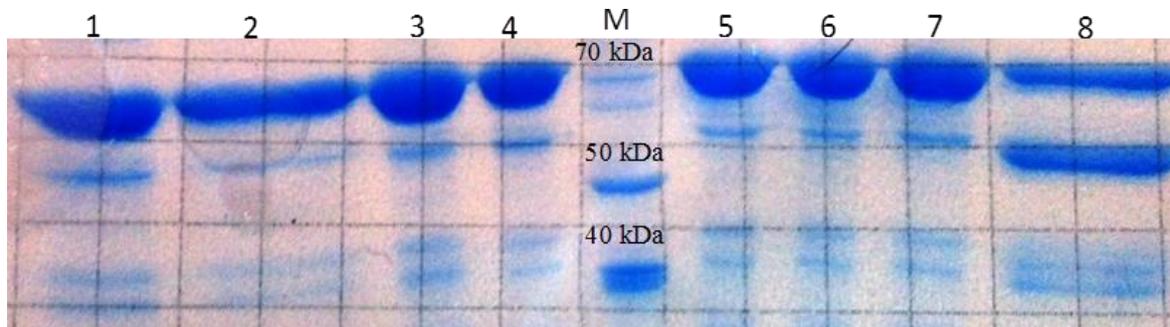


Figure 49. Gel d'électrophorèse après six heures d'incubation, 1: partie C-terminale d'AnaG ,  
2: partie C-terminale d'AnaG + cocktail commercial d'inhibiteurs, 3: partie C-terminale  
d'AnaG + Pefabloc, 4: partie C-terminale d'AnaG + Pepstatin, 5: partie C-terminale d'AnaG  
+ EDTA, 6: partie C-terminale d'AnaG + Aprotinine, 7: partie C-terminale d'AnaG + bestatin,  
8: partie C-terminale d'AnaG dégradée, dépôt de 15  $\mu$ L à 1,07 mg.mL<sup>-1</sup>

Au bout de six heures, la dégradation observée reste en deçà de ce que nous avons pu observer par le passé. Dans le cas d'une autoprolyse, la dégradation aurait été plus importante, nous pensons donc que la dégradation observée est due à un contaminant. Nous avons aussi remarqué que le cocktail commercial d'inhibiteurs reste le plus efficace et nous avons donc décidé de continuer à l'utiliser. J'ai réalisé toutes les purifications suivantes de la partie C-terminale d'AnaG avec le cocktail commercial d'inhibiteurs de protéases dans les tampons de lyse, de lavage et d'élution ce qui nous permet d'obtenir la protéine non dégradée. Après purification, j'ai réalisé un gel d'électrophorèse avec les fractions en sortie de colonne de nickel comme montré sur la Figure 50.

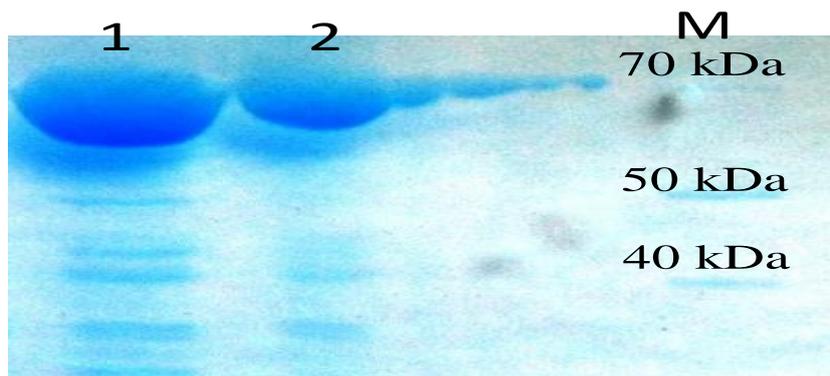


Figure 50. 1: Fraction purifiée sur colonne d'affinité concentrée de la partie C-terminale d'AnaG, dépôt de 10  $\mu\text{L}$  à 16,8  $\text{mg.mL}^{-1}$ , 2: fraction purifiée sur colonne d'affinité de la partie C-terminale d'AnaG, dépôt de 10  $\mu\text{L}$  à 5,9  $\text{mg.mL}^{-1}$

Nous avons remarqué que la partie C-terminale d'AnaG n'était pas totalement pure. Pour la purifier, j'ai donc entamé une purification sur colonne échangeuse d'anions. Ceci nous a permis d'obtenir la partie C-terminale d'AnaG pure comme montré sur la Figure 51.

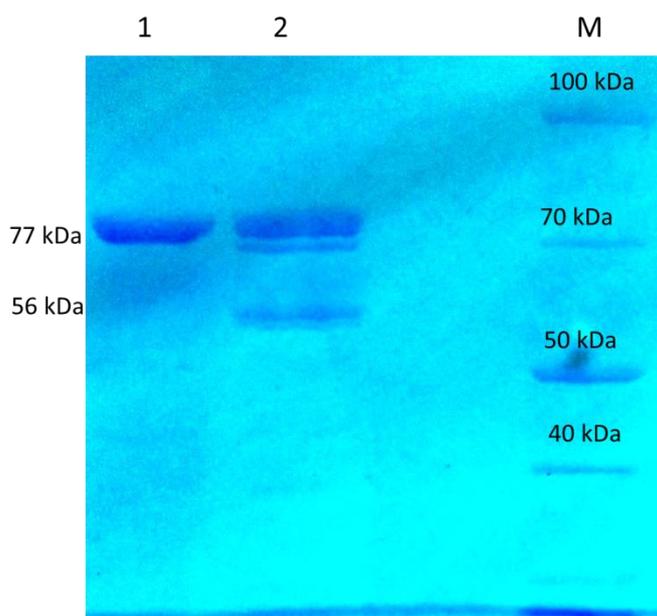


Figure 51. 1: Fraction purifiée sur colonne échangeuse d'ions de la partie C-terminale d'AnaG, dépôt de 15  $\mu\text{L}$  à 1,55  $\text{mg.mL}^{-1}$ , 2: partie C-terminale d'AnaG dégradée, dépôt de 10  $\mu\text{L}$  à 1,07  $\text{mg.mL}^{-1}$

Finalement, j'ai rassemblé deux fractions concentrées que nous avons passées en chromatographie gel filtration en collaboration avec Karine Moncoq à l'IBPC comme montré

sur la Figure 52. La protéine est en condition native et il est donc possible de voir si elle est sous forme monomérique ou sous forme de complexe. La masse du premier pic est aux alentours de 480 kDa et la masse du second pic est aux alentours de 120 kDa. Le premier pic correspond sans doute à un amas protéique. Il semble peu probable que nous obtenions un octamère. Le deuxième pic correspond à un dimère ce qui est cohérent avec ce que nous savons de la structure des PKSs.

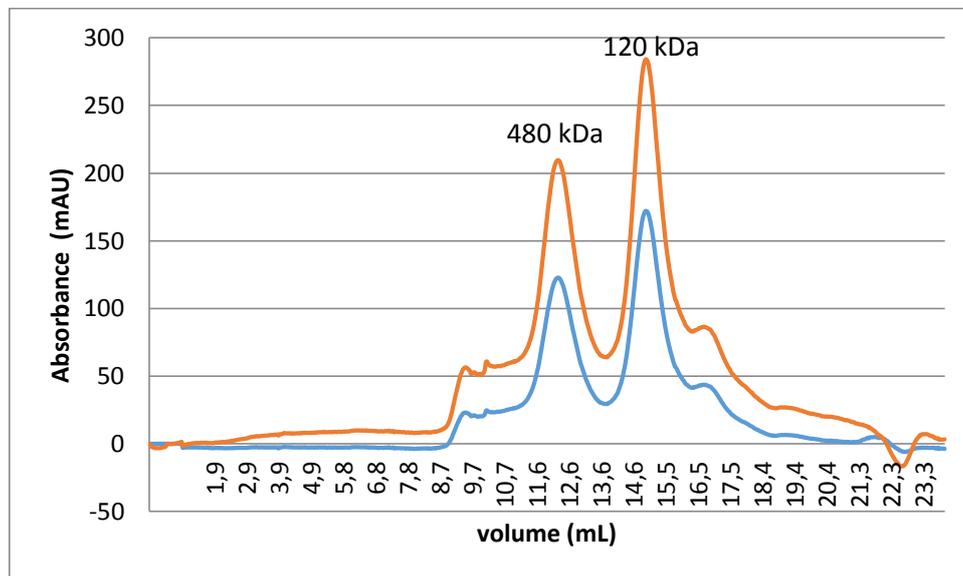


Figure 52. Résultat du gel filtration sur une colonne Superdex 200, orange : suivi de l'absorbance à 260 nm, bleu : suivi de l'absorbance à 280 nm

### III.2.1.3) Obtention du domaine ACP d'AnaG

Une première construction du domaine ACP d'AnaG avait été effectuée au laboratoire avant le début de mon doctorat par la doctorante Rabia Mazouz pour donner le plasmide pET28a(+)(Kanr)-ACP 1 anaG Histag N terminal. La séquence nucléotidique choisie pour ce clonage va du codon correspondant à la leucine 1536 d'AnaG jusqu'au codon stop de la même protéine.

Nous avons utilisé un logiciel de prédiction de domaine (antiSMASH 3.0) ainsi qu'un logiciel de prédiction de structure secondaire (Pspred) pour fixer les limites de ce domaine pour ce clonage. Il est souvent difficile de dire où s'arrête exactement un domaine du PKS en se basant sur des méthodes prédictives mais nous ne possédons pas de structure tridimensionnelle de ce domaine.

Nous n'avons pas réussi à obtenir le domaine ACP pur, ce dernier étant toujours co-élué avec d'autres protéines sans être majoritaire. Pour nous assurer de sa présence, nous avons effectué une analyse en spectrométrie de masse. Cette analyse nous a permis de confirmer la présence du domaine ACP de l'AnaG en très faible quantité. Nous pensons que l'His-tag n'est pas accessible et donc qu'il a du mal à se fixer sur la colonne. En observant la prédiction de structure secondaire de la séquence clonée comme montré sur la Figure 53, nous avons remarqué que le domaine ACP cloné commençait à l'intérieur d'une hélice  $\alpha$ . L'His-tag pour cette construction doit donc être peut accessible car incorporé dans une structure secondaire, ici, l'hélice  $\alpha$ . Nous avons donc décidé de refaire le clonage du domaine ACP en allant plus loin pour inclure l'hélice  $\alpha$  complète. La séquence du deuxième clonage part du codon de la lysine 1525.

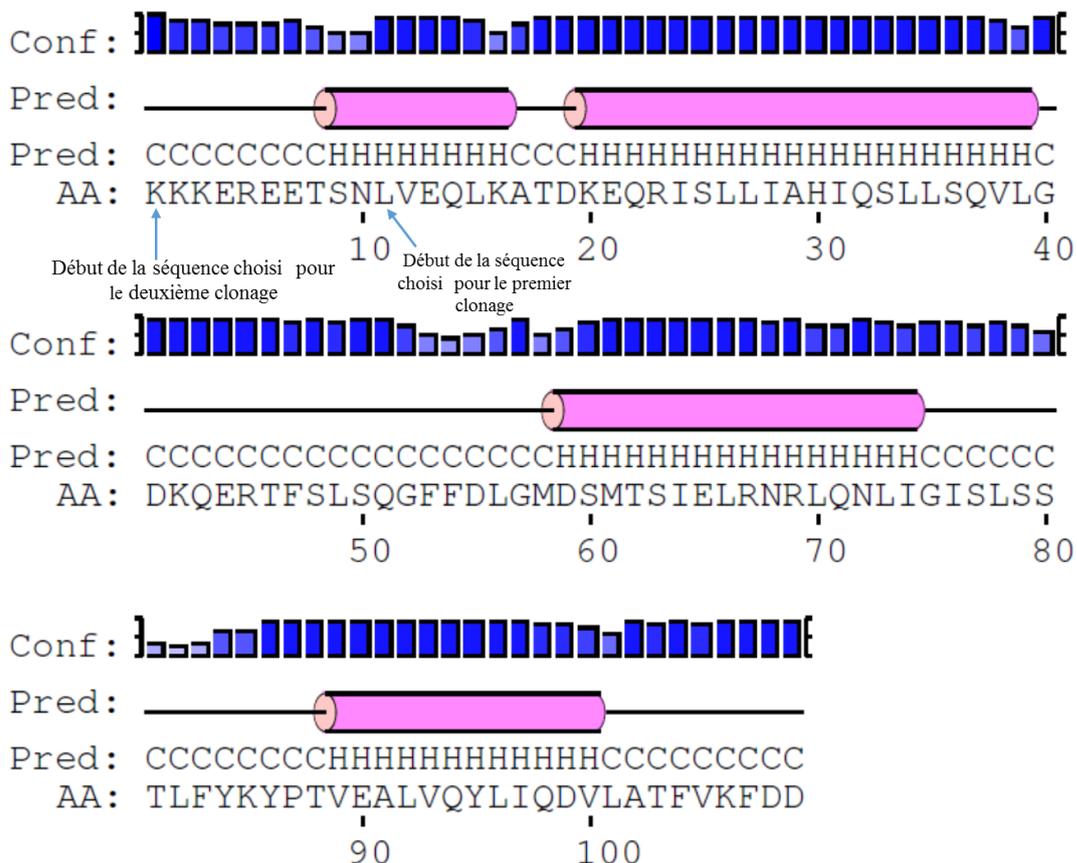


Figure 53. Séquence peptidique du domaine ACP avec prédiction de structure secondaire obtenue avec le logiciel PSIPRED et les deux séquences choisies pour le clonage, les cylindres roses représentent les hélices  $\alpha$  prédites. Les barres bleues correspondent à la probabilité de la prédiction.

Nous avons choisi de réaliser trois constructions avec cette nouvelle séquence avec un placement différent des His-tags : un clone avec un His-tag en C-terminal, un en N-terminal et un clone avec les His-tags en C-terminal et N-terminal. J'ai donc réalisé ces trois clones pour choisir celui qui donnera le meilleur résultat en termes de purification de protéine. J'ai utilisé les sites de restriction NcoI (amorce GP33, voir chapitre VI.3) et XhoI (amorce GP31, voir chapitre VI.3) pour obtenir le plasmide pET28a(+)(Kan<sup>r</sup>)-ACP 2 anaG Histag C terminal. J'ai utilisé les sites de restriction NdeI (amorce GP30, voir chapitre VI.3) et BamHI (amorce GP32, voir chapitre VI.3) pour obtenir le plasmide pET28a(+)(Kan<sup>r</sup>)-ACP 2 anaG Histag N terminal. J'ai finalement utilisés les sites de restriction NdeI (amorce GP30, voir chapitre VI.3) et XhoI (amorce GP31, voir chapitre VI.3) pour obtenir le plasmide pET28a(+)(Kan<sup>r</sup>)-ACP 2 anaG Histag N terminal et C terminal.

J'ai réalisé des purifications avec chacune des 3 protéines obtenues et il s'est avéré que la protéine portant le domaine ACP avec l'His-tag en C-terminal était celle dont l'expression et la purification étaient les plus aisées. J'ai donc utilisé la souche ayant le plasmide pET28a(+)(Kan<sup>r</sup>)-ACP 2 anaG Histag C terminal pour la purification du domaine ACP d'AnaG. La protéine obtenue doit avoir un poids moléculaire de 13,7 kDa.

Après la purification, j'ai rassemblé les fractions les plus pures, concentrées et évaluées par gel. La protéine est pure comme montré sur la Figure 54. Nous n'avons pas observé d'autres protéines, même après la concentration. Sachant que la protéine a bien été retenue par la colonne, nous pouvons penser que l'His-tag en C-terminal est intacte et que la protéine n'a donc pas subi de protéolyse en C-terminal. Nous pouvons voir que le poids moléculaire de la protéine est similaire à celui que nous attendions, nous pouvons donc penser que la protéine n'a pas subi de protéolyse en N-terminal.

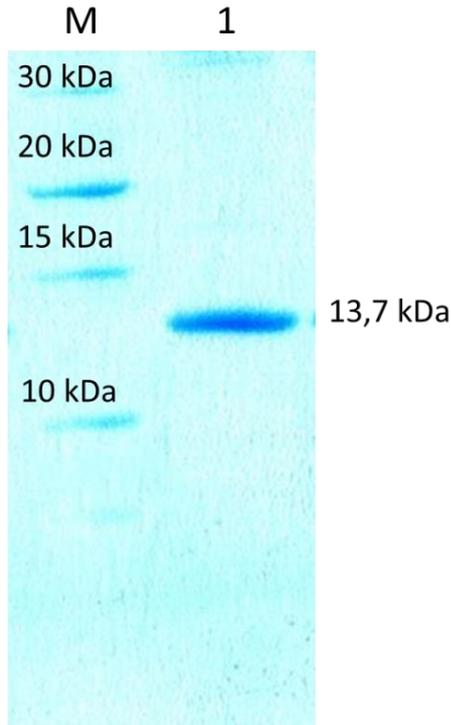


Figure 54. 1 : Domaine ACP d'AnaG après purification, dépôt de 5  $\mu\text{L}$  à  $0,58 \text{ mg.mL}^{-1}$

Il nous reste à effectuer la réaction de greffage du bras phosphopantéthéine sur le domaine ACP d'AnaG. Nous ne l'avons pas effectué car nous voulions d'abord le réaliser sur la partie C-terminale d'AnaG contenant les domaines D1, MT et D2.

#### III.2.1.4) Greffage du bras phosphopantéthéine

Le bras phosphopantéthéine est chargé sur la sérine présente dans le site actif du domaine ACP par la phosphopantéthéinyl transférase à partir du Coenzyme A, ce qui ajoute 340 Da au poids de la protéine. L'apo-AnaG devient l'holo-AnaG après ce greffage.

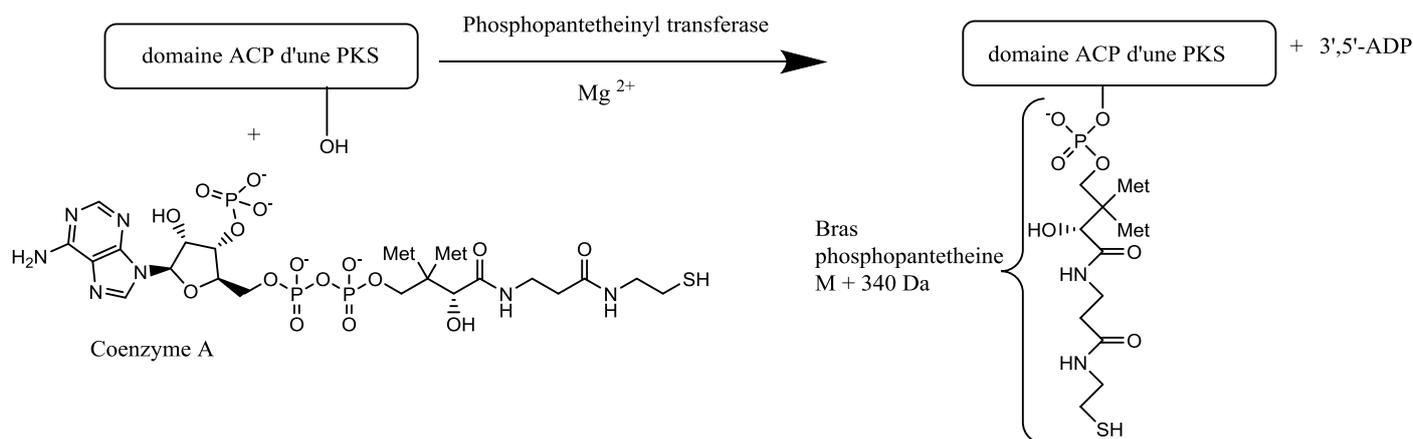


Figure 55. Chargement du bras phosphopantéthéine

Pour vérifier que la partie C-terminale de l'AnaG contenant les domaines D1, MT, D2 et ACP est bien purifiée et structurée, nous avons tenté de charger un bras phosphopantéthéine sur son domaine ACP. Pour ce faire, j'ai réalisé un mélange contenant la protéine purifiée, la PPTase provenant d'*Oscillatoria* purifiée au laboratoire appelé Ospt<sup>101,100</sup> et le Coenzyme A. En parallèle, j'ai réalisé une solution témoin qui ne contient pas l'Ospt et qui n'est pas incubée. Après incubation, nous avons effectué un MALDI-TOF au laboratoire de spectrométrie de masse protéomique de l'Institut Curie des deux solutions comme montré sur la Figure 56.

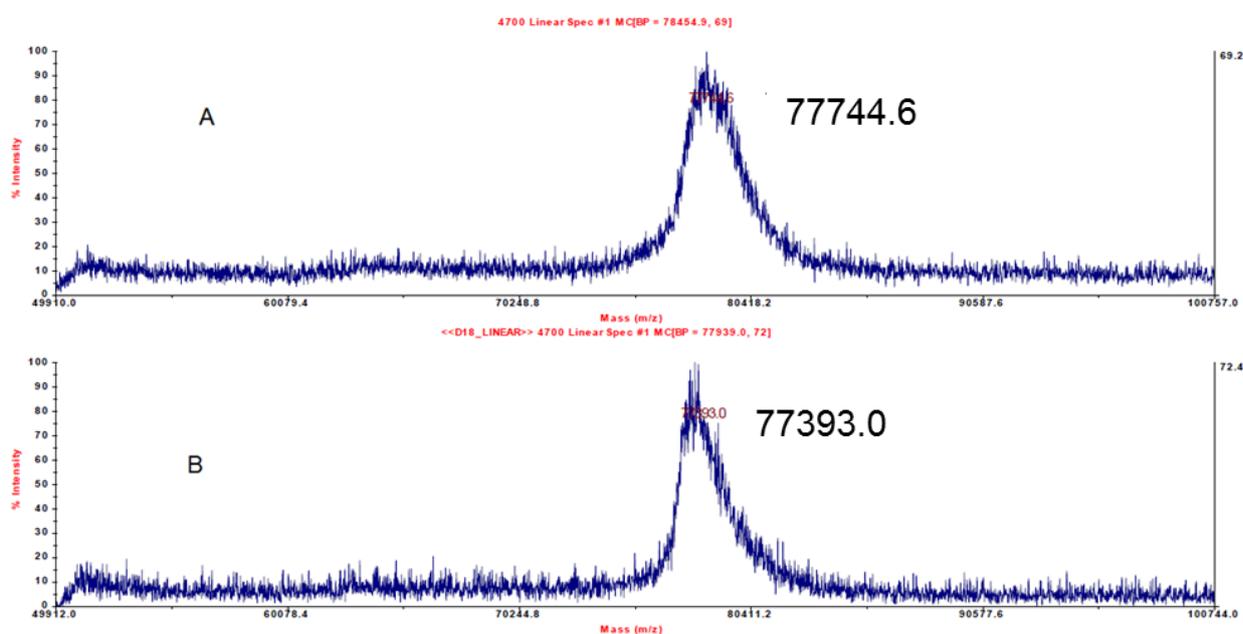


Figure 56. Spectre de MALDI-TOF A: mélange contenant l'Ospt et B: mélange témoin sans l'Ospt

La masse attendue pour la protéine sans le bras est de 77371 Da ce qui est proche de la masse obtenue chez le témoin en tenant compte de l'incertitude des mesures. Pour le mélange contenant l'Ospt, nous avons une masse de 77744 Da ce qui correspond à un déplacement des pics de 350 Da. En prenant en compte l'incertitude de ces mesures, nous pouvons penser qu'il y a bien une différence de masse entre les deux protéines et que le bras phosphopantéthéine a donc bien été chargé sur la partie C-terminale d'AnaG contenant les domaines D1, MT, D2 et ACP.

### III.2.1.5) Obtention d'AnaG synthétique

Face à l'impossibilité d'obtenir AnaG en entier, nous avons réfléchi à une façon détournée d'obtenir un AnaG plus court pour favoriser son expression chez *E. coli*. Nous voulions essayer d'obtenir une protéine AnaG plus courte ne contenant que les domaines KS, AT et ACP. En effet, nous pensons que nous pouvons laisser seulement les domaines dont l'activité est essentielle à la fonction de la PKS et retirer les domaines D1, D2 et MT. Avec la publication du génome de *Cylindrospermum stagnale* PCC 7417<sup>105</sup>, nous avons la preuve que la protéine AnaG peut être fonctionnelle avec seulement les domaines KS, AT et ACP.

Nous avons donc inséré la séquence du domaine ACP d'AnaG entier que nous avons déjà utilisé pour le deuxième clonage du domaine ACP dans le plasmide pET28a(+)(Kanr)-anaG pGP02. J'ai positionné la séquence du domaine ACP à la suite de la séquence de la partie N-terminale en utilisant les sites de restriction HindIII (amorce GP36, voir chapitre VI.3) et XhoI (amorce GP37, voir chapitre VI.3) comme présenté sur la Figure 57. J'ai obtenu le plasmide pET28a(+)(Kanr)-anaG pGP03. Nous obtenons donc une séquence codant pour le domaine KS, le domaine AT et le domaine ACP à la suite sur une seule et même protéine avec un His-tag en C-terminal. La protéine doit avoir un poids moléculaire de 118 kDa. Nous espérons que sa taille réduite par rapport à l'AnaG entière nous permettrait une expression facilitée.

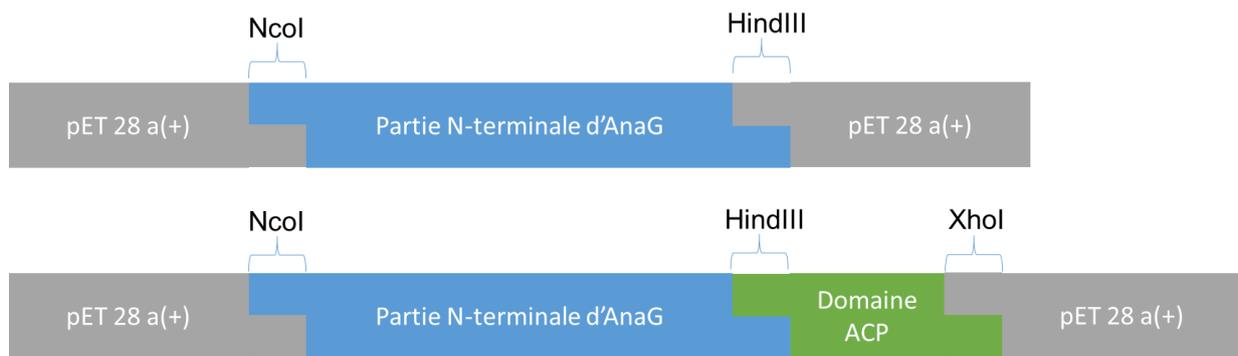


Figure 57. A: Séquence du pET28a(+)(Kan<sup>r</sup>)-anaG pGP02 , B: séquence du pET28a(+)(Kan<sup>r</sup>)-anaG pGP03, XhoI et HindIII correspondent au site de l'enzyme de restriction utilisé pour ce clonage

Je n'ai malheureusement pas réussi à purifier la protéine malgré l'ajout d'inhibiteurs de protéase et la mise au point des conditions de culture.

### III.2.2) Obtention d'AnaA

Pour AnaA, nous avons cloné la séquence nucléotidique correspondant à la séquence peptidique entière de l'AnaA avec un His-tag en N-terminal. J'ai utilisé les sites de restriction EcoRI (amorce GP40, voir chapitre VI.3) et HindIII (amorce GP41, voir chapitre VI.3) pour obtenir le plasmide pET28a(+)(Kan<sup>r</sup>)-anaA. La protéine devrait avoir un poids moléculaire de 32 kDa.

Je n'ai pas eu de problèmes d'expression lors de l'obtention d'AnaA et n'ai pas donc eu besoin de mises au point supplémentaires pour les conditions de purification. J'ai rassemblé les fractions les plus pures, concentrées et évaluées par gel. Sachant que la protéine a bien été retenue par la colonne, nous pouvons penser que l'His-tag en N-terminal est intacte et que la protéine n'a donc pas subi de protéolyse en N-terminal. Nous pouvons voir que le poids moléculaire de la protéine est similaire à celui que nous attendions, nous pouvons donc penser que la protéine n'a pas subi de protéolyse en C-terminal.

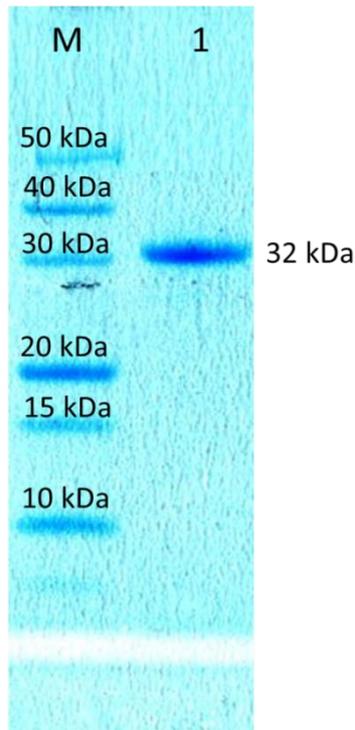


Figure 58. 1 : AnaA après purification, dépôt de 20  $\mu\text{L}$  de protéine à  $0,14 \text{ mg.mL}^{-1}$

Même si AnaA est largement majoritaire, après concentration, il apparaît deux protéines minoritaires, à un plus haut poids moléculaire. Il n'a pas été possible d'éliminer ces protéines par la purification par colonne d'affinité. En effet, il arrive que certaines protéines aient une affinité pour la colonne de nickel proche de celle d'une protéine portant un His-tag à cause d'une séquence peptidique riche en histidine<sup>167</sup>. Étant donné la faible concentration de ces protéines, il n'est pas nécessaire de prévoir d'étapes de purification supplémentaires.

### III.2.3) Discussion

Lors de ces expériences, nous sommes parvenus à obtenir les protéines nécessaires à la poursuite de l'étude de la biosynthèse de l'anatoxine-a. Nous avons cependant rencontré des problèmes dus à l'obtention d'AnaG tronqué. Nous avons réussi à régler une partie de ces problèmes en utilisant des inhibiteurs de protéase. Cette dégradation peut être due à plusieurs facteurs. Le premier peut être que la séquence peptidique d'AnaG contienne une partie flexible aisément hydrolysable. Le deuxième peut être la présence d'une protéase qui de par son activité protéolytique va couper l'AnaG à certains endroits. Le troisième peut être due à un défaut de traduction du ribosome d'*E. coli* s'arrêtant avant la fin de la séquence peptidique. Les résultats que nous avons obtenus semblent indiquer qu'une protéolyse est à l'œuvre étant donné l'effet des inhibiteurs de protéases sur l'expression d'AnaG. Cependant, ceci

n'explique pas le fait que nous n'arrivions pas à obtenir l'AnaG entier avec des inhibiteurs de protéases. Nous pouvons penser que la protéolyse peut être l'un des facteurs de la non-obtention d'AnaG mais cela ne peut pas être le seul.

