

UNIVERSITE DE LILLE I
SCIENCES ET TECHNOLOGIES

THESE

Présentée par **AURELIEN VANVLASSEN BROECK**

Pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITE DE LILLE 1

Ecole doctorale science de la matière, du rayonnement et de l'environnement

Filière : Ingénierie des fonctions biologiques

Etude expérimentale et *in silico* du potentiel de synthèse
NRPS chez les *Pseudomonas* fluorescents

N° Ordre 40848

Préparée au Laboratoire ProBioGEM, UPRES EA 1026

Polytech'Lille, France

Soutenance le 17 juillet 2012 devant le jury composé de :

Philippe Jacques , Professeur (ProBioGEM, Université de Lille I)	Directeur
Marie-Dominique Devignes , Chargée de recherche CNRS (LORIA, Université de Nancy)	Rapporteur
Herbert Budzikiewicz , Professeur émérite (Université de Cologne)	Rapporteur
Sandra Matthijs , Chargée de recherche (IRMW, Université libre de Bruxelles)	Examinatrice
Bernard Wathelet , Professeur (UCBI, Faculté Univ. de Gembloux)	Examineur
Marc Ongena , Chercheur qualifié FNRS (AgroBioTech, Faculté Univ. de Gembloux)	Examineur
Valérie Leclère , Maître de conférences-HDR (ProBioGEM, Université de Lille I)	Examinatrice
Maude Pupin , Maître de conférences (LIFL-INRIA, Université de Lille I)	Examinatrice

Remerciements

Je tiens tout d'abord à remercier Messieurs Pascal Dhulster et Philippe Jacques, directeur et co-directeur du laboratoire ProBioGEM pour m'avoir permis de réaliser ce travail au sein de leur laboratoire. Ainsi que Monsieur Didier Guillochon, directeur du laboratoire ProBioGEM au cours de mon master 2 recherche.

Je remercie plus particulièrement Monsieur Philippe Jacques, qui a dirigé cette thèse, pour son expérience scientifique, son esprit critique, sa disponibilité, son enthousiasme et sa patience. Je remercie Madame Valérie Leclère pour ses conseils, son enthousiasme, ses connaissances en analyse bioinformatique et sa disponibilité. Je tiens à remercier également Madame Maude Pupin pour ses conseils en bioinformatique, sa disponibilité et pour la partie bioinformatique de ce travail. D'une manière générale, je remercie tous mes encadrants qui m'ont permis de mener à bien ce travail.

Je remercie également madame Marie-Dominique Devignes et Monsieur Herbert Budzikiewicz pour avoir accepté d'être rapporteur de ce travail ainsi que Madame Sandra Matthijs, et Messieurs Marc Ongena et Bernard Wathelet pour avoir accepté d'être examinateurs.

Je remercie particulièrement Monsieur Bernard Wathelet pour l'accueil qu'il m'a réservé au sein de son laboratoire de la faculté universitaire de Gembloux pour l'utilisation des installations de spectrométrie de masse, pour sa gentillesse et le temps qu'il a consacré à interpréter les résultats obtenus en ma présence ainsi que les analyses qu'il a réalisées lui-même.

Je tiens à remercier l'ensemble des membres du laboratoire pour leur aide et leurs conseils durant ce travail et plus particulièrement à Monsieur Max Béchet pour sa disponibilité et son aide au laboratoire. Un grand merci aux jeunes docteurs ou doctorants et étudiants que j'ai cotoyé pendant ces années, pour les discussions et les bons moments passés avec eux. Je remercie également l'ensemble de ma famille.

Je dédie cette thèse à mes grands parents partis trop vite ...

Sommaire

Liste des figures	7
Liste des tableaux	11
Liste des abréviations.....	13
Introduction générale.....	15
Introduction bibliographique.....	19
I. Le mécanisme NRPS.....	20
I.1. Généralités.....	20
I.2. Les domaines des synthétases NRPS.....	20
I.2.1. Les domaines principaux.....	21
I.2.2. Les domaines secondaires.....	25
I.3. Organisation modulaire des synthétases peptidiques non-ribosomiales.....	30
II. La biosynthèse NRPS.....	30
II.1. La biosynthèse linéaire.....	30
II.2. La biosynthèse itérative : exemple de la synthèse de l'entérobactine.....	31
II.3. La biosynthèse non linéaire : exemple de la synthèse de la vibriobactine.....	32
III. Le mécanisme PKS	32
IV. Les systèmes hybrides NRPS/PKS.....	34
V. Structures et activités biologiques des peptides non ribosomiaux.....	35
V.1. Les structures des NRPs.....	35
V.2. Les activités biologiques des NRPs.....	37
VI. Les NRPs produits par <i>Pseudomonas</i>	41
VI.1. Les sidérophores.....	41
VI.1.1. Structure générale.....	46
VI.1.1.1. La chaîne peptidique.....	46
VI.1.1.2. Le chromophore.....	47
VI.1.1.3. La chaîne latérale du chromophore.....	48
VI.1.2. Biosynthèse des pyoverdines.....	48
VI.1.2.1. Biosynthèse de la chaîne peptidique.....	49
VI.1.2.2. Biosynthèse du chromophore.....	50
VI.1.3. Régulation de la synthèse des pyoverdines.....	50
VI.2. Les lipopeptides cycliques de <i>Pseudomonas</i>	51
VII. Outils <i>in silico</i> d'analyse des NRPs.....	57
 II. Matériel et Méthodes.....	61
II.1. Origine des produits.....	62

II.2. Milieux de culture.....	62
II.3. Les souches bactériennes utilisées.....	63
II.4. Méthodes d'analyse.....	65
II.4.1. Méthodes d'analyse biologique.....	65
II.4.1.1. Concentration des échantillons.....	65
II.4.1.2. Spectrométrie de masse MALDI-TOF.....	65
II.4.2. Méthodes d'analyse bioinformatique.....	67
II.4.2.1. Protocole d'analyse des génomes de <i>Pseudomonas</i>	67
II.4.2.2. Outils bioinformatiques d'analyse des synthétases et de prédiction de la composition du peptide final.....	68
II.4.2.3. Banques de données.....	70
II.4.2.4. Outils de comparaison de séquences.....	71
II.4.2.5. Outil des résultats de spectrométrie de masse	72
Résultats.....	73
III. Analyse des peptides d'origine non-ribosomiale produits par les <i>Pseudomonas</i> fluorescents.....	74
III.1. Les pyoverdines.....	74
III.2. Les lipopeptides cycliques.....	77
III.3. Comparaison de la composition des CLPs et des pyoverdines.....	79
IV. Analyse des synthétases de pyoverdines.....	81
IV.1. Introduction.....	81
IV.2. Présentation de l'outil bio-informatique Pyomass.....	82
IV.3. Caractérisation des gènes qui codent pour des synthétases de pyoverdines connues.....	83
IV.4. Détermination de nouvelles signatures du code NRPS spécifiques aux pyoverdine.....	92
IV.5. Etude des synthétases de pyoverdines et de leurs produits pour les autres souches de <i>Pseudomonas</i>	96
V. Amplification de la biodiversité au sein des pyoverdines par feeding.....	108
V.1. Stratégie d'analyse.....	108
V.2. Application de pyomass à l'étude des résultats de masse du feeding.....	109
V.3. Analyse des résultats obtenus par spectrométrie de masse MALDI-TOF...	111
V. 3.1. <i>Ps. aeruginosa</i> PA7.....	111
V. 3.2. <i>Ps. syringae</i> B728a	113
V. 3.3. <i>Ps. entomophila</i> L48.....	117
V.4. Analyse MS2 du cas de <i>Pseudomonas entomophila</i> L48 cultivée en présence de sérine dans le milieu de culture.....	119

VI. Analyse des autres NRPS de <i>Pseudomonas</i>	122
VI. 1. Sidérophore produit par une synthétase hybride NRPS/PKS.....	122
VI. 2. Synthétases de lipopeptides.....	123
VI.2.1. Synthétase de la syringafactine A.....	123
VI.2.2. Synthétase de la syringomycine.....	125
VI.2.3. Synthétase de la syringopeptine.....	126
VI.2.4. Synthétase hybride NRPS/PKS.....	130
VI.2.5. Synthétase de lipopeptide chez <i>Pseudomonas fluorescens</i> Pf5.....	131
VI.3. Analyse de peptides inconnus.....	132
VII. Discussion.....	138
VII. 1. Bilan des synthétases de <i>Pseudomonas</i> fluorescents et de leurs produits	138
VII. 2. Caractérisation de mécanismes de biosynthèse particuliers.....	139
VII. 3. Développement des outils bioinformatiques.....	140
VII. 4. Perspectives	141
Références bibliographiques.....	142
Résumé.....	155

Liste des figures :

Figure 1: Réaction d'activation du monomère sélectionné par le domaine A. (Stachelhaus *et al.*, 2002).

Figure 2: Représentation schématique des sites de fixation du substrat spécifique à la sélection de l'aspartate (a), de l'ornithine (b) et de la valine (c). (Stachelhaus *et al.*, 1999).

Figure 3: Fixation covalente de l'aminoacyl adénylate sur la synthétase par le biais d'un bras phosphopantéthénique. (Stachelhaus *et al.*, 2002)

Figure 4: Formation d'une liaison peptidique entre deux acides aminés liés par une liaison thioester aux domaines T1 et T2 (Stachelhaus *et al.*, 2002).

Figure 5: Schéma de la libération de la daptomycine de la synthétase et de sa cyclisation sous l'action du domaine TE (Grünewald et Marahiel, 2006).

Figure 6: Localisation des domaines secondaires au sein d'un module (Stachelhaus *et al.*, 2002).

Figure 7: Schéma de la réaction d'épimérisation de l'acide aminé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

Figure 8: Schéma de la réaction de cyclisation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

Figure 9: Réaction de réduction de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

Figure 10: Réaction d'oxydation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

Figure 11: Réaction de méthylation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

Figure 12: Réaction de formylation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

Figure 13: Biosynthèse linéaire de l'ACV (Mootz *et al.*, 2002).

Figure 14: Biosynthèse itérative de l'entérobactine (Mootz *et al.*, 2002).

Figure 15: Biosynthèse non linéaire de la vibriobactine (Mootz *et al.*, 2002).

Figure 16: Structure de la yersiniabactine synthétase selon Miller *et al.*, (2002).

Figure 17 : Représentation schématique de la répartition des NRPs recensés dans Norine (1122 peptides) en fonction de la structure des peptides.

Figure 18: Exemple de glycopeptide : la vancomycine.

Figure 19 : Représentation schématique de la répartition des activités biologiques des 1122 NRPs recensés dans Norine. A noter qu'un même NRP peut posséder deux activités biologiques différentes.

Figure 20: Structure de différents NRPs classés en fonction de leur activité biologique (Schwarzer *et al.*, 2003).

Figure 21: Structure de la pyoverdine produite par *Pseudomonas aeruginosa* PAO1 (Meyer *et al.*, 1997).

Figure 22: Représentation schématique du chromophore de pyoverdine avec localisation des atomes de carbones sur lesquels se fixent les chaines peptidiques.

Figure 23: Structures des différents chromophores de sidérophores (Budzikiewicz *et al.*, 2006).

Figure 24 : Organisation des clusters de gènes de la pyoverdine synthétase chez différentes espèces de *Pseudomonas* (Ravel et Cornelis, 2003).

Figure 25: Représentation de l'opéron de gènes de synthèse de l'orfamide A et de la synthétase codée par ce cluster (Loper et Gross, 2007).

Figure 26: Représentation du cluster de gènes de la syringafactine A, de la structure de la synthétase et de la composition de la syringafactine A (NOR01075) (Berti *et al.*, 2007).

Figure 27: Stratégie générale d'exploration des génomes.

Figure 28: Exemple de résultat obtenu par Weblogo à partir d'un alignement multiple de séquences protéiques.

Figure 29: Répartition des molécules de pyoverdines en fonction de la longueur de la chaîne peptidique. Etude réalisée sur 82 molécules de pyoverdines.

Figure 30: Proportions des monomères retrouvés dans les 61 molécules de pyoverdines dont l'isomérisation des monomères est connue.

Figure 31: Répartition des molécules de lipopeptides en fonction de la longueur de la chaîne peptidique.

Figure 32: Pourcentage des monomères retrouvés dans les molécules de lipopeptides cycliques.

Figure 33: Capture d'écran du logiciel « Pyomass ».

Figure 34: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas putida* KT2440.

Figure 35: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas entomophila* L48.

Figure 36: Arbre de distance issu d'un alignement de séquences de domaines A, de synthétases de pyoverdine, (environ 500 aa) incorporant des acides aminés connus .

Figure 37: Spectre MS2 du pic correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48.

Figure 38: Représentation sous la forme d'un arbre de distances de l'alignement des séquences des domaines A des synthétases de pyoverdines dont la séquence est connue.