C'est pour cela que nous avons pensé à tenter l'expression hétérologue de l'AnaG provenant de la souche *Cylindrospermum stagnale* PCC 7417 chez *E. coli*. Un premier clonage a déjà été réalisé par un autre membre de l'équipe au cours de mon doctorat mais nous nous sommes heurtés à des problèmes d'expression et de purification. Il nous faudra préparer d'autres clones avec un placement des His-tags différents.

Concernant le greffage du bras phosphopantéthéine, nous pouvons aussi développer une autre approche. Nous pouvons transformer une souche exprimant déjà la partie C-terminale d'AnaG avec un plasmide portant le gène codant pour l'Ospt. Ceci nous permettrait de réaliser la réaction de greffage directement chez *E. coli* et de purifier une protéine ayant déjà le bras phosphopantéthéine. Des travaux de clonage de l'Ospt chez une souche d'*E. coli* exprimant l'AnaD ont été effectués récemment au laboratoire par un autre membre de l'équipe. L'AnaD est un domaine ACP isolé qui intervient dans la biosynthèse de l'anatoxine-a. Ces expériences de clonage se sont révélées positives et des clones exprimant à la fois l'AnaD et l'Ospt ont été obtenus. Ceci nous a permis de voir si le bras phosphopantéthéine était bien chargé sur le domaine ACP d'AnaD, ce qui est le cas. Par contre nous obtenons un mélange d'apo- et d'holo-AnaD, le chargement du bras ne s'effectue pas chez tous les AnaD présents. Si nous appliquons cette technique pour le chargement du bras phosphopantéthéine sur AnaG, il faudra nous assurer que le chargement a bien eu lieu sur toutes les protéines AnaG.

### **III.3) Conclusion**

Nous avons pu obtenir une partie des éléments nécessaires à l'étude des dernières étapes de la biosynthèse de l'anatoxine-a. Nous savons qu'il est possible d'obtenir ces éléments en poursuivant les expériences que nous avons menées au laboratoire. Une fois que nous aurons obtenus ces éléments, nous pourrions valider notre proposition de voie de biosynthèse et confirmer le mécanisme réactionnel des étapes finales impliquant l'AnaG. Ceci nous permettra aussi de clarifier le rôle d'AnaA dans la voie de biosynthèse.

# Chapitre IV : Etude de la biosynthèse de la dihydroanatoxine-a

## IV.1) La dihydroanatoxine-a

### IV.1.1) Synthèse de la *cis* et *trans* dihydroanatoxine-a

La dihydroanatoxine-a (dhATX) a été synthétisée en 1978 par Bates et al. en mélange d'isomères *cis* et *trans* lors des premières voies de synthèse chimique pour donner l'anatoxine-a<sup>126</sup>. A l'époque, la dhATX n'avait pas encore été détectée en milieu naturel. Sa LD<sub>50</sub> a été déterminée à 2.5 mg.kg<sup>-1</sup> en injection chez la souris. La dhATX serait donc dix fois moins toxique que l'anatoxine-a.

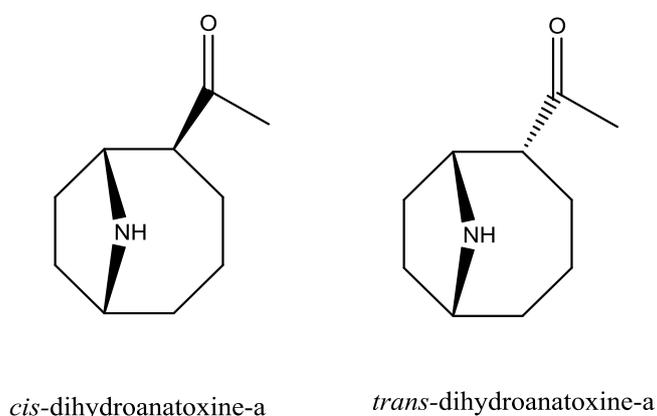


Figure 59. Structure de deux isomères de la dihydroanatoxine-a

### IV.1.2) Toxicité de la dihydroanatoxine-a

Une étude de la toxicité des analogues de l'anatoxine-a a été réalisée en 1991 par Wonnacott et al.<sup>168</sup>. Cette étude a permis de confirmer la toxicité de la dhATX et de déterminer sa constante d'inhibition (K<sub>i</sub>). La K<sub>i</sub> de la dhATX est de 27 nM pour les récepteurs à l'acétylcholine composés des sous-unités  $\alpha 4$  et  $\beta 2$ . Pour les récepteurs à l'acétylcholine composés des sous-unités  $\alpha 7$ , la K<sub>i</sub> de la dhATX est de 1,2  $\mu$ M. Nous retrouvons ici aussi un facteur dix entre ces K<sub>i</sub> et ceux de l'anatoxine-a.

### **IV.1.3) Détection dans l'environnement**

La dihydroanatoxine-a a été détectée pour la première fois dans l'environnement en 1986 après un épisode d'empoisonnement d'animaux au Canada<sup>169</sup>. Il faut cependant rester prudent quant à cette découverte car, à l'époque, seule la spectrométrie de masse a été utilisée pour caractériser les échantillons naturels, sans standards analytiques. Les spectres de masse obtenus présentent 2 isomères pouvant correspondre à la dhATX. A l'époque, il a été proposé que ce soient les formes bateau et chaise de la dhATX. Cependant nous pensons que ce sont les isomères *cis* et *trans*.

### **IV.1.4) Dégradation de l'anatoxine-a**

Par ailleurs, il a été proposé que la dhATX soit un produit de dégradation de l'anatoxine-a<sup>97</sup>. Dans la littérature, il a souvent été avancé que l'anatoxine-a était photosensible et qu'elle se dégradait rapidement dans les conditions naturelles tout en montrant des données parfois contradictoires<sup>170,171</sup>. Une étude récente montre que l'anatoxine-a est stable en présence des conditions suivantes : un pH acide, des radiations et une température supérieure à 20 °C<sup>172</sup>. En revanche, l'anatoxine-a est beaucoup plus sensible lorsqu'elle est à pH neutre et encore plus à pH basique.

Aucun argument sérieux n'a été avancé pour le moment pour supporter l'hypothèse que la dhATX soit un produit de dégradation de l'anatoxine-a.

## **IV.2) Synthèse des standards dihydroanatoxine-a et dihydrohomoanatoxine-a**

Des standards de la dhATX et de la dihydrohomoanatoxine-a (dhHATX) ont déjà été synthétisés et utilisés par le passé par notre équipe<sup>97</sup>. Ces standards avaient déjà été obtenus par hydrogénation catalytique et le produit obtenu comporte donc les deux isomères sans qu'il soit possible de dire lequel est l'isomère *cis* et lequel est l'isomère *trans*. Dans l'optique de déterminer la configuration du composé présent en milieu naturel, l'équipe a réfléchi à une façon d'obtenir et de séparer les deux isomères de façon à pouvoir les caractériser. L'équipe a développé une voie de synthèse<sup>98</sup> en se basant sur la voie développée par Parson comme montré sur la Figure 60. La séparation des isomères *cis* et *trans* de la dhATX et de la dhHATX a été possible. A chaque fois, l'un des deux isomères était majoritaire. La

configuration du Boc-dhATX majoritaire a été confirmée par diffraction au rayon X comme étant l'isomère *cis*. L'autre isomère obtenu est donc le *trans*. Les configurations de la Boc-dhHATX ont été attribuées par analogie des spectres RMN <sup>1</sup>H et <sup>13</sup>C entre les deux composés. Lors de cette synthèse, nous avons remarqué que l'épimérisation entre les deux isomères était possible.

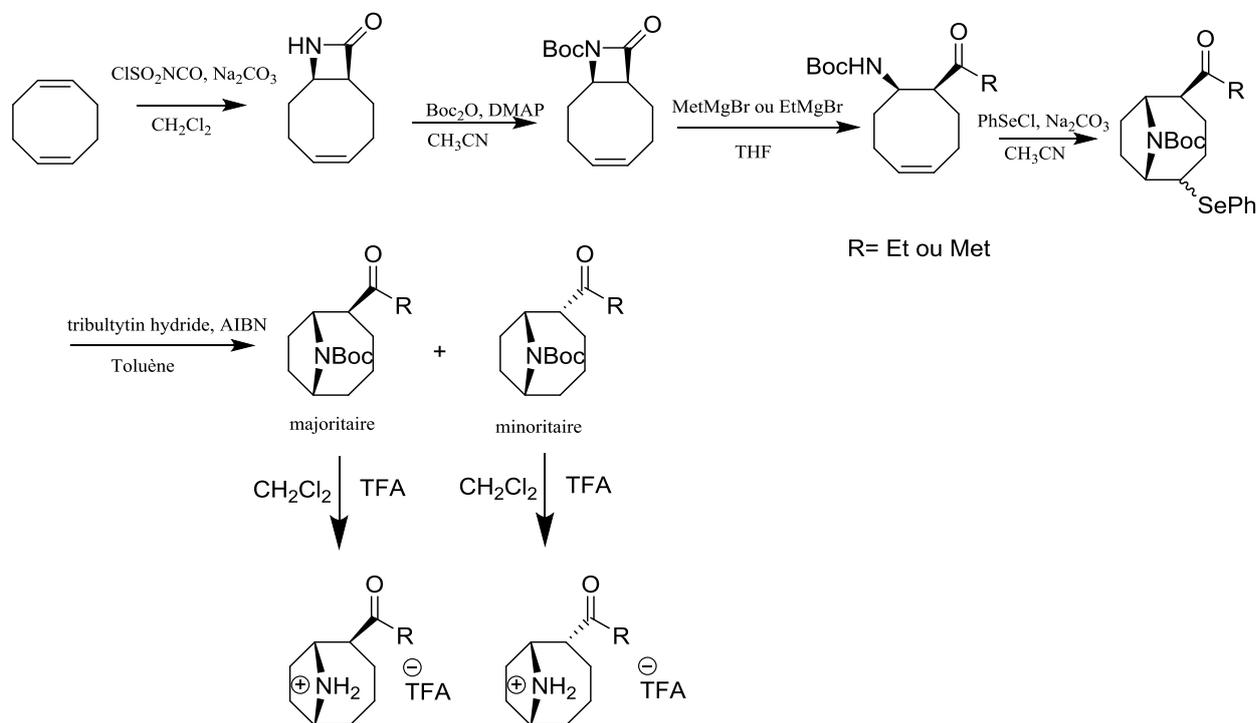


Figure 60. Voie de synthèse des dhATXs et dhHATXs

Après déprotection par le retrait du groupement Boc avec du TFA, ces standards ainsi que des standards de l'anatoxine-a et de l'homoanatoxine-a ont été utilisés en LC-MS<sup>2</sup> pour quantifier la présence de ces toxines dans des échantillons naturels.

### IV.3) Souche productrice de dihydroanatoxine-a

Il existe une souche, *Cylindrospermum stagnale* PCC 7417, qui produit exclusivement la dhATX ce qui a été prouvé par l'équipe à travers des expériences de LC-MS<sup>119</sup> avec les standards dhATX synthétisés par l'équipe<sup>98</sup>. D'après des analyses récentes d'extraits naturels par spectrométrie de masse, il doit cependant exister d'autres souches productrices de dhATX dans l'environnement<sup>124,173,174</sup>. Nous n'avons pas observé d'anatoxine-a chez *Cylindrospermum stagnale* PCC 7417 en LC-MS. Ceci n'est pas consistant avec l'hypothèse

que la dhATX soit un produit de dégradation de l'anatoxine-a. En effet, nous aurions dû au moins observer la présence d'anatoxine-a. L'équipe s'est donc orientée vers la possibilité que la dhATX soit un produit de sa propre voie de biosynthèse chez cette souche. Nous n'avons cependant pas pu déterminer si la dhATX, trouvée chez cette souche, est en configuration *cis* ou *trans* car la LC-MS ne nous permettait pas de séparer les deux isomères.

La publication du génome annoté de *Cylindrospermum stagnale* PCC 7417 nous a amené à penser, en nous basant sur la séquence annotée de ce génome, que le cluster *ana* est responsable de la biosynthèse de la dihydroanatoxine-a<sup>105</sup>.

## **IV.4) Etude de la dihydroanatoxine-a en spectrométrie de masse**

### **IV.4.1) But de l'étude**

L'étude de la dihydroanatoxine-a est, pour nous, de la même importance que celle de l'anatoxine-a. En effet, la dihydroanatoxine-a est, certes, dix fois moins toxique que l'anatoxine-a<sup>95,175,126</sup> mais elle est présente en grande quantité dans de nombreux échantillons naturels et représente donc un danger pour l'homme au même titre que l'anatoxine-a<sup>68,98,124,173,174</sup>.

Nous voulions dans un premier temps confirmer les résultats que nous avons obtenus chez la souche *Cylindrospermum stagnale* PCC 7417 et arriver à déterminer quel isomère est produit en milieu naturel. Pour cela, il faut nous appuyer sur les standards pour différencier en spectrométrie de masse l'isomère *cis* de l'isomère *trans*.

Nous voulons confirmer avec cette étude que la voie de biosynthèse est bien similaire à celle de l'anatoxine-a et qu'elle utilise donc la proline comme produit initial. C'est pour cette raison que nous allons réaliser des expériences d'incorporation isotopiques par feeding en utilisant la même approche que pour l'étude des premières étapes de la voie de biosynthèse de l'anatoxine-a menée au laboratoire<sup>100</sup>.

### **IV.4.2) Analyse des standards en LC-MS**

Nous avons d'abord réexaminé plus en détails toutes les données obtenues en LC-MS en 2011 lors de la synthèse des dhATXs et des dhHATXs<sup>98</sup>. Nous avons donc observé les spectres LC-MS obtenus avec la dhATX et la dhHATX. Ces composés protégés par une

fonction boc sont de configurations connues. Ils sont ensuite déprotégés avec de l'acide trifluoroacétique (TFA) et analysés en LC-MS. En 2011, l'équipe avait visualisé seulement un seul pic par chromatogramme ce qui nous avait amené à penser que les deux isomères n'étaient pas séparables dans ces conditions. En reprenant les données de 2011, nous avons obtenu plusieurs chromatogrammes présentés sur la Figure 61.

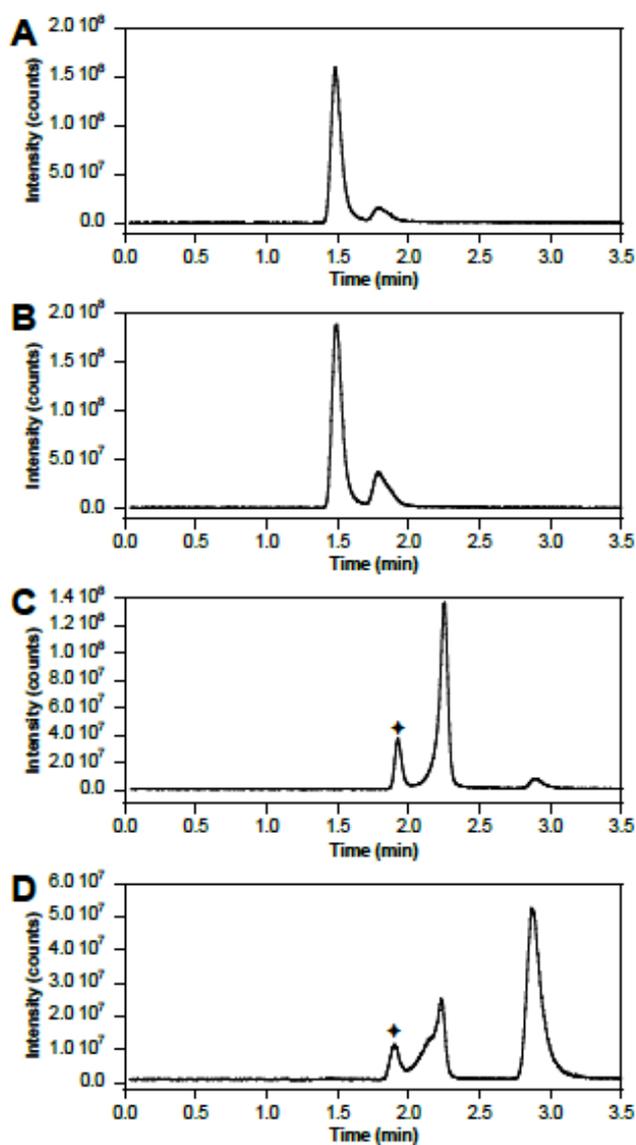


Figure 61. Chromatogrammes obtenus en LC-MS sur une colonne C18 effectué par Mélanie Cohen en collaboration avec l'équipe de Valérie Pichon à l'ESPCI; (A) cis-dhATX, (B) trans-dhATX, (C) cis-dhHATX, (D) trans-dhHATX, \* impureté

Pour la dhATX, les chromatogrammes de l'isomère *cis* et *trans* sont quasiment identiques. Nous remarquons juste que, sur le chromatogramme de la *trans*-dhATX, le pic minoritaire est un peu plus important que chez la *cis*-dhATX. Les chromatogrammes des deux isomères de la dhHATX sont, en revanche, très différents. Les pics majoritaires ne sont pas les mêmes entre les deux isomères. Le pic à 2,2 minutes est majoritaire chez la *cis*-dhHATX et minoritaire chez la *trans*-dhHATX. Le pic à 2,9 minutes est majoritaire chez la *trans*-dhHATX et minoritaire chez la *cis*-dhHATX.

Nous pensons que le fait que nous ayons deux pics à chaque fois correspond à un phénomène d'épimérisation entre l'isomère *cis* et *trans*. L'épimérisation doit être plus lente chez la dhHATX que chez la dhATX. Nous avons donc un pic majoritaire que nous attribuons à l'isomère correspondant au produit avant déprotection et un pic minoritaire que nous attribuons à l'autre isomère. Nous pouvons donc en déduire que le premier isomère à apparaître est l'isomère *cis* et le deuxième est l'isomère *trans*. Par extrapolation, nous pouvons penser que le premier pic obtenu en C18 chez la dhATX est l'isomère *cis* et le deuxième l'isomère *trans*.

Au vu des résultats obtenus concernant l'épimérisation entre les deux isomères *cis* et *trans* de la dhATX et de la dhHATX, nous en avons déduit que l'isomère *cis* est le plus stable. Nous pensons que cela est dû à la formation d'une liaison hydrogène entre l'oxygène de la fonction carbonyle et un hydrogène porté par l'azote. Nous proposons donc un mécanisme d'épimérisation entre les deux isomères avec un passage par une forme énol, montré sur la Figure 62.

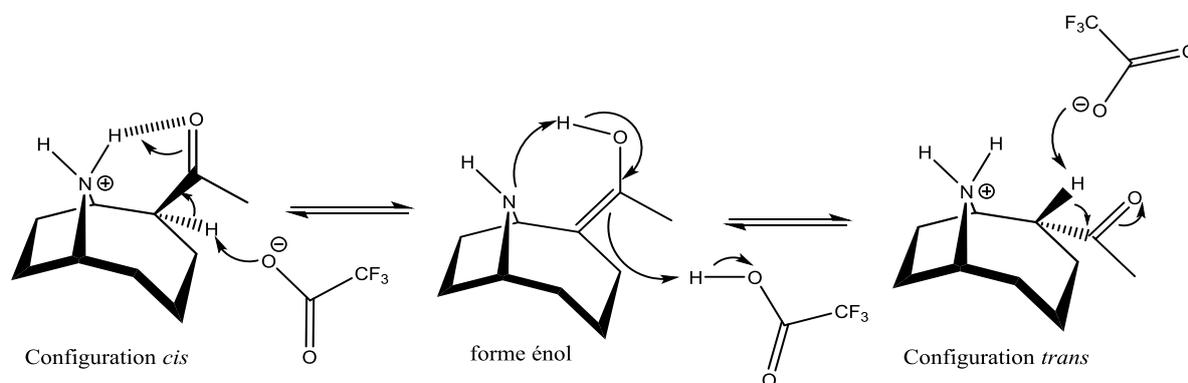


Figure 62. Mécanisme d'épimérisation entre l'isomère *cis* et l'isomère *trans*

Nous avons voulu ensuite vérifier que nous obtenions bien les mêmes résultats en GC-MS. La GC-MS nous permettait aussi d'avoir accès aux fragments ce qui nous permettrait de proposer une voie de fragmentation pour la dhATX.

#### IV.4.3) Analyse des standards en GC-MS

Nous avons réalisé ces travaux à l'école polytechnique en collaboration avec Stéphane Bouchonnet. Nous avons utilisé de la GC-MS en CI MS<sup>2</sup> pour obtenir des chromatogrammes avec une meilleure résolution et de l'EI pour obtenir les spectres de masse avec fragmentation.

Nous avons d'abord injecté en GC-MS la *cis*-dhATX et la *trans*-dhATX que nous avons déprotégées en retirant le groupement protecteur Boc avec du TFA. Lors de l'injection des deux isomères, nous avons obtenu des chromatogrammes relativement similaires.

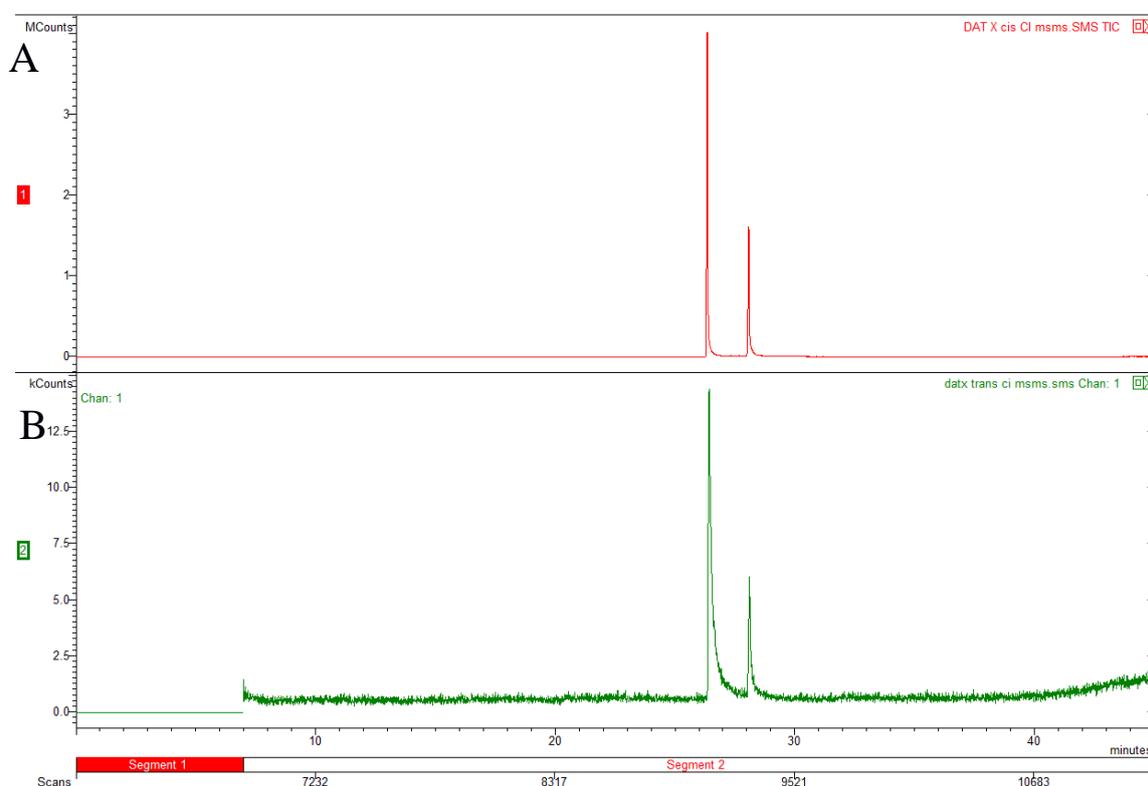


Figure 63. Chromatogramme en CI MS<sup>2</sup> de la *cis*-dhATX (A) et de la *trans*-dhATX (B)

Les spectres de masse de ces deux pics ont un profil similaire. Nous observons un pic majoritaire à  $m/z$  168,0 qui peut correspondre à la dhATX. Nous sommes donc en présence des deux isomères *cis* et *trans* de cette molécule. Les conditions acides utilisées pour la déprotection induisent une réaction d'épimérisation qui conduit aux mêmes proportions des

deux isomères dans les deux cas. Il est aussi possible que les conditions utilisées lors des expériences de GC-MS conduisent aussi à une épimérisation. Il nous faut d'abord vérifier que ce n'est pas le cas.

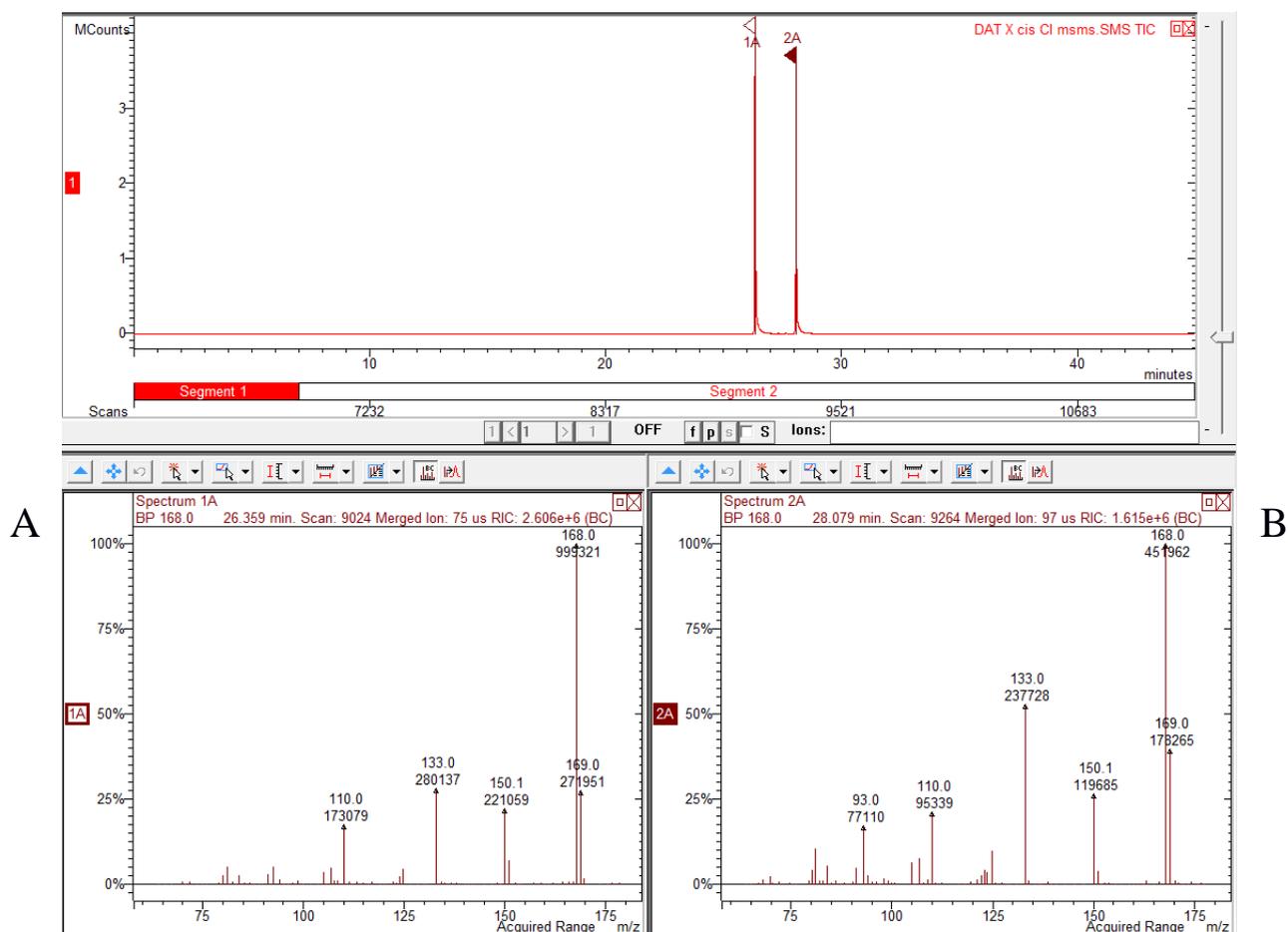


Figure 64. Spectres de masse en CI MS<sup>2</sup> du premier pic à 26,36 minutes (A) et deuxième pic à 28 minutes (B) de la *cis*-dhATX

Pour vérifier que l'épimérisation ne provient pas des conditions utilisées en GC-MS, nous avons donc décidé d'isoler l'un des deux isomères par une purification sur HPLC avec une colonne C18. Nous sommes parvenus à isoler l'un des isomères à partir du mélange obtenu après le retrait du groupement Boc de la *cis*-dhATX. Lorsque ce composé est passé en GC-MS, nous n'observons qu'un seul pic au temps de rétention de l'isomère *cis* avec une masse correspondante à la dhATX, comme montré sur la Figure 65. Nous observons également l'apparition d'un second pic au cours du temps. Sa masse correspond, elle aussi, à la dhATX. Au bout de plusieurs heures à température ambiante, nous avons le même profil chromatographique que celui de la Figure 65. Ceci nous montre qu'il y a bien une épimérisation au cours du temps qui conduit à un mélange des deux isomères dans les mêmes

proportions. Nous avons cependant montré que les conditions utilisées en HPLC ou en GC-MS ne conduisent pas à une épimérisation mais que l'épimérisation a bien lieu au cours du temps.

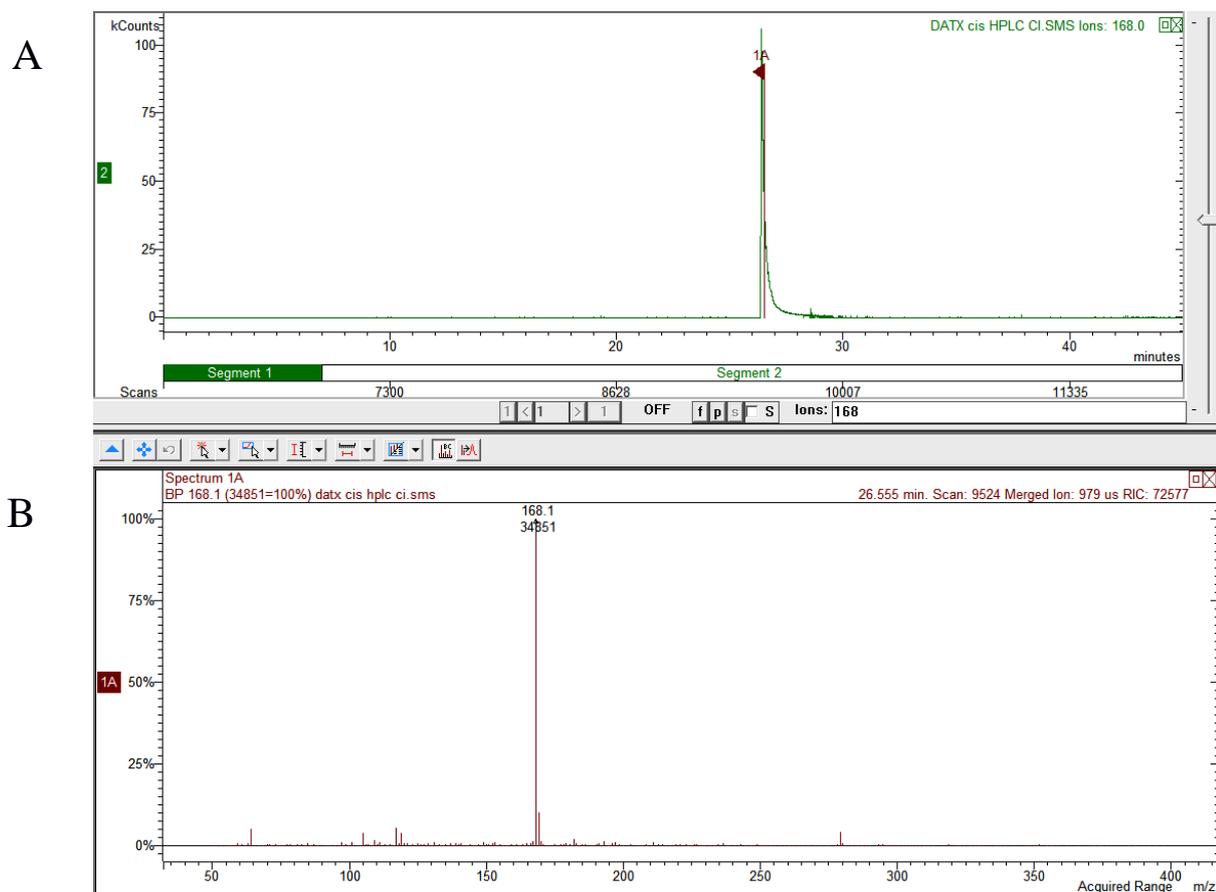


Figure 65. Chromatogramme en CI MS<sup>2</sup> de la *cis*-dhATX purifié en HPLC (A), spectre de masse du pic à 26,5 minutes (B)

Nous avons donc décidé de voir si nous obtenions un résultat similaire à ceux obtenus en LC-MS avec la dhHATX. Les spectres de masse obtenus comportent eux aussi deux pics pouvant correspondre à la dhHATX. En effet, les spectres de masse des pics correspondent à la masse attendue pour la dhHATX. Nous remarquons que les deux chromatogrammes des isomères de la dhHATX ne sont pas similaires contrairement à ceux de la dhATX. Dans les deux chromatogrammes, nous observons, à chaque fois, un pic majoritaire et un pic minoritaire qui varie chez chacun des deux isomères. Lorsque nous injectons les deux isomères de la dhHATX en HPLC nous observons, là aussi, un profil différent, comme montré en Figure 66. Ceci nous confirme bien les résultats de LC-MS et le fait que l'épimérisation est plus lente chez la dhHATX que la dhATX.

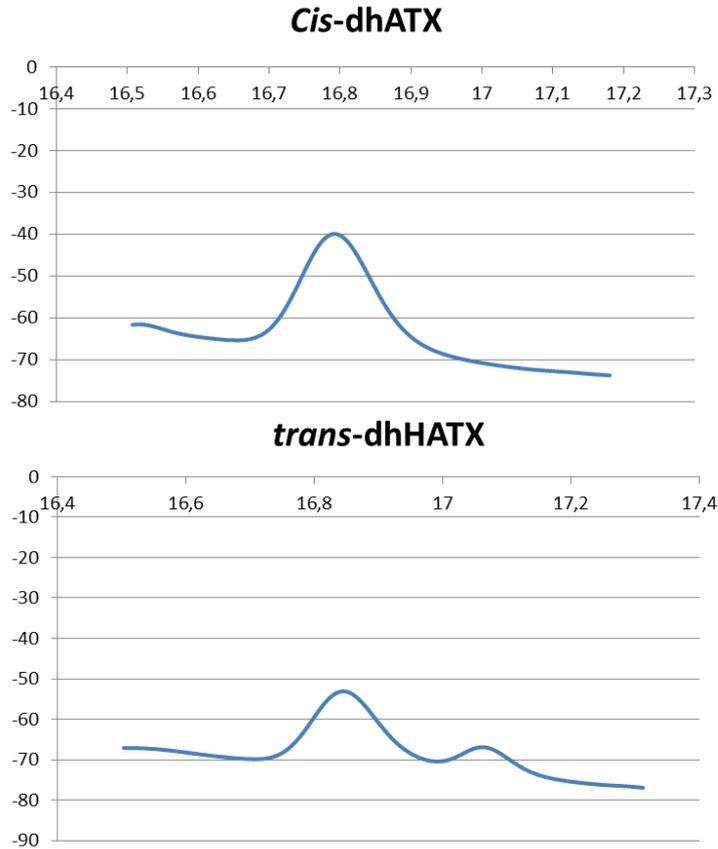


Figure 66. Injection sur HPLC analytique avec en abscisse le temps en minutes et en ordonnée l'absorbance à 230 nm, la colonne utilisé est une C18 analytique avec un volume d'injection de 20  $\mu\text{L}$

Nous avons aussi analysé les spectres de masse obtenus par impact électronique (EI) de la dhATX, comme montré sur la Figure 67. Comme pour l'anatoxine-a et l'homoanatoxine-a<sup>176</sup>, la voie de fragmentation prévalente est celle de la coupure en  $\alpha$  de la cétone ce qui donne des ions fils caractéristiques à des  $m/z$  de 124, 96 et 68. Nous nous baserons sur ces spectres de masse lors des expériences de feeding.

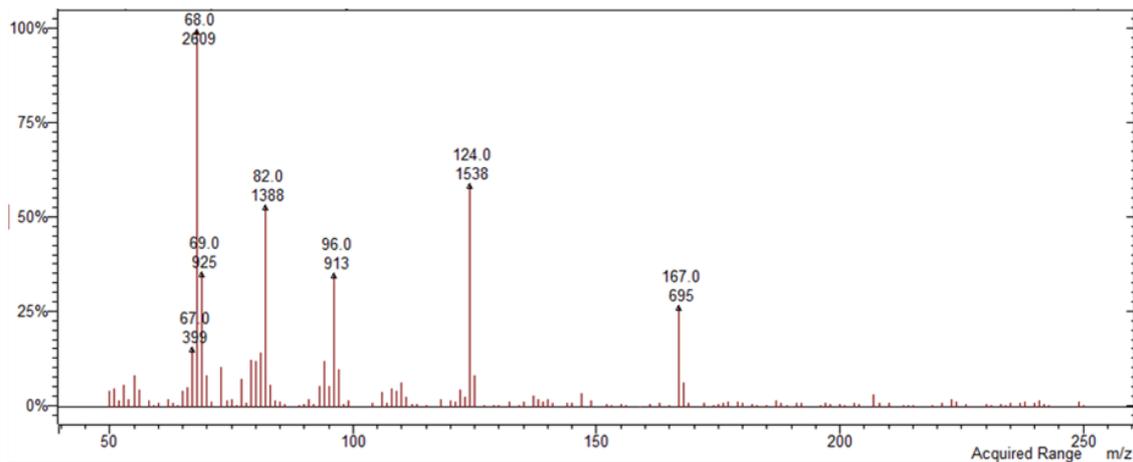


Figure 67. Spectres de masses obtenus en EI pour la dhATX

#### IV.4.4) Extrait de la *Cylindrospermum stagnale* PCC7417

Nous avons mis en culture au laboratoire des souches *Cylindrospermum stagnale* PCC 7417 de façon à pouvoir analyser leur surnageant en GC-MS pour confirmer la production de la dhATX par cette cyanobactérie. Après mise au point des conditions expérimentales, nous obtenons un chromatogramme où nous retrouvons les deux pics correspondant à la *cis*- et *trans*-dihydroanatoxine-a. Nous observons un spectre de masse similaire à celui observé chez les standards, confirmant bien la présence de la toxine. Nous retrouvons, ici aussi, les deux isomères. Ceci indique que l'épimérisation a donc également lieu chez la cyanobactérie. Le rapport entre les deux pics n'est pas le même que celui des standards, nous sommes ici en présence de deux pics de taille égale. Est-ce dû au fait que la culture s'effectue à un pH différent de celui de la réaction de déprotection qui nous permet d'obtenir les standards ? Nous pouvons aussi nous demander si AnaK est stéréosélective et si dans ce cas, l'épimérisation a lieu après la réaction avec AnaK dans la bactérie ou dans le milieu de culture.

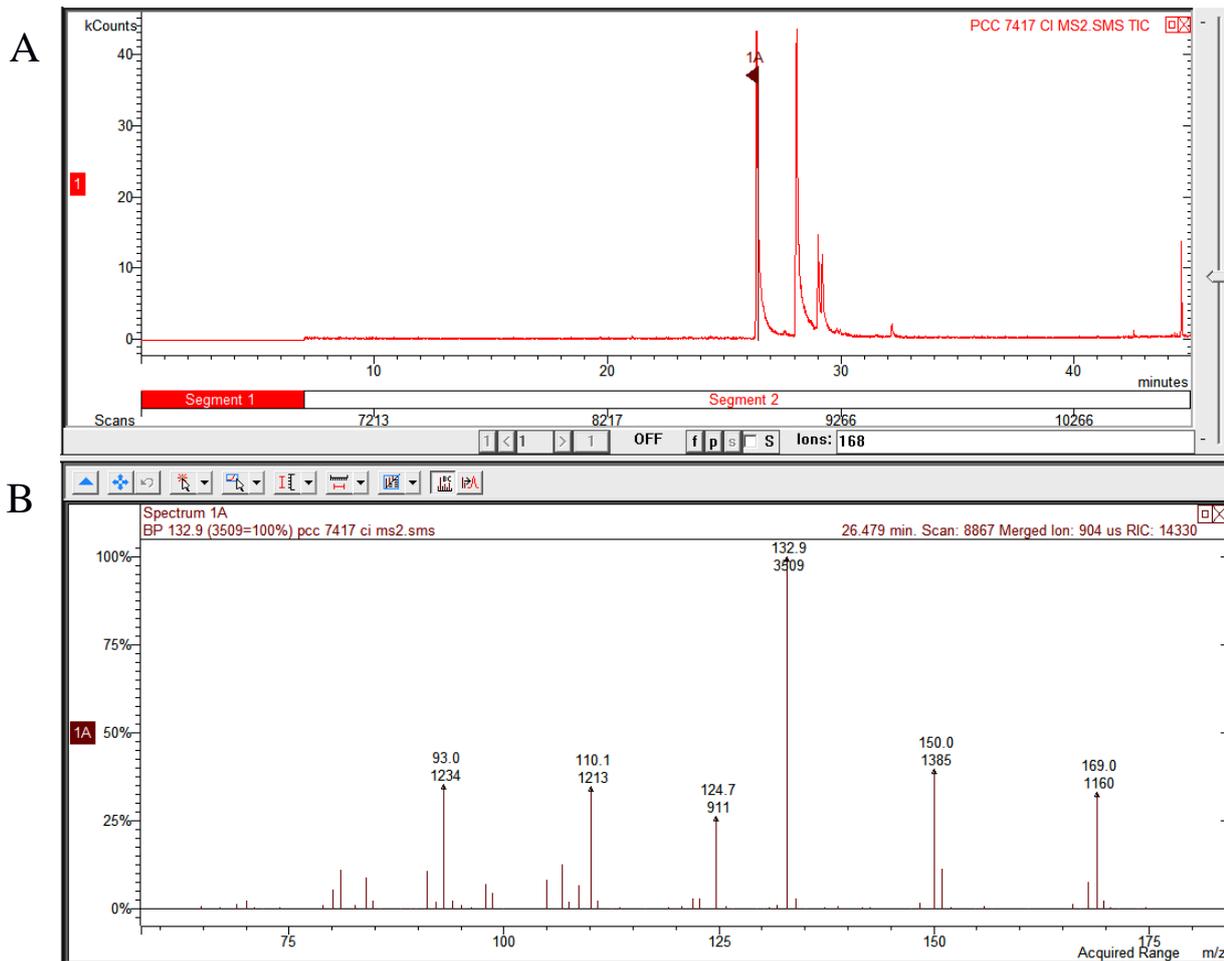


Figure 68. Chromatogramme en CI MS<sup>2</sup> d'un extrait naturel de *Cylindrospermum stagnale* PCC 7417 (A), spectre de masse de la dhATX présente dans l'échantillon (B)

#### IV.4.5) Expérience d'incorporation isotopique par feeding

Maintenant que nous avons confirmé la présence de la dhATX chez la souche *Cylindrospermum stagnale* PCC 7417, nous voulons nous assurer que cette toxine est bien produite par une voie de biosynthèse similaire à celle de l'anatoxine-a. Plus exactement, nous voudrions savoir si la proline est bien le produit de départ de la voie de biosynthèse de la dihydroanatoxine-a. Pour cela, nous avons effectué une incorporation de <sup>13</sup>C et <sup>15</sup>N par feeding dans une culture de *Cylindrospermum stagnale* PCC7417. Si la voie de biosynthèse est similaire à celle de l'anatoxine-a, alors il y aura incorporation de ces molécules marquées dans la dhATX comme montré sur la Figure 69 .Nous nous sommes placés dans les mêmes conditions expérimentales que pour les injections de surnageant de *Cylindrospermum stagnale* PCC 7417 réalisées précédemment.

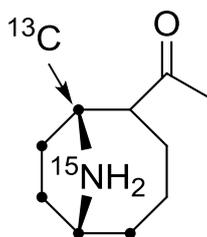


Figure 69. Incorporation des carbones marqués au C<sub>13</sub> et de l'azote N<sub>15</sub> dans la dhATX selon la voie de biosynthèse proposée

Nous observons à nouveau deux pics et nous avons ensuite analysé les spectres de masses de ces deux pics. Les deux spectres de masse sont similaires et il y a une apparition de pics avec une *m/z* plus élevée que chez la *Cylindrospermum stagnale* PCC 7417. Leur *m/z* correspond à une masse +6 par rapport aux pics de masse correspondant à la dihydroanatoxine-a.

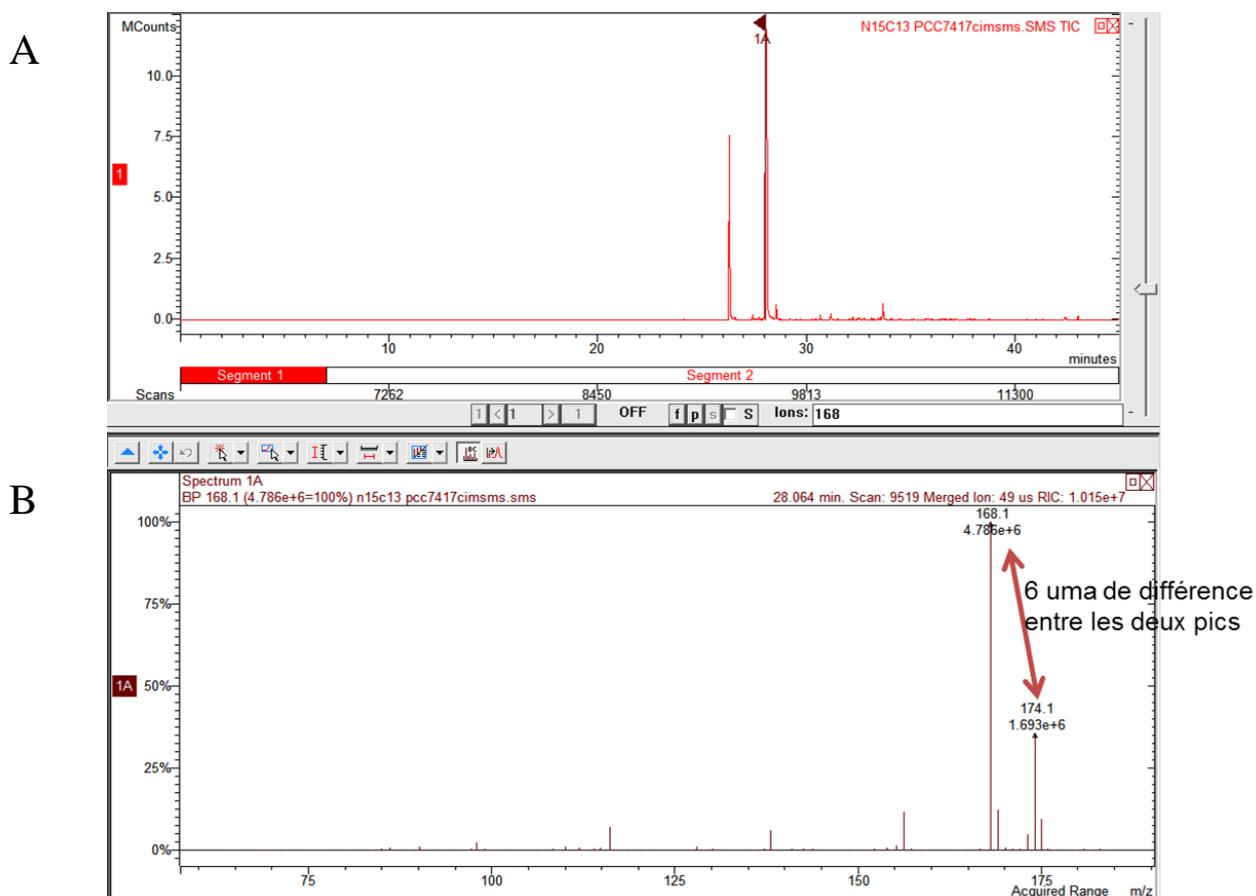


Figure 70. Chromatogramme en CI MS<sup>2</sup> d'un extrait naturel de *Cylindrospermum stagnale* PCC 7417 avec incorporation de Proline marquée (A), spectre de masse de la dihydroanatoxine-a présente dans l'échantillon (B)

Nous avons aussi analysé le spectre de masse obtenu en EI comme montré sur la Figure 71 pour le comparer avec le spectre EI de la dhATX. Nous observons l'apparition de pics à une  $m/z$  de 130 et 73 ce qui correspond à une masse + 6 et +5 pour les ions fils de  $m/z$  124 et 68. Ces résultats sont cohérents avec ceux que l'équipe a obtenus lors d'expériences précédentes d'incorporation isotopiques avec l'homoanatoxine-a<sup>100</sup>. Tous les carbones et l'azote marqués se retrouvent bien dans la dihydroanatoxine-a ce qui montre que la proline est bien le produit de départ de la voie de biosynthèse de la dihydroanatoxine-a.

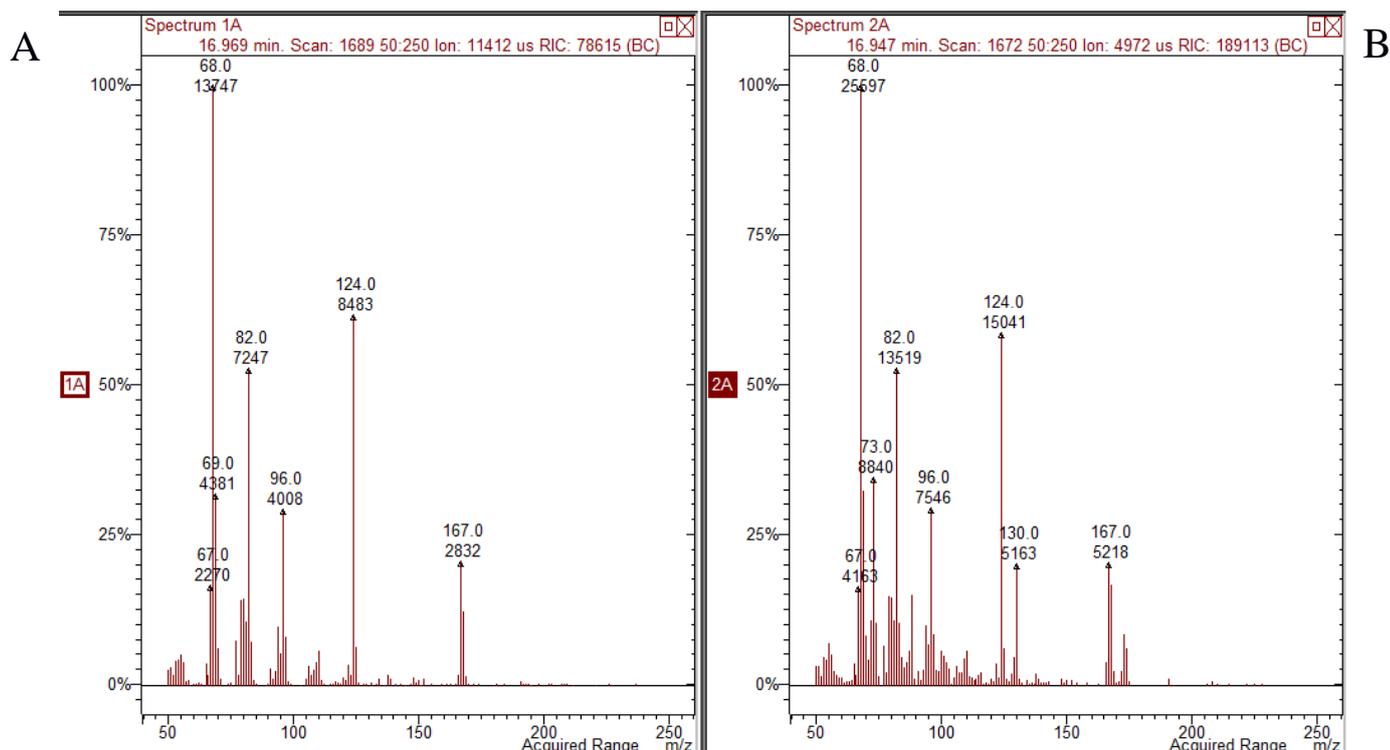


Figure 71. Spectres de masse obtenue en EI (A) pour la dhATX sans incorporation isotopique et (B) pour la dhATX avec incorporation isotopique chez la *Cylindrospermum stagnale* PCC7417

Ces ion-fils montrent que nous sommes face à un mécanisme de fragmentation par clivage en  $\alpha$  de la fonction cétone. En nous basant sur les  $m/z$  des ions caractéristiques à la dhATX et les voies de fragmentation de l'anatoxine-a et l'homoanatoxine-a<sup>176</sup>, nous proposons une voie de fragmentation de la dhATX. Cette voie de fragmentation est représentée sur la Figure 72.

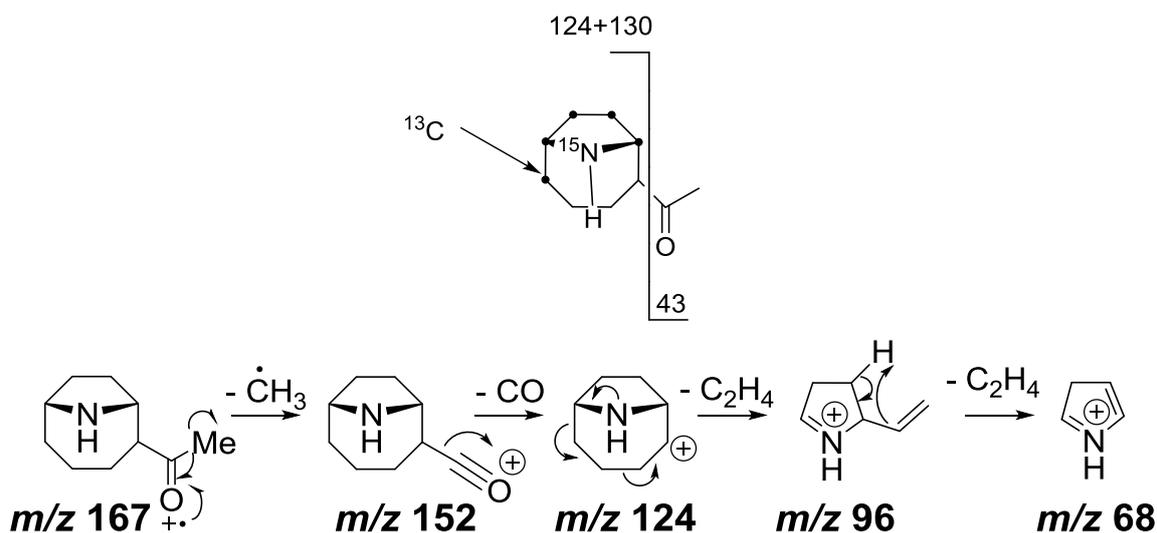


Figure 72. Voie de fragmentation et présence des C<sub>13</sub> et du N<sub>15</sub> sur la dhATX

#### IV.4) Discussions

Nous voulions aussi réaliser une expérience d'incorporation isotopique avec une souche productrice d'anatoxine-a. Notre choix s'est donc porté sur la souche *Oscillatoria* sp. PCC 9107 car c'est une souche productrice d'anatoxine-a qui est aisément disponible. Cette expérience n'a jamais été réalisée et nous voulions pouvoir comparer les spectres de l'anatoxine-a, de l'homoanatoxine et de la dihydroanatoxine-a avec et sans incorporation de Proline marquée.

Concernant la voie de biosynthèse de la dhATX, les résultats des expériences d'incorporation isotopique et la présence du cluster ana chez la *Cylindrospermum stagnale* PCC 7417 nous montrent bien que nous sommes en présence de la même voie de biosynthèse que celle de l'anatoxine-a et de l'homoanatoxine-a. Les premières étapes de ces deux voies doivent donc être similaires. Nous ne pouvons cependant pas donner de réponses concernant la configuration de la dhATX obtenue lors de la biosynthèse à cause du phénomène d'épimérisation.

## IV.5) Présence de réductase dans le cluster *ana* de la *Cylindrospermum stagnale* PCC 7417

La présence de la réductase *anaK* dans le cluster *ana* est l'une des différences de la voie de biosynthèse de la dhATX par rapport à celle de l'anatoxine-a. Cette enzyme serait responsable de la réduction de la toxine ou de l'un de ses précurseurs et pourrait expliquer la formation de la dhATX. Il serait intéressant de connaître le mécanisme de cette enzyme et de savoir à quel moment de la voie de biosynthèse elle intervient. Pour cela, il nous faut exprimer la réductase chez *E. coli* de façon à ensuite purifier AnaK comme nous l'avons déjà fait pour les autres protéines issues du cluster *ana*. Un mécanisme d'action pour cette réductase peut cependant être proposé en se basant sur les mécanismes connus pour ce type de reductase<sup>120</sup> comme montré sur la Figure 73.

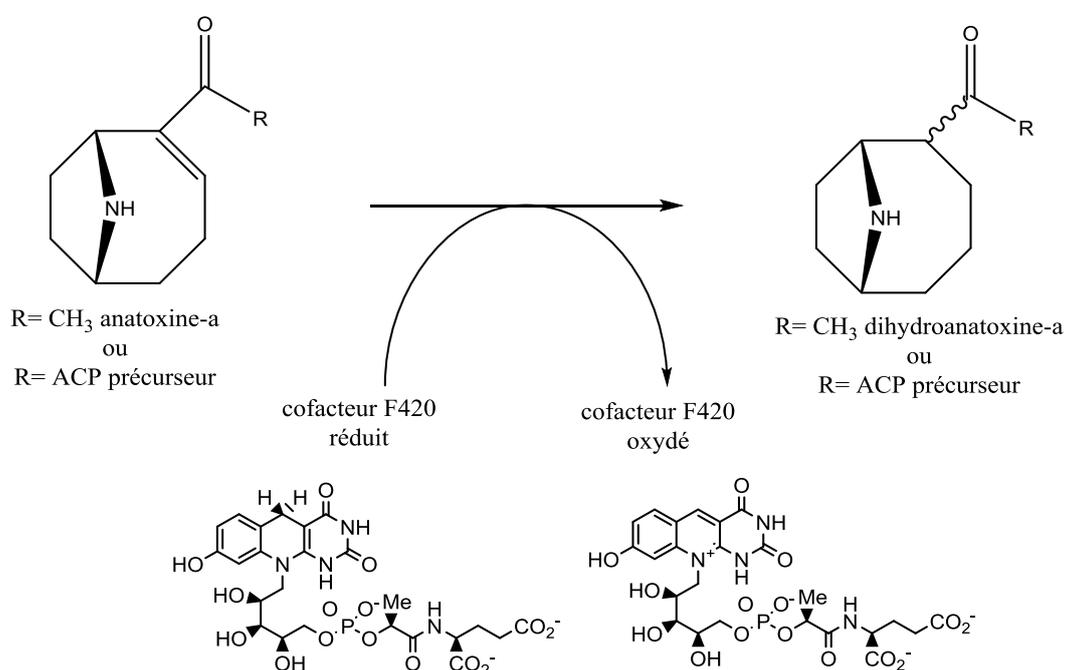


Figure 73. Mécanisme de réaction proposée pour la réductase

Cette réductase nécessite cependant un cofacteur, le F420, pour être fonctionnelle. Il nous faudrait donc obtenir ce cofacteur pour étudier le mécanisme de cette enzyme. Nous devons purifier le cofacteur et nous assurer qu'il est dans son état réduit pour être utilisé avec AnaK. Il existe de nombreuses méthodes de purification de ce cofacteur dans la littérature<sup>177-181</sup>. Il est aussi possible que l'enzyme ne puisse être purifiée qu'en présence de son cofacteur, celui-ci aidant à son repliement.