Figure 39: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas syringae*.

Figure 40: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas putida* W619.

Figure 41: Représentation sous la forme d'un arbre de distance de l'alignement des séquences des domaines A des synthétases de pyoverdines dont la séquence est connue avec les domaines A dont l'analyse bioinformatique n'a pas permis la détermination du substrat (Xn).

Figure 42: Capture d'écran du logiciel Pyomass présentant l'interface de calcul qui permet de calculer la masse d'une pyoverdine avec des substitutions dans la chaîne peptidique.

Figure 43: Capture d'écran du tableau de résultat de Pyomass qui confronte dans cet exemple les masses de la molécule entrée comme requête avec la même molécule à laquelle on a substitué le premier acide aminé (D-Ser → Thr).

Figure 44: Spectre de masse MALDI-TOF en mode réflectron d'une culture de *Pseudomonas aeruginosa* PA7 produisant une molécule de pyoverdine témoin (A), en présence de sérine dans le milieu de culture (B).

Figure 45: Fenêtre de résultat du logiciel pyomass.

Figure 46: Spectre de masse MALDI-TOF en mode réflectron d'une culture de *Pseudomonas syringae* B728a produisant une molécule de pyoverdine témoin (A), en présence de thréonine dans le milieu de culture (B) et en présence de sérine (C).

Figure 47: Fenêtre de résultat du logiciel pyomass. Le tableau 1 reprend le calcul des masses moléculaires de la pyoverdine de *P. syringae* B728a en fonction de la chaîne latérale portée par le chromophore. Le tableau 2 reprend le même calcul mais avec la substitution de Lys en position 1 par une thréonine. Et les tableaux 3 et 4 reprennent le calcul avec la substitution de la thréonine en position 3 et 4 respectivement par une sérine.

Figure 48: Spectre de masse MALDI-TOF en mode réflectron d'une culture de *Pseudomonas entomophila* L48 produisant une molécule de pyoverdine témoin (A), en présence de sérine dans le milieu de culture (B).

Figure 49: Fenêtre de résultat du logiciel pyomass. Le tableau supérieur reprend le calcul des masses moléculaires de la pyoverdine de *P. entomophila* L48 en fonction de la chaîne latérale portée par le chromophore et le tableau inférieur reprend le même calcul mais avec la substitution de aThr en position 8 par une sérine.

Figure 50: Spectre MS2 du pic correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 cultivée en présence de sérine.

Figure 51: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas syringae* phaseolicola 1448A et *syringae* DC3000.

Figure 52: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par le cluster de la syringafactine A chez *Pseudomonas syringae* pv. *Syringae* B728a.

Figure 53: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par le cluster de la syringomycine synthétase.

Figure 54: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par le cluster de la syringopeptine.

Figure 55: Analyse phylogénétique des domaines C de la syringopeptine A (sypA_Cx) et de la syringomycine E (symE_Cx).

Figure 56: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par cette synthétase hybride NRPS/PKS.

Figure 57: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par *Pseudomonas fluorescens* Pf5.

Liste des tableaux :

Tableau 1 : Critères pour le choix du genre bactérien modèle quatre grands genres bactériens connus pour leur grande capacité de production de peptides d'origine non-ribosomique (Cyanobactéries, *Streptomyces*, *Pseudomonas*, *Bacillus*).

Tableau 2 : Liste des motifs conservés dans les séquences des différents domaines de la synthétase NRPS (Schwarzer *et al.*, 2003).

Tableau 3: Liste des 82 molécules de pyoverdine dont la séquence peptidique est connue

Tableau 4: Liste des lipopeptides produits par les bactéries du genre *Pseudomonas*.

Tableau 5: Tableau des souches de *Pseudomonas* fluorescents utilisées dans cette étude.

Tableau 6: Composition en peptides de la solution de calibration du spectromètre de masse.

Tableau 7: Répartition des molécules de pyoverdines connues en fonction des cinq classes structurales définies.

Tableau 8: Liste des génomes de *Pseudomonas* entièrement séquencés issus de la banque de données MBGD.

Tableau 9: Liste des sept souches de *Pseudomonas* dont la structure de la pyoverdine est connue.

Tableau 10: Masses calculées par Pyomass et masses expérimentales déterminées par spectrométrie de masse MALDI-TOF pour les pyoverdines connues.

Tableau 11: Séquences des pyoverdines produites par les bactéries du groupe des *Pseudomonas* fluorescents (séquences décrites dans la littérature au-dessus et séquences issues de l'analyse bioinformatique des différents génomes en-dessous, souche par souche).

Tableau 12: Résultats de l'étude LC-MS-MS du spectre correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 .

Tableau 13: Liste des signatures extraites à partir des domaines A des sept souches dont la structure de la pyoverdine est connue.

Tableau 14: Séquences des pyoverdines produites par les bactéries du groupe des *Pseudomonas* fluorescents (séquences prédites par les logiciels d'analyse des NRPS au-dessus et séquences issues de l'analyse des signatures en-dessous, souche par souche).

Tableau 15: Références des monomères notés 'Xn' qui vont être utilisés dans l'analyse phylogénétique des domaines A.

Tableau 16: Identification des monomères prédits 'X' par les outils d'analyse des synthétases disponibles par étude de la séquence du domaine A entier en comparant avec des séquences de domaines A dont le substrat est déjà déterminé.

Tableau 17: Liste des séquences prédites pour les pyoverdines à partir de l'étude des signatures des domaines A en regard de la séquence la plus proche issue de la littérature.

Tableau 18: Conditions de feeding testées en fonction des souches de *Pseudomonas* disponibles.

Tableau 19: Résultats de l'étude LC-MS-MS du spectre correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 en présence de sérine dans le milieu de culture.

Tableau 20: Tableau d'identification des gènes YbtE et YbtU chez *Pseudomonas syringae* DC3000 et *phaseolicola* 1448A par TblastN à partir de la séquence peptidique correspondant aux gènes YbtE et YbtU de *Yersinia pestis*.

Tableau 21: Alignement de la structure prédite du peptide final (1) avec la structure de la syringopeptine décrite par Scholz-Schroeder (2). Le DhAbu est un dérivé de la thréonine.

Tableau 22: Analyse globale du potentiel de synthèse NRPS des génomes du genre *Pseudomonas*.

Liste des abréviations

A: adénylation

Å: Angström

ACP: acyl carrier protein

AT: acétyltransférase

ATP: Adénosine-5'-triphosphate

BLAST: Basic Local Alignment Search Tool

C: condensation

CHCA: α -cyano-4-hydroxy-cinnamic acid

Chr: chromophore

CLPs: cyclic lipopeptides

Cy: cyclisation

Da: Dalton

DH: déhydratase

domaine C/E: domaine condensation/épimérisation

E: épimérisation

ER: ényltransférase

FMN: flavin mononucleotide

KR: kétéductase

KS: kétoynthase

Kv: kilovolt

LC-MS-MS: Liquid Chromatography – Mass Spectrometry – Mass Spectrometry

M: méthylation

MALDI-TOF: Matrix Assisted Laser Desorption Ionization – Time Of Flight

NADPH: nicotinamide adenine dinucleotide phosphate-oxidase

NRPS: Non Ribosomal Peptide Synthetase

NRPS/PKS: Non Ribosomal Peptide Synthetase/Polyketide Synthase

NRPs: Non Ribosomal Peptides

Ox: oxydation

PCP: peptidyl carrier protein

PKS: PolyKetide Synthase

Ppant: 4'-phosphopantéthéine

Re: réduction

RMN: resonance magnétique nucléaire

SAM: S-adénosyle méthionine

T: thiolation

Te: thioestérase

THF: tétrahydrofolate

v/v: volume/volume

TE: thioestérase de type II

Introduction générale

Certaines bactéries et champignons produisent des peptides nécessaires à leur développement et à leur survie, par un mécanisme de synthèse différent de la synthèse ribosomiale. En effet, cette voie de synthèse est indépendante de l'intervention des ribosomes, elle utilise des complexes multi-enzymatiques qui permettent une synthèse peptidique dite non-ribosomiale. Ces enzymes sont appelées en anglais « Non Ribosomal Peptide Synthetase » (NRPS). Cette voie de synthèse peptidique originale est caractérisée par la diversité des activités biologiques de ces produits. La synthèse peptidique non-ribosomiale est caractéristique des micro-organismes, ce type de synthèse n'a pas encore été démontré chez les organismes supérieurs.

Ce travail s'inscrit dans un cadre plus large de développement d'outils bioinformatiques destinés à mieux maîtriser la synthèse non-ribosomiale et à mettre en évidence de nouveaux peptides actifs synthétisés par ce mécanisme. Cette recherche est le fruit d'une collaboration étroite entre le laboratoire ProBioGEM et le Laboratoire d'Informatique Fondamentale de Lille (LIFL). Dans le cadre de cette collaboration, les chercheurs des deux laboratoires ont mis en place une banque de données (NORINE), unique au monde, qui rassemble la majorité des peptides synthétisés par la voie non-ribosomiale (NonRibosomal Peptides ou NRPs). Les chercheurs ont également développés une banque de données de synthétases (DORIS) et différents outils destinés à prédire la structure et les propriétés des NRPs en partant de données de séquences génomiques de micro-organismes. Dans ce travail, nous avons souhaité développer et exploiter ces outils bioinformatiques pour l'analyse du potentiel de synthèse de NRPs par un genre bactérien donné. Différents genres bactériens tels que *Bacillus*, *Streptomyces* ou *Pseudomonas* ainsi que les cyanobactéries sont connus pour leur capacité à produire une grande diversité de NRPs, possédant des activités biologiques variées. Ils présentent donc un potentiel d'étude intéressant.

Le choix du modèle s'est basé sur 4 grands critères qui sont résumés dans le tableau 1 pour les 4 grands genres/groupes bactériens définis ci-dessus : le nombre de NRPs produits, le nombre de familles différentes de peptides non ribosomiaux, le nombre de génomes séquencés et les compétences du laboratoire. Le genre *Bacillus* est largement étudié au laboratoire notamment les lipopeptides qu'il produit, néanmoins ce genre bactérien présente une diversité limitée dans ces produits NRPS ce qui ne correspond pas à nos attentes pour le choix d'un bon modèle d'étude. Ensuite, les cyanobactéries nécessitent des conditions de

culture particulières et nous ne sommes pas spécialisés dans ce genre de cultures, il y a néanmoins 40 génomes de cyanobactéries complètement séquencés. Le genre bactérien *Streptomyces* est un fort producteur de peptides d'origine NRPS, ayant tous des structures spécifiques non comparables; il y a en fait peu de famille de peptides. Les compétences du laboratoire sur ces souches restent toutefois limitées. Le choix du modèle s'est donc porté vers le genre *Pseudomonas* et plus particulièrement vers les *Pseudomonas* fluorescents qui sont caractérisés par la production d'un sidérophore fluorescent (la pyoverdine) et de lipopeptides cycliques (CLPs) ayant une activité surfactante qui améliore la miscibilité entre les phases aqueuses et organiques. Les sidérophores et les lipopeptides cycliques sont les deux molécules NRPS principalement produites par les *Pseudomonas*. De plus, 22 génomes de *Pseudomonas* sont complètement séquencés et accessibles en 2012. Les pyoverdines présentent l'avantage de constituer une famille de 80 molécules présentant donc une grande diversité autour d'une structure de base similaire. Les lipopeptides de *Pseudomonas* sont plutôt répartis en différentes familles dont le nombre et la biodiversité au sein de chaque famille est en constante évolution. De plus des travaux réalisés au laboratoire sur la souche *Pseudomonas aeruginosa* 7NSK2 ont pu mettre en évidence la production de variants de la molécule de pyoverdine produite par cette souche en fonction des conditions de culture.

	Nb de NRPs	Nb de famille	Nb génomes séquencés	Maîtrise au laboratoire
Cyanobactéries	106	16	40	-
Streptomyces	95	25	8	+
Pseudomonas	135	14	22	++
Bacillus	127	15	39	+++

Tableau 1 : Critères pour le choix du genre bactérien modèle parmi quatre grands genres/groupes bactériens connus pour leur grande capacité de production de peptides d'origine non-ribosomique (Cyanobactéries, *Streptomyces*, *Pseudomonas*, *Bacillus*).

Le travail entrepris sur le potentiel de synthèse NRPS chez les *Pseudomonas* a pour objectif d'acquérir une meilleure connaissance de ce mécanisme de synthèse ainsi que d'améliorer les outils de prédictions existants et de construire des outils spécifiques à l'analyse des NRPs produits par les souches de ce genre bactérien. Pour se faire, le travail va être réalisé en plusieurs étapes :

Dans un premier temps, le travail consistera à caractériser le potentiel de synthèse NRPS chez les *Pseudomonas* par un relevé des molécules décrites dans la littérature. L'étude

des NRPs chez *Pseudomonas* sera entreprise sur les molécules de pyoverdines et les lipopeptides cycliques. Cette étude portera sur la taille des molécules, leur composition et une structure modèle des pyoverdines sera proposée.

Les vingt génomes de *Pseudomonas* séquencés et disponibles sur le site du NCBI en 2012 seront ensuite analysés à la recherche d'un potentiel de synthèse NRPS grâce aux outils bioinformatiques disponibles. La composition en modules et domaines des synthétases identifiées par cette analyse sera prédite, ainsi que la séquence du produit de la synthétase. Le travail sera dans un premier temps consacré aux pyoverdines. La masse moléculaire des pyoverdines sera définie par des analyses par spectrométrie de masse sur des cellules en culture sur milieu solide. Ces informations sur la séquence et la masse moléculaire des peptides seront comparées aux séquences et masses moléculaires des pyoverdines décrites dans la littérature. La prédiction de la séquence des pyoverdines produites sera améliorée par l'étude des signatures des 10 acides aminés de Stachelhaus présents au site actif des domaines d'adénylation (domaine A) des synthétases. Ces signatures permettent de définir l'acide aminé reconnu et activé par ces domaines d'adénylation. De nouvelles signatures spécifiques des synthétases de pyoverdines pourront être ainsi définies. Elles permettront d'améliorer les logiciels existants et notamment de différencier les différents dérivés d'un acide aminé présents dans la chaîne peptidique des pyoverdines tel que l'ornithine retrouvée sous différentes formes dans la composition de ces dernières (Orn, cOHOrn, AcOHOrn, D-FoHOrn, D-AcOHOrn, OHButOrn, D-OHOrn et OHOrn).

Par ailleurs, le caractère spécifique de domaines A de synthétases de pyoverdines NRPS sera étudié par des expériences de « feeding ». Ces expériences consistent à apporter un acide aminé en excès dans le milieu de culture du micro-organisme afin de favoriser l'incorporation de cet acide aminé dans la synthèse d'un NRP et notamment d'une pyoverdine. Ces expériences ont pour objectif de caractériser des domaines A, qui ont une tolérance quant à l'incorporation d'un acide aminé particulier, et qui sont alors qualifiés de permissifs. Cette partie du travail sera associée à une étude en spectrométrie de masse MALDI-TOF des molécules de pyoverdines de différentes souches de *Pseudomonas* afin de déterminer des modifications dans la composition de la chaîne peptidique des pyoverdines et donc la présence de domaines A permissifs dans la synthétase de la molécule produite. Des domaines dits permissifs ont déjà été mis en évidence chez certains lipopeptides du genre *Bacillus* par différents auteurs, mais restent des cas particuliers qui permettent d'augmenter la biodiversité structurale des NRPs.

Pour terminer, le potentiel d'autres NRPs que les pyoverdines sera étudié. Ceci permettra de mettre en évidence d'éventuelles nouvelles molécules potentiellement produites par ces souches.

Introduction bibliographique

I. Le mécanisme NRPS

I.1. Généralités

La biosynthèse peptidique non ribosomiale est réalisée, chez les bactéries et les champignons, par de grands complexes multi-enzymatiques issus de la voie classique de synthèse et appelé en anglais « Non Ribosomal Peptide Synthetases » (NRPS). Cette voie de biosynthèse a été mise en évidence pour la première fois par Lipmann *et al.* en 1971. Ces auteurs ont démontré que la synthèse d'un polypeptide antibiotique, la gramicidine S, par *Bacillus brevis* était toujours possible après traitement des extraits cellulaires à la ribonucléase. Ils en ont déduit que la synthèse de cet antibiotique était réalisée à partir d'une ou plusieurs protéines déjà existantes sans l'intervention des ribosomes. Les NRPS sont codées de manière générale par plusieurs gènes structurés en opéron ou clusters au niveau du génome. Ces synthétases sont organisées en modules eux même divisés en domaines portant chacun une activité enzymatique particulière. La taille des protéines peut varier entre un et plus d'une vingtaine de modules. Les synthétases de certains lipopeptides cycliques chez des *Pseudomonas* ont été décrites comme comportant jusqu'à 25 modules, c'est le cas de la syringopeptine 25A (Raaijmakers *et al.*, 2006). Le produit des NRPS est notamment caractérisé par la présence d'acides aminés non protéogéniques et/ou modifiés. Les peptides non ribosomiaux peuvent contenir des acides aminés sous forme D ou formylés par exemple. Dans la suite du travail, nous utiliserons plutôt le terme « monomères » pour parler des acides aminés impliqués dans la synthèse non ribosomiale. La partie qui va suivre reprend l'organisation modulaire des NRPS avec la description de l'activité enzymatique de chaque domaine ainsi que des différents modes de synthèse des peptides non ribosomiaux.

I.2. Les domaines des synthétases NRPS

Les domaines peuvent être classés en domaines principaux (indispensables à la synthèse peptidique) et en domaines secondaires (domaines facultatifs, qui modifient le produit final de la synthétase en agissant spécifiquement sur l'un ou l'autre des monomères incorporés).

I.2.1. Les domaines principaux

Les domaines principaux qui composent les modules sont les domaines de condensation (C), d'adénylation (A), de thiolation (T) et de thioestérase (TE). Chacun de ces domaines porte une activité enzymatique nécessaire à la biosynthèse du peptide par la voie non ribosomiale.

- Le domaine d'adénylation (A)

Le domaine d'adénylation (A) est impliqué dans la sélection et l'activation d'un acide aminé. Ce domaine est composé d'environ 550 acides aminés et joue un rôle primordial dans la reconnaissance des monomères. L'acide aminé sélectionné après réaction avec une molécule d'ATP en présence de Mg^{2+} donne un aminoacyl-adénylate (forme active du monomère) et un pyrophosphate. La figure 1 décrit la réaction d'activation du monomère.

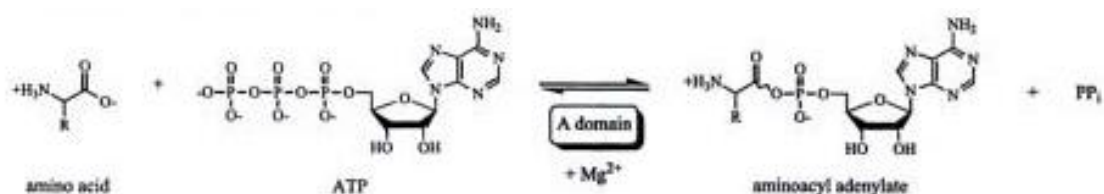


Figure 1: Réaction d'activation du monomère sélectionné par le domaine A. (Stachelhaus *et al.*, 2002).

En 2003, Schwarzer *et al.* ont défini dix motifs conservés (notés A1 à A10) que l'on retrouve dans les séquences protéiques des domaines d'adénylation (voir tableau 2). Conti *et al.* (1997) ont identifié 10 résidus responsables de la fixation du substrat. La position de ces 10 acides aminés dans la séquence du domaine A a été identifiée grâce à l'étude structurale du domaine A incorporant la phénylalanine A (Phe A) au cours de la synthèse de la gramicidine S. Il s'est avéré que 9 de ces résidus sont localisés entre les motifs A4 et A5 définis par Schwarzer *et al.*, et qu'un dixième est en dehors de cette région. Ces 10 acides aminés sont localisés dans un rayon de $5,5\text{\AA}$ autour du site de fixation du substrat (figure 2) définissant ainsi un code appelé « code Stachelhaus » ou code NRPS. Ce code reprend les séquences de 10 acides aminés impliqués dans la sélection d'un substrat au niveau du domaine A de la synthétase. Plus tard Rausch *et al.* (2005) ont démontré que les 34 acides aminés situés dans un rayon de $8\text{\AA}$ autour du site de fixation du substrat peuvent aussi influencer cette dernière.

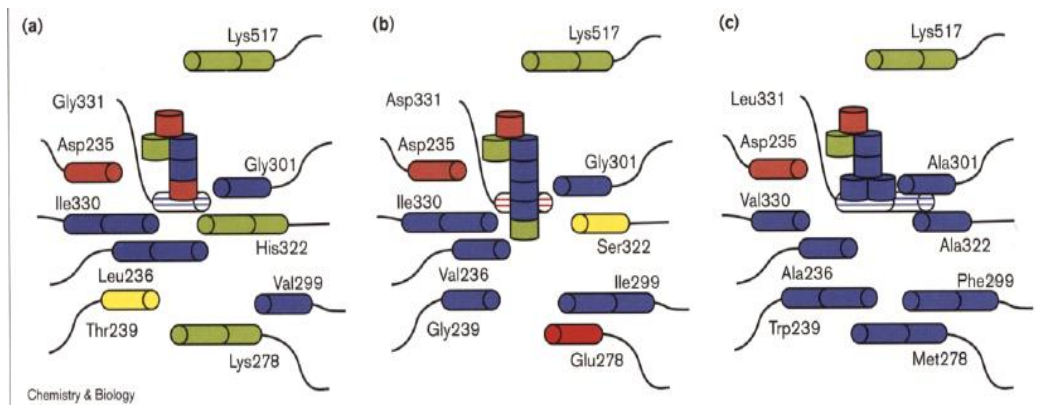


Figure 2: Représentation schématique des sites de fixation du substrat spécifique à la sélection de l'aspartate (a), de l'ornithine (b) et de la valine (c). (Stachelhaus *et al.*, 1999).

La sélection d'un monomère particulier par le domaine A est donc spécifique en fonction de la composition de la poche d'incorporation du substrat définie à la fois par les 10 acides aminés du code Stachelhaus et par les 34 acides aminés de Rausch *et al.* (2005) (comportant les 10 acides aminés du code Stachelhaus). Néanmoins, certains auteurs ont montré une souplesse dans la sélection du monomère par certains domaines A. La surfactine est un heptapeptide sur lequel des travaux ont montré qu'une substitution de la leucine en position 7 par une valine ou une isoleucine était possible, ainsi qu'une substitution de la valine en position 4 par un résidu plus hydrophobe comme la leucine ou l'isoleucine (Baumgart *et al.*, 1991 ; Bonmatin *et al.*, 1995). Plus de 30 variants de la microcystine, un heptapeptide cyclique hépatotoxique, ont été identifiés chez *Microcystis*, ces variants diffèrent par l'activation d'acides aminés différents au moment de la synthèse au niveau des modules 2 et 4 de la microcystine synthétase (Kurmayer *et al.*, 2001). En 1997, Van Dörhen *et al.* ont montré l'obtention de variants de la cyclosporine, un undecapeptide à activité immunosuppresseur, en fonction du milieu de culture utilisé. Ces différentes observations montrent le caractère « permissif » de certains domaines A dans l'incorporation de monomères lors de la synthèse peptidique non-ribosomiale.

- Le domaine de thiolation (T)

Le domaine de thiolation (T) est aussi retrouvé dans la littérature sous le nom de PCP (peptidyl carrier protein). Ce domaine est composé d'environ 100 acides aminés. Il intervient dans la fixation de l'acide aminé, activé par le domaine A, sur la synthétase de manière

covalente par un lien thioester, par l'intermédiaire d'un cofacteur phosphopantéthénique (Ppant) (figure 3). La transformation du domaine PCP de sa forme inactive (apo-PCP) à sa forme active (holo-PCP), contenant ce cofacteur, est catalysée par une phosphopantéthéine transférase.

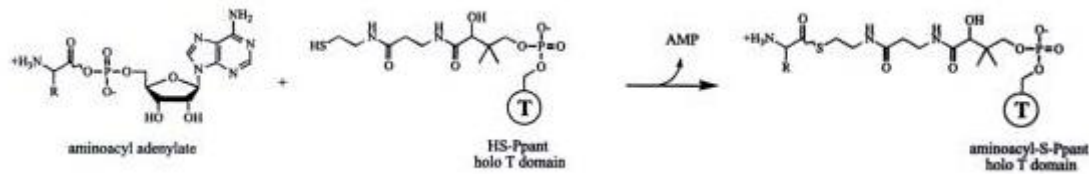


Figure 3: Fixation covalente de l'aminocyl adénylate sur la synthétase par le biais d'un bras phosphopantéthénique (Stachelhaus *et al.*, 2002).