## **IV.6) Conclusion**

Grâce aux expériences que nous avons effectuées, nous pouvons maintenant affirmer que la dhATX est bien le produit d'une voie de biosynthèse propre et non pas un produit de dégradation obtenu à partir de l'anatoxine-a. Nous espérons que les informations que nous avons pu apporter concernant la dhATX permettront de mieux identifier sa présence dans les milieux naturels et de découvrir de nouvelles souches productrices de cette toxine.

# Chapitre V : Conclusion

## V.1) Bilan général

Au cours de ce travail de recherche, nous avons pu aborder plusieurs aspects de l'étude de la voie de biosynthèse de l'anatoxine-a, de l'homoanatoxine-a et de la dihydroanatoxine-a. Nous avons d'abord pu, avec les clusters *ana* séquencés, établir avec précision la synténie de ce cluster et mettre en avant les différences qui pourraient être le sujet d'un futur projet de recherche. En particulier, les différences de séquence chez la protéine AnaG semblent particulièrement intéressantes pour apporter plus d'information sur la voie de biosynthèse. Nous ne sommes cependant pas en mesure d'apporter des réponses concernant l'impact de ces différences sur la voie de biosynthèse. Nous ne pouvons pas non plus donner des éléments de réponses au niveau de leurs impacts sur le fonctionnement d'AnaG.

Pour les dernières étapes de la voie de biosynthèse de l'anatoxine-a, nous avons réussi à isoler la protéine AnaA ainsi que les parties C et N-terminale de la protéine AnaG ainsi que son domaine ACP d'AnaG. Nous avons pu charger le bras phosphopantéthéine sur la partie C-terminale par une réaction enzymatique. Nous n'avons pas réussi à isoler l'analogue du substrat d'AnaG sous forme pure par synthèse chimique mais nous avons obtenu un mélange de produits. Les expériences de caractérisation que nous avons effectuées sur ce mélange nous laissent penser qu'il contient bien l'analogue recherché. Il nous reste cependant à travailler sur les dernières étapes de la synthèse chimique de façon à pouvoir l'obtenir.

Pour la dhATX et la dhHATX, nous avons pu différencier l'isomère *cis* de l'isomère *trans* grâce à nos expériences de GC-MS et de LC-MS. Grâce aux standards de dihydroanatoxine-a synthétisés par l'équipe, nous avons pu confirmer que *Cylindrospermum stagnale* PCC 7417 est bien une souche productrice de dhATX. Grâce aux expériences d'incorporation isotopique chez *Cylindrospermum stagnale* PCC 7417, nous avons pu confirmer que la dhATX est bien produite à partir de la proline par la même voie de biosynthèse que l'anatoxine-a et l'homoanatoxine-a.

## V.2) Perspectives

L'anatoxine-a est une neurotoxine majeure chez les cyanobactéries. Elle reste pourtant peu étudiée malgré le risque qu'elle représente pour les activités humaines. Il nous reste cependant beaucoup de choses à découvrir sur la voie de biosynthèse de l'anatoxine-a ainsi que ses dérivés.

Il nous faudrait d'abord continuer à poursuivre la synthèse de l'analogue de substrat d'AnaG. C'est l'élément qui nous manque pour pouvoir étudier les dernières étapes de la voie de biosynthèse de l'anatoxine-a. Nous pourrions ainsi obtenir des informations précieuses sur le rôle d'AnaA dans cette voie de biosynthèse. Nous pourrions aussi savoir si AnaA est bien responsable de l'hydrolyse du composé attaché à AnaG. La compréhension du rôle d'AnaA nous donnerait aussi des informations sur le fonctionnement des PKSs et plus particulièrement sur leur domaine TE II. Il nous faut aussi travailler sur la régulation du domaine MT d'AnaG pour nous permettre d'identifier quels sont les facteurs qui favorisent la production d'anatoxine-a ou d'homoanatoxine-a. Nous pourrions aussi nous assurer que le mécanisme de d'AnaG chez *Cylindrospermum stagnale* PCC 7417 est similaire au mécanisme de cette protéine chez *Oscillatoria* sp. PCC 6506. Pour obtenir cette information, il nous faudra purifier la protéine AnaG chez *Cylindrospermum stagnale* PCC 7417 et mettre au point un analogue de substrat similaire à celui préparé pour l'AnaG de *Oscillatoria* sp. PCC 6506.

Avec les standards de dhATX ou de dhHATX, nous pourrions identifier d'autres souches productrices de dhATX ou de dhHATX parmi les échantillons naturels et les isoler. En effet, même si cette toxine est 10 fois moins toxique que l'anatoxine-a et l'homoanatoxine-a, elle est souvent présente en grande quantité dans les échantillons naturels et son impact sur l'environnement ne doit pas être sous-estimé. Il serait par ailleurs intéressant de vérifier que la souche *Cuspidothrix issatschenkoi* CHABD3 ne possède pas le gène *anaK* dans son génome. Si c'est le cas, il serait intéressant de vérifier que cette souche produit ou non de la dihydroanatoxine-a. Nous avons des outils en main pour lancer une étude plus approfondie sur la biosynthèse de la dihydroanatoxine-a.

Concernant AnaK, l'équipe compte étudier son mécanisme et son rôle dans la voie de biosynthèse de la dihydroanatoxine-a. Il nous faudrait alors purifier la protéine AnaK en utilisant une approche similaire à celle que nous avons employée pour la purification des autres protéines impliquées dans la voie de biosynthèse de l'anatoxine-a. Il nous faudra aussi

purifier le coenzyme F420 sous sa forme réduite pour que la réductase soit fonctionnelle. Nous pourrions ensuite savoir si cette enzyme peut réduire l'anatoxine-a pour donner de la dihydroanatoxine-a ou si elle agit sur un précurseur de l'anatoxine-a. Dans ce cas, il nous faudra utiliser à nouveau l'analogue de substrat d'AnaG de façon à nous assurer qu'AnaK intervienne bien à cette étape dans la voie de biosynthèse de la dihydroanatoxine-a.

Concernant la comparaison des différentes séquences des clusters *ana*, nous avons besoin de séquences supplémentaires provenant de cluster d'autres souches avant de pouvoir avoir une vision globale de l'organisation du cluster. Avec ces nouvelles données, nous pourrions établir un arbre phylogénétique du cluster *ana* et le comparer avec l'arbre phylogénétique des souches de cyanobactéries. Nous pouvons espérer qu'avec les progrès constant des techniques de séquençage et l'intérêt actuel pour les cyanobactéries de nouvelles séquences de cluster *ana* seront publiées dans les années à venir. Une meilleure connaissance du génome des cyanobactéries nous permettrait de découvrir et d'identifier de nouveaux métabolites secondaires et ainsi de pouvoir identifier des nouveaux composés toxiques afin de mieux nous en prémunir.

# Partie Expérimentale de Biologie

## VI.1) Souches de cyanobactéries

Les souches axéniques de cyanobactéries *Oscillatoria* PCC 6407 (ATCC 27906), PCC 6506 (ATCC 29081 ou UTEX 1547), et *Cylindrospermum stagnale* PCC 7417 ont été achetées à la collection de l'Institut Pasteur (PCC). Toutes les cultures des souches *Oscillatoria* et *Cylindrospermum* sont cultivées à 20 °C sous lumière constante dans du milieu BG-11. La pureté des cultures est vérifiée régulièrement<sup>182</sup>. Les cultures des cyanobactéries au laboratoire ont été effectuées et mises au point par Annick Méjéan.

## VI.2) Culture d'*Escherichia coli*

Les souches qui ont été utilisées au laboratoire sont : *E. coli* JM109 (PROMEGA) utilisée pour le clonage et la production des plasmides et *E. coli* BL21(+)(DE3)-RIL(Cm<sup>R</sup>) (STRATAGENE) utilisée pour la surinduction des gènes clonés<sup>183-185</sup>.

Ces souches sont cultivées à 37 °C sous une agitation de 200 rpm dans le milieu Lysogeny Broth (LB)<sup>186,187</sup>. En plus du milieu LB, nous utilisons au laboratoire du milieu Super Optimal Broth with catabolite repression (SOC)<sup>188</sup> dans les mêmes conditions.

Après solubilisation des différents composants du milieu dans de l'eau distillée, le pH est ajusté à 7,0 puis les milieux sont soumis à un autoclavage pendant 20 min à 120 °C à une pression de 1 bar.

Les cultures d'*E. coli* ont été réalisées à l'aide d'un incubateur infors HT Minitron pour les cultures en milieux liquides et d'un incubateur Labnet Mini Incubator pour les cultures en milieux solides.

## VI.3) PCR

Les PCRs ayant pour but d'obtenir des inserts pour un clonage ou pour le séquençage d'un gène ou d'un contig sont réalisées avec une polymérase fidèle et processive : la Phusion

fournie par New England Biolabs. Au contraire, les PCRs permettant un « screen » ou un contrôle sont réalisées avec une polymérase conventionnelle Taq Polymerase (Promega). Pour des matrices riches en GC et structurées, c'est la polymérase LongAmp (NEB) qui est utilisée. Les amorces utilisées synthétisées par Eurogentec sont listées dans le Tableau 3. Le tampon et les dNTPs sont fournis par les fournisseurs pour chaque polymérase. Les amplifications sont faites selon les recommandations des fournisseurs. Les amorces sont dessinées avec l'aide du logiciel Serial Cloner 2.6.

La vérification des produits des PCR se fait par dépôt sur gel d'agarose de 0,7 à 1% (Agarose LE, Analytical Grade, Promega) dans du tampon TAE 1X. L'intercalant utilisé pour les gels d'agarose est le GelRed™ Nucleic Acid Gel Stain d'Interchim. La migration se fait à 100 V à ampérage constant pendant 30 à 45 minutes.

Les PCRs sont effectuées sur un appareil de PCR Eppendorf™ MasterCycler Personal Thermal Cycler.

| Séquence amplifiée | Application           | Nom de l'amorce | Séquence de l'amorce de 5' en 3' | Sites de restriction |
|--------------------|-----------------------|-----------------|----------------------------------|----------------------|
| Cluster <i>ana</i> | Séquençage du cluster | GP1             | CCTTGCTTAAAATTATAGCAAGCTAC       |                      |
|                    |                       | GP2             | CAGAATCGCCTGTAACCAACCTCTT        |                      |
|                    |                       | GP3             | TACTAGCACGACCATACATTTGTCG        |                      |
|                    |                       | GP4             | GTGAGGATTATGATTCTGTTAAGGC        |                      |
|                    |                       | GP5             | CAGCTATTTAATTGTAAGATTTCCGG       |                      |
|                    |                       | GP6             | CCAAGACGTTTCGAGCATCAAAAAG        |                      |
|                    |                       | GP7             | CGGTTAGTTATAGTTCGCTGCTGCG        |                      |
|                    |                       | GP8             | GATTGCTGACTTGGCATGATTGTGG        |                      |
|                    |                       | GP9             | GGTCAATGGCAGGGAGATCGAGCTC        |                      |
|                    |                       | GP10            | TGGGAGGATTCCAGGTTGGTGGACG        |                      |
|                    |                       | GP11            | TTTAGCACAAGGAGATATTCTCCGC        |                      |
|                    |                       | GP12            | ATTATACTGAGCTGGAAGTATGAGG        |                      |
|                    |                       | GP13            | AATAGCAACTTGGCTGAACGATCCG        |                      |
|                    |                       | GP14            | GTCTGACATCGCACTCGCAACCC          |                      |
|                    |                       | GP15            | AAGGTTGCGGCATAGTCGTTCTCA         |                      |
|                    |                       | GP16            | TGAGTGAAATGTTACTTCCTCCACC        |                      |
|                    |                       | GP17            | AGTACGATCCGCTCCAGTTG             |                      |
|                    |                       | GP18            | CGGGGCTATTACGATCTGATTATTGC       |                      |
|                    |                       | GP19            | TCGTTTGAAAAACGACTGT              |                      |
|                    |                       | GP20            | GAAGGAGTTGATAACCCAGT             |                      |
|                    |                       | GP21            | CGCCAACAGTATCGTATCTCGGC          |                      |
|                    |                       | GP22            | AACAGTATCGTATCTCGGCTATACC        |                      |

|                            |                     |      |   |         |
|----------------------------|---------------------|------|---|---------|
|                            |                     | GP23 | TTCAAACCCTTGTGTAATAACCC                                 |         |
|                            |                     | GP24 | AAATTCAGCGATTGAAGCGGAAAAA                               |         |
|                            |                     | GP25 | CGCTCCATTTATCGTTACTAAATGC                               |         |
|                            |                     | GP26 | TTCACCTGAAAATATTAATAGTATTG                              |         |
|                            |                     | GP27 | AATTGACTTTTTTCTGATATGTTTAC                              |         |
|                            |                     | GP28 | TAATACGACTCACTATAGGG                                    |         |
| T7                         | Criblage des clones | GP29 | GCTAGTTATTGCTCAGCGG                                     |         |
| Domaine ACP d' <i>anaG</i> | Clonage             | GP30 | GCC <u>CATATG</u> AAAAAGAAAGAGAGAGAAGAACTTC             | Nde I   |
|                            |                     | GP31 | CGCT <u>CGAG</u> ATCATCAAACCTTTACAAAAGTA                | Xho I   |
|                            |                     | GP32 | GCC <u>GATCC</u> TTAATCATCAAACCTTTACAAAAGT              | BamHI   |
|                            |                     | GP33 | GCC <u>CCATGGG</u> GAAAAAGAAAGAGAGAGAAGAAACTTC          | NcoI    |
| <i>anaG</i>                | Clonage             | GP34 | GCC <u>CATGGG</u> CAGTGAAAATAACGAACAACGC                | NcoI    |
|                            |                     | GP35 | GCC <u>AAGCTT</u> GCCTTGTGTTGGCAAGGGTTG                 | HindIII |
|                            |                     | GP36 | CGG <u>AAGCTT</u> CAAGAAAAGCCTGGGGCG                    | HindIII |
|                            |                     | GP37 | CGCT <u>CGAG</u> ATCATCAAACCTTTACAAAAGTA                | Xho I   |
| <i>osppt</i>               | Clonage             | GP38 | GGCT <u>CGAGT</u> AAGAAGGAGATATACCATGAATGTTTCTGATTGCCTC | XhoI    |
|                            |                     | GP39 | GCCT <u>CGAGT</u> CAGCTTAAGATATAATTCTCTG                | XhoI    |
| <i>anaA</i>                | Clonage             | GP40 | GCG <u>AATTC</u> AAAACAGAAATACCTAATCTTTGG               | EcoRI   |
|                            |                     | GP41 | CGC <u>AAGCTT</u> TTTATAAATAGTCTTTTAGCGTTTC             | HindIII |
| <i>anaD</i>                | clonage             | GP42 | GCCG <u>CATATG</u> ACAGAAGAAAAAGTCGAAC                  | NdeI    |
|                            |                     | GP43 | CGC <u>GAATTC</u> CATACCTGTGACACCAATTAC                 | EcoRI   |

Tableau 3 : Amorces utilisées pour les PCRs. Les sites de restriction sont en gras et soulignés

La matrice pour les PCR est de l'ADN génomique extrait et purifié par Annick Méjean à partir des différentes souches de cyanobactéries cultivées au laboratoire et de l'ADN plasmidique issus des souches *E. Coli* purifiés avec le kit QIAGEN Plasmid Mini Kit selon le protocole fourni (QIAGEN).

Les produits de PCR sont purifiés avec le kit Wizzard® PCR Preps DNA purification System selon le protocole fourni (Promega) ou par le kit QIAquick® PCR Purification Kit selon le protocole fourni (QIAGEN).

Les mesures de concentration en ADN ont été effectuées à l'aide d'un nanodrop NanoVue Plus.

Le séquençage des produits PCR est réalisé par la société GATC, selon le protocole de Sanger<sup>189,190</sup>. Les chromatogrammes obtenus ont été lus avec le logiciel Chromas Lite 2.1.1.

## **VI.4) Préparation de bactéries compétentes**

### **VI.4.1) Préparation d'*E. coli* JM109 électrocompétentes**

On inocule 400 mL de LB (avec le ou les antibiotique(s) adapté(s)) à 1% vol/vol avec une préculture. La culture est incubée jusqu'à  $DO_{550nm} \sim 0,7$ . La culture est transférée dans deux pots de centrifugation de 500 mL, conservée à 4 °C, puis centrifugée à 4 000 g pendant 15 minutes, à 4 °C. Chaque culot est repris avec 100 mL d'une solution aqueuse, contenant 10% de glycérol, conservée à 4 °C, puis centrifugé à 4 000 g pendant 15 minutes à 4 °C. Ce lavage est répété une deuxième fois. Chaque culot est repris avec 100 mL d'une solution aqueuse, contenant 10% de glycérol conservée à 4 °C. Les deux solutions sont ensuite réunies puis centrifugées à 4 000 g pendant 20 minutes, à 4 °C. Le surnageant est délicatement enlevé et le volume est ajusté à 1 mL à l'aide d'une solution aqueuse contenant 10% de glycérol à 4 °C. La DO à 550nm du culot est mesurée en diluant 3  $\mu$ L du culot dans 1 mL de LB. Cette DO doit être comprise entre 0,6 et 0,75. Les cellules sont ensuite aliquotées par 50  $\mu$ L par microtubes de 1,5 mL puis congelées dans de l'azote liquide. Les cellules sont conservées à -80 °C. Toutes ces manipulations sont réalisées en condition stérile avec du matériel stérilisé préalablement à l'autoclave.

### **VI.4.2) Préparation d'*E. coli* BL21(+)(DE3)-RIL(Cm<sup>R</sup>) chimiocompétentes**

On inocule 10 mL de LB (avec le ou les antibiotique(s) adapté(s)) à 1% vol/vol avec une préculture. La culture est incubée jusqu'à  $DO_{600nm} \sim 0,5$ . La culture est ensuite transférée dans deux tubes falcon de 50 mL froids et la culture est refroidie 10 minutes dans la glace. La culture est centrifugée à 4 000 g pendant 10 minutes à 4 °C puis le culot est repris avec 5 mL d'une solution aqueuse de CaCl<sub>2</sub> à une concentration de 0,1 M. Les solutions sont ensuite centrifugées à 4 000 g pendant 10 minutes à 4 °C. Le culot est repris avec 2 mL d'une solution aqueuse de CaCl<sub>2</sub> à une concentration de 0,1 M puis les deux solutions sont réunies. Les cellules sont ensuite aliquotées par 100  $\mu$ L par microtubes de 1,5 mL, congelées dans de l'azote liquide, puis sont conservées à -80 °C. Toutes ces manipulations sont réalisées en condition stérile avec du matériel stérilisé préalablement à l'autoclave ou stérilisé par filtration dans le cas de la solution de CaCl<sub>2</sub>.

## **VI.5) Clonage**

### **VI.5.1) Digestion enzymatique**

Les digestions par des enzymes de restrictions sont réalisées selon les recommandations des fournisseurs (NEB et Fermentas). Pour les doubles digestions, si les tampons de digestion des enzymes ne sont pas compatibles entre eux, les digestions sont effectuées séquentiellement. Les incubations d'enzymes de restriction ont été effectuées à l'aide d'un bain à température contrôlée Julabo ED. La température d'incubation suit les recommandations des fournisseurs pour les enzymes de restriction.

Les produits de digestion sont purifiés avec le kit Wizzard® PCR Preps DNA purification System selon le protocole fourni (Promega) ou par le kit QIAquick® PCR Purification Kit selon le protocole fourni (QIAGEN).

### **VI.5.2) Déphosphorylation**

Lorsque des grands inserts sont utilisés ou lorsque le clonage a déjà échoué ou lorsqu'un seul site de restriction est utilisé, afin d'améliorer le taux de transformants une déphosphorylation du plasmide après digestion est réalisée par la phosphatase alcaline intestinale de veau (CIAP, NEB) dans les conditions décrites par les fournisseurs.

### **VI.5.3) Ligation**

Une fois digérés, le vecteur et l'insert sont incubés (avec en moyenne un rapport molaire insert/vecteur compris entre 3 et 10, et une masse de vecteur compris entre 10 et 50 ng) à 50 °C, pendant 5 minutes de façon à dissocier les complexes vecteur-insert, insert-insert ou vecteur-vecteur qui auraient pu apparaître. Une fois que la solution est revenue à température ambiante, il est ajouté de la T4 DNA ligase à la concentration de 20 U/μL finale et son tampon 1X (NEB) pour un volume final de 20 μL. Le mélange complet est incubé pendant 16h, à 16 °C puis 10 minutes à 65 °C pour inactiver la ligase.

Pour des clonages difficiles, le rapport molaire insert/vecteur peut être augmenté dans la solution pour obtenir un résultat optimal.

## **VI.5.4) Transformation bactérienne**

### **VI.5.4.1) Transformation par électroporation**

La transformation des bactéries compétentes de souche *E. coli* JM 109 par les produits de ligation est réalisée par électroporation. Le produit de ligation est d'abord dessalé pendant 30 minutes sur une pastille de dialyse reposant sur de l'eau milliporée. 50µL de cellules compétentes sont ensuite incubées en présence de 5 à 10 ng de produit de la ligation dans la glace pendant 2 minutes, puis elles sont transférées dans un tube d'électroporation maintenu à 4 °C dans la glace. Les cellules sont alors soumises à un choc électrique à 2.5KV, 25µF et 200W dans l'électroporateur. Après une heure d'incubation à 37 °C sous agitation en présence de 1 mL de SOC, le produit de transformation est étalé sur des boîtes de pétri contenant du LB contenant 37,5% agar, auquel a été ajouté les antibiotiques adaptés, et incubé pendant la nuit, à 37 °C.

### **VI.5.4.2) Transformation par choc thermique**

On transforme une souche dédiée à l'expression de protéines, *E. coli* BL21(+)(DE3)-RIL(Cm<sup>R</sup>), par une transformation chimique. 200 µL de cellules compétentes sont transférées dans un microtube stérile auquel le produit de ligation est ajouté (<10µL et <50ng). Le mélange est incubé 10 minutes dans la glace puis 90 s, à 42 °C au bain-marie et à nouveau 2 minutes dans la glace. On ajoute 1 mL de SOC et incube 1h, à 37 °C sous agitation. Le produit de transformation est étalé sur des boîtes de pétri contenant du LB contenant 37,5% agar, auquel a été ajouté les antibiotiques adaptés, et incubé pendant la nuit à 37 °C.

### **VI.5.5) Criblage des clones**

Après la transformation, pour les clones obtenus sur boîte, les constructions sont vérifiées par PCR sur colonies grâce aux amorces spécifiques du plasmide qui encadre l'insert. Dans le cas du plasmide pET 28a, les amorces s'hybrident aux promoteurs et terminateur T7 du plasmide. Les conditions de la PCR sur colonie utilisée, ici, sont identiques aux PCRs décrites précédemment avec pour seule différence de remplacer le microlitre d'ADN par une colonie de bactéries issue de la transformation, prélevée à l'aide d'un cône. Le séquençage de ces produits PCR nous permet de confirmer la transformation.

Différents plasmides ont été obtenus par l'équipe grâce à cette technique. Ceux qui ont été utilisés lors de mon doctorat sont listés dans le Tableau 4. Sur tous ces plasmides, deux plasmides (pET28a(+)(Kan<sup>r</sup>)-anaG, pET28a(+)(Kan<sup>r</sup>)-anaG V954-stop et pET28a(+)(Kan<sup>r</sup>)-ACP 1 anaG Histag N terminal ) n'ont pas été obtenus par moi-même mais sont l'œuvre de la doctorante précédente Rabia Mazmouz. Toutes les constructions ont été obtenues à partir du pET28a(+)(Kan<sup>r</sup>) (NOVAGEN)<sup>191</sup>.

| Plasmide  | Caractéristiques  | Origine     |
|---|---|-------------|
| pET28a(+)(Kan <sup>r</sup> )-anaA                                       | Vecteur d'expression contenant anaA   | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-ACP 1 anaG Histag N terminal               | Vecteur d'expression contenant le domaine ACP d'anaG issu du premier clonage avec Histag en N terminal                | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-ACP 2 anaG Histag C terminal               | Vecteur d'expression contenant le domaine ACP d'anaG issu du deuxième clonage avec Histag en C terminal               | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-ACP 2 anaG Histag N terminal               | Vecteur d'expression contenant le domaine ACP d'anaG issu du deuxième clonage avec Histag en N terminal               | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-ACP 2 anaG Histag N terminal et C terminal | Vecteur d'expression contenant le domaine ACP d'anaG issu du deuxième clonage avec Histag en N terminal et C terminal | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-anaG                                       | Vecteur d'expression contenant anaG   | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-anaG V954-stop                             | Vecteur d'expression contenant la partie C-terminal d'anaG  | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-anaG pGP02                                 | Vecteur d'expression contenant la partie N-terminal d'anaG  | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-anaG pGP03                                 | Vecteur d'expression contenant la séquence de l'anaG synthétique  | Laboratoire |
| pET28a(+)(Kan <sup>r</sup> )-anaD                                       | Vecteur d'expression contenant la séquence de l'anaD  | Laboratoire |

Tableau 4 : Récapitulatif des plasmides utilisés au laboratoire

## VI.6) Expression et purification de protéines chez *E. coli*

Toutes les expressions et purifications de protéines ont été réalisées chez la souche d'*E. coli* BL21(+)(DE3)-RIL(Cm<sup>R</sup>). Ces manipulations suivent un schéma identique.

### VI.6.1) Préculture

Un volume de LB, auquel est ajouté des solutions d'antibiotiques adaptés comme montré sur le Tableau 5, estensemencé avec une souche de *E. coli* mise en collection à -80 °C. La culture est incubée à 37 °C sous agitation à 200 rpm pendant une nuit.

| Souche utilisée   | Antibiotique avec concentration finale |
|---|--|
| <i>E. coli</i> JM 109                                       | Pas d'antibiotique                     |
| <i>E. coli</i> JM 109-[pET 28a]                             | Kan 50 µg/ml                           |
| <i>E. coli</i> BL21(+)(DE3)-RIL(Cm <sup>R</sup> )           | Cm 30 µg/ml                            |
| <i>E. coli</i> BL21(+)(DE3)-RIL(Cm <sup>R</sup> )-[pET 28a] | Cm 30 µg/ml + Kan 50 µg/ml             |

Tableau 5 : Antibiotiques utilisés avec les différentes souches d'*E.coli*

### VI.6.2) Culture et induction

Un volume de LB, auquel est ajouté des solutions d'antibiotiques adaptés comme montré dans le Tableau 5, estensemencé à 1% vol/vol avec une préculture de la veille. La culture est incubée à 37 °C sous agitation à 200 rpm jusqu'à une DO<sub>600</sub> d'environ 0,6. A ce moment, de l'Isopropyl-beta-D-Thiogalactopyranoside est ajouté pour atteindre une concentration finale de 200 µM afin d'induire l'expression du gène cloné sous le contrôle du promoteur T7. La culture est refroidie dans un bain de glace jusqu'à 15 °C et est incubée à 15 °C toute la nuit sous agitation à 200 rpm, puis les cellules sont centrifugées à 4000 g pendant 20 minutes, à 4 °C. Le culot est ensuite repris dans 10 mL d'une solution aqueuse tamponnée à pH 7,4 de NaH<sub>2</sub>PO<sub>4</sub> à une concentration 0,02 M. La solution est ensuite centrifugée à 4000 g pendant 20 minutes, à 4 °C. Le culot est stocké à - 20 °C.

### VI.6.3) Sonication de la culture

Le culot cellulaire est repris dans 10 mL de tampon A (20 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7,4 ; 500 mM NaCl et 20 mM d'imidazole) par litre de culture. Les cellules sont ensuite cassées par sonication, dans la glace, par une succession de cycles de 15 s de sonication à 500 W et 1 minute de repos, jusqu'à 4 minutes de sonication totale. Le lysat est centrifugé à 13 500 g pendant 30 minutes, le surnageant est récupéré puis centrifugé à 13 500 g pendant 15 minutes, à 4 °C.

### VI.6.4) Purification des protéines sur colonne d'affinité

Les purifications de protéines sont effectuées à l'aide de l'ÄKTA prime<sup>TM</sup> de GE Healthcare. Les colonnes d'affinité utilisées avec cet appareil sont des colonnes HisTrap 1 mL ou 5 mL GE Healthcare ou des colonnes Protino Ni-NTA 1 mL Macherey-Nagel. Les colonnes échangeuses d'anions utilisées avec cet appareil sont des colonnes HiTrap QFF 1 mL

ou 5 mL GE Healthcare.

Le surnageant obtenu (entre 5 et 10 mL en fonction de la taille de la culture) par centrifugation est injecté sur une colonne d'affinité au nickel qui a été préalablement équilibré par un tampon A. Dans le cas de protéines fragiles comme AnaG et AnaG V954-Stop, nous avons ajouter des inhibiteurs de protéases au tampon pour éviter la protéolyse. Nous utilisons au laboratoire des pastilles cOmplete, EDTA-free (ROCHE).

Après un lavage avec 35 mL du tampon A, les protéines sont éluées par un gradient de tampon B (20 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7,4 ; 500 mM NaCl et 500 mM d'imidazole). Le gradient est linéaire de 0 à 20 minutes pour atteindre 30% en tampon B de façon à éliminer les protéines retenues de façon non spécifique sur la colonne. Ensuite, il est appliqué un gradient isocratique à 30% de B pendant 20 minutes. Finalement, il est appliqué un gradient linéaire de 30 à 100% de tampon B en 20 minutes. L'ensemble de la purification a lieu à débit constant à 1 ou 5 mL.minute<sup>-1</sup> respectivement pour les colonnes de 1 ou 5 mL. Le suivi de la purification se fait par spectrométrie UV avec mesure de la DO à 280 nm.

#### **VI.6.5) Gel d'électrophorèse SDS PAGE**

Les fractions concentrées en protéines sont analysées sur gel d'électrophorèse en conditions dénaturantes (SDS-PAGE)<sup>192,193</sup>. 4 à 15 µL de chaque fraction sont ajoutés à 6 µL de tampon de charge (tampon Tris-HCl 60 mM, pH 6,8 ; Glycérol 10% ; SDS 2% ; Bleu de bromophénol 44% (m/v) ; DTT 46 mM) puis chauffés à 95 °C pendant 5 à 10 minutes. Le volume de fraction ajouté avec le tampon est fonction de la concentration en protéine de cette fraction.

Chaque fraction, ainsi qu'un marqueur de poids moléculaire, est ensuite déposée sur gel SDS-PAGE. Les marqueurs utilisés au laboratoire sont le PageRuler Broad Range Unstained Protein Ladder et le Spectra Multicolor Broad Range Protein Ladder (Thermo scientific). Deux types de tampon sont utilisés pour faire des gels d'électrophorèse. Le gel de Tris-Tricine est adapté à des protéines de moins de 30 kDa<sup>194</sup> comme AnaA ou le domaine ACP d'AnaG tandis que le gel de Tris-glycine est adapté à des protéines plus volumineuses<sup>195</sup> comme AnaG ou des parties d'AnaG. Le pourcentage d'acrylamide dans le gel de résolution varie en fonction de la taille attendue de la protéine purifiée. En effet, lorsque la protéine est de petite taille, nous utilisons un gel avec un haut pourcentage d'acrylamide, de 10 à 15%, et

pour des protéines de taille plus importante, nous utilisons un gel avec un pourcentage de 8%.

La migration dure entre 30 et 60 minutes avec un ampérage constant entre 20 mA et 40 mA. Après la migration, le gel est coloré pendant 30 à 45 minutes dans le tampon de coloration (acide acétique 10%, éthanol 30% et bleu de bromophénol 0,2% [m/v]) puis décoloré dans le tampon de décoloration (acide acétique 10%) jusqu'à l'apparition des bandes.

Les fractions les plus concentrées, et contenant les protéines les plus pures, sont ensuite rassemblées.

Les manipulations de trypsinolyse sur les gels d'électrophorèses ont été réalisées à la plateforme de spectrométrie de masse protéique de l'Institut Curie.

#### **VI.6.6) Dessalage et concentration**

Une solution protéique est obtenue après réunion des fractions d'intérêt. Une fraction correspond à 1 mL. Cette solution peu concentrée et fortement chargée en sel, est dessalée sur une colonne PD10 (*GE Healthcare*) en utilisant le même tampon pour l'équilibration et l'élution. Le tampon d'élution est spécifique d'une purification à une autre car il doit être adapté à la protéine purifiée.

Après avoir été dessalée, la solution protéique est concentrée en utilisant un système sur filtre (Amicon Ultra-15 centrifugal filter unit 10k ou 30k, *Millipore*).

Les protéines sont conservées dans une solution tamponnée de  $\text{NaH}_2\text{PO}_4$  contenant du glycérol à une concentration comprise entre 10 et 30% (v/v) à  $-80\text{ }^\circ\text{C}$ .

#### **VI.6.7) Quantification des protéines**

Toutes les concentrations protéiques sont estimées par le kit Protein assay (Bio-Rad) qui est basé sur le dosage de Bradford.

#### **VI.6.8) Récapitulatif des purifications effectuées**

J'ai effectué plusieurs expressions et purifications de différentes protéines du cluster *ana*. Pour l'AnaG V954-Stop, des étapes supplémentaires ont cependant été nécessaires. Les

différentes protéines portaient toutes une étiquette Polyhistidine-tag pour permettre leur purification sur une colonne d'affinité au nickel. Toutes les conditions d'expression et de purification des protéines sont décrites dans le Tableau 6.

| Protéine                  | Position du Polyhistidine-tag sur la protéine | Conditions d'expression   | Gradient de purification  |
|---------------------------|---|---|---|
| AnaA                      | 6 His en N-ter                                | LB + Kan 50µg.ml <sup>-1</sup> +<br>Cm30 µg.ml <sup>-1</sup><br><br>37 °C jusqu'à<br>DO <sub>600nm</sub> =0,6<br><br>Induction à 200µM<br>d'IPTG<br><br>15 °C O/N<br><br>Conditions similaires<br>mais ajout d'inhibiteurs<br>de protéase dans les<br>tampons | Equilibration 10 ml<br><br>Lavage 35 ml<br><br>Gradient 1 TpB (0-30%)<br><br>Gradient 2 isocratique TpB (30%)<br><br>Gradient 3 TpB (30-100%) 20 ml<br><br>Lavage 100% TpB 5 ml |
| Domaine ACP d'AnaG        | 6 His en N-ter                                |   |   |
| Domaine ACP d'AnaG        | 6 His en N-ter et C-ter                       |   |   |
| Partie N-terminale d'AnaG | 6 His en C-ter                                |   |   |
| Domaine ACP d'AnaG        | 6 His en C-ter                                |   |   |
| AnaG entier               | 6 His en N-ter et C-ter                       |   |   |
| Partie C-terminale d'AnaG | 6 His en C-ter                                |   |   |
| AnaG synthétique          | 6 His en C-ter                                |   |   |

Tableau 6 : Conditions de purification des protéines

#### VI.6.8.1) Obtention d'AnaA :

La protéine AnaA a été obtenue en suivant le protocole décrit ci-dessus. Nous avons pu obtenir 250 µL de protéine à une concentration de 1,43 mg/mL.

#### VI.6.8.2) Obtention de la partie N-terminal d'AnaG (Domaine KS+AT) :

La partie N-terminal d'AnaG a été obtenue en suivant le protocole décrit ci-dessus. Nous avons pu obtenir 420 µL de protéine à une concentration de 1,135 mg/mL.

#### VI.6.8.3) Obtention d'AnaG :

La protéine AnaG entière n'a pas pu être obtenue en suivant le protocole décrit ci-dessus.

#### VI.6.8.4) Obtention d'AnaG synthétique :

La protéine AnaG synthétique entière n'a pas pu être obtenue en suivant le protocole décrit ci-dessus.

#### **VI.6.8.5) Obtention du domaine ACP d'AnaG :**

Le domaine ACP d'AnaG a été obtenu avec l'étiquette polyhistidine en C-terminal, en N-terminal et sur ces deux positions à la fois. La protéine avec l'étiquette polyhistidine en C-terminal était celle qui s'exprimait le mieux et qui était le mieux retenue par la colonne d'affinité. Elle a été obtenue en deux fractions de 500  $\mu$ L à une concentration respective de 0,84 et 0,193 mg/mL.

#### **VI.6.8.6) Obtention de la partie C-terminal d'AnaG :**

Pour l'obtention de la partie C-terminal d'AnaG avec le greffage du bras phosphopantéthéine, nous avons dû effectuer des manipulations supplémentaires par rapport au protocole décrit ci-dessus.

##### **VI.6.8.6.1) Chromatographie sur colonne échangeuse d'anions**

Après le passage sur colonne d'affinité, il faut faire passer l'échantillon sur une colonne échangeuse d'anions car la protéine n'est pas encore pure. Les fractions récupérées sont injectées sur une colonne échangeuse d'anions qui a été préalablement équilibrée par un tampon A (pH 7,8 ; Tris-HCl 20 mM). Après un lavage avec 40 mL du tampon A, les protéines sont éluées par un gradient de tampon B (pH 7,8 ; 1 M NaCl et Tris-HCl 20 mM). Le gradient est linéaire de 0 à 20 minutes pour atteindre 100% en tampon B. Ensuite, il est appliqué un gradient isocratique à 100% de B pendant 20 minutes. L'ensemble de la purification a lieu à débit constant à 1 ou 5 mL.minute<sup>-1</sup> respectivement pour les colonnes de 1 ou 5 mL). Nous avons pu obtenir 500  $\mu$ L de protéine à une concentration de 0,55 mg.mL<sup>-1</sup>.

##### **VI.6.8.6.2) Gel filtration**

L'expérience de gel filtration est effectuée à l'IBPC dans le laboratoire de « Biologie Physico-Chimique des Protéines Membranaires » en collaboration avec Karine Moncoq. Environ ~350  $\mu$ g de protéines sont chargés sur une colonne Superdex 200 équilibrée avec le tampon 20 mM Tris-HCl pH 7,5 ; 150 mM NaCl et 2 mM 2-ME) 0,5 mL.minute<sup>-1</sup>, et la détection se fait à 220, 260 et 280 nm.

##### **VI.6.8.6.3) Test de greffage enzymatique du bras phosphopantéthéine d'AnaG :**

Une solution de 200 $\mu$ l MgCl<sub>2</sub> 10mM, TCEP 5mM, CoASH 0,06mM, AnaG purifiée 5

$\mu\text{M}$ , PPTase 0,2 $\mu\text{M}$  dans 100mM Tris HCl buffer, pH 7,5 est préparée. Cette même solution, sans la Pptase, est préparée et gardée à 4° C pour servir de témoin à la manipulation. La première solution est incubée à 28° C, pendant deux heures et la réaction est arrêtée en ramenant la solution à 4 °C. La solution et le témoin sont analysés par une technique de MALDI-TOF. La partie C-terminal d'AnaG avec le bras phosphatéthéine est obtenue en une fraction de 200  $\mu\text{L}$  à une concentration de 0,39 mg/mL.

#### **VI.6.8.7) Western blot avec Hisprobe sur l'AnaG et sur AnaA :**

La culture d'*E. coli* est réalisée dans des conditions similaires à celles nécessaires pour l'expression de protéines jusqu'à l'étape de sonication. Le culot est alors resuspendu dans 4 ml de tampon 20 mM  $\text{NaH}_2\text{PO}_4$ , 500 mM NaCl et des inhibiteurs de protéases COMplete EDTA Free (ROCHE). Les cellules sont ensuite cassées par sonication dans la glace par une succession de cycles de 15 s de sonication à 500 W et 40 s de repos jusqu'à 2 minutes de sonication totale. Le lysat est centrifugé à 13 500 g pendant 15 minutes.

En fonction de la concentration en protéines, 1 à 10  $\mu\text{l}$  d'échantillon sont déposés sur un gel SDS-PAGE. Après la migration, les protéines sont transférées sur une membrane de nitrocellulose (Bio-Rad) à 100V constant pendant 1h dans une cuve de transfert de type Bio-Rad. La membrane est bloquée dans du tampon « Tris Buffered Saline Tween-20 0,05% » (TBST) complété avec 5% BSA, pendant 1 heure, à température ambiante ou pendant une nuit à 4° C, sous agitation. La membrane est ensuite lavée deux fois pendant 15 minutes au TBST puis incubée avec l'Hisprobe-HRP working solution (Pierce Biotechnology Inc.), pendant 1 heure à température ambiante, sous agitation. La membrane est de nouveau lavée quatre fois pendant 10 minutes avec du TBST. Les protéines sont détectées par chemiluminescence à l'aide du kit SuperSignal West HisProbe (Pierce Biotechnology Inc.) sur film auto radiographique sur chemidoc<sup>196,197</sup>.

### **VI.7) Etude de la *Cylindrospermum stagnale* PCC 7417**

#### **VI.7.1) Expérience d'incorporation isotopique par feeding**

500 mL de culture de *Cylindrospermum stagnale* PCC7 417 sont préparées jusqu'à ce que la DO à 750 nm atteigne une valeur de 0,3 à 0,4. De la [ $U$ - $^{13}\text{C}$ , 15N]-L-proline (Cambridge Isotope Laboratoire), diluée dans 40 mL de milieu BG11, est alors ajoutée au

milieu de culture. Cette culture est ensuite maintenue dans les mêmes conditions pendant 3 jours. La concentration finale de [ $U$ - $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-L-proline est de 90  $\mu\text{M}$ . Les toxines sont ensuite extraites du culot cellulaire ou du milieu extracellulaire pour être analysées en GC-MS.

#### **VI.7.2) Préparation d'extrait de milieu de culture de cyanobactéries pour analyse GC-MS**

Une solution d'extrait de milieu de culture est séchée à l'évaporateur rotatif. Le milieu doit prendre une teinte jaune trouble. L'extrait est ensuite centrifugé à 10000  $g$ , pendant 30 minutes. Le surnageant est récupéré et son pH est ajusté à une valeur supérieure à 9 par ajout d'une solution d'ammoniac.

Une colonne Sepac C-18 est équilibrée avec 2 mL de Méthanol et 6 mL d'eau milliporée avant que la solution soit ajoutée. Elle est ensuite lavée avec 6mL d'eau milliporée et 2 mL d'une solution de Méthanol 10%. Le produit est élué avec l'ajout de 6 mL de Méthanol sur la colonne. Le solvant est ensuite évaporé à l'évaporateur rotatif et le produit est transféré dans de l'acétonitrile. L'acétonitrile est ensuite évaporé à l'évaporateur rotatif et le produit est encore repris deux fois dans de l'acétonitrile. Le produit est conservé dans l'acétonitrile à 4 °C.

# Partie Expérimentale de Chimie

## VII.1) Matériels

Les spectres RMN  $^1\text{H}$  1D, 2D et  $^{13}\text{C}$  sont réalisés sur un spectromètre Bruker Avance 300 ou un Bruker Avance 400. Les déplacements chimiques sont exprimés en ppm.

Les chromatographies sur couche mince (CCM) analytiques ont été réalisées sur plaque de silice Merck TLC Silica gel 60 F<sub>254</sub>. Les produits possédant une fonction aromatique sont visualisés par UV et les autres produits par révélation avec une solution d'acide phosphomolybdique (25 g.L<sup>-1</sup> dans EtOH) ou avec une solution de ninhydrine (0,02 g.mL<sup>-1</sup> dans une solution heptane/EtOH : 6/1) suivi d'un chauffage à 120 °C.

Les purifications par chromatographie flash sur gel de silice ont été effectuées sur silice Merck Geduran® Silica Gel 60.

Les réactifs proviennent de Sigma-Aldrich et les solvants proviennent de WVR.

Différentes méthodes sont utilisées pour l'obtention des solvants secs. Après ajout d'hydrure de calcium, l'acétonitrile ou la triéthylamine ou le dichlorométhane sont laissés une nuit sous agitation. Le mélange est ensuite distillé. Le N,N-Diméthylformamide est, lui, laissé une nuit sur du tamis moléculaire 4 Å. Le toluène et le tétrahydrofurane ont été laissés sous agitation avec du sodium à 1 g.L<sup>-1</sup> et de la benzophénone à 10 g.L<sup>-1</sup>. Une fois que la couleur de la solution a viré au bleu, cette dernière est distillée. Tous ces solvants secs obtenus sont conservés sur des tamis moléculaires (Sigma-Aldrich).

Les expériences d'HPLC menées au laboratoire utilisent un appareil HPLC Agilent 1220 Infinity LC Systems. Les colonnes utilisées sont, soit une colonne Cosmosil 5µm C18 120 Å 250×4,6 mm, soit une colonne Luna 5µm C18 100 Å 250×4,6 mm.

Les expériences d'HPLC menées au laboratoire des Biomolécules à l'ENS utilisent un appareil HPLC semi-préparative et un appareil HPLC analytique. Les colonnes utilisées sont, soit une colonne Cosmosil 5µm C18 120 Å 250×4,6 mm pour l'analytique, soit une colonne µBondapal 10 µm C18 125 Å 300×7,8 mm pour la semi-préparative.

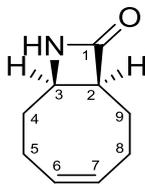
Les expériences de LC-MS menées au Laboratoire Sciences Analytiques à L'ESPCI utilisent un appareil chromatographique UHPLC Agilent 1200 series couplé à un spectromètre de masse Agilent 6410. La colonne utilisée est une colonne Agilent Poroshell 120 2,7  $\mu\text{m}$  EC-C18 50 $\times$ 2,1 mm.

Les expériences de GC-MS menées au Laboratoire des Mécanismes Réactionnels à l'école Polytechnique utilisent un appareil GC-MS de marque Varian GC-450 couplé à une trappe ionique Varian 240-MS. La colonne utilisée est une colonne Varian VF-X MS 0,25  $\mu\text{m}$  60000 $\times$ 0,25 mm.

Les solvants proviennent de Carlo Erba Reagents et de WVR. Les réactifs proviennent de Sigma-Aldrich et de Carlo Erba Reagents. Les solvants sont filtrés sur des membranes de 0,22  $\mu\text{m}$ . L'eau utilisée est milliporée sur un appareil Millipore Milli-Q® Direct 8.

## VII.2) Synthèse d'analogue de substrat d'AnaG

### VII.2.1) *rac*-(2*S*,3*R*)-3-Azabicyclo[6.2.0]dec-7-ene-2-one (1)



1

Formule brute: C<sub>9</sub>H<sub>13</sub>NO

24 mL de cycloocta-1,5-diène (195,1 mmol, 1 éq), 3,1 g de Na<sub>2</sub>CO<sub>3</sub> (29,2 mmol, 0,15 éq) et 10 mL de CH<sub>2</sub>Cl<sub>2</sub> sont introduits dans un ballon sec sous argon, puis le ballon est refroidi dans un bain de glace. 17 mL de chlorosulfonyl isocyanate (195,1 mmol, 1 éq) sont ajoutés au mélange goutte à goutte sous agitation pendant une heure. Le mélange est laissé à 0 °C pendant deux heures, puis laissé à température ambiante pendant une nuit. La réaction étant exothermique avec un risque d'explosion, la température de la solution doit être maintenue à TA. Le mélange est ensuite versé sur une solution de 340 mL de CHCl<sub>3</sub>, 408 mL d'H<sub>2</sub>O, 71,4g de Na<sub>2</sub>SO<sub>4</sub> et 79,56g de Na<sub>2</sub>HPO<sub>4</sub>, sous agitation. Ce mélange est transféré dans une ampoule à décanter et la phase organique est récupérée. La phase aqueuse est lavée deux fois avec 100 mL de CH<sub>2</sub>Cl<sub>2</sub>. Les phases organiques sont ensuite rassemblées et séchées avec du MgSO<sub>4</sub>. Après évaporation du CH<sub>2</sub>Cl<sub>2</sub>, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 100:0 à 70 :30). 5,66 g d'une poudre blanche sont obtenus.

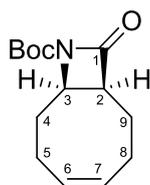
Rendement(Rdt)= 20%

Rapport frontal (Rf) = 0,46 dans cyclohexane :acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique (Rf = 0,31 dans acétate d'éthyle dans les données de la littérature<sup>140</sup>)

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.83-2.05 (m, 6H, CH<sub>2</sub>), 2.3 (m, 2H, CH<sub>2</sub>), 3.27-3.32 (m, 1H, CHCO), 3.8 (m, 1H, CHN), 5.65 (m, 2H, CH=CH), 6.18 (m, 1H, NH)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 23.0 (CH<sub>2</sub>), 25.9 (2CH<sub>2</sub>), 30.6 (CH<sub>2</sub>), 53.4 et 54.2 (CH), 128.49 et 131.2 (C=C), 171.1 (CO) données cohérentes avec celles de la littérature<sup>140</sup>

## VII.2.2) *rac*-(2*S*,3*R*)-3-(tert-Butoxycarbonyl)-3-azabicyclo[6.2.0]dec-7-ene-2-one (2)



**2**

Formule brute : C<sub>14</sub>H<sub>21</sub>NO<sub>3</sub>

0,67 g d'4-Diméthylaminopyridine (5,5 mmol, 0,1 éq) et 8,3 g de (1) (55 mmol, 1 éq) sont introduits dans un ballon sec sous argon, puis 79 mL d'acétonitrile sec sont ajoutés dans le ballon sous agitation. 24 g de Boc<sub>2</sub>O (110 mmol, 2 éq) sont introduits dans un autre ballon sec sous argon, puis 40 ml d'acétonitrile sec sont ajoutés sur le Boc<sub>2</sub>O sous agitation. Le ballon contenant le premier mélange est refroidi dans un bain de glace, puis la solution de Boc<sub>2</sub>O est ajoutée goutte à goutte. Le mélange est laissé une nuit à TA, puis 110 mL de dichlorométhane sont ajoutés. Le mélange est transféré dans une ampoule à décanter puis 30 mL d'une solution de NaHCO<sub>3</sub> saturé sont ajoutés. La phase organique est lavée avec 30 mL d'une solution de NaCl saturée, puis elle est récupérée et séchée avec du MgSO<sub>4</sub>. Après évaporation du solvant, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 100:0 à 50 :50). 10,5 g d'une poudre blanche sont obtenus.

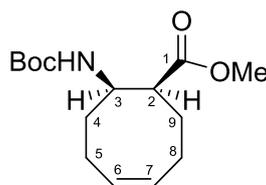
Rdt= 78%

Rf=0,6 dans cyclohexane :acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique (Rf = 0,4 dans éther :éther de pétrole/2 :3 dans les données de la littérature<sup>132</sup>)

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.50 (s, 9H, CH<sub>3</sub>), 1.85 (m, 1H, CH<sub>2</sub>), 2.09 (m, 4H, CH<sub>2</sub>), 2.50 (m, 3H, CH<sub>2</sub>), 3.35 (m, 1H, CHCO), 4.13 (m, 1H, CHN), 5.70 (m, 2H, C=C)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 24,4(CH<sub>2</sub>) , 25.5(CH<sub>2</sub>), 26.1(CH<sub>2</sub>), 29.8 (CH<sub>2</sub>), 29.9 (CH<sub>3</sub>), 54.7 et 58.2 (CH), 83.7 ( C(CH<sub>3</sub>)), 132.4 et 132.6 (CH=CH), 150.0 (CO Boc), 170.1 (CO), données cohérentes avec celles de la littérature<sup>132</sup>

### VII.2.3) *rac*-(2*S*,3*R*)-3-(( *tert*-Butoxycarbonyl)amino)cyclooct-7-ene-2-oate de méthyle (3)



**3**

Formule brute : C<sub>15</sub>H<sub>25</sub>NO<sub>4</sub>

2 g de (2) (8 mmol, 1 éq) et 0,52 g de NaN<sub>3</sub> (8 mmol, 1 éq) sont ajoutés dans un ballon sec sous argon sous agitation, puis 0,6 mL de méthanol (16 mmol, 2 éq) et 24 mL de N,N-Diméthylformamide sont ajoutés dans le ballon. Le mélange est chauffé à 64 °C, au reflux, pendant une nuit, puis le mélange est dilué par l'ajout de 30 mL de NaCl saturé et lavé deux fois avec 30 ml d'éther. Après évaporation du solvant, le résidu organique est purifié sur colonne de silice (Cyclohexane : Acétate d'éthyle/ 100:0 à 70 :30). 1,68 g d'une huile jaune sont obtenus.

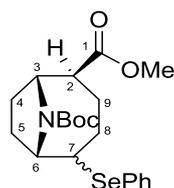
Rdt= 85%

Rf= 0,45 dans cyclohexane : acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.39 (s, 9H, CH<sub>3</sub>) , 1.74 (m, 1H, CH<sub>2</sub>), 2.08 (m, 2H, CH<sub>2</sub>) , 2.10-2.25 (m, 3H, CH<sub>2</sub>), 2.4 (m, 1H, CH<sub>2</sub>) , 2.82 (m, 1H, CH<sub>2</sub>), 3.7 (s, 3H, OCH<sub>3</sub>) , 4.09 (m, 1H, CH), 5.03 (m, 1H, CH), 5.57-5.63 (m, 2H, CH=CH)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 23.6(CH<sub>2</sub>) , 24.3(CH<sub>2</sub>), 26.8(CH<sub>2</sub>), 27.1 (CH<sub>2</sub>), 28.3 (CH<sub>3</sub>Boc) , 32.8 (CH<sub>2</sub>), 47.8 et 50.1 (CH) , 51.4 (OCH<sub>3</sub>), 79.1 ( C(CH<sub>3</sub>)) ,129 et 130 (C=C), 154.9 (CO Boc), 174.5 (CO), données cohérentes avec celles de la littérature<sup>139</sup>

**VII.2.4)rac-(2*S*,3*S*,6*R*)-10-(tert-butoxycarbonyl)-6-(phenylselanyl)-3-azabicyclo[4.2.1]nonane-2-oate de methyle (4)**



**4**

Formule brute : C<sub>21</sub>H<sub>29</sub>NO<sub>4</sub>Se

1,6 g de **(3)** (5,9 mmol, 1 éq), 0,62 g de Na<sub>2</sub>CO<sub>3</sub> (5,9 mmol, 1 éq) et 53 mL d'acétonitrile sec sont ajoutés dans un ballon sec sous argon, puis le ballon est refroidi dans un bain de glace sous agitation. 1,1 g de PhSeCl (5,9 mmol, 1 éq) est solubilisé dans 14 mL d'acétonitrile sec, puis cette solution est ajoutée dans le ballon goutte à goutte sur une heure. Le mélange est ensuite laissé sous agitation, à 0 °C, pendant une nuit. Après évaporation de l'acétonitrile, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 90:10 à 80 :20). 2,24 g d'une huile jaune claire sont obtenus.

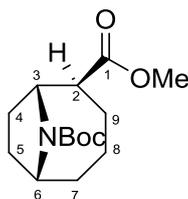
Rdt= 86%

Rf= 0,26 dans cyclohexane : acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique et au UV

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.34 (s, 9H, CH<sub>3</sub>) , 1.64 (m, 1H, SeCH), 1.90-1.84 (m, 4H, CH<sub>2</sub> cycle) , 2.04 (m, 2H, CH<sub>2</sub> cycle), 2.26 (m, 2H, CH<sub>2</sub> cycle) , 3.47 (m, 1H, HCCO cycle), 3.60 (s, 3H, OCH<sub>3</sub>) , 4.46 (m, 1H, NC<sub>6</sub>H cycle), 4.64-4.48 (d, 1H, NC<sub>3</sub>H cycle) , 7.16 (s, 3H, CH aromatique), 7.49 (s, 2H, CH aromatique)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 23.1(CH<sub>2</sub>cycle), 24.1(CH<sub>2</sub>cycle), 28.1 (CH<sub>3</sub>Boc), 29.1 (CH<sub>2</sub>cycle), 35.0 (CH<sub>2</sub> cycle), 46.7 (CH cycle), 51.4 (OCH<sub>3</sub>), 52.7 (CH cycle), 56.6 (CH cycle), 59.3 (CH cycle), 79.1 (C(CH<sub>3</sub>)), 127.3 et 126.9 (C aromatique), 128.9 (CH aromatique), 134 et 133.9 (CH aromatique), 154.9 (CO Boc), 174.5 (CO), composé obtenu dans la littérature sans données de caractérisation <sup>139</sup>

VII.2.5) *rac*-(2*S*,3*S*,6*S*)-10-(*tert*-butoxycarbonyl)-3-azabicyclo[4.2.1]nonane-2-oate de methyle (5)



5

Formule brute : C<sub>15</sub>H<sub>25</sub>NO<sub>4</sub>

24 mg d'Azobisisobutyronitrile (0,14 mmol, 0,05 éq) sont ajoutés dans un ballon bicol sec sous argon. 1,28 g de (4) sont solubilisés dans 16,5 ml de toluène sec puis ils sont ajoutés à l'AIBN. Le mélange est chauffé à 120 °C au reflux sous agitation. 1,5 mL de Bu<sub>3</sub>SnH (5,8 mmol, 2 éq) sont ajoutés au mélange et la réaction est suivie par CCM. Une fois la réaction terminée et après évaporation du toluène, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 90:10 à 80 :20). 0,75 g d'une huile jaune claire est obtenu.

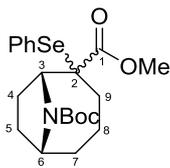
Rdt= 89%

Rf= 0,38 dans cyclohexane:acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique (Rf = 0,36 dans hexane:acétate d'éthyle/3 :1 dans les données de la littérature<sup>198</sup>)

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.43 (s, 9H, CH<sub>3</sub>), 1.67-1.95 (m, 10H, CH<sub>2</sub>cycle), 2.38 (m, 1H, CHCO), 3.62 (s, 3H, OCH<sub>3</sub>), 4.15- 4.30 (m, 1H, NC<sub>6</sub>H cycle), 4.44-4.63 (m, 1H, NC<sub>3</sub>H cycle)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 21.7(CH<sub>2</sub>cycle), 26.3(CH<sub>2</sub>cycle), 28.1 (CH<sub>3</sub>Boc) , 32.5 (CH<sub>2</sub>cycle), 35.3 (CH<sub>2</sub> cycle) , 48.7 (CH cycle), 52.1 (OCH<sub>3</sub>), 53.1 (CH cycle) , 56.6 (CH cycle) , 79.2 ( C(CH<sub>3</sub>)), 153.1 (CO Boc), 174.4 (CO), données cohérentes avec celles de la littérature<sup>198</sup>

**VII.2.6)rac-(3*S*,6*S*)-10-(tert-butoxycarbonyl)-2-(phenylselanyl)-3-azabicyclo[4.2.1]nonane-2-oate de methyle (6)**



**6**

Formule brute : C<sub>21</sub>H<sub>29</sub>NO<sub>4</sub>Se

3,6 mL de tétrahydrofurane sec sont ajoutés dans un ballon sec sous argon, puis le ballon est refroidi dans un bain d'azote liquide maintenu à -78 °C. 1,3ml de LDA (2,7 mmol, 1,4 éq) est ajouté dans le ballon. 0,55 g de (5) (1,96 mmol, 1 éq) est solubilisé dans 6,2 mL de THF sec, puis cette solution ajoutée goutte à goutte au mélange sur 15 minutes. 0,57g de PhSeBr (2,45 mmol, 1,25 éq) est solubilisé dans 2,6 mL de THF. 20 minutes après le précédent ajout, le PhSeBr est ajouté dans le ballon goutte à goutte pendant 20 minutes. Le mélange est laissé sous agitation avec suivi CCM. La réaction est arrêtée par ajout de 30 mL de KH<sub>2</sub>PO<sub>4</sub> 1M puis la phase organique est extraite avec deux fois 30 mL d'acétate éthyle/cyclohexane. La phase organique est séchée avec MgSO<sub>4</sub>. Après évaporation du THF, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 90:10 à 85 :15). 0,49 g d'une huile jaune claire est obtenu.