- Le domaine de condensation (C)

Le domaine C est composé d'environ 450 acides aminés, il se situe au début de chaque module de la synthétase hormis pour le module d'initiation, exception faite des modules incorporant des acides gras. Il intervient dans la formation de la liaison peptidique entre deux acides aminés sélectionnés et activés par les domaines A de deux modules consécutifs (appelés modules 1 et 2 dans l'exemple). Le domaine C catalyse l'attaque nucléophile de l'acide aminé fixé sur le domaine T1 du module 1 par l'acide aminé fixé sur le domaine T2 du module 2. (figure 4)

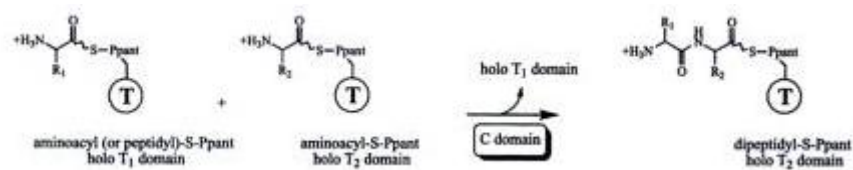


Figure 4: Formation d'une liaison peptidique entre deux acides aminés liés par une liaison thioester aux domaines T1 et T2 (Stachelhaus *et al.*, 2002).

Différents domaines de condensation ont été mis en évidence par des auteurs (Rausch *et al.*, 2007) travaillant sur les séquences protéiques de ces domaines. Quatre types de domaines C ont ainsi été décrits :

- Le domaine $^L C_L$ intervenant dans la formation d'une liaison peptidique entre deux acides aminés sous forme L.
- Le domaine $^D C_L$ intervenant dans la formation d'une liaison peptidique entre un acide aminé sous forme D et un acide aminé sous forme L (Clugston *et al.*, 2003).

Ce type de domaine C est retrouvé dans un module précédé par un domaine d'épimérisation (la fonction de ce domaine est décrite dans la partie sur les domaines secondaires des NRPS).

- Le domaine C/E (condensation/épimérisation) a été identifié dans des NRPS dont le produit final contient des acides aminés sous forme D alors que la NRPS ne possède pas de domaine spécifique ayant une activité propre d'épimérisation. Cette observation a été faite sur diverses NRPS telles que celles produisant l'arthrofactine, la syringomycine ou encore la syringopeptine (Balibar *et al.*, 2005).
- Le domaine « C starter » identifié comme intervenant dans la liaison entre un acide gras et un acide aminé sous forme L d'une chaîne peptidique issue de la voie non ribosomiale (Rausch *et al.*, 2007). Ce type de domaine C est présent au début de la synthétase d'où son nom.

- **Le domaine de thioestérase (TE)**

Le domaine TE joue un rôle dans la libération du peptide néoformé du complexe multi-enzymatique par hydrolyse de la liaison thioester qui lie le peptide au domaine T du dernier module de la synthétase. Ce domaine TE est composé d'environ 280 acides aminés et se situe à l'extrémité C-terminale de la synthétase. Outre son rôle dans la libération du peptide, ce domaine peut aussi intervenir dans la cyclisation de celui-ci. Il est courant de retrouver à la fin des synthétases de divers lipopeptides cycliques produits par des *Pseudomonas*, deux domaines TE en tandem. Le deuxième domaine TE interviendrait dans la libération des domaines de thiolation qui aurait été préalablement acétylés de façon erronée. La présence de ce tandem TE-TE à la fin des synthétases de lipopeptides cycliques est un des critères utilisé par Raaijmakers pour aller chercher de nouveaux lipopeptides cycliques (CLPs) dans les génomes de *Pseudomonas*. Lors de la synthèse d'un antibiotique, la daptomycine (Figure 5), un seul domaine TE permet la libération du peptide néoformé de la synthétase ainsi que sa cyclisation par l'intermédiaire d'une liaison peptidique entre le dernier acide aminé de la chaîne et un acide aminé au cœur de cette dernière.

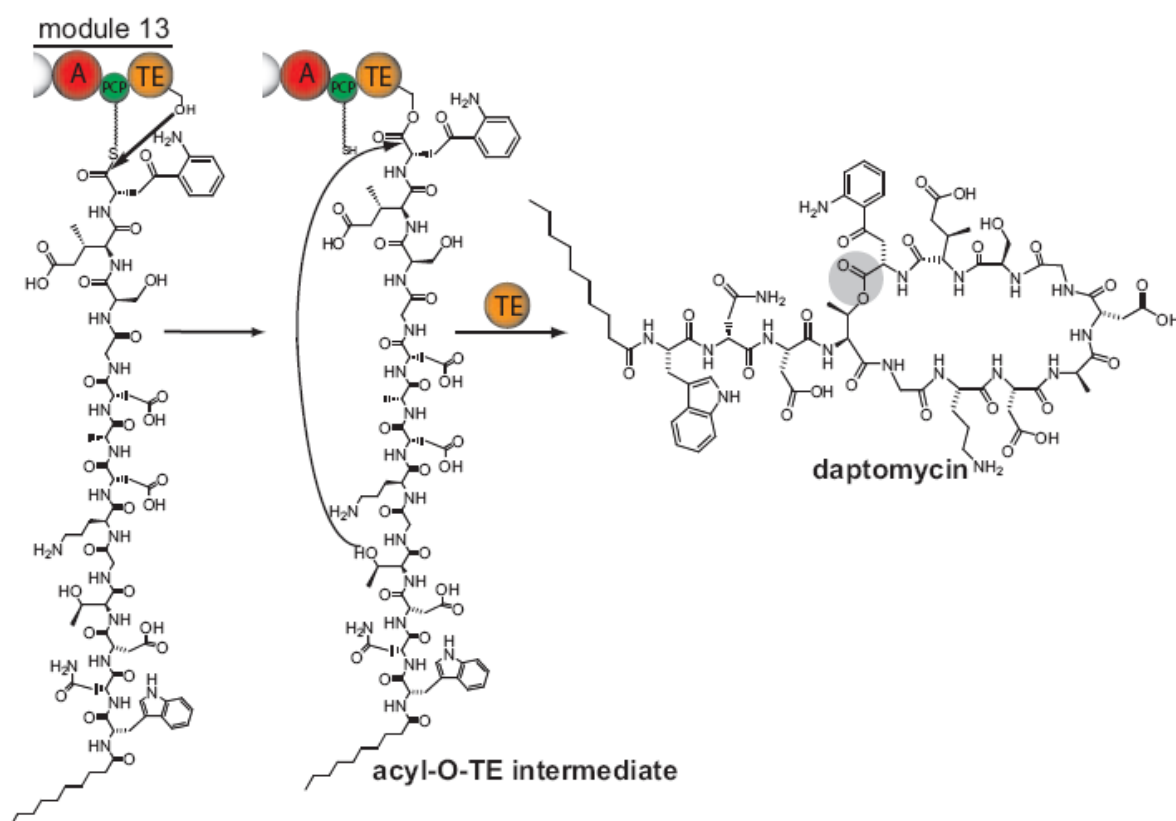


Figure 5: Schéma de la libération de la daptomycine de la synthétase et de sa cyclisation sous l'action du domaine TE (Grünwald et Marahiel, 2006).

I.2.2. Les domaines secondaires

Certains modules contiennent des domaines secondaires. Les plus courants sont les domaines d'épimérisation (E), de méthylation (M), de cyclisation (Cy), de formylation (F), d'oxydation (Ox) ou encore de réduction (Re) (figure 6). Ils permettent d'obtenir des variations structurales au niveau des monomères contenus dans le peptide formé et ainsi d'offrir le panel important d'activités biologiques qui caractérise les molécules d'origine non ribosomiales.

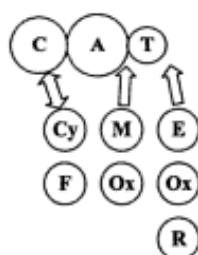


Figure 6: Localisation des domaines secondaires au sein d'un module (Stachelhaus *et al.*, 2002).

- Le domaine d'épimérisation (E)

Le domaine d'épimérisation transforme un acide aminé de forme L en son isomère de forme D (figure 7). Lors de la biosynthèse non ribosomiale des acides aminés sous forme D sont rarement présents dans le pool de substrats disponibles, la réaction d'épimérisation est donc un moyen d'obtenir des acides aminés de forme D dans le produit final de la synthétase. Comme cité précédemment la présence de domaines C/E peut aussi engendrer la formation de d'acide aminé de forme D. L'existence de ce type de domaine C/E a été mise en évidence pour la première fois dans la synthétase de l'arthrofactine (Balibar *et al.*, 2005). En effet l'arthrofactine compte 6 monomères sous leur forme D sur les 11 qui la compose alors que la synthétase ne présente aucun domaine d'épimérisation. Les auteurs ont ainsi démontré que cette activité d'épimérisation était incluse dans le domaine de condensation.



Figure 7: Schéma de la réaction d'épimérisation de l'acide aminé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

- Le domaine de cyclisation (Cy)

Le domaine Cy forme des cycles au sein des peptides non ribosomiaux, en formant une liaison entre la chaîne latérale de la sérine, la cystéine ou la thréonine avec la chaîne peptidique principale (figure 8). Le produit de cette cyclisation est une thiazoline si l'acide aminé impliqué est une cystéine, ou une oxazoline si l'acide aminé impliqué dans la cyclisation est une thréonine ou une sérine. Ainsi dans la synthétase de la vibriobactine, le premier domaine Cy intervient dans la cyclisation de la molécule par la formation d'un noyau oxazoline entre une thréonine et un acide hydroxybenzoïque et le second catalyse une condensation (Silakowski *et al.*, 2000).

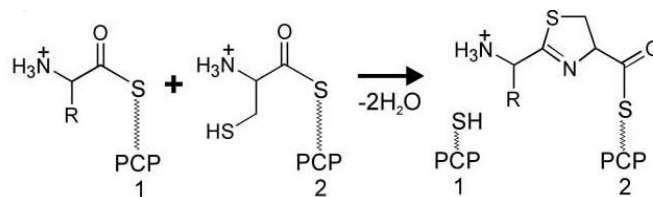


Figure 8: Schéma de la réaction de cyclisation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

- Le domaine de réduction (R)

Le domaine de réduction réduit l'acide aminé fixé sur le domaine PCP voisin en modifiant sa structure initiale grâce à l'action du coenzyme NADPH (Figure 9).

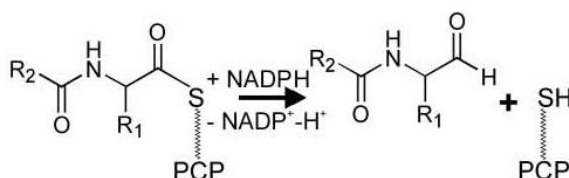


Figure 9: Réaction de réduction de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

- Le domaine d'oxydation (Ox)

Le domaine Ox forme des cycles thiazoles et oxazoles, respectivement à partir des thiazolines et des oxazolines sous l'action de la flavin mononucleotide (FMN).

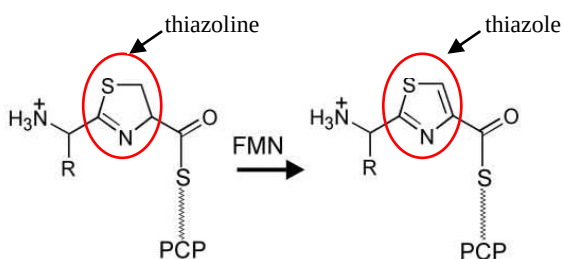


Figure 10: Réaction d'oxydation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

- Le domaine de méthylation (M)

Deux types de méthylation sont connus, la N-méthylation et la C-méthylation. La N-méthylation méthyle un atome d'azote de l'acide aminé par l'intermédiaire de la S-adenosyl méthionine (SAM) (figure 11). Ce type de méthylation est par exemple observé dans la molécule de cyclosporine qui est composée de 11 monomères dont 7 sont N-méthylés (Schwarzer *et al.*, 2003). Le deuxième type de méthylation est la C-méthylation qui ajoute un groupement méthyle sur un noyau thiazol. Par exemple, la yersiniabactine contient une cystéine C-méthylée par une méthyltransférase (Miller *et al.*, 2001).

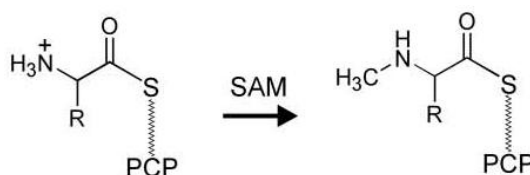


Figure 11: Réaction de méthylation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

- Le domaine de formylation (F)

Le domaine F ou Fo permet l'ajout d'un groupement formyle sur l'atome d'azote de l'acide aminé par l'intermédiaire du tétrahydrofolate (THF) qui est un cofacteur donneur du groupement formyle.

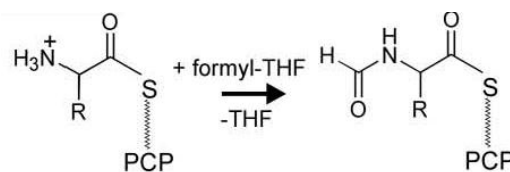


Figure 12: Réaction de formylation de l'acide aminé fixé sur le domaine PCP (ou domaine T) (Schwarzer *et al.*, 2003).