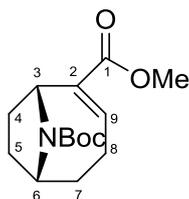
Rdt= 74%

Rf= 0,35 dans cyclohexane : acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique et au UV

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.39 (s, 9H, CH<sub>3</sub>) , 2.27-1.83 (m, 8H, CH<sub>2</sub>cycle), 3.37 (m, 2H, CH<sub>2</sub>COSe) , 3.57 (s, 3H, OCH<sub>3</sub>), 4.46-4.28 (m, 1H, NC<sub>6</sub>H cycle) , 4.95-4.68 (m, 1H, NC<sub>3</sub>H), 7.17 (m, 3H, CH aromatique), 7.45 (m, 2H, CH aromatique)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 21.7(CH<sub>2</sub>cycle), 28.1 (CH<sub>3</sub>Boc), 31.3 (CH<sub>2</sub>cycle), 32.5 (CH<sub>2</sub>cycle), 35.3 (CH<sub>2</sub> cycle), 51.3 (OCH<sub>3</sub>), 56.9, 56.1 et 53.0 (CH cycle), 79.2 (C(CH<sub>3</sub>)) ,127.1 (C aromatique), 137 et 128.4 (CH aromatique), 153.1 (CO Boc), 174.4 (CO), composé obtenu dans la littérature sans données de caractérisation <sup>139</sup>

**VII.2.7) (3*S*,6*S*)-10-(Butoxycarbonyl-tert) - tert-3-azabicyclo[4.2.1]non-9-ene-2-oate de methyle (7)**



**7**

Formule brute : C<sub>15</sub>H<sub>23</sub>NO<sub>4</sub>

0,5 g de (6) (1,13 mmol, 1 éq) et 5 mL de THF sec sont ajoutés dans un ballon sec sous argon, puis le ballon est refroidi dans un bain de glace. 0,6 g d'acide métachloroperbenzoïque (3,4 mmol, 2,5 éq) et 5 mL de THF sec sont ajoutés dans un ballon sec sous argon, puis cette solution est ajoutée au premier ballon goutte à goutte sur 20 minutes. Le mélange est laissé sous agitation, à -10 °C, pendant 2 heures, puis 0,7 mL de triéthylamine (4,55 mmol, 4 éq) y est ajouté. Le mélange est ensuite laissé sous agitation, à -10 °C pendant une heure, puis il est ramené à TA, sous agitation, pendant une heure. Après évaporation du THF, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 90:10 à 80 :20). 0,3 g d'une huile jaune claire est obtenu.

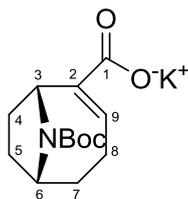
Rdt= 90%

Rf= 0,27 dans Cyclohexane : Acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique (Rf = 0,31 dans hexane:acétate d'éthyle/3 :1 dans les données de la littérature<sup>198</sup>)

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.41 (s, 9H, CH<sub>3</sub>), 1.50-1.81 (m, 4H, CH<sub>2</sub> cycle), 1.96-2.45 (m, 4H, CH<sub>2</sub> cycle), 3.76 (s, 3H, OCH<sub>3</sub>), 4.46 (m, 1H, NC<sub>6</sub>H cycle), 5.06 (m, 1H, NC<sub>3</sub>H), 6.9 (t, 1H, C=CH)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 24.4 (CH<sub>2</sub>cycle), 28.3 (CH<sub>3</sub>Boc), 29.1 (CH<sub>2</sub>cycle), 30.5 (CH<sub>2</sub>cycle), 31.3 (CH<sub>2</sub> cycle), 51.2 (OCH<sub>3</sub>), 54.0 (CH cycle), 55.7 (CH cycle), 79.2 (C(CH<sub>3</sub>)), 139.3-141.5 (C=C), 153.1 (CO Boc), 168.4 (CO), données cohérentes avec celles de la littérature<sup>198</sup>

## VII.2.8)(3*S*,6*S*)-Acide 10-(tert-butoxycarbonyl)-9-azabicyclo[4.2.1]non-9-ene –2-oïque(8)



**8**

Formule brute : C<sub>14</sub>H<sub>21</sub>NO<sub>4</sub>

0,4 g de (7) (1,4 mmol, 1 éq), 10 mL de EtOH 96% et 2,5 mL de KOH 5M sont ajoutés dans un ballon sec. Le mélange est chauffé à 100 °C au reflux pendant 3 heures, puis est acidifié par ajout de HCl 5M jusqu'à atteindre un pH de 2-3. Le mélange est ensuite dilué avec 10 ml d'H<sub>2</sub>O. Les phases organiques sont extraites du mélange avec deux fois 10 mL de CH<sub>2</sub>Cl<sub>2</sub>, puis elles sont séchées avec du MgSO<sub>4</sub>. Après évaporation du solvant, le résidu organique est purifié sur colonne de silice (Cyclohexane : Acétate d'éthyle/ 50:50 à 10 :90 puis Dichlorométhane :EtOH/ 90 :10). 0,23 g d'une huile jaune claire est obtenu.

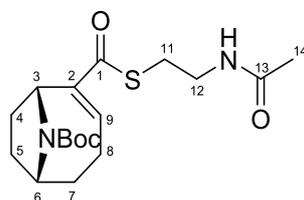
Rdt= 80%

Rf= 0,37 dans cyclohexane : acétate d'éthyle/4 :6 révélation à l'acide phosphomolybdique(Rf = 0,4 dans Chloroforme:Méthanol/9 :1 dans les données de la littérature<sup>198</sup>)

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.44 (s, 9H, CH<sub>3</sub>), 1.58-1.74 (m, 3H, CH<sub>2</sub>cycle), 2.03-2.34 (m, 5H, CH<sub>2</sub>cycle), 4.32-4.41 (m, 1H, NC<sub>6</sub>H cycle), 4.85-5.02 (m, 1H, NC<sub>3</sub>H), 7.01-7.05 (t, 1H, C=CH)

<sup>13</sup>C RMN (75 MHz, CDCl<sub>3</sub>) δ ppm 24.2 (CH<sub>2</sub>cycle), 28.2 (CH<sub>3</sub>Boc), 28.7 (CH<sub>2</sub>cycle), 29.6 (CH<sub>2</sub> cycle), 31.6 (CH<sub>2</sub> cycle), 53.3 (CH cycle), 54.9 (CH cycle), 79.2 (C(CH<sub>3</sub>)), 139.2-142.7 (C=C), 154.1 (CO Boc), 171.9(CO), données cohérentes avec celles de la littérature <sup>198</sup>

**VII.2.9)(3*S*,6*S*)-10-(tert-butoxycarbonyl)-9-azabicyclo[4.2.1]thionon-9-ene-2-oate de N-ethyl-éthanamide (9)**



**9**

Formule brute : C<sub>18</sub>H<sub>28</sub>N<sub>2</sub>O<sub>4</sub>

80 mg de (8) (0,3 mmol, 1 éq) et 2 mL de CH<sub>2</sub>Cl<sub>2</sub> sec sont ajoutés dans un ballon sec sous argon, puis le ballon est refroidi dans un bain de glace. 40 mg de N-acetylcysteamine (0,36 mmol, 1,21 éq) dans 1 mL de CH<sub>2</sub>Cl<sub>2</sub> sec sont ajoutés dans le ballon. Puis 10 mg de DMAP (0,09 mmol, 0,25 éq) et 70 mg de EDCI (0,36 mmol, 1,21 éq) dans 1 mL de CH<sub>2</sub>Cl<sub>2</sub> sec sont ajoutés dans le ballon. Le mélange est laissé sous agitation, à 0 °C, pendant deux heures, puis est ramené à TA et laissé sous agitation une nuit. Après évaporation du CH<sub>2</sub>Cl<sub>2</sub>, le résidu organique est purifié sur colonne de silice (cyclohexane : acétate d'éthyle/ 80:20 à 40 :60). 10 mg d'une huile vert sombre est obtenu.

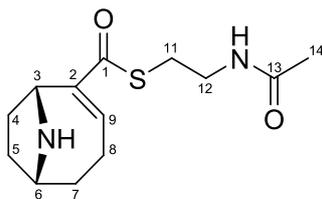
Rdt= 40%

Rf= 0,27 dans cyclohexane : acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique et à la ninhydrine

<sup>1</sup>H RMN (300 MHz, CDCl<sub>3</sub>) δ ppm 1.35 (s, 9H, CH<sub>3</sub>), 1.44- 1.75 (m, 6H, CH<sub>2</sub>cycle) , 1.95 (s, 3H, COCH<sub>3</sub>), 2.0-2.4 (m, 2H, CH<sub>2</sub>cycle) 2.99 (m, 2H, CH<sub>2</sub>S), 3.41 (m, 2H, CH<sub>2</sub>NH), 4.9 (m, 1H, NC<sub>3</sub>H cycle), 5.65 (m, 1H, NC<sub>6</sub>H), 7.0 (t, 1H, C=CH), 7.9 (s, 1H, NH)

<sup>13</sup>C RMN (100 MHz, CDCl<sub>3</sub>) δ ppm 23.7 et 24.2 (CH<sub>2</sub>cycle), 28.2 (CH<sub>3</sub>Boc), 29.7 (CH<sub>3</sub>CO), 29.6 (CH<sub>2</sub> cycle), 34.6 (CH<sub>2</sub> cycle), 58.7 (NCH cycle), 60.7 (NCH cycle), 67.9 (SCH<sub>2</sub>), 79.6 (C(CH<sub>3</sub>)), 119.1-130.0 (C=C), 153.7 (CO Boc), 170.4(CO), 201.4 (SCO). Les données spectroscopiques ont été obtenues à partir d'un mélange contenant des impuretés et elles sont donc soumises à caution.

**VII.2.10) (3*S*,6*S*)-9-azabicyclo[4.2.1]thionon-9-ene-2-oate de N-ethyl-éthanamide (10)**



**10**

Formule brute : C<sub>13</sub>H<sub>20</sub>N<sub>2</sub>O<sub>2</sub>S

60 mg de (9) (0,16 mmol, 1 éq), 0,75 mL de CH<sub>2</sub>Cl<sub>2</sub> sec et 0,25 mL de TFA (3,26 mmol, 20 éq) sont ajoutés dans un ballon sec sous argon. Le mélange est laissé sous agitation, à TA, pendant 15 minutes. Après évaporation du CH<sub>2</sub>Cl<sub>2</sub>, 1 mL de CH<sub>2</sub>Cl<sub>2</sub> est ajouté et le CH<sub>2</sub>Cl<sub>2</sub> est à nouveau évaporé. Après avoir répété cette étape deux fois, 60 mg d'une huile marron sombre est obtenu.

Rdt= quantitatif

Rf= 0,27 dans cyclohexane : acétate d'éthyle/8 :2 révélation à l'acide phosphomolybdique et à la ninhydrine

Masse théorique: 268,38

Masse expérimentale d'un composé obtenu pouvant être le produit recherché (M-H): 269,39

### VII.2.11) Analyse de (10) en LC-MS

La phase mobile est constituée d'un mélange hydro-organique. Le gradient utilisé pour la séparation est décrit dans le Tableau 7. Le débit est de 1 mL.minute<sup>-1</sup> et le volume d'injection est de 8 µL. La détection se fait par spectrométrie de masse avec un mode d'ionisation par source d'ionisation à électrospray (ESI). Le produit injecté est le résultat de la dernière étape de la synthèse chimique de l'analogue de substrat après passage sur colonne de silice. Le composé recherché a une *m/z* aux alentours de 269,39.

| Temps        | Tampon A : Eau + 0,1% de TFA | Tampon B : Méthanol |
|--------------|------------------------------|---------------------|
| 0-12 minutes | 80% (linéaire)               | 20% (linéaire)      |

Tableau 7 : Gradient utilisé lors de l'analyse de l'analogue de substrat

Les pics à 2,08 minutes et à 2,72 minutes peuvent correspondre.

Ces analyses de LC-MS sont mises au point et effectuées à l'ESPCI ParisTech dans le Laboratoire Sciences Analytiques, Bioanalytiques et Miniaturisation.

### VII.2.12) Purification de (10) par HPLC analytique

La phase mobile est constituée d'un mélange hydro-organique. Le gradient utilisé pour la séparation est décrit dans le Tableau 8. Le débit est de 1 mL.minute<sup>-1</sup> et le volume d'injection est de 40 µL. La détection se fait à une longueur d'onde de 230 nm et à une longueur d'onde de 280 nm. Le produit injecté est le produit de la dernière étape de la synthèse chimique de l'analogue de substrat après passage sur colonne de silice.

| Temps        | Tampon A : Eau + 0,1% de TFA | Tampon B : Acétonitrile |
|--------------|------------------------------|-------------------------|
| 0-30 minutes | 95% à 70%                    | 5% à 30%                |

Tableau 8 : Gradient utilisé lors de la purification de l'analogue en analytique

Les pics majoritaires sont collectés dans des tubes ependorfs.

Ces purifications par HPLC sont mises au point et effectuées à l'ENS dans le laboratoire des Biomolécules.

### VII.2.13) Purification de (10) par HPLC semi-preparative

La phase mobile est constituée d'un mélange hydro-organique. Le gradient utilisé pour la séparation est décrit dans le Tableau 9. Le débit est de 2 mL.minute<sup>-1</sup> et le volume d'injection est de 600 µL. La détection se fait à une longueur d'onde de 230 nm et une longueur d'onde de 280 nm. Le produit injecté est le produit de la dernière étape de la synthèse chimique de l'analogue de substrat après passage sur colonne de silice.

| Temps        | Tampon A : Eau + 0,1% de TFA | Tampon B : Acétonitrile |
|--------------|------------------------------|-------------------------|
| 0-35 minutes | 95% à 80%                    | 5% à 20%                |

Tableau 9 : Gradient utilisé lors de la purification de l'analogue en semi-préparative

Les deux pics majoritaires sont collectés dans des tubes ependorfs. Plusieurs runs successifs sont effectués pour purifier tout le produit et les deux fractions collectées sont ensuite évaporées à l'évaporateur rotatif.

Ces purifications par HPLC sont mises au point et effectuées à l'ENS dans le laboratoire des Biomolécules.

### VII.2.14) Analyse de (10) purifié par HPLC en LC-MS

La phase mobile est constituée d'un mélange hydro-organique. Le gradient utilisé pour la séparation est décrit dans le Tableau 10. Le débit est de 1 ml.minute<sup>-1</sup> et le volume d'injection est de 8 µL. La détection se fait par ESI. Les produits injectés sont les fractions provenant de la purification sur HPLC semi-préparative du produit de la dernière étape de la synthèse chimique de l'analogue de substrat après passage sur colonne de silice. Le composé recherché a une  $m/z$  aux alentours de 269,39.

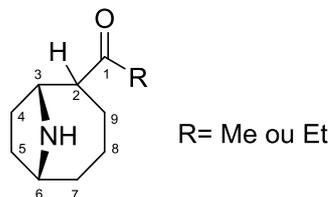
| Temps        | Tampon A : Eau + 0,1% de TFA | Tampon B : Méthanol |
|--------------|------------------------------|---------------------|
| 0-10 minutes | 80% (linéaire)               | 20%(linéaire)       |

Tableau 10 : Gradient utilisé lors de l'analyse de l'analogue de substrat purifié par HPLC

Ces analyses de LC-MS sont mises au point et effectuées à l'ESPCI ParisTech dans le Laboratoire Sciences Analytiques, Bioanalytiques et Miniaturisation.

## VII.3) Etude de la biosynthèse de la dihydroanatoxine-a

### VII.3.1) Déprotection des analogues dihydro d'anatoxine-a et d'homoanatoxine-a



Formule brute :  $C_{10}H_{17}NO$  ou  $C_{11}H_{19}NO$

10 mg de de la Boc *cis*- ou *trans*-dihydroanatoxine-a ou 5 mg de la Boc *cis*- ou *trans*-dihydrohomoanatoxine-a, 0,75 mL de  $CH_2Cl_2$  sec et 0,25 mL de TFA sont ajoutés dans un ballon sec. Le mélange est laissé sous agitation à TA, pendant 15 minutes. Après évaporation du  $CH_2Cl_2$ , 1 mL de  $CH_2Cl_2$  est ajouté et le  $CH_2Cl_2$  est à nouveau évaporé. Après avoir répété cette étape deux fois, 10 mg d'une huile marron sombre est obtenu.

### VII.3.2) Analyse des standards par GC-MS

La phase mobile est constituée d'hélium gazeux. Le gradient utilisé pour la séparation est décrit dans le Tableau 11. Le débit est de 1,4 mL.minute<sup>-1</sup> et le volume d'injection est de 10 µL. La détection se fait par spectrométrie de masse avec un mode d'ionisation chimique (CI) avec du méthanol ou un mode d'ionisation par impact électronique (EI) à 70 eV avec un courant d'ionisation de 20 µA. Les produits injectés sont des standards purifiés et déprotégés de *cis, trans*-dihydroanatoxine-a et de *cis, trans*-dihydrohomoanatoxine-a. Le composé recherché a une *m/z* aux alentours de 168 pour la dihydroanatoxine-a et de 182 pour la dihydrohomoanatoxine-a.

| Temps        | Température (°C) |
|--------------|------------------|
| 0-45 minutes | 280 (linéaire)   |

Tableau 11 : Condition de température lors de l'analyse des standards

Ces analyses de GC-MS sont mises au point et effectuées à l'école Polytechnique dans le Laboratoire des Mécanismes Réactionnels.

### VII.3.3) Purification de la *cis*-dihydroanatoxine-a

La phase mobile est constituée d'un mélange hydro-organique. Le gradient utilisé pour la séparation est décrit dans le Tableau 12. Le débit est de 1 mL.minute<sup>-1</sup> et le volume d'injection est de 20 µL. La détection se fait à une longueur d'onde de 230 nm. Le produit injecté est le produit de la déprotection de la *cis* boc-dihydroanatoxine-a.

| Temps         | Tampon A : Eau + 0,1% d'acide formique | Tampon B : Méthanol |
|---------------|--|---------------------|
| 0-20 minutes  | 100% à 35%                             | 0% à 65%            |
| 20-21 minutes | 35% à 0%                               | 65% à 100 %         |
| 21-26 minutes | 0% (linéaire)                          | 100%(linéaire)      |
| 26-27 minutes | 0% à 100%                              | 100% à 0%           |
| 27-30 minutes | 100% (linéaire)                        | 0%(linéaire)        |

Tableau 12 : Gradient utilisé lors de la purification du *cis* dihydroanatoxine-a

Le pic à 8,4 minutes est collecté dans un tube ependorf et stocké à 4 °C.

### VII.3.4) Analyse de la *cis*-dihydroanatoxine-a

La phase mobile est constituée d'hélium gazeux. Le gradient utilisé pour la séparation est décrit dans le Tableau 13. Le débit est de 1,4 mL.minute<sup>-1</sup> et le volume d'injection est de

10  $\mu\text{L}$ . La détection se fait par spectrométrie de masse par CI avec du méthanol ou par EI à 70 eV, avec un courant d'ionisation de 20  $\mu\text{A}$ . Les produits injectés sont des standards purifiés et déprotégés de *cis*-dihydroanatoxine-a purifié par HPLC. Le composé recherché a une  $m/z$  aux alentours de 168.

| Temps        | Température ( $^{\circ}\text{C}$ ) |
|--------------|------------------------------------|
| 0-45 minutes | 280 (linéaire)                     |

Tableau 13 : Condition de température lors de l'analyse de la *cis*-dihydroanatoxine-a

Ces analyses de GC-MS sont mises au point et effectuées à l'école Polytechnique dans le Laboratoire des Mécanismes Réactionnels.

### VII.3.5) Analyse d'extrait de *Cylindrospermum stagnale* PCC 7417

La phase mobile est constituée d'hélium gazeux. Le gradient utilisé pour la séparation est décrit dans le Tableau 14. Le débit est de 1,4  $\text{mL}\cdot\text{minute}^{-1}$  et le volume d'injection est de 10  $\mu\text{L}$ . La détection se fait par spectrométrie de masse par CI avec du méthanol ou par EI à 70 eV, avec un courant d'ionisation de 20  $\mu\text{A}$ . Les produits injectés sont des extraits de culture de *C. stagnale* PCC7417 avec et sans incorporation isotopique de carbone 13 et d'azote 15. Le composé recherché a une  $m/z$  aux alentours de 168 pour la dihydroanatoxine-a et de 174 pour la dihydrohomoanatoxine-a ayant incorporé le carbone 13 et l'azote 15

| Temps        | Température ( $^{\circ}\text{C}$ ) |
|--------------|------------------------------------|
| 0-45 minutes | 280 (linéaire)                     |

Tableau 14 : Condition de température lors de l'analyse d'extrait de *Cylindrospermum stagnale* PCC 7417

Ces analyses de GC-MS sont mises au point et effectuées à l'école Polytechnique dans le Laboratoire des Mécanismes Réactionnels.

## Annexes

### Annotation des gènes en 5' et 3' du cluster *ana*

| Désignation du gène   | Position sur le contig 69 d' <i>Oscillatoria</i> sp. PCC 6407 | Fonction putative |
|-----------------------|---|-------------------|
| <i>ALVI_v1_690001</i> | 6593267 - 6593452   | Fonction inconnue |
| <i>ALVI_v1_690002</i> | 6593488 - 6593979   | Transposase       |
| <i>ALVI_v1_690003</i> | 6594075 - 6594983   | Fonction inconnue |
| <i>ALVI_v1_690004</i> | 6595159 - 6595803   | Transposase       |
| <i>ALVI_v1_690005</i> | 6595844 - 6596311   | Transposase       |
| <i>ALVI_v1_690006</i> | 6596577 - 6597335   | Fonction inconnue |
| <i>ALVI_v1_690009</i> | 6600104 - 6600397   | Fonction inconnue |
| <i>ALVI_v1_690010</i> | 6600568 - 6600705   | Fonction inconnue |
| <i>ALVI_v1_690011</i> | 6600849 - 6601061   | Fonction inconnue |
| <i>ALVI_v1_690013</i> | 6601648 - 6601764   | Fonction inconnue |
| <i>ALVI_v1_690022</i> | 6625096 - 6626094   | Transposase       |

Tableau 15 : Annotation des gènes en 5' et en 3' du cluster *ana*, les gènes du cluster *ana* n'y figurent pas ainsi que les fragments de gènes

### Annotation des gènes chez la souche *Anabaena* sp. strain 37

| Désignation du gène | Position sur la séquence nucléotidique | Fonction putative  |
|---------------------|--|--------------------|
| <i>orf00001</i>     | 2181 – 2840                            | Acétyltransférase  |
| <i>orf00002</i>     | 3032 - 4255                            | Protéine de Rieske |
| <i>orf00003</i>     | 4512 - 5072                            | Oxydoréductase     |

Tableau 16 : Annotation des gènes de la souche *Anabaena* sp. strain 37 compris entre le nucléotide 20587 et le nucléotide 27525

## Liste des Figures

Figure 1. *Lyngbya* sp. (gauche) et *Anabaena azollae* (droite) observées au microscope au grossissement  $\times 400$ . <http://www-cyanosite.bio.purdue.edu/images/images.html> ..... 9

Figure 2. PKS responsable de la biosynthèse de l'érythromycine, LDD correspond au module d'initiation

[http://proteopedia.org/wiki/index.php/6-deoxyerythronolide\\_B\\_synthase\\_%28DEBS%29...](http://proteopedia.org/wiki/index.php/6-deoxyerythronolide_B_synthase_%28DEBS%29...) 13

|  |    |
|--|----|
| Figure 3. Fonctionnement des domaines KS, AT et ACP avec un monomère de malonyl-CoA.   | 14 |
| Figure 4. Réaction catalysée par le domaine KR.  | 15 |
| Figure 5. Réaction catalysée par le domaine DH   | 15 |
| Figure 6. Réaction catalysée par le domaine ER.  | 16 |
| Figure 7. Mécanisme d'un domaine TE.   | 17 |
| Figure 8. A: mécanisme classique de condensation , B: Condensation erronée   | 19 |
| Figure 9. Fonctionnement d'une NRPS  | 22 |
| Figure 10. Différentes structures des cyanotoxines les plus connues  | 23 |
| Figure 11. Structure de l'anatoxine-a et de ses analogues  | 24 |
| Figure 12. Micrographie de bactéries de la souche <i>Oscillatoria</i> sp. PC 6506 cultivée au laboratoire (grossissement $\times 400$ )  | 25 |
| Figure 13. Le cluster <i>ana</i> responsable de la biosynthèse de l'anatoxine-a  | 26 |
| Figure 14. Schéma de biosynthèse de l'anatoxine-a (R=H) ou de l'homoanatoxine-a (R=Me)   | 27 |
| Figure 15. Mécanisme de la réaction catalysée par AnaB   | 28 |
| Figure 16. Elongation de la chaîne sur AnaE  | 29 |
| Figure 17. Mécanisme proposé pour la cyclisation sur AnaF  | 29 |
| Figure 18. Etapes catalysées sur AnaG  | 30 |
| Figure 19. Mécanisme proposé pour la décarboxylation de la carboxy-anatoxine-a   | 31 |
| Figure 20. Comparaison des clusters <i>ana</i> d' <i>Oscillatoria</i> sp. PCC 6506 et PCC 6407 avant le travail de reséquençage par l'équipe. Les nombres en dessous des gènes désignent les numéros des contigs, les deux barres verticales bleues délimitent la zone à l'intérieur de laquelle les séquences nucléotidiques sont identiques  | 35 |
| Figure 21. A : Contig 407 chez la souche <i>Oscillatoria</i> sp. PCC 6506 (version 2010). B : Assemblage des différents contigs à cet endroit du génome chez la souche <i>Oscillatoria</i> sp. PCC 6506 (version 2015). Les barres en rouge représentent les parties qui ont été reséquencées à partir d'amplification PCR. Les nombres en gras correspondent au numéro des contigs dans sa version 2010. Les caractères ' indiquent des différences par rapport à la séquence du contig disponible en 2010. Les quadrilatères en pointillés représentent des séquences qui étaient absents du génome dans sa version 2010. C : Contig 69 chez la souche <i>Oscillatoria</i> sp. PCC 6407. | 37 |
| Figure 22. Comparaison de la synténie des clusters <i>ana</i> de souches productrices d'anatoxine-a et d'homoanatoxine-a. Les traits en pointillés montrent que nous avons accès aux séquences nucléotidiques encadrant le cluster. Les barres verticales représentent des clusters dont le génome n'est pas disponible et les séquences nucléotidiques disponibles sont donc limitées en taille. L'alignement a été réalisé à partir du gène <i>anaB</i> .  | 41 |
| Figure 23. Comparaison des domaines présents dans la protéine AnaG dans les souches productrices d'anatoxine-a et ses dérivés. Les limites des domaines sont issues du logiciel de prédiction de domaine antiSMASH 3.0 et sont donc à envisager avec prudence  | 43 |
| Figure 24. Séquence peptidique du domaine D2 chez la souche <i>Oscillatoria</i> sp. PCC 6506 allant de l'acide aminé 1320 à 1559 avec prédiction de structure secondaire obtenue avec le logiciel PSIPRED, les cylindres roses représentent les hélices $\alpha$ prédites, les flèches jaunes les brins $\beta$ . Les barres bleues correspondent à la probabilité de la prédiction.   | 45 |
| Figure 25. Voie de synthèse en série racémique de l'anatoxine-a développée par Parson <sup>132</sup> ...   | 49 |
| Figure 26. Structure de l'analogue et du substrat putatif d'AnaG.  | 50 |
| Figure 27. Voie de synthèse de l'analogue du substrat putatif d'AnaG   | 51 |
| Figure 28. Mécanisme proposé pour l'ouverture du $\beta$ -lactame  | 52 |
| Figure 29. Mécanisme proposé de formation du produit sélénié   | 53 |
| Figure 30. Mécanisme de la SN2   | 54 |

|   |    |
|---|----|
| Figure 31. Mécanisme de la formation du composé (5).....  | 55 |
| Figure 32. Réaction avec le LDA et le PhSeBr .....  | 56 |
| Figure 33. Mécanisme d'élimination du sélénioxyde .....   | 57 |
| Figure 34. Spectre de masse du mélange après thioesterification en MH+ en CI (ionisation chimique) par le Dr Stéphane Mann au service de spectrométrie de masse de Chimie Paristech .....   | 58 |
| Figure 35. A : Produits attendus lors de la réaction de thioestérification ; B : Sous-produit obtenu lors de la réaction de thioesterification, les masses sont en g.mol <sup>-1</sup> .....  | 59 |
| Figure 36. Déprotection du produit de thioestérification .....  | 60 |
| Figure 37. Chromatogramme de la LC-MS du produit, la colonne utilisé est une C18 analytique avec un volume d'injection de 8 µL avec un gradient isocratique 20% méthanol/ 80% d'eau milliporée avec 0,1% en volume de TFA.....  | 60 |
| Figure 38. Spectre de masse en ESI du deuxième pic.....   | 61 |
| Figure 39. Spectre de masse en ESI du troisième pic .....   | 61 |
| Figure 40. Chromatogrammes des injections LC-MS. A : Première fraction. B : Deuxième fraction. La colonne utilisée est une C18 analytique avec un volume d'injection de 8 µL avec un gradient isocratique 20% Méthanol/ 80% d'eau milliporée avec 0,1% en volume de TFA. ....   | 62 |
| Figure 41. Spectre de masse en ESI du premier pic de la première fraction.....  | 63 |
| Figure 42. Spectre de masse en ESI du deuxième pic.....   | 63 |
| Figure 43. Structure proposée de l'impureté.....  | 64 |
| Figure 44. AnaG avec les domaines connus.....   | 66 |
| Figure 45. Domaines présents chez AnaG .....  | 68 |
| Figure 46. Récapitulatifs des étapes prévues pour l'étude des dernières étapes de la voie de biosynthèse .....  | 69 |
| Figure 47. 1 : Partie N-terminale d'AnaG après purification, dépôt de 15 µL à 0,126 mg.mL <sup>-1</sup> .....   | 70 |
| Figure 48. 1: Partie C-terminale d'AnaG après purification, dépôt de 15 µL à 0,88 mg.mL <sup>-1</sup>   | 71 |
| Figure 49. Gel d'électrophorèse après six heures d'incubation, 1: partie C-terminale d'AnaG , 2: partie C-terminale d'AnaG + cocktail commercial d'inhibiteurs, 3: partie C-terminale d'AnaG + Pefabloc, 4: partie C-terminale d'AnaG + Pepstatin, 5: partie C-terminale d'AnaG + EDTA,6: partie C-terminale d'AnaG + Aprotinine, 7: partie C-terminale d'AnaG + bestatin, 8: partie C-terminale d'AnaG dégradée, dépôt de 15 µL à 1,07 mg.mL <sup>-1</sup> ..... | 72 |
| Figure 50. 1: Fraction purifiée sur colonne d'affinité concentrée de la partie C-terminale d'AnaG, dépôt de 10 µL à 16,8 mg.mL <sup>-1</sup> , 2: fraction purifiée sur colonne d'affinité de la partie C-terminale d'AnaG, dépôt de 10 µL à 5,9 mg.mL <sup>-1</sup> .....  | 73 |
| Figure 51. 1: Fraction purifiée sur colonne échangeuse d'ions de la partie C-terminale d'AnaG, dépôt de 15 µL à 1,55 mg.mL <sup>-1</sup> , 2: partie C-terminale d'AnaG dégradée, dépôt de 10 µL à 1,07 mg.mL <sup>-1</sup> .....   | 73 |
| Figure 52. Résultat du gel filtration sur une colonne Superdex 200, orange : suivi de l'absorbance à 260 nm, bleu : suivi de l'absorbance à 280 nm.....   | 74 |
| Figure 53. Séquence peptidique du domaine ACP avec prédiction de structure secondaire obtenue avec le logiciel PSIPRED et les deux séquences choisies pour le clonage, les cylindres roses représentent les hélices α prédites. Les barres bleues correspondent à la probabilité de la prédiction.....  | 75 |
| Figure 54. 1 : Domaine ACP d'AnaG après purification, dépôt de 5 µL à 0,58 mg.mL <sup>-1</sup> .....  | 77 |
| Figure 55. Chargement du bras phosphopantéthéine .....  | 78 |
| Figure 56. Spectre de MALDI-TOF A: mélange contenant l'Osppt et B: mélange témoin sans l'Osppt.....   | 78 |