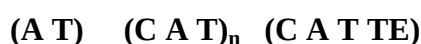
Une étude bioinformatique des séquences a permis de mettre en évidence des motifs conservés au sein des différents domaines principaux et secondaires cités ci-dessus (Schwarzer *et al.*, 2003). Le tableau 2 reprend les différents motifs ainsi que leur séquence protéique.

Domaines NRPS	motif	Séquence de la signature
Domaine A	A1	L(TS)YxEL
	A2	LKAGxAYL(VL)P(LI)D
	A3	LAYxxYTSG(ST)TGxPKG
	A4	FDxS
	A5	NxYGPTE
	A6	GELxJGx(VL)ARGYL
	A7	Y(RK)TGDL
	A8	GRxPxQVKIRGxRIELGEIE
	A9	LPxYM(IV)P
	A10	NGK(VL)DR
Domaine PCP	T	LGG(DH)SL
Domaine C	C1	SxAQxR(LM)(WY)XI
	C2	RHExLRTxF
	C3	MHHxISDG(WV)S
	C4	YxD(FY)AVW
	C5	(IV)GxFVNT(QL)(CA)xR
	C6	HNQD(YD)PFE
	C7	RDxSRNPL
Domaine TE	Te	GxSxG
Domaine E	E1	PIQxWF
	E2	HHxISDG(WV)S
	E3	DxLLxAxG
	E4	EGHGRE
	E5	RTVGWFTxxYG(YV)PFE
	E6	PxxGxGYG
	E7	FNYLG(QR)
Domaine Cy	Cy1	FPL(TS)xxQxAYxxGR
	Cy2	RHx(IM)L(PAL)x(ND)GxQ
	Cy3	LPxxPxLPLxxxP
	Cy4	(TS)(PA)3x(LAF)6x(IVT)LxxW
	Cy5	(GA)DFTxLxLL
	Cy6	PVVFTSxL
	Cy7	(ST)(QR)TPQVx(LI)D13xWD
Domaine Ox	Ox1	KYxYxSxGxxY(PG)VQ
	Ox2	GxxxG(LV)xxGxYYY(HD)P
	Ox3	IxxxYG
Domaine N-Mt	M1	VL(DE)xGxGxG
	M2	NELxYRYxAV
	M3	VExSxARQxGxLD
Domaine R	R1	V(L)(L)TG(A)TG(F)(L)GxxLL
	R2	Vx(L)(L)VR(A)
	R3	GPL(G)x(P)x(L)GL
	R4	V(Y)PYxYLxx(P)NVxxT
	R5	GYxxSKW(A)(A)E
	R6	R(P)G
	R7	YxxxxG(LF)LxxP

Tableau 2 : Liste des motifs conservés dans les séquences des différents domaines impliqués dans la synthèse non ribosomiale. Dans le motif, les acides aminés entre parenthèses correspondent aux possibilités à une position donnée et le 'x' représente n'importe quel acide aminé. (Schwarzer *et al.*, 2003).

I.3. Organisation modulaire des synthétases peptidiques non-ribosomiales

Dans les synthétases NRPS, les domaines décrits ci-dessus sont organisés en modules. Trois types de modules sont généralement observés (Mootz *et al.*, 2002). La synthétase NRPS débute par un module d'initiation qui se compose au minimum d'un domaine d'adénylation (A) et d'un domaine de thiolation (T). Ensuite cette synthétase comporte « n » modules d'élongation qui se composent au minimum d'un domaine de condensation (C), d'un domaine A et d'un domaine T et elle se termine par un module de terminaison comprenant au minimum un domaine C, un domaine A, un domaine T et un domaine de thioestérase (TE). La structure modulaire générique des NRPS est donc la suivante :



II. La biosynthèse NRPS

L'organisation des synthétases en modules offre une flexibilité qui permet trois mécanismes de synthèse qui ont été décrits par Mootz *et al.* en 2002.

II.1. La biosynthèse linéaire

La biosynthèse linéaire correspond à un assemblage des monomères activés suivant l'ordre des modules de la synthétase. La séquence du peptide final est donc gouvernée par le nombre et l'ordre des modules dans la synthétase. Le δ -Aminoadipyl-cysteinyl-D-valine (ACV), précurseur de la pénicilline et de la céphalosporine, est synthétisé de manière linéaire comme le montre la figure 13.

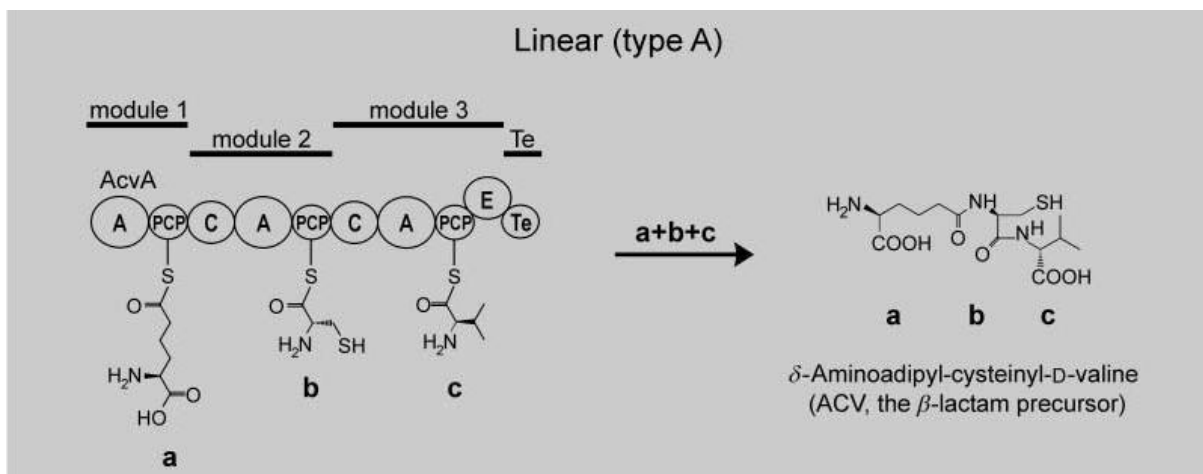


Figure 13: Biosynthèse linéaire de l'ACV (Mootz et al., 2002).

II.2. La biosynthèse itérative : exemple de la synthèse de l'entérobactine

Lors de la biosynthèse itérative chaque module peut être utilisé plusieurs fois dans l'assemblage d'un même produit. C'est le cas de l'entérobactine, un sidérophore produit par *Escherichia coli* lors d'une carence en fer dans le milieu environnant. Ce peptide est composé de 6 monomères alors que la synthétase ne comporte que 2 modules. L'entérobactine est donc construite par 3 répétitions du di-peptide produit par la synthétase. La figure 14 montre le mécanisme moléculaire itératif mis en œuvre par les différents domaines dans la synthèse de l'entérobactine.

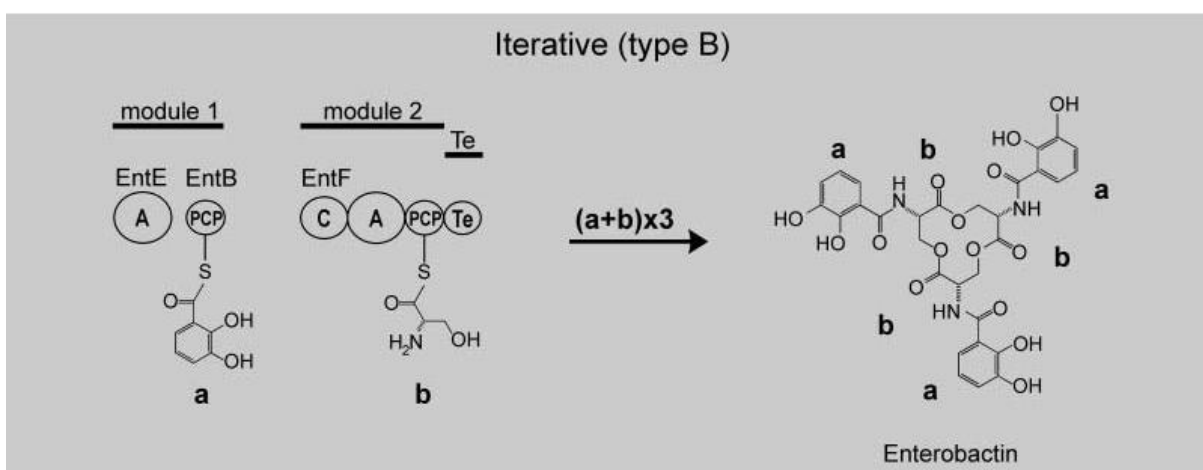


Figure 14: Biosynthèse itérative de l'entérobactine (Mootz et al., 2002). L'opéron de l'entérobactine code pour 4 synthétases (EntE, B, F et D) regroupées en deux modules permettant la reconnaissance et l'activation de deux monomères (l'acide 2,3-dihydroxybenzoïque et la L-sérine).

II.3. La biosynthèse non linéaire : exemple de la synthèse de la vibriobactine

La biosynthèse non linéaire est caractérisée par un agencement différent des monomères par rapport à l'ordre d'apparition des modules dans la synthétase. La figure 15 présente l'organisation de la synthétase de la vibriobactine et des différents domaines. La vibriobactine est un sidérophore hybride NRPS/PKS composé de trois monomères (un acide 2,3 dihydroxybenzoïque, une thréonine et une norspermidine). Mootz *et al.* (2002) ont décrit le fonctionnement non linéaire des gènes VibE, B, H et F dans la synthèse de la molécule par une condensation des monomères cités ci-dessus.

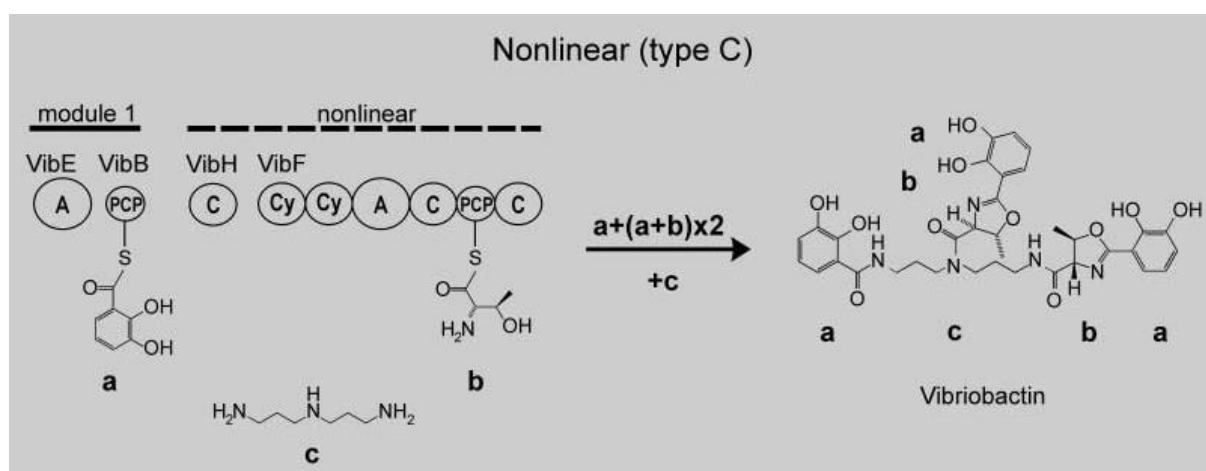


Figure 15: Biosynthèse non linéaire de la vibriobactine (Mootz *et al.*, 2002). Il n'y a pas de corrélation entre l'organisation des modules et des domaines et la structure du produit formé. L'assemblage $a + (a+b) \times 2$ conduit à la formation de la molécule finale de vibriobactine.

III. Le mécanisme PKS

Outre le mécanisme NRPS, il existe un autre mécanisme de synthèse modulaire faisant intervenir des complexes enzymatiques appelés les polykétides synthases (PKS).

Les polykétides synthases sont, comme les NRPS, des protéines multienzymatiques organisées en modules et domaines produisant des molécules actives appelées polykétides ou polycétides. Les PKS activent de acides carboxyliques de petite taille (acétate ou malonate) au contraire des NRPS qui eux sélectionnent et activent essentiellement des acides aminés. Les polykétides synthases ont été divisées en trois familles en fonction de leur mécanisme de fonctionnement et de leur organisation structurale: type I, II et III.

- **PKS de type I**

Les PKS de type I sont divisés en deux groupes : les PKS de type I itératives, qui réutilisent des domaines et les PKS de type I linéaires qui possèdent une organisation modulaire comparable aux NRPS. Les deux groupes de PKS de type I utilisent quatre domaines enzymatiques principaux :

- Le domaine acétyltransférase (AT) qui reconnaît le substrat et l'active
- Le domaine kétosynthase (KS) qui condense le substrat sur la chaîne
- Le domaine porteur d'acyl (ACP) qui porte la chaîne en cours d'élongation
- Le domaine thioestérase (TE) qui permet la libération du polykétide néosynthétisé de la synthase et peut permettre sa cyclisation

Comme pour le mécanisme NRPS, il existe des domaines secondaires qui modifient le substrat

- Le domaine kétoréductase (KR) qui réduit la fonction cétone en un alcool
- Le domaine déhydratase (DH) qui élimine une molécule d'eau et crée une double liaison
- Le domaine énoylréductase (ER) qui réduit une double liaison

- **PKS de type II**

Le système de type II consiste en un complexe de protéine mono-fonctionnel. Chez les PKS de type II, on trouve les domaines KS et ACP mais pas le domaine AT. Dans ce cas, le domaine ACP joue le rôle du domaine AT dans la reconnaissance et l'activation du substrat. Le fonctionnement de ces domaines est itératif c'est-à-dire qu'un même module peut condenser plusieurs fois le substrat. Enfin le composé ainsi synthétisé est libéré par l'action du domaine TE (Hutchinson, 2003).

- **PKS de type III**

Les PKS de type III sont des enzymes itératives qui sont capables de synthétiser des polykétides, cycliques ou non, à partir de thioester-CoA comme le malonyl-CoA (Shen, 2003). Hormis le domaine ACP, les PKS de type III possèdent les mêmes domaines que les PKS de type II.

IV. Les systèmes hybrides NRPS/PKS

Les synthétases hybrides NRPS/PKS possèdent des domaines catalytiques conservés des mécanismes de synthèse NRPS et PKS, ces synthétases portent un domaine kétoacyl-synthase (KS) unique qui permet d'accepter des peptides et non des groupements acyl. Les produits de ces synthétases hybrides sont donc composés à la fois de monomères (issus du mécanisme NRPS) et de chaînes d'acides carboxyliques (issues du mécanisme PKS). La yersiniabactine est un métabolite chélateur de fer produit par une synthétase hybride NRPS/PKS (Pfeifer *et al.*, 2003). La figure suivante reprend la structure de la yersiniabactine synthétase.

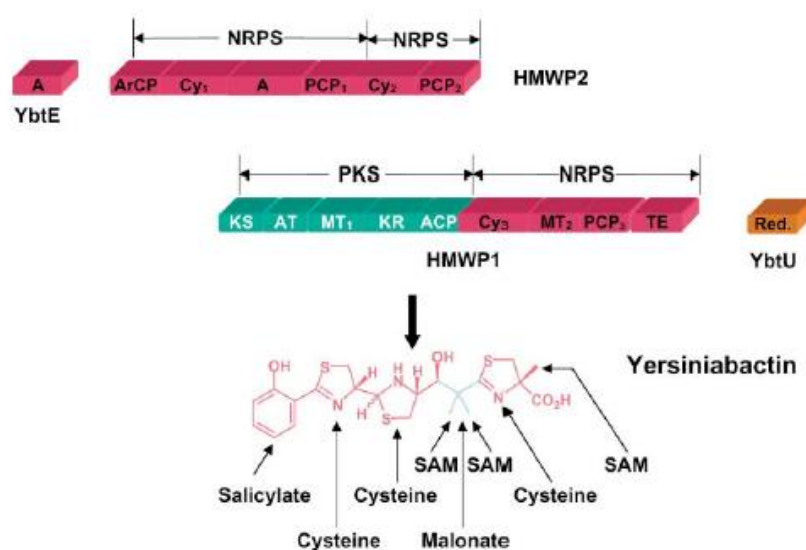


Figure 16: Structure de la yersiniabactine synthétase selon Miller *et al.*, (2002). La synthétase de la yersiniabactine est composée de 4 protéines : YbtE, HMWP2, HMWP1 et YbtU. Les abbreviations des domaines sont: A, adenylation; ArCP, aryl carrier protein; Cy, cyclization; PCP, peptidyl carrier protein; KS, ketosynthase; AT, acyltransferase; MT, methyltransferase; KR, ketoreductase; ACP, acylcarrier protein; TE, thioesterase. La yersiniabactine synthétase comporte quatre protéines, YbtE, HMWP1, HMWP2 et YbtU, codant pour dix-sept domaines fonctionnels, dont douze domaines catalytiques et cinq domaines T, qui sélectionnent, activent et incorporent le salicylate, trois cystéines, et un groupement malonyl dans la yersiniabactine.

V. Structures et activités biologiques des peptides non ribosomiaux

V.1. Les structures des NRPs

Les peptides synthétisés par la voie non ribosomiale peuvent présenter des structures particulières (linéaire, cyclique, partiellement cyclique, branchée etc...), d'une part générées par le mode de synthèse et de l'autre par la présence de domaines secondaires dans la synthétase capables de modifier la structure du monomère incorporé dans le produit final. Le nombre de monomères identifiés sur l'ensemble des peptides non ribosomiaux (NRPs) est supérieur à 500 (Caboche *et al.*, 2010) contrairement à la synthèse peptidique ribosomiale où les peptides sont composés à partir de 20 (étendus à 21 avec l'ornithine) acides aminés protéogéniques.

Les peptides non ribosomiaux ou NRPs peuvent être divisés en cinq classes en fonction de la composition du peptide :

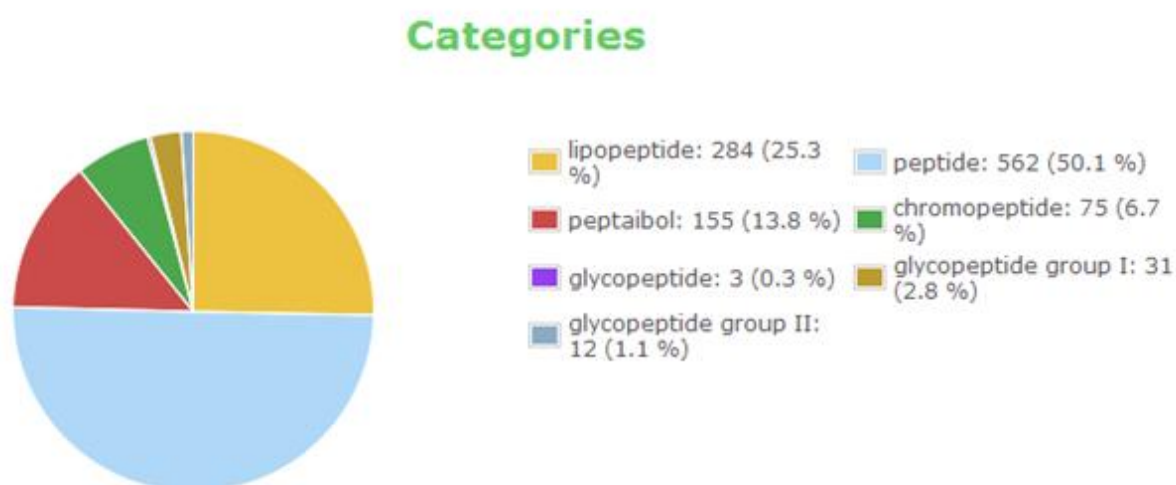


Figure 17 : Représentation schématique de la répartition des NRPs recensés dans Norine (1122 peptides) en fonction de la structure des peptides.

La figure 17 représente la répartition des NRPs contenus dans la banque de données Norine (Caboche *et al.*, 2008) en fonction de la composition des peptides. Cette banque contient 1122 NRPs. Chaque type de NRPs est détaillé dans les paragraphes suivants :

- **Les peptides**

Les peptides sont uniquement composés d'acides aminés, protéogéniques ou non, c'est le cas par exemple de l'ACV-tripeptide, peptide linéaire, qui est un précurseur de deux antibiotiques : la pénicilline et de la céphalosporine.

- **Les lipopeptides**

Les lipopeptides sont composés d'une chaîne peptidique issue du mécanisme de synthèse non ribosomal lié par une liaison amide ou ester à un acide gras provenant du pool intracellulaire ou synthétisé en partie par les hybrides NRPS/PKS. Les lipopeptides les plus étudiés sont produits par *Bacillus* (tels que la surfactine, un puissant biosurfactant ou la fengycine, une molécule à activité antifongique) et *Pseudomonas* (arthrofactine ou tolaasine par exemple).

- **Les glycopeptides**

Les glycopeptides présent dans la base de données Norine sont des molécules complexes constituées d'une partie peptidique avec plusieurs cycles (en général 3) sur laquelle viennent se greffer des sucres. Les glycopeptides sont des molécules à activité antibiotique, par inhibition de la synthèse du peptidoglycane de la paroi bactérienne. La vancomycine est le premier glycopeptide non ribosomal à avoir été isolé, il est produit par *Streptomyces orientalis*. Les sucres greffés sur sa partie peptidique sont le glucose et la vancosamine (Hubbard et Walsh, 2003) (figure 18). Il existe différents lipoglycopeptides semi-synthétiques à activité antibiotique utilisés dans le milieu médical tels que l'oritavancine (Billstein *et al.*, 2011), la dalbavancine ou encore la telavancine, un dérivé de la vancomycine.

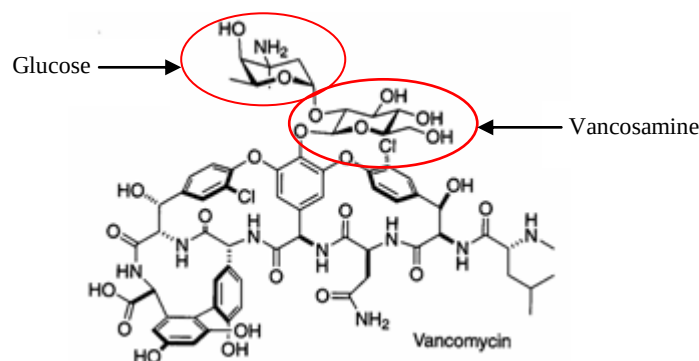


Figure 18: Exemple de glycopeptide : la vancomycine.

- **Les chromopeptides**

Les chromopeptides sont composés d'une partie peptidique de 6 à 15 acides aminés sur laquelle se fixe un chromophore qui confère à cette molécule des propriétés de fluorescence. Les pyoverdines, molécules chélatrices du fer, représentent une grande partie de la diversité des chromopeptides. Ces molécules seront davantage décrites dans la partie suivante.

- **Les peptaibols**

Les peptaibols sont des peptides linéaires contenant entre 5 et 20 acides aminés pour la plupart non protéogéniques. Ils sont caractérisés par la forte présence de l'acide α -aminoisobutyric (Aib) dans la chaîne peptidique ainsi que par le fait que l'acide aminé en position C-terminale du peptide est hydroxylé. Ces molécules à activité antibiotique, ne sont produites, *a priori*, que par des champignons.

V.2. Les activités biologiques des NRPs

Les molécules issues de la voie non ribosomiale couvrent une large gamme d'activités biologiques (Schwarzer *et al.*, 2003). Ces activités sont fortement liées à leur structure tertiaire (Peypoux *et al.*, 1999) et à leur composition en monomères (Caboche *et al.*, 2010). On peut classer les activités biologiques en cinq groupes principaux (figure 20):

- **Les antibiotiques et précurseurs d'antibiotiques**

L'activité antibiotique concerne presque la moitié des peptides référencés dans Norine (512 sur 1122). La gramicidine S (Byford *et al.*, 1997) et la tyrocidine A (Marahiel *et al.*, 1997) sont deux antibiotiques cycliques parmi les plus connus. La vancomycine est un glycopeptide à activité antibiotique (Mootz *et al.*, 1997). L'iturine et la fengycine (Jacques, 2011) sont deux lipopeptides à activité antifongique. L'ACV-tripeptide est un précurseur de deux antibiotiques, la pénicilline et la céphalosporine (Hori *et al.*, 1989).

- **Les agents immunosuppresseurs**

Les immunosuppresseurs sont utilisés pour inhiber l'activité du système immunitaire en cas de greffe d'organes pour éviter les phénomènes de rejet. Norine référence 20 peptides porteurs de cette activité. Le plus connu est la cyclosporine A produite par *Tolypocladium inflatum*. Cette molécule est utilisée pour soigner les maladies auto-immunes grâce à son action sur les cellules T (Green, 1981).

- **Les agents cytostatiques**

Les agents cytostatiques (63 sont référencés dans Norine) sont utilisés dans les traitements contre le cancer, ils jouent un rôle d'inhibiteur de la division cellulaire en endommageant ou en détruisant les cellules. L'épothilone est utilisé dans les traitements chimiothérapeutiques en inhibant la fonction des microtubules intracellulaires qui jouent un rôle essentiel dans la division cellulaire (Nettles *et al.*, 2004). La bléomycine A2 a également une activité antitumorale (Brahim *et al.*, 2008).