|  |    |
|--|----|
| Figure 57. A: Séquence du pET28a(+)(Kanr)-anaG pGP02 , B: séquence du pET28a(+)(Kanr)-anaG pGP03, XhoI et HindIII correspondent au site de l'enzyme de restriction utilisé pour ce clonage.....  | 80 |
| Figure 58. 1 : AnaA après purification, dépôt de 20 µL de protéine à 0,14 mg.mL <sup>-1</sup> .....  | 81 |
| Figure 59. Structure de deux isomères de la dihydroanatoxine-a .....   | 83 |
| Figure 60. Voie de synthèse des dhATXs et dhHATXs.....   | 85 |
| Figure 61. Chromatogrammes obtenus en LC-MS sur une colonne C18 effectué par Mélanie Cohen en collaboration avec l'équipe de Valérie Pichon à l'ESPCI; (A) cis-dhATX, (B) trans-dhATX, (C) cis-dhHATX, (D) trans-dhHATX, * impureté .....  | 87 |
| Figure 62. Mécanisme d'épimérisation entre l'isomère <i>cis</i> et l'isomère <i>trans</i> .....  | 88 |
| Figure 63. Chromatogramme en CI MS <sup>2</sup> de la <i>cis</i> -dhATX (A) et de la <i>trans</i> -dhATX (B).....  | 89 |
| Figure 64. Spectres de masse en CI MS <sup>2</sup> du premier pic à 26,36 minutes (A) et deuxième pic à 28 minutes (B) de la <i>cis</i> -dhATX .....   | 90 |
| Figure 65. Chromatogramme en CI MS <sup>2</sup> de la <i>cis</i> -dhATX purifié en HPLC (A), spectre de masse du pic à 26,5 minutes (B) .....  | 91 |
| Figure 66. Injection sur HPLC analytique avec en abscisse le temps en minutes et en ordonnée l'absorbance à 230 nm, la colonne utilisé est une C18 analytique avec un volume d'injection de 20 µL .....                                    | 92 |
| Figure 67. Spectres de masses obtenus en EI pour la dhATX .....  | 93 |
| Figure 68. Chromatogramme en CI MS <sup>2</sup> d'un extrait naturel de <i>Cylindrospermum stagnale</i> PCC 7417 (A), spectre de masse de la dhATX présente dans l'échantillon (B).....  | 94 |
| Figure 69. Incorporation des carbones marqués au C <sub>13</sub> et de l'azote N <sub>15</sub> dans la dhATX selon la voie de biosynthèse proposée .....   | 95 |
| Figure 70. Chromatogramme en CI MS <sup>2</sup> d'un extrait naturel de <i>Cylindrospermum stagnale</i> PCC 7417 avec incorporation de Proline marquée (A), spectre de masse de la dihydroanatoxine-a présente dans l'échantillon (B)..... | 95 |
| Figure 71. Spectres de masse obtenue en EI (A) pour la dhATX sans incorporation isotopique et (B) pour la dhATX avec incorporation isotopique chez la <i>Cylindrospermum stagnale</i> PCC7417 .....  | 96 |
| Figure 72. Voie de fragmentation et présence des C <sub>13</sub> et du N <sub>15</sub> sur la dhATX.....   | 97 |
| Figure 73. Mécanisme de réaction proposée pour la réductase.....   | 98 |

## Liste des Tableaux

|  |     |
|--|-----|
| Tableau 1 : Souches de cyanobactéries productrices d'anatoxine-a et de ses dérivés connues et dont le cluster <i>ana</i> a été séquencé. ....  | 39  |
| Tableau 2 : Comparaison des différents gènes <i>ana</i> entre les différents clusters. La première valeur correspond à celle obtenus avec un BLASTN et la deuxième avec un BLASTP. * : identité obtenue avec une couverture de seulement 29% de la séquence. # : identité obtenue sur les parties de séquence présentes dans le cluster..... | 40  |
| Tableau 3 : Amorces utilisées pour les PCRs. Les sites de restriction sont en gras et soulignés .....  | 105 |
| Tableau 4 : Récapitulatif des plasmides utilisés au laboratoire.....   | 109 |
| Tableau 5 : Antibiotiques utilisés avec les différentes souches d' <i>E.coli</i> .....   | 110 |
| Tableau 6 : Conditions de purification des protéines .....   | 113 |
| Tableau 7 : Gradient utilisé lors de l'analyse de l'analogue de substrat .....   | 129 |
| Tableau 8 : Gradient utilisé lors de la purification de l'analogue en analytique .....   | 129 |
| Tableau 9 : Gradient utilisé lors de la purification de l'analogue en semi-préparative .....   | 130 |
| Tableau 10 : Gradient utilisé lors de l'analyse de l'analogue de substrat purifié par HPLC .   | 130 |

|  |     |
|--|-----|
| Tableau 11 : Condition de température lors de l'analyse des standards .....  | 132 |
| Tableau 12 : Gradient utilisé lors de la purification du cis dihydroanatoxine-a .....  | 132 |
| Tableau 13 : Condition de température lors de l'analyse de la cis-dihydroanatoxine-a .....   | 133 |
| Tableau 14 : Condition de température lors de l'analyse d'extrait de <i>Cylindrospermum stagnale</i> PCC 7417.....   | 133 |
| Tableau 15 : Annotation des gènes en 5' et en 3' du cluster <i>ana</i> , les gènes du cluster <i>ana</i> n'y figurent pas ainsi que les fragments de gènes ..... | 134 |
| Tableau 16 : Annotation des gènes de la souche <i>Anabaena</i> sp. strain 37 compris entre le nucléotide 20587 et le nucléotide 27525.....                       | 134 |

## Bibliographie

- (1) Seckbach, D. J.; Oren, A. Oxygenic Photosynthetic Microorganisms in Extreme Environments. In *Algae and Cyanobacteria in Extreme Environments*; Seckbach, D. J., Ed.; Cellular Origin, Life in Extreme Habitats and Astrobiology; Springer Netherlands, 2007; pp 3–25.
- (2) Paerl, H. W.; Fulton, R. S.; Moisaner, P. H.; Dyble, J. Harmful Freshwater Algal Blooms, With an Emphasis on Cyanobacteria. *The Scientific World Journal* **2001**, *1*, 76–113.
- (3) Marinelli, F. Chapter 2. From Microbial Products to Novel Drugs That Target a Multitude of Disease Indications. *Meth. Enzymol.* **2009**, *458*, 29–58.
- (4) Gao, Q.; Garcia-Pichel, F. Microbial Ultraviolet Sunscreens. *Nat. Rev. Microbiol.* **2011**, *9* (11), 791–802.
- (5) Méjean, A.; Ploux, O. A Genomic View of Secondary Metabolite Production in Cyanobacteria. In *Genomics of Cyanobacteria*; Advances in Botanical Research; Academic Press, 2013; Vol. 65, pp 189–234.
- (6) Kehr, J.-C.; Gatte Picchi, D.; Dittmann, E. Natural Product Biosyntheses in Cyanobacteria: A Treasure Trove of Unique Enzymes. *Beilstein Journal of Organic Chemistry* **2011**, *7*, 1622–1635.
- (7) Nunnery, J. K.; Mevers, E.; Gerwick, W. H. Biologically Active Secondary Metabolites from Marine Cyanobacteria. *Curr. Opin. Biotechnol.* **2010**, *21* (6), 787–793.
- (8) Hellier, P.; Purton, S.; Ladommatos, N. Molecular Structure of Photosynthetic Microbial Biofuels for Improved Engine Combustion and Emissions Characteristics. *Front Bioeng Biotechnol* **2015**, *3*, 49.
- (9) Wackett, L. P. Engineering Microbes to Produce Biofuels. *Curr. Opin. Biotechnol.* **2011**, *22* (3), 388–393.
- (10) Wackett, L. P. Microbial-Based Motor Fuels: Science and Technology. *Microb Biotechnol* **2008**, *1* (3), 211–225.
- (11) van Apeldoorn, M. E.; van Egmond, H. P.; Speijers, G. J. A.; Bakker, G. J. I. Toxins of Cyanobacteria. *Mol Nutr Food Res* **2007**, *51* (1), 7–60.
- (12) Smith, S.; Tsai, S.-C. The Type I Fatty Acid and Polyketide Synthases: A Tale of Two Megasyntases. *Nat Prod Rep* **2007**, *24* (5), 1041–1072.
- (13) Fischbach, M. A.; Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chem. Rev.* **2006**, *106* (8), 3468–3496.
- (14) Kumar, P.; Li, Q.; Cane, D. E.; Khosla, C. Intermodular Communication in Modular Polyketide Synthases: Structural and Mutational Analysis of Linker Mediated Protein–Protein Recognition. *J. Am. Chem. Soc.* **2003**, *125* (14), 4097–4102.

- (15) Tsuji, S. Y.; Cane, D. E.; Khosla, C. Selective Protein–Protein Interactions Direct Channeling of Intermediates between Polyketide Synthase Modules†. *Biochemistry* **2001**, *40* (8), 2326–2331.
- (16) Gokhale, R. S.; Tsuji, S. Y.; Cane, D. E.; Khosla, C. Dissecting and Exploiting Intermodular Communication in Polyketide Synthases. *Science* **1999**, *284* (5413), 482–485.
- (17) Liang, Z.-X. Complexity and Simplicity in the Biosynthesis of Eneidyne Natural Products. *Nat Prod Rep* **2010**, *27* (4), 499–528.
- (18) Lim, J.; Kong, R.; Murugan, E.; Ho, C. L.; Liang, Z.-X.; Yang, D. Solution Structures of the Acyl Carrier Protein Domain from the Highly Reducing Type I Iterative Polyketide Synthase CalE8. *PLoS ONE* **2011**, *6* (6), e20549.
- (19) Minto, R. E.; Townsend, C. A. Enzymology and Molecular Biology of Aflatoxin Biosynthesis. *Chem. Rev.* **1997**, *97* (7), 2537–2556.
- (20) Weissman, K. J. The Structural Biology of Biosynthetic Megaenzymes. *Nat Chem Biol* **2014**, *11* (9), 660–670.
- (21) Stinear, T. P.; Mve-Obiang, A.; Small, P. L. C.; Frigui, W.; Pryor, M. J.; Brosch, R.; Jenkin, G. A.; Johnson, P. D. R.; Davies, J. K.; Lee, R. E.; et al. Giant Plasmid-Encoded Polyketide Synthases Produce the Macrolide Toxin of *Mycobacterium Ulcerans*. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (5), 1345–1349.
- (22) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of *Saccharopolyspora Erythraea*. *Nature* **1990**, *348* (6297), 176–178.
- (23) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. Modular Organization of Genes Required for Complex Polyketide Biosynthesis. *Science* **1991**, *252* (5006), 675–679.
- (24) Xue, Y.; Zhao, L.; Liu, H.; Sherman, D. H. A Gene Cluster for Macrolide Antibiotic Biosynthesis in *Streptomyces Venezuelae*: Architecture of Metabolic Diversity. *PNAS* **1998**, *95* (21), 12111–12116.
- (25) Kittendorf, J. D.; Sherman, D. H. The Methymycin/Pikromycin Biosynthetic Pathway: A Model for Metabolic Diversity in Natural Product Biosynthesis. *Bioorg Med Chem* **2009**, *17* (6), 2137–2146.
- (26) Keatinge-Clay, A. T. The Structures of Type I Polyketide Synthases. *Natural Product Reports* **2012**.
- (27) Tang, Y.; Kim, C.-Y.; Mathews, I. I.; Cane, D. E.; Khosla, C. The 2.7-Angstrom Crystal Structure of a 194-kDa Homodimeric Fragment of the 6-Deoxyerythronolide B Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (30), 11124–11129.
- (28) Keatinge-Clay, A. T.; Stroud, R. M. The Structure of a Ketoreductase Determines the Organization of the Beta-Carbon Processing Enzymes of Modular Polyketide Synthases. *Structure* **2006**, *14* (4), 737–748.
- (29) 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. Structure of a Modular Polyketide Synthase. *Nature* **2014**, *510* (7506), 512–517.
- (30) Kwan, D. H.; Schulz, F. The Stereochemistry of Complex Polyketide Biosynthesis by Modular Polyketide Synthases. *Molecules* **2011**, *16* (7), 6092–6115.
- (31) Weissman, K. J.; Timoney, M.; Bycroft, M.; Grice, P.; Hanefeld, U.; Staunton, J.; Leadlay, P. F. The Molecular Basis of Celmer’s Rules: The Stereochemistry of the Condensation Step in Chain Extension on the Erythromycin Polyketide Synthase. *Biochemistry* **1997**, *36* (45), 13849–13855.

- (32) Garg, A.; Khosla, C.; Cane, D. E. Coupled Methyl Group Epimerization and Reduction by Polyketide Synthase Ketoreductase Domains. Ketoreductase-Catalyzed Equilibrium Isotope Exchange. *J. Am. Chem. Soc.* **2013**, *135* (44), 16324–16327.
- (33) 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. Molecular Basis of Celmer's Rules: Stereochemistry of Catalysis by Isolated Ketoreductase Domains from Modular Polyketide Synthases. *Chemistry & Biology* **2005**, *12* (10), 1145–1153.
- (34) Caffrey, P. Conserved Amino Acid Residues Correlating With Ketoreductase Stereospecificity in Modular Polyketide Synthases. *ChemBioChem* **2003**, *4* (7), 654–657.
- (35) Li, Y.; Fiers, W. D.; Bernard, S. M.; Smith, J. L.; Aldrich, C. C.; Fecik, R. A. Polyketide Intermediate Mimics as Probes for Revealing Cryptic Stereochemistry of Ketoreductase Domains. *ACS Chem. Biol.* **2014**, *9* (12), 2914–2922.
- (36) Akey, D. L.; Razelun, J. R.; Tehranisa, J.; Sherman, D. H.; Gerwick, W. H.; Smith, J. L. Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* **2010**, *18* (1), 94–105.
- (37) Palaniappan, N.; Alhamadsheh, M. M.; Reynolds, K. A. Cis- $\Delta$ 2,3-Double Bond of Phoslactomycins Is Generated by a Post-PKS Tailoring Enzyme. *J. Am. Chem. Soc.* **2008**, *130* (37), 12236–12237.
- (38) Persson, B.; Zigler, J. S.; Jörnvall, H. A Super-Family of Medium-Chain Dehydrogenases/reductases (MDR). Sub-Lines Including Zeta-Crystallin, Alcohol and Polyol Dehydrogenases, Quinone Oxidoreductase Enoyl Reductases, VAT-1 and Other Proteins. *Eur. J. Biochem.* **1994**, *226* (1), 15–22.
- (39) Maier, T.; Jenni, S.; Ban, N. Architecture of Mammalian Fatty Acid Synthase at 4.5 Å Resolution. *Science* **2006**, *311* (5765), 1258–1262.
- (40) Kwan, D. H.; Leadlay, P. F. Mutagenesis of a Modular Polyketide Synthase Enoylreductase Domain Reveals Insights into Catalysis and Stereospecificity. *ACS Chem. Biol.* **2010**, *5* (9), 829–838.
- (41) Kwan, D. H.; Sun, Y.; Schulz, F.; Hong, H.; Popovic, B.; Sim-Stark, J. C. C.; Haydock, S. F.; Leadlay, P. F. Prediction and Manipulation of the Stereochemistry of Enoylreduction in Modular Polyketide Synthases. *Chemistry & Biology* **2008**, *15* (11), 1231–1240.
- (42) Dunn, B. J.; Khosla, C. Engineering the Acyltransferase Substrate Specificity of Assembly Line Polyketide Synthases. *J R Soc Interface* **2013**, *10* (85).
- (43) Patel, K.; Piagentini, M.; Rascher, A.; Tian, Z.-Q.; Buchanan, G. O.; Regentin, R.; Hu, Z.; Hutchinson, C. R.; McDaniel, R. Engineered Biosynthesis of Geldanamycin Analogs for Hsp90 Inhibition. *Chem. Biol.* **2004**, *11* (12), 1625–1633.
- (44) Biosynthesis and Attachment of Novel Bacterial Polyketide Synthase Starter Units. *Natural Product Reports* **2002**, *19* (1), 70–99.
- (45) Keating, T. A.; Ehmann, D. E.; Kohli, R. M.; Marshall, C. G.; Trauger, J. W.; Walsh, C. T. Chain Termination Steps in Nonribosomal Peptide Synthetase Assembly Lines: Directed Acyl-S-Enzyme Breakdown in Antibiotic and Siderophore Biosynthesis. *ChemBioChem* **2001**, *2* (2), 99–107.
- (46) Khosla, C. Harnessing the Biosynthetic Potential of Modular Polyketide Synthases. *Chem. Rev.* **1997**, *97* (7), 2577–2590.
- (47) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase. *Chemistry & Biology* **1999**, *6* (2), 117–125.
- (48) Krätzschmar, J.; Krause, M.; Marahiel, M. A. Gramicidin S Biosynthesis Operon Containing the Structural Genes *grsA* and *grsB* Has an Open Reading Frame

- Encoding a Protein Homologous to Fatty Acid Thioesterases. *J. Bacteriol.* **1989**, *171* (10), 5422–5429.
- (49) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. Characterisation of 3-Methylorcinaldehyde Synthase (MOS) in *Acremonium Strictum*: First Observation of a Reductive Release Mechanism during Polyketide Biosynthesis. *Chem. Commun. (Camb.)* **2007**, No. 39, 4053–4055.
- (50) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. Identification of Claisen Cyclase Domain in Fungal Polyketide Synthase WA, a Naphthopyrone Synthase of *Aspergillus Nidulans*. *Chemistry & Biology* **2001**, *8* (2), 189–197.
- (51) Awakawa, T.; Yokota, K.; Funo, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. Physically Discrete  $\beta$ -Lactamase-Type Thioesterase Catalyzes Product Release in Atrochryson Synthesis by Iterative Type I Polyketide Synthase. *Chemistry & Biology* **2009**, *16* (6), 613–623.
- (52) Frank, B.; Wenzel, S. C.; Bode, H. B.; Scharfe, M.; Blöcker, H.; Müller, R. From Genetic Diversity to Metabolic Unity: Studies on the Biosynthesis of Aurafurones and Aurafuron-like Structures in Myxobacteria and Streptomycetes. *Journal of Molecular Biology* **2007**, *374* (1), 24–38.
- (53) Sun, Y.; Zhou, X.; Dong, H.; Tu, G.; Wang, M.; Wang, B.; Deng, Z. A Complete Gene Cluster from *Streptomyces Nanchangensis* NS3226 Encoding Biosynthesis of the Polyether Ionophore Nanchangmycin. *Chemistry & Biology* **2003**, *10* (5), 431–441.
- (54) Heathcote, M. L.; Staunton, J.; Leadlay, P. F. Role of Type II Thioesterases: Evidence for Removal of Short Acyl Chains Produced by Aberrant Decarboxylation of Chain Extender Units. *Chemistry & Biology* **2001**, *8* (2), 207–220.
- (55) Kim, B. S.; Cropp, T. A.; Beck, B. J.; Sherman, D. H.; Reynolds, K. A. Biochemical Evidence for an Editing Role of Thioesterase II in the Biosynthesis of the Polyketide Pikromycin. *J. Biol. Chem.* **2002**, *277* (50), 48028–48034.
- (56) Liu, T.; Lin, X.; Zhou, X.; Deng, Z.; Cane, D. E. Mechanism of Thioesterase-Catalyzed Chain Release in the Biosynthesis of the Polyether Antibiotic Nanchangmycin. *Chem. Biol.* **2008**, *15* (5), 449–458.
- (57) Walsh, C. T.; Fischbach, M. A. Natural Products Version 2.0: Connecting Genes to Molecules. *J. Am. Chem. Soc.* **2010**, *132* (8), 2469–2493.
- (58) Lin, S.; Huang, T.; Shen, B. Tailoring Enzymes Acting on Carrier Protein-Tethered Substrates in Natural Product Biosynthesis. In *Methods in Enzymology*; Elsevier, 2012; Vol. 516, pp 321–343.
- (59) Andersen, J. F.; Tatsuta, K.; Gunji, H.; Ishiyama, T.; Hutchinson, C. R. Substrate Specificity of 6-Deoxyerythronolide B Hydroxylase, a Bacterial Cytochrome P450 of Erythromycin A Biosynthesis. *Biochemistry* **1993**, *32* (8), 1905–1913.
- (60) Lambalot, R. H.; Cane, D. E.; Aparicio, J. J.; Katz, L. Overproduction and Characterization of the Erythromycin C-12 Hydroxylase, EryK. *Biochemistry* **1995**, *34* (6), 1858–1866.
- (61) Staunton, J.; Wilkinson, B. Biosynthesis of Erythromycin and Rapamycin. *Chemical reviews* **1997**, *97* (7), 2611–2630.
- (62) Thompson, T. B.; Katayama, K.; Watanabe, K.; Hutchinson, C. R.; Rayment, I. Structural and Functional Analysis of Tetracenomycin F2 Cyclase from *Streptomyces Glaucescens* A TYPE II POLYKETIDE CYCLASE. *J. Biol. Chem.* **2004**, *279* (36), 37956–37963.
- (63) Hutchinson, C. R. Biosynthetic Studies of Daunorubicin and Tetracenomycin C. *Chem. Rev.* **1997**, *97* (7), 2525–2536.

- (64) Yu, D.; Xu, F.; Zeng, J.; Zhan, J. Type III Polyketide Synthases in Natural Product Biosynthesis. *IUBMB Life* **2012**, *64* (4), 285–295.
- (65) Calteau, A.; Fewer, D. P.; Latifi, A.; Coursin, T.; Laurent, T.; Jokela, J.; Kerfeld, C. A.; Sivonen, K.; Piel, J.; Gugger, M. Phylum-Wide Comparative Genomics Unravel the Diversity of Secondary Metabolism in Cyanobacteria. *BMC Genomics* **2014**, *15* (1), 977.
- (66) Du, L.; Sánchez, C.; Shen, B. Hybrid Peptide–Polyketide Natural Products: Biosynthesis and Prospects toward Engineering Novel Molecules. *Metabolic Engineering* **2001**, *3* (1), 78–95.
- (67) Nishizawa, T.; Asayama, M.; Fujii, K.; Harada, K.; Shirai, M. Genetic Analysis of the Peptide Synthetase Genes for a Cyclic Heptapeptide Microcystin in *Microcystis* Spp. *J Biochem* **1999**, *126* (3), 520–529.
- (68) Méjean, A.; Paci, G.; Gautier, V.; Ploux, O. Biosynthesis of Anatoxin-a and Analogues (anatoxins) in Cyanobacteria. *Toxicon* **2014**, *91*, 15–22.
- (69) Mazmouz, R.; Chapuis-Hugon, F.; Mann, S.; Pichon, V.; Méjean, A.; Ploux, O. Biosynthesis of Cylindrospermopsin and 7-Epicylindrospermopsin in *Oscillatoria* Sp. Strain PCC 6506: Identification of the *Cyr* Gene Cluster and Toxin Analysis. *Appl. Environ. Microbiol.* **2010**, *76* (15), 4943–4949.
- (70) Tidgewell, K.; Engene, N.; Byrum, T.; Media, J.; Doi, T.; Valeriote, F. A.; Gerwick, W. H. Evolved Diversification of a Modular Natural Product Pathway: Apratoxins F and G, Two Cytotoxic Cyclic Depsipeptides from a Palmyra Collection of *Lyngbya Bouillonii*. *ChemBioChem* **2010**, *11* (10), 1458–1466.
- (71) Gugger, M.; Lenoir, S.; Berger, C.; Ledreux, A.; Druart, J.-C.; Humbert, J.-F.; Guette, C.; Bernard, C. First Report in a River in France of the Benthic Cyanobacterium *Phormidium Favosum* Producing Anatoxin-a Associated with Dog Neurotoxicosis. *Toxicon* **2005**, *45* (7), 919–928.
- (72) Cadel-Six, S.; Peyraud-Thomas, C.; Brient, L.; Marsac, N. T. de; Rippka, R.; Méjean, A. Different Genotypes of Anatoxin-Producing Cyanobacteria Coexist in the Tarn River, France. *Appl. Environ. Microbiol.* **2007**, *73* (23), 7605–7614.
- (73) Matsushima, R.; Yoshizawa, S.; Watanabe, M. F.; Harada, K.; Furusawa, M.; Carmichael, W. W.; Fujiki, H. In Vitro and in Vivo Effects of Protein Phosphatase Inhibitors, Microcystins and Nodularin, on Mouse Skin and Fibroblasts. *Biochem. Biophys. Res. Commun.* **1990**, *171* (2), 867–874.
- (74) Sivonen, K.; Jones, G. J. Cyanobacterial toxins. I. Chorus and J. Bartram 1999.
- (75) Norris, R. L. G.; Eaglesham, G. K.; Shaw, G. R.; Senogles, P.; Chiswell, R. K.; Smith, M. J.; Davis, B. C.; Seawright, A. A.; Moore, M. R. Extraction and Purification of the Zwitterions Cylindrospermopsin and Deoxycylindrospermopsin from *Cylindrospermopsis Raciborskii*. *Environ. Toxicol.* **2001**, *16* (5), 391–396.
- (76) Banker, R.; Teltsch, B.; Sukenik, A.; Carmeli, S. 7-Epicylindrospermopsin, a Toxic Minor Metabolite of the Cyanobacterium *Aphanizomenon Ovalisporum* from Lake Kinneret, Israel. *J. Nat. Prod.* **2000**, *63* (3), 387–389.
- (77) Harada, K. I.; Ohtani, I.; Iwamoto, K.; Suzuki, M.; Watanabe, M. F.; Watanabe, M.; Terao, K. Isolation of Cylindrospermopsin from a Cyanobacterium *Umezakia Natans* and Its Screening Method. *Toxicon* **1994**, *32* (1), 73–84.
- (78) Banker, R.; Carmeli, S.; Hadas, O.; Teltsch, B.; Porat, R.; Sukenik, A. IDENTIFICATION OF CYLINDROSPERMOPSIN IN APHANIZOMENON OVALISPORUM (CYANOPHYCEAE) ISOLATED FROM LAKE KINNERET, ISRAEL. *Journal of Phycology* **1997**, *33* (4), 613–616.

- (79) Preussel, K.; Stüken, A.; Wiedner, C.; Chorus, I.; Fastner, J. First Report on Cylindrospermopsin Producing *Aphanizomenon Flos-Aquae* (Cyanobacteria) Isolated from Two German Lakes. *Toxicon* **2006**, *47* (2), 156–162.
- (80) Li, R.; Carmichael, W. W.; Brittain, S.; Eaglesham, G. K.; Shaw, G. R.; Liu, Y.; Watanabe, M. M. FIRST REPORT OF THE CYANOTOXINS CYLINDROSPERMOPSIN AND DEOXYCYLINDROSPERMOPSIN FROM RAPHIDIOPSIS CURVATA (CYANOBACTERIA). *Journal of Phycology* **2001**, *37* (6), 1121–1126.
- (81) Spoo, L.; Berg, K. A.; Rapala, J.; Lahti, K.; Lepistö, L.; Metcalf, J. S.; Codd, G. A.; Meriluoto, J. First Observation of Cylindrospermopsin in *Anabaena Lapponica* Isolated from the Boreal Environment (Finland). *Environ. Toxicol.* **2006**, *21* (6), 552–560.
- (82) Seifert, M.; McGregor, G.; Eaglesham, G.; Wickramasinghe, W.; Shaw, G. First Evidence for the Production of Cylindrospermopsin and Deoxy-Cylindrospermopsin by the Freshwater Benthic Cyanobacterium, *Lyngbya Wollei* (Farlow Ex Gomont) Zpeziale and Dyck. *Harmful algae* **6** (1), 73–80.
- (83) Mazmouz, R.; Chapuis-Hugon, F.; Pichon, V.; Méjean, A.; Ploux, O. The Last Step of the Biosynthesis of the Cyanotoxins Cylindrospermopsin and 7-Epi-Cylindrospermopsin Is Catalysed by CyrI, a 2-Oxoglutarate-Dependent Iron Oxygenase. *Chembiochem* **2011**, *12* (6), 858–862.
- (84) Méjean, A.; Mazmouz, R.; Mann, S.; Calteau, A.; Médigue, C.; Ploux, O. The Genome Sequence of the Cyanobacterium *Oscillatoria* Sp. PCC 6506 Reveals Several Gene Clusters Responsible for the Biosynthesis of Toxins and Secondary Metabolites. *J. Bacteriol.* **2010**, *192* (19), 5264–5265.
- (85) *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring, and Management*; Chorus, I., Bartram, J., Eds.; E & FN Spon: London ; New York, 1999.
- (86) Strichartz, G.; Rando, T.; Hall, S.; Gitschier, J.; Hall, L.; Magnani, B.; Bay, C. H. On the Mechanism by Which Saxitoxin Binds to and Blocks Sodium Channels. *Ann. N. Y. Acad. Sci.* **1986**, *479*, 96–112.
- (87) Humpage, A. R.; Rositano, J.; Baker, P. D.; Nicholson, B. C.; Steffensen, D. A.; Bretag, A. H.; Brown, R. K. Paralytic Shellfish Poisons from Freshwater Blue-Green Algae. *Med. J. Aust.* **1993**, *159* (6), 423.
- (88) Wonnacott, S.; Gallagher, T. The Chemistry and Pharmacology of Anatoxin-a and Related Homotropenes with Respect to Nicotinic Acetylcholine Receptors. *Marine Drugs* **2006**, *4* (3), 228–254.
- (89) Devlin, J. P.; Edwards, O. E.; Gorham, P. R.; Hunter, N. R.; Pike, R. K.; Stavric, B. Anatoxin-A, a Toxic Alkaloid from *Anabaena Flos-Aquae* NRC-44h. *Can. J. Chem.* **1977**, *55* (8), 1367–1371.
- (90) Rogers, E. H.; Hunter, E. S.; Moser, V. C.; Phillips, P. M.; Herkovits, J.; Muñoz, L.; Hall, L. L.; Chernoff, N. Potential Developmental Toxicity of Anatoxin-A, a Cyanobacterial Toxin. *J Appl Toxicol* **2005**, *25* (6), 527–534.
- (91) Devlin, J. P.; Edwards, O. E.; Gorham, P. R.; Hunter, N. R.; Pike, R. K.; Stavric, B. Anatoxin-A, a Toxic Alkaloid from *Anabaena Flos-Aquae* NRC-44h. *Can. J. Chem.* **1977**, *55* (8), 1367–1371.
- (92) Campbell, H. F.; Edwards, O. E.; Kolt, R. Synthesis of nor-Anatoxin-a and Anatoxin-A. *Canadian Journal of Chemistry* **1977**, *55* (8), 1372–1379.
- (93) Carmichael, W. W.; Biggs, D. F.; Gorham, P. R. Toxicology and Pharmacological Action of *Anabaena Flos-Aquae* Toxin. *Science* **1975**, *187* (4176), 542–544.

- (94) Spivak, C. E.; Witkop, B.; Albuquerque, E. X. Anatoxin-a: A Novel, Potent Agonist at the Nicotinic Receptor. *Mol Pharmacol* **1980**, *18* (3), 384–394.
- (95) Wonnacott, S.; Swanson, K. L.; Albuquerque, E. X.; Huby, N. J.; Thompson, P.; Gallagher, T. Homoanatoxin: A Potent Analogue of Anatoxin-A. *Biochem. Pharmacol.* **1992**, *43* (3), 419–423.
- (96) Skulberg, O. M.; Skulberg, R.; Carmichael, W. W.; Andersen, R. A.; Matsunaga, S.; Moore, R. E. Investigations of a Neurotoxic Oscillatorialean Strain (Cyanophyceae) and Its Toxin. Isolation and Characterization of Homoanatoxin-A. *Environmental Toxicology and Chemistry* **1992**, *11* (3), 321–329.
- (97) James, K. J.; Furey, A.; Sherlock, I. R.; Stack, M. A.; Twohig, M.; Caudwell, F. B.; Skulberg, O. M. Sensitive Determination of Anatoxin-A, Homoanatoxin-a and Their Degradation Products by Liquid Chromatography with Fluorimetric Detection. *J Chromatogr A* **1998**, *798* (1-2), 147–157.
- (98) Mann, S.; Cohen, M.; Chapuis-Hugon, F.; Pichon, V.; Mazmouz, R.; Méjean, A.; Ploux, O. Synthesis, Configuration Assignment, and Simultaneous Quantification by Liquid Chromatography Coupled to Tandem Mass Spectrometry, of Dihydroanatoxin-a and Dihydrohomoanatoxin-a Together with the Parent Toxins, in Axenic Cyanobacterial Strains and in Environmental Samples. *Toxicon* **2012**, *60* (8), 1404–1414.
- (99) Cadel-Six, S.; Iteman, I.; Peyraud-Thomas, C.; Mann, S.; Ploux, O.; Mejean, A. Identification of a Polyketide Synthase Coding Sequence Specific for Anatoxin-a-Producing Oscillatoria Cyanobacteria. *Applied and Environmental Microbiology* **2009**, *75* (14), 4909–4912.
- (100) Méjean, A.; Mann, S.; Maldiney, T.; Vassiliadis, G.; Lequin, O.; Ploux, O. Evidence That Biosynthesis of the Neurotoxic Alkaloids Anatoxin-a and Homoanatoxin-a in the Cyanobacterium Oscillatoria PCC 6506 Occurs on a Modular Polyketide Synthase Initiated by L-Proline. *J. Am. Chem. Soc.* **2009**, *131* (22), 7512–7513.
- (101) Méjean, A.; Mann, S.; Vassiliadis, G.; Lombard, B.; Loew, D.; Ploux, O. In Vitro Reconstitution of the First Steps of Anatoxin-a Biosynthesis in Oscillatoria PCC 6506: From Free L-Proline to Acyl Carrier Protein Bound Dehydroproline. *Biochemistry* **2010**, *49* (1), 103–113.
- (102) Moncoq, K.; Regad, L.; Mann, S.; Méjean, A.; Ploux, O. Structure of the Prolyl-Acyl Carrier Protein Oxidase Involved in the Biosynthesis of the Cyanotoxin Anatoxin-A. *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69* (Pt 12), 2340–2352.
- (103) The Biosynthesis of the Aromatic Myxobacterial Electron Transport Inhibitor Stigmatellin Is Directed by a Novel Type of Modular Polyketide Synthase <http://www.jbc.org/content/277/15/13082.long> (accessed Aug 5, 2015).
- (104) Selwood, A. I.; Holland, P. T.; Wood, S. A.; Smith, K. F.; McNabb, P. S. Production of Anatoxin-a and a Novel Biosynthetic Precursor by the Cyanobacterium *Aphanizomenon flos-aquae*. *Environ. Sci. Technol.* **2007**, *41* (2), 506–510.
- (105) Shih, P. M.; Wu, D.; Latifi, A.; Axen, S. D.; Fewer, D. P.; Talla, E.; Calteau, A.; Cai, F.; Marsac, N. T. de; Rippka, R.; et al. Improving the Coverage of the Cyanobacterial Phylum Using Diversity-Driven Genome Sequencing. *PNAS* **2013**, *110* (3), 1053–1058.
- (106) Rothberg, J. M.; Leamon, J. H. The Development and Impact of 454 Sequencing. *Nat Biotech* **2008**, *26* (10), 1117–1124.
- (107) King, C.; Scott-Horton, T. Pyrosequencing: A Simple Method for Accurate Genotyping. *J Vis Exp* **2008**, No. 11.

- (108) Margulies, M.; Egholm, M.; Altman, W. E.; Attiya, S.; Bader, J. S.; Bembien, L. A.; Berka, J.; Braverman, M. S.; Chen, Y.-J.; Chen, Z.; et al. Genome Sequencing in Microfabricated High-Density Picolitre Reactors. *Nature* **2005**, *437* (7057), 376–380.
- (109) Megléczy, E.; Pech, N.; Gilles, A.; Martin, J.-F.; Gardner, M. G. A Shot in the Genome: How Accurately Do Shotgun 454 Sequences Represent a Genome? *BMC Research Notes* **2012**, *5* (1), 259.
- (110) Vallenet, D.; Belda, E.; Calteau, A.; Cruveiller, S.; Engelen, S.; Lajus, A.; Le Fèvre, F.; Longin, C.; Mornico, D.; Roche, D.; et al. MicroScope--an Integrated Microbial Resource for the Curation and Comparative Analysis of Genomic and Metabolic Data. *Nucleic Acids Res.* **2013**, *41* (Database issue), D636–D647.
- (111) Hvorup, R. N.; Winnen, B.; Chang, A. B.; Jiang, Y.; Zhou, X.-F.; Saier, M. H. The Multidrug/oligosaccharidyl-Lipid/polysaccharide (MOP) Exporter Superfamily. *European Journal of Biochemistry* **2003**, *270* (5), 799–813.
- (112) Kuroda, T.; Tsuchiya, T. Multidrug Efflux Transporters in the MATE Family. *Biochim. Biophys. Acta* **2009**, *1794* (5), 763–768.
- (113) Brown, M. H.; Paulsen, I. T.; Skurray, R. A. The Multidrug Efflux Protein NorM Is a Prototype of a New Family of Transporters. *Mol. Microbiol.* **1999**, *31* (1), 394–395.
- (114) Rantala-Ylinen, A.; Känä, S.; Wang, H.; Rouhiainen, L.; Wahlsten, M.; Rizzi, E.; Berg, K.; Gugger, M.; Sivonen, K. Anatoxin-a Synthetase Gene Cluster of the Cyanobacterium *Anabaena* Sp. Strain 37 and Molecular Methods To Detect Potential Producers. *Appl. Environ. Microbiol.* **2011**, *77* (20), 7271–7278.
- (115) Namikoshi, M.; Murakami, T.; Fujiwara, T.; Nagai, H.; Niki, T.; Harigaya, E.; Watanabe, M. F.; Oda, T.; Yamada, J.; Tsujimura, S. Biosynthesis and Transformation of Homoanatoxin-a in the Cyanobacterium *Raphidiopsis Mediterranea* Skuja and Structures of Three New Homologues. *Chem. Res. Toxicol.* **2004**, *17* (12), 1692–1696.
- (116) Namikoshi, M.; Murakami, T.; Watanabe, M. F.; Oda, T.; Yamada, J.; Tsujimura, S.; Nagai, H.; Oishi, S. Simultaneous Production of Homoanatoxin-A, Anatoxin-A, and a New Non-Toxic 4-Hydroxyhomoanatoxin-a by the Cyanobacterium *Raphidiopsis Mediterranea* Skuja. *Toxicon* **2003**, *42* (5), 533–538.
- (117) Jiang, Y.; Song, G.; Pan, Q.; Yang, Y.; Li, R. Identification of Genes for Anatoxin-a Biosynthesis in *Cuspidothrix Issatschenkoi*. *Harmful Algae* **2015**, *46*, 43–48.
- (118) Sivonen, K.; Himberg, K.; Luukkainen, R.; Niemela, S. I.; Poon, G. K.; Codd, G. A. Preliminary Characterization of Neurotoxic Cyanobacteria Blooms and Strains from Finland. *Toxicity Assessment* **1989**, *4* (3), 339–352.
- (119) Méjean, et al. Résultats Non Publiés.
- (120) Bashiri, G.; Squire, C. J.; Moreland, N. J.; Baker, E. N. Crystal Structures of F420-Dependent Glucose-6-Phosphate Dehydrogenase FGD1 Involved in the Activation of the Anti-Tuberculosis Drug Candidate PA-824 Reveal the Basis of Coenzyme and Substrate Binding. *Journal of Biological Chemistry* **2008**, *283* (25), 17531–17541.
- (121) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Brucoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; et al. antiSMASH 3.0—a Comprehensive Resource for the Genome Mining of Biosynthetic Gene Clusters. *Nucl. Acids Res.* **2015**, *43* (W1), W237–W243.
- (122) Buchan, D. W. A.; Minneci, F.; Nugent, T. C. O.; Bryson, K.; Jones, D. T. Scalable Web Services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research* **2013**, *41* (W1), W349–W357.
- (123) Shams, S.; Capelli, C.; Cerasino, L.; Ballot, A.; Dietrich, D. R.; Sivonen, K.; Salmaso, N. Anatoxin-a Producing *Tychonema* (Cyanobacteria) in European Waterbodies. *Water Res.* **2015**, *69*, 68–79.

- (124) Wood, S. A.; Smith, F. M. J.; Heath, M. W.; Palfroy, T.; Gaw, S.; Young, R. G.; Ryan, K. G. Within-Mat Variability in Anatoxin-a and Homoanatoxin-a Production among Benthic Phormidium (Cyanobacteria) Strains. *Toxins* **2012**, *4* (10), 900–912.
- (125) Stucken, K.; Murillo, A. A.; Soto-Liebe, K.; Fuentes-Valdés, J. J.; Méndez, M. A.; Vásquez, M. Toxicity Phenotype Does Not Correlate with Phylogeny of *Cylindrospermopsis Raciborskii* Strains. *Syst. Appl. Microbiol.* **2009**, *32* (1), 37–48.
- (126) Bates, H. A.; Rapoport, H. Synthesis of Anatoxin a via Intramolecular Cyclization of Iminium Salts. *J. Am. Chem. Soc.* **1979**, *101* (5), 1259–1265.
- (127) Petersen, J. S.; Fels, G.; Rapoport, H. Chiroselective Syntheses of (+)- and (-)-Anatoxin a. *J. Am. Chem. Soc.* **1984**, *106* (16), 4539–4547.
- (128) Koskinen, A. M. P.; Rapoport, H. Synthetic and Conformational Studies on Anatoxin-a: A Potent Acetylcholine Agonist. *J. Med. Chem.* **1985**, *28* (9), 1301–1309.
- (129) Sardina, F. J.; Howard, M. H.; Morningstar, M.; Rapoport, H. Enantiodivergent Synthesis of (+)- and (-)-Anatoxin from L-Glutamic Acid. *J. Org. Chem.* **1990**, *55* (17), 5025–5033.
- (130) Magnus, N. A.; Ducry, L.; Rolland, V.; Wonnacott, S.; Gallagher, T. Direct C-11 Functionalisation of Anatoxin-A. Application to the Synthesis of New Ligand-Based Structural Probes. *J. Chem. Soc., Perkin Trans. 1* **1997**, No. 16, 2313–2318.
- (131) Mansell, H. L. Synthetic Approaches to Anatoxin-A. *Tetrahedron* **1996**, *52* (17), 6025–6061.
- (132) Parsons, P. J.; Camp, N. P.; Edwards, N.; Ravi Sumoreeah, L. Synthesis of (±)-Anatoxin-a and Analogues. *Tetrahedron* **2000**, *56* (2), 309–315.
- (133) Sandy, M.; Zhu, X.; Rui, Z.; Zhang, W. Characterization of AntB, a Promiscuous Acyltransferase Involved in Antimycin Biosynthesis. *Org. Lett.* **2013**, *15* (13), 3396–3399.
- (134) Jenner, M.; Frank, S.; Kampa, A.; Kohlhaas, C.; Pöplau, P.; Briggs, G. S.; Piel, J.; Oldham, N. J. Substrate Specificity in Ketosynthase Domains from Trans-AT Polyketide Synthases. *Angew. Chem. Int. Ed.* **2013**, *52* (4), 1143–1147.
- (135) Vergnolle, O.; Hahn, F.; Baerga-Ortiz, A.; Leadlay, P. F.; Andexer, J. N. Stereoselectivity of Isolated Dehydratase Domains of the Borrelidin Polyketide Synthase: Implications for Cis Double Bond Formation. *ChemBioChem* **2011**, *12* (7), 1011–1014.
- (136) Gupta, S.; Lakshmanan, V.; Kim, B. S.; Fecik, R.; Reynolds, K. A. Generation of Novel Pikromycin Antibiotic Products through Mutasynthesis. *Chembiochem* **2008**, *9* (10), 1609–1616.
- (137) Alhamadsheh, M. M.; Palaniappan, N.; DasChouduri, S.; Reynolds, K. A. Modular Polyketide Synthases and Cis-Double Bond Formation: Establishment of Activated Cis-3-Cyclohexylpropenoic Acid as the Diketide Intermediate in Phoslactomycin Biosynthesis. *J Am Chem Soc* **2007**, *129* (7), 1910–1911.
- (138) Wu, J.; Kinoshita, K.; Khosla, C.; Cane, D. E. Biochemical Analysis of the Substrate Specificity of the Beta-Ketoacyl-Acyl Carrier Protein Synthase Domain of Module 2 of the Erythromycin Polyketide Synthase. *Biochemistry* **2004**, *43* (51), 16301–16310.
- (139) Marc, M.; Outurquin, F.; Renard, P.-Y.; Créminon, C.; Franck, X. Synthesis of a (+)-Anatoxin-a Analogue for Monoclonal Antibodies Production. *Tetrahedron Letters* **2009**, *50* (31), 4554–4557.
- (140) Parsons, P. J.; Camp, N. P.; Mark Underwood, J.; Harvey, D. M. Tandem Reactions of Anions: A Short and Efficient Route to ±anatoxin-A. *Tetrahedron* **1996**, *52* (35), 11637–11642.
- (141) Clive, D. L. J.; Farina, V.; Singh, A.; Wong, C. K.; Kiel, W. A.; Menchen, S. M. Cyclofunctionalization of Olefinic Urethanes Using Benzeneselenenyl Chloride in the

- Presence of Silica Gel: A General Route to Nitrogen Heterocycles. *J. Org. Chem.* **1980**, *45* (11), 2120–2126.
- (142) RajanBabu, T. V. (Babu); Bulman Page, P. C.; Buckley, B. R. Tri-N-Butylstannane. In *Encyclopedia of Reagents for Organic Synthesis*; John Wiley & Sons, Ltd, 2001.
- (143) Reich, H. J.; Wollowitz, S. Preparation of  $\alpha,\beta$ -Unsaturated Carbonyl Compounds and Nitriles by Selenoxide Elimination. In *Organic Reactions*; John Wiley & Sons, Inc., 2004.
- (144) Callant, P.; Ongena, R.; Vandewalle, M. Iridoids : Novel Total Synthesis of ( $\pm$ )-Isoiridomyrmecin and of ( $\pm$ )-Verbenalol. *Tetrahedron* **1981**, *37* (11), 2085–2089.
- (145) Suo, Z.; Chen, H.; Walsh, C. T. Acyl-CoA Hydrolysis by the High Molecular Weight Protein 1 Subunit of Yersiniabactin Synthetase: Mutational Evidence for a Cascade of Four Acyl-Enzyme Intermediates during Hydrolytic Editing. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (26), 14188–14193.
- (146) Ehmann, D. E.; Trauger, J. W.; Stachelhaus, T.; Walsh, C. T. Aminoacyl-SNACs as Small-Molecule Substrates for the Condensation Domains of Nonribosomal Peptide Synthetases. *Chem. Biol.* **2000**, *7* (10), 765–772.
- (147) Kusebauch, B.; Brendel, N.; Kirchner, H.; Dahse, H.-M.; Hertweck, C. Assessing Oxazole Bioisosteres as Mutasythons on the Rhizoxin Assembly Line. *Chembiochem* **2011**, *12* (15), 2284–2288.
- (148) Mann, S.; Lombard, B.; Loew, D.; Méjean, A.; Ploux, O. Insights into the Reaction Mechanism of the Prolyl–Acyl Carrier Protein Oxidase Involved in Anatoxin-a and Homoanatoxin-a Biosynthesis. *Biochemistry* **2011**, *50* (33), 7184–7197.
- (149) Hochuli, E.; Bannwarth, W.; Döbeli, H.; Gentz, R.; Stüber, D. Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Nat Biotech* **1988**, *6* (11), 1321–1325.
- (150) Porath, J. Immobilized Metal Ion Affinity Chromatography. *Protein Expr. Purif.* **1992**, *3* (4), 263–281.
- (151) Hengen, P. N. Purification of His-Tag Fusion Proteins from Escherichia Coli. *Trends in Biochemical Sciences* **1995**, *20* (7), 285–286.
- (152) Rosano, G. L.; Ceccarelli, E. A. Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. *Front. Microbiol.* **2014**, *5*, 172.
- (153) Bell, P. A. E. Coli Expression Systems. *Molecular Biology Problem Solver: A Laboratory Guide* **2001**, 461–490.
- (154) Rosano, G. L.; Bruch, E. M.; Ceccarelli, E. A. Insights into the CLP/HSP100 Chaperone System from Chloroplasts of Arabidopsis Thaliana. *J. Biol. Chem.* **2011**, *286* (34), 29671–29680.
- (155) Shestakov, S. V.; Khyen, N. T. Evidence for Genetic Transformation in Blue-Green Alga Anacystis Nidulans. *Molec. Gen. Genet.* **1970**, *107* (4), 372–375.
- (156) Mermet-Bouvier, P.; Chauvat, F. A Conditional Expression Vector for the Cyanobacteria Synechocystis Sp. Strains PCC6803 and PCC6714 or Synechococcus Sp. Strains PCC7942 and PCC6301. *Curr. Microbiol.* **1994**, *28* (3), 145–148.
- (157) Chen, A. Y.; Cane, D. E.; Khosla, C. Structure-Based Dissociation of a Type I Polyketide Synthase Module. *Chem. Biol.* **2007**, *14* (7), 784–792.
- (158) Burgess, R. R.; Deutscher, M. P. *Guide to Protein Purification*; Academic Press, 2009; Vol. 463.
- (159) *Proteolytic Enzymes: A Practical Approach*, Second Edition.; Beynon, R., Bond, J. S., Eds.; Practical Approach Series 247, 2001.
- (160) Powers, J. C.; Asgian, J. L.; Ekici, Ö. D.; James, K. E. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.* **2002**, *102* (12), 4639–4750.

- (161) Marcinišzyn, J.; Hartsuck, J. A.; Tang, J. Mode of Inhibition of Acid Proteases by Pepstatin. *J. Biol. Chem.* **1976**, *251* (22), 7088–7094.
- (162) Mahdy, A. M.; Webster, N. R. Perioperative Systemic Haemostatic Agents. *Br. J. Anaesth.* **2004**, *93* (6), 842–858.
- (163) Umezawa, H.; Aoyagi, T.; Suda, H.; Hamada, M.; Takeuchi, T. Bestatin, an Inhibitor of Aminopeptidase B, Produced by Actinomycetes. *J. Antibiot.* **1976**, *29* (1), 97–99.
- (164) Muskardin, D. T.; Voelkel, N. F.; Fitzpatrick, F. A. Modulation of Pulmonary Leukotriene Formation and Perfusion Pressure by Bestatin, an Inhibitor of Leukotriene A4 Hydrolase. *Biochemical Pharmacology* **1994**, *48* (1), 131–137.
- (165) Sekine, K.; Fujii, H.; Abe, F. Induction of Apoptosis by Bestatin (ubenimex) in Human Leukemic Cell Lines. *Leukemia* **1999**, *13* (5), 729–734.
- (166) Auld, D. S. Removal and Replacement of Metal Ions in Metalloproteinases. *Meth. Enzymol.* **1995**, *248*, 228–242.
- (167) Bolanos-Garcia, V. M.; Davies, O. R. Structural Analysis and Classification of Native Proteins from E. Coli Commonly Co-Purified by Immobilised Metal Affinity Chromatography. *Biochim. Biophys. Acta* **2006**, *1760* (9), 1304–1313.
- (168) Wonnacott, S.; Jackman, S.; Swanson, K. L.; Rapoport, H.; Albuquerque, E. X. Nicotinic Pharmacology of Anatoxin Analogs. II. Side Chain Structure-Activity Relationships at Neuronal Nicotinic Ligand Binding Sites. *J. Pharmacol. Exp. Ther.* **1991**, *259* (1), 387–391.
- (169) Smith, R. A.; Lewis, D. A Rapid Analysis of Water for Anatoxin A, the Unstable Toxic Alkaloid from Anabaena Flos-Aquae, the Stable Non-Toxic Alkaloids Left after Bioreduction and a Related Amine Which May Be Nature's Precursor to Anatoxin a. *Vet Hum Toxicol* **1987**, *29* (2), 153–154.
- (170) Afzal, A.; Oppenländer, T.; Bolton, J. R.; El-Din, M. G. Anatoxin-a Degradation by Advanced Oxidation Processes: Vacuum-UV at 172 Nm, Photolysis Using Medium Pressure UV and UV/H<sub>2</sub>O<sub>2</sub>. *Water Research* **2010**, *44* (1), 278–286.
- (171) Stevens, D. K.; Krieger, R. I. Stability Studies on the Cyanobacterial Nicotinic Alkaloid Snatoxin-A. *Toxicon* **1991**, *29* (2), 167–179.
- (172) Kaminski, A.; Bober, B.; Lechowski, Z.; Bialczyk, J. Determination of Anatoxin-a Stability under Certain Abiotic Factors. *Harmful Algae* **2013**, *28*, 83–87.
- (173) Harland, F.; Wood, S.; Moltchanova, E.; Williamson, W.; Gaw, S. Phormidium Autumnale Growth and Anatoxin-a Production under Iron and Copper Stress. *Toxins* **2013**, *5* (12), 2504–2521.
- (174) Wood, S. a.; Heath, M. w.; Kuhajek, J.; Ryan, K. g. Fine-Scale Spatial Variability in Anatoxin-a and Homoanatoxin-a Concentrations in Benthic Cyanobacterial Mats: Implication for Monitoring and Management. *Journal of Applied Microbiology* **2010**, *109* (6), 2011–2018.
- (175) Swanson, K. L.; Aronstam, R. S.; Wonnacott, S.; Rapoport, H.; Albuquerque, E. X. Nicotinic Pharmacology of Anatoxin Analogs. I. Side Chain Structure-Activity Relationships at Peripheral Agonist and Noncompetitive Antagonist Sites. *J. Pharmacol. Exp. Ther.* **1991**, *259* (1), 377–386.
- (176) Furey, A.; Crowley, J.; Lehane, M.; James, K. J. Liquid Chromatography with Electrospray Ion-Trap Mass Spectrometry for the Determination of Anatoxins in Cyanobacteria and Drinking Water. *Rapid Commun. Mass Spectrom.* **2003**, *17* (6), 583–588.
- (177) Bashiri, G.; Rehan, A. M.; Greenwood, D. R.; Dickson, J. M. J.; Baker, E. N. Metabolic Engineering of Cofactor F420 Production in Mycobacterium Smegmatis. *PLoS ONE* **2010**, *5* (12), e15803.

- (178) Decamps, L.; Philmus, B.; Benjdia, A.; White, R.; Begley, T. P.; Berteau, O. Biosynthesis of F<sub>0</sub>, Precursor of the F<sub>420</sub> Cofactor, Requires a Unique Two Radical-SAM Domain Enzyme and Tyrosine as Substrate. *Journal of the American Chemical Society* **2012**, *134* (44), 18173–18176.
- (179) Isabelle, D.; Simpson, D. R.; Daniels, L. Large-Scale Production of Coenzyme F<sub>420</sub>-5,6 by Using Mycobacterium Smegmatis. *Appl. Environ. Microbiol.* **2002**, *68* (11), 5750–5755.
- (180) Purwantini, E.; Daniels, L. Purification of a Novel Coenzyme F<sub>420</sub>-Dependent Glucose-6-Phosphate Dehydrogenase from Mycobacterium Smegmatis. *J. Bacteriol.* **1996**, *178* (10), 2861–2866.
- (181) Jacobson, F. S.; Daniels, L.; Fox, J. A.; Walsh, C. T.; Orme-Johnson, W. H. Purification and Properties of an 8-Hydroxy-5-Deazaflavin-Reducing Hydrogenase from Methanobacterium Thermoautotrophicum. *J. Biol. Chem.* **1982**, *257* (7), 3385–3388.
- (182) Rippka, R. [1] Isolation and Purification of Cyanobacteria. In *Methods in Enzymology*; Lester Packer, A. N. G., Ed.; Cyanobacteria; Academic Press, 1988; Vol. 167, pp 3–27.
- (183) Yanisch-Perron, C.; Vieira, J.; Messing, J. Improved M13 Phage Cloning Vectors and Host Strains: Nucleotide Sequences of the M13mp18 and pUC19 Vectors. *Gene* **1985**, *33* (1), 103–119.
- (184) Studier, F. W.; Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *Journal of Molecular Biology* **1986**, *189* (1), 113–130.
- (185) Studier, F. W.; Rosenberg, A. H.; Dunn, J. J.; Dubendorff, J. W. Use of T7 RNA Polymerase to Direct Expression of Cloned Genes. *Meth. Enzymol.* **1990**, *185*, 60–89.
- (186) Bertani, G. Studies on Lysogenesis. I. The Mode of Phage Liberation by Lysogenic Escherichia Coli. *J. Bacteriol.* **1951**, *62* (3), 293–300.
- (187) Bertani, G. Lysogeny at Mid-Twentieth Century: P1, P2, and Other Experimental Systems. *J. Bacteriol.* **2004**, *186* (3), 595–600.
- (188) Hanahan, D. Studies on Transformation of Escherichia Coli with Plasmids. *Journal of Molecular Biology* **1983**, *166* (4), 557–580.
- (189) Sanger, F.; Nicklen, S.; Coulson, A. R. DNA Sequencing with Chain-Terminating Inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74* (12), 5463–5467.
- (190) Sanger, F.; Coulson, A. R. A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase. *Journal of Molecular Biology* **1975**, *94* (3), 441–448.
- (191) Mierendorf, R.; Yeager, K.; Novy, R. The pET System: Your Choice for Expression. *inNovations* **1994**, *1*, 1–12.
- (192) Raymond, S.; Weintraub, L. Acrylamide Gel as a Supporting Medium for Zone Electrophoresis. *Science* **1959**, *130* (3377), 711–711.
- (193) Davis, B. J.; Ornstein, L. A New High Resolution Electrophoresis Method. *Delivered at The Society for the Study of Blood at the New York Academy of Medicine* **1959**.
- (194) Schägger, H. Tricine–SDS–PAGE. *Nature Protocols* **2006**, *1* (1), 16–22.
- (195) Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227* (5259), 680–685.
- (196) Sykaluk, L. K.; Lipton, W. E.; Savage, M. D.; Hermanson, G. T.; KingSpengler, T. Nickel Chelated to a Novel Tridentate Chelating Group Binds Histidine-Tagged Fusion Proteins: Applications of a Nickel-Activated Horseradish Peroxidase Probe and a Nickel-Activated Microtiter Plate for Assays. *Faseb J.* **1997**, *11* (3), 3204–3204.

- (197) Chu, R.; et, al. SuperSignal® HisProbe™ Western Blotting Kit: Rapid Detection of Polyhistidine-Tagged Fusion Protein with Enhanced Specificity, Versatility, and High Sensitivity. *Previews* **1997**, *1* (3).
- (198) Howard, M. H.; Sardina, F. J.; Rapoport, H. Chirospecific Syntheses of Nitrogen and Side-Chain-Modified Anatoxin Analogs. Synthesis of (1R)-Anatoxinal and (1R)-Anatoxinic Acid Derivatives. *The Journal of Organic Chemistry* **1990**, *55* (9), 2829–2838.