- **Les sidérophores**

Les sidérophores représentent une large gamme de molécules capables de capter le fer à l'état ferrique dans l'environnement des microorganismes. Ces sidérophores sont sécrétés par le microorganisme dans le cas d'une carence en fer qui est nuisible aux besoins métaboliques auxquels le fer est lié. Une fois libérés, les sidérophores captent le fer et réincorporés par la cellule par le biais de récepteurs membranaires spécifiques. Ces molécules représentent une grande diversité dans leur composition, 89 peptides sont connus dans Norine comme ayant une activité de sidérophore. On peut citer la yersiniabactine (Pfeifer *et al.*, 2003), un sidérophore NRPS/PKS ou encore l'entérobactine produite par *Escherichia coli* (O'Brien *et al.*, 1970). Le sidérophore le plus répandu est la pyoverdine, un sidérophore produit par les bactéries du genre *Pseudomonas* appartenant au groupe des fluorescents (voir chapitre VI.).

- Autres

Les NRPs peuvent avoir une fonction de pigment comme l'indigoïdine (pigment bleu) produit par *Dickeya dadantii* ou encore être des toxines (216 sont référencés dans Norine) comme l'anabaenopeptilide 90-A ou les microcystines produites par des cyanobactéries (Rouhiadinen *et al.* , 2000).

La figure 19, issue de Norine, représente la répartition des activités biologiques des 1122 NRPs recensés dans la banque de données. Cette figure montre que près de la moitié des NRPs connus possèdent une activité antibiotique.

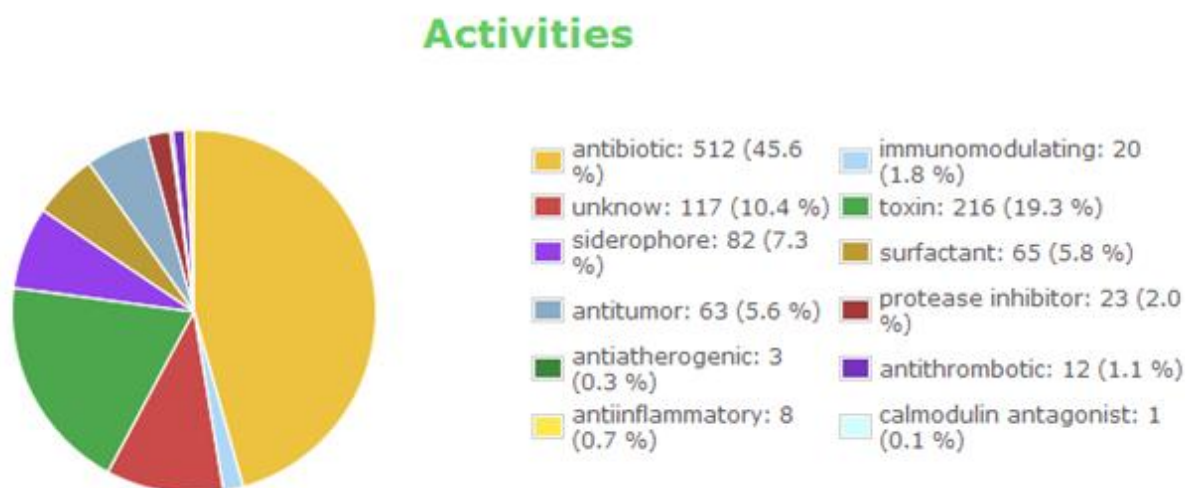
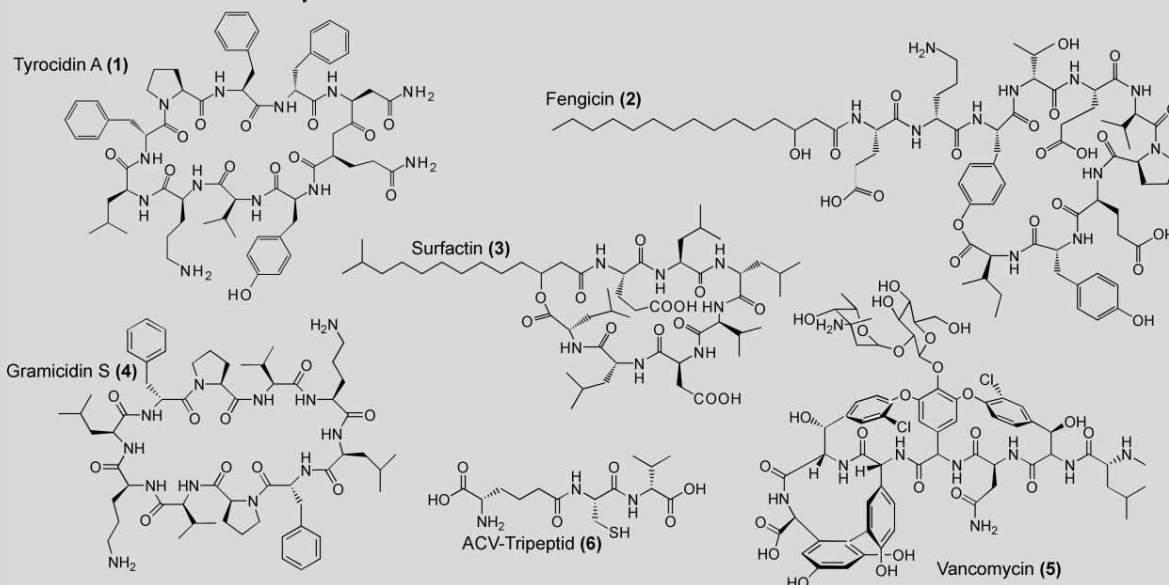
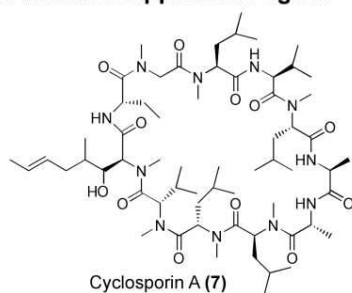


Figure 19 : Représentation schématique de la répartition des activités biologiques des 1122 NRPs recensés dans Norine. A noter qu'un même NRP peut posséder deux activités biologiques différentes.

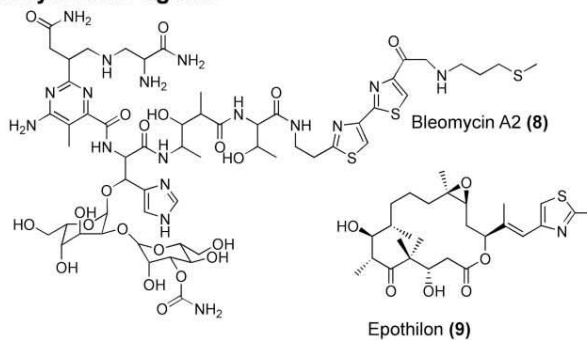
A. antibiotics/ antibiotic precursors



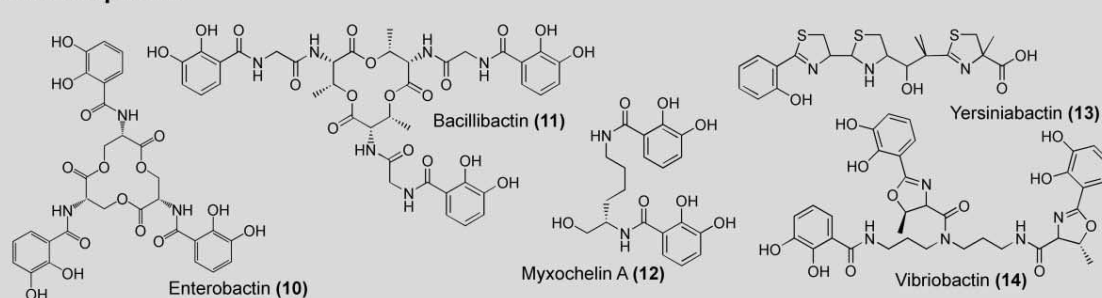
B. immunosuppressive agent



C. cytostatic agents



D. siderophores



E. others

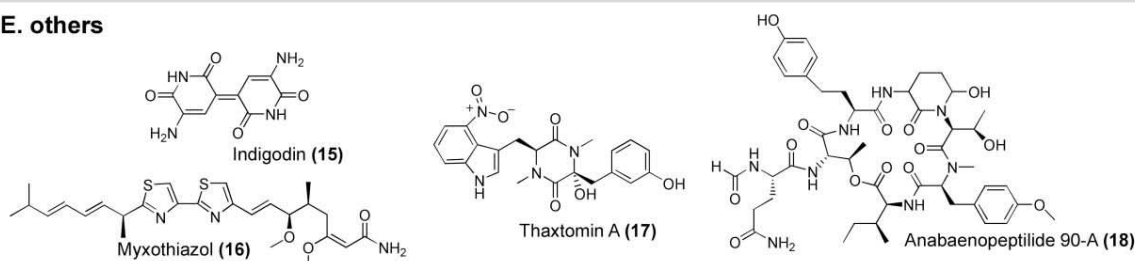


Figure 20: Structure de différents NRPs classés en fonction de leur activité biologique (Schwarzer *et al.*, 2003).

VI. Les NRPs produits par *Pseudomonas*

Les *Pseudomonas* sont des bactéries en forme de bâtonnets, Gram-négatives, aérobies, mobiles grâce à plusieurs flagelles. Parmi les *Pseudomonas*, les *Pseudomonas* fluorescents sont caractérisés par la production d'un sidérophore jaune-vert, appelé pyoverdine, lorsque le fer biodisponible dans leur environnement est insuffisant pour satisfaire leurs besoins métaboliques.

VI.1. Les sidérophores

Les pyoverdines sont les siderophores les plus produits par les *Pseudomonas* fluorescents, néanmoins ces bactéries produisent aussi des sidérophores additionnels tels que la pyocheline (Cox *et al.*, 1981), la thioquinolobactine (Matthijs *et al.*, 2007) ou la pseudomonine (Anthoni *et al.*, 1995), ou capturent le fer par des agents chélateurs exogènes (Cornelis et Matthijs, 2002). La faculté à produire la pyoverdine, sidérophore à haute affinité pour le fer (de l'ordre de 10^{25}M^{-1} à pH7) (Neilands, 1995), en plus des autres sidérophores d'affinité moins importante fait des *Pseudomonas* du genre fluorescent des compétiteurs importants pour le fer vis-à-vis des autres micro-organismes (Ravel et Cornelis, 2003). Le tableau 3 reprend l'ensemble des molécules de pyoverdines dont la séquence de la chaîne peptidique est connue.

Espèce productrice	Souche	chaîne peptidique de la pyoverdine	référence
<i>Pseudomonas putida</i>	GS43*	R-ChrP-Lys-OHAsp-Ser-Ser-Ser-cOHOrn	Meyer <i>et al.</i> , 2008
	Thai*	R-ChrP-(Ser-Dab)-Thr-Ser-AOHOrn-cOHOrn	Ruangviriyachai <i>et al.</i> , 2004
	BTP1 ou 90-40	R-ChrP-Asp-Ala-Asp-D-AcOHOrn-Ser-cOHOrn	Jacques <i>et al.</i> , 1995
	G4R	R-ChrP-Asp-Orn-(D-OHAsp-Dab)-Gly-Ser-cOHOrn	Salah-el-din <i>et al.</i> , 1997
	11370	R-ChrP-Asp-εLys-D-OHAsp-D-Ser-Ala-D-Ser-cOHOrn	Budzikiewicz <i>et al.</i> , 1999
	ATCC 39167	R-ChrP-D-Ser-D-AcOHOrn-Ala-Gly-(D-Ser-Ala-OHAsp-Thr)	Uria-Fernandez <i>et al.</i> , 2003
	90-33	R-ChrP-Asp-Lys-Thr-D-OHAsp-Thr-D-aThr-cOHOrn	Sultana <i>et al.</i> , 2001
	90-51	R-ChrP-Asp-εLys-D-OHAsp-Ser-Gly-D-aThr-Lys-cOHOrn	Sultana <i>et al.</i> , 2000
	G172*	R-ChrP-Ala-Lys-Dab-OHAsp-(Thr-Gly-OHAsp-Gly-Thr-Thr)	Meyer <i>et al.</i> , 2008
	ATCC12633	R-ChrP-Asp-εLys-D-OHAsp-Ser-Thr-D-Ala-D-Glu-Ser-cOHOrn	Persmark <i>et al.</i> , 1990
	L1	R-ChrP-Asp-εLys-D-OHAsp-Ser-D-aThr-Ala-Thr-D-Lys-cOHOrn	Uria-Fernandez <i>et al.</i> , 2003
	3b	R-ChrP-Asp-(AOHOrn-Dab)-Thr-D-Ala-Thr-Thr-D-Gln-cOHOrn	Budzikiewicz <i>et al.</i> , 2004
	putC	R-ChrP-Asp-D-OHbutOHOrn-D-Dab-Thr-Gly-D-Ser-Ser-OHAsp-Thr	Seinsche <i>et al.</i> , 1993
	90-44	R-ChrP-Asp-Lys-D-AcOHOrn-Thr-Ser-Ser-Gly-Ser-Ser-D-cOHOrn	Sultana <i>et al.</i> , 2001
	90-136 ou G168*	R-ChrP-Ser-Lys-Ser-Ser-Thr-Thr-AcOHOrn-Ser-Ser-cOHOrn	Meyer <i>et al.</i> , 2008
	C	R-ChrP-Asp-D-OHOrn-C4:0-OH(3)-D-Dab-Thr-Gly-D-Ser-Ser-OhAsp-Thr	Seinsche <i>et al.</i> , 1993
	BTP7 ou 95-275	R-ChrP-D-ser-ser-FoHOrn-ser-ser-lys-FoHOrn-lys-ser	Sultana <i>et al.</i> , 2000
	BTP16	R-ChrP-Asp-D-OHOrn-C4:0-OH(3)-D-Dab-Thr-Gly-Ser-Ser-OHAsp-aThr	Ongena <i>et al.</i> , 1997
	2461	R-ChrP-Asp-Lys-D-OHAsp-Ser-D-aThr-D-Ala-Thr-D-Lys-cOHOrn	Uria-Fernandez <i>et al.</i> , 2003
	pseudobactin 37167-I*	R-ChrP-Ser-Ala-AcOHOrn-Gly-Ala-OHAsp-Ser-Thr	Khalil-Rizvi <i>et al.</i> , 1997
	pseudobactin 37167-II*	R-ChrP-Ser-Ala-AcOHOrn-Gly-Ala-OHAsp-Ser-Thr	Khalil-Rizvi <i>et al.</i> , 1997
	pseudobactin 589A	R-ChrP-Asp-Lys-D-OHAsp-Ser-Thr-D-Ala-D-Glu-Ser-cOHOrn	Persmark <i>et al.</i> , 1990
	pseudobactin A214	R-ChrP-D-Ser-D-AcOHOrn-Ala-Gly-D-Ser-Ala-OHAsp-Thr	Uria-Fernandez <i>et al.</i> , 2003
	KT2440*	R-ChrP-Asp-Orn-(OHAsp-Dab)-Gly-Ser-cOHOrn	Matthijs <i>et al.</i> , 2009
	9AW	R-ChrP-D-Ser-εLys-OHHis -D-aThr-Ser-cOHOrn	Budzikiewicz <i>et al.</i> , 1997
	Ps4a	R-ChrP-D-Ala-Lys-Thr-D-Ser-AOHOrn-cOHOrn	Budzikiewicz <i>et al.</i> , 1992
	BTP2	R-ChrP-D-Ser-Val-OHAsp-Gly-Thr-D-Ser-cOHOrn	Ongena <i>et al.</i> , 2001

<i>Pseudomonas fluorescens</i>	PL8	R-ChrP-D-Lys-D-AcOHOrn-Ala-Gly-D-aThr-Ser-cOHOrn	Barelmann <i>et al.</i> , 2002
	PL9 ou P19	R-ChrP-D-Ser-D-AcOHOrn-Ala-Gly-(D-Ser-Ser-D-OHAsp-Thr)	Uria-Fernandez <i>et al.</i> , 2003
	Pfl13525	R-ChrP-D-Ser-Lys-Gly-FOHOrn-(Lys-D-FOHOrn-Ser)	Hohlneicher <i>et al.</i> , 1995
	G173	R-ChrP-D-Ser-Ala-AOHOrn-(Orn-D-Asp-AOHOrn-Ser)	Uria-Fernandez <i>et al.</i> , 2003
	Pfl W	R-ChrP-Ser-Dab-Gly-Ser-D-OHAsp-D-Ala-Gly-Ala-Gly-cOHOrn	Demange <i>et al.</i> , 1990
	Pfl1.3	R-ChrP-D-Ala-D-Lys-Gly-Gly-OHAsp-(D-Gln-Dab)-Gly-Ser-cOHOrn	Georgias <i>et al.</i> , 1999
	Pfl "ng"	R-ChrP-Asn-D-FOHOrn-Lys-(Thr-Ala-Ala-D-FOHOrn-Lys)	Poppe <i>et al.</i> , 1987
	CHA0	R-ChrP-Asp-D-FOHOrn-Lys-(Thr-Ala-Ala-D-FOHOrn-Lys)	Wong-Lung-Sang <i>et al.</i> , 1996
	G153*	R-ChrP-Ser-Lys-Ala-Ser-Ser-AcOHOrn-Ser-Ser-cOHOrn	Meyer <i>et al.</i> , 2008
	Pfl17400	R-ChrP-D-Ala-D-Lys-Gly-Gly-OHAsp-(D-Gln-Dab)-Ser-D-Ala-cOHOrn	Demange <i>et al.</i> , 1990
	A6	R-ChrP-D-Lys-D-AcOHOrn-Gly-D-aThr-Thr-Gln-Gly-D-Ser-D-cOHOrn	Beiderbeck <i>et al.</i> , 1999
	95-275	R-ChrP-D-Ser-Ser-FOHOrn-Ser-Ser-(Lys-FOHOrn-Lys-Ser)	Meyer <i>et al.</i> , 2008
	51W	R-ChrP-D-Ala-D-Lys-Gly-Gly-D-OHAsp-D-Gln-D-Ser-Ala-Gly-D-aThr-cOHOrn	Voss <i>et al.</i> , 1999
	DSM50106*	R-ChrP-Ser-Lys-Gly-FOHOrn-Ser-Ser-Gly-(Orn-FOHOrn-Ser)	Meyer <i>et al.</i> , 2008
	Pf0-1*	R-ChrP-Ala-AcOHOrn-Orn-Ser-Ser-Ser-Arg-OHAsp-Thr	Meyer <i>et al.</i> , 2008
	Pfl18.1	R-ChrP-D-Ser-Lys-Gly-FOHOrn-Ser-D-Ser-Gly-(Lys-D-FOHOrn-Ser)	Amann <i>et al.</i> , 2000
	Gwose ou 1.2	R-ChrP-Ser-Thr-D-Ser-Orn-OHAsp-(D-Gln-Dab)-Ser-D-aThr-cOHOrn	Gwose et Taraz, 1992
	Pflii ou GM	R-ChrP-D-Ala-D-Lys-Gly-Gly-D-OHAsp-D-Gln-D-Ser-Ala-D-Ala-D-Ala-Ala-cOHOrn	Mohn <i>et al.</i> , 1990
	Pfl12	R-ChrP-D-Ser-Lys-Gly-FoOHOrn-D-Ser-Ser-Gly-(Lys-D-FoOHOrn-Glu-Ser)	Geisen <i>et al.</i> , 1992
	Pf5*	R-ChrP-Asp-FoOHOrn-Lys-(Thr-Ala-Ala-FOHOrn-Lys)	Paulsen <i>et al.</i> , 2005
	TR133	R-ChrP-Asp-D-FoOHOrn-Lys-Thr-Ala-Ala-D-FoOHOrn-Ala	Barelmann <i>et al.</i> , 2003
	2798	R-ChrP-Ser-Dab-Gly-Ser-D-OHAsp-Ala-Gly-D-Ala-cOHOrn	Demange <i>et al.</i> , 1990
	SBW25	R-ChrP-D-Ser-Lys-Gly-FOHOrn-(Lys-Ser-D-FoOHOrn)	Moon <i>et al.</i> , 2008
	2392	R-ChrP-D-Lys-D-AcOHOrn-Gly-D-aThr-Thr-Gln-Gly-D-Ser-cOHOrn	Beiderbeck <i>et al.</i> , 1999
	pseudobactin A	R-ChrD-Lys-D-OHAsp-Ala-D-aThr-Ala-D-cOHOrn	Teintze <i>et al.</i> , 1981
	pseudobactin A225-I*	R-ChrP-Ser-Ala-AcOHOrn-Gly-Ser-OHAsp-Ser-Thr	Khalil-Rizvi <i>et al.</i> , 1997
	pseudobactin A225-II*	R-ChrP-Ser-Ala-AcOHOrn-Gly-Ser-OHAsp-Ser-Thr	Khalil-Rizvi <i>et al.</i> , 1997
	1547	R-ChrP-D-Ser-Lys-D-Ala-AOHOrn-Thr-D-Ala-Gly-D-Gln-D-Ala-Ser-Ser-cOHOrn	Ruangviriyachai <i>et al.</i> , 2000

<i>Pseudomonas aeruginosa</i>	R'	R-ChrP-(D-Ser-Dab)-FOHOrn-D-Gln-FOHOrn-Gly	Ruangviriyachai <i>et al.</i> , 2001
	Pa27853 ou TII	R-ChrP-D-Ser-D-FOHOrn-Orn-Gly-D-aThr-Ser-cOHOrn	Tappe <i>et al.</i> , 1993
	Pa6 "R"	R-ChrP-(D-Ser-Dab)-FOHOrn-Gln-D-Gln-D-FOHOrn-Gly	Gipp <i>et al.</i> , 1991
	I-III	R-ChrP-Asn-D-FoHOrn-Lys-Thr-Ala-Ala-D-FoHOrn-Lys	Briskott <i>et al.</i> , 1986
	PA7	R-ChrP-D-Ser-OHOrn-Orn-Gly-D-aThr-Ser-cOHOrn	Budzikiewicz <i>et al.</i> , 2004
	PA01 ou 15692 ou C-E	R-ChrP-D-Ser-Arg-D-Ser-FOHOrn-(Lys-FOHOrn-Thr-Thr)	Briskott <i>et al.</i> , 1989
<i>Pseudomonas sp.</i>	B10	R-ChrP-εLys-D-OHAsp-Ala-D-aThr-Ala-D-cOHOrn	Teintze <i>et al.</i> , 1981
	2908*	R-ChrP-Ser-Orn-OHAsp-Ser-Ser-Ser-cOHOrn	Vossen et Taraz, 1999
	Ps 6-10	R-ChrP-D-Ala-Orn-OHAsp-Dab-AOHOrn-Lys	Budzikiewicz <i>et al.</i> , 2006
	G85*	R-ChrP-Ser-Lys-OHAsp-Ser-Orn-Ser-cOHOrn	Meyer <i>et al.</i> , 2008
	96-312	R-ChrP-D-Ser-Ser-FOHOrn-Lys-FOHOrn-Lys-Ser)	Schlegel <i>et al.</i> , 2001
	D47	R-ChrP-D-Ser-Orn-FOHOrn-(Lys-D-FOHOrn-Glu-Ser)	Schäfer <i>et al.</i> , 2006
	96-188	R-ChrP-D-Ser-Lys-FOHOrn-(Lys-D-FOHOrn-D-Glu-D-Ser)	Weber <i>et al.</i> , 2000
	G76*	R-ChrP-Ser-Ser-FOHOrn-Ser-Ser-(Lys-Ser-FOHOrn)	Meyer <i>et al.</i> , 2008
	HR6*	R-ChrP-Asp-εLys-OHAsp-Ser-Ser-Thr-Thr-Thr-cOHOrn	Meyer <i>et al.</i> , 2008
	LBSA1*	R-ChrP-Asp-Arg-AOHOrn-Lys-Ser-Asp-cOHOrn	Meyer <i>et al.</i> , 2008
	96-318	R-ChrP-D-Ser-Orn-FOHOrn-D-Ser-Ser-(Lys-FOHOrn-Ser)	Schlegel <i>et al.</i> , 2001
	IB3	R-ChrP-Ser-Ala-Thr-Lys-Orn-AcOHOrn-Thr-Thr-Ala-Ser-Thr-Ala-Ala-cOHOrn	Meyer <i>et al.</i> , 2008
<i>Pseudomonas tolaasii</i>	2192 ou Tol	R-ChrP-D-Ser-Lys-Ser-D-Ser-Thr-D-Ser-AcOHOrn-Thr-D-Ser-cOHOrn	Demange <i>et al.</i> , 1990
<i>Pseudomonas syringae</i>	19310	R-ChrP-Lys-D-OHAsp-Thr-(Thr-Ser-D-OHAsp-Ser)	Jülich <i>et al.</i> , 2001
<i>Pseudomonas cichorii</i>	PaB ou cic	R-ChrP-Lys-D-OHAsp-Thr-(Thr-Gly-D-OHAsp-Ser)	Bultreys <i>et al.</i> , 2004
<i>Pseudomonas aureofaciens</i>	Pau	R-ChrP-D-Ser-D-Ac-OHOrn-Gly-D-aThr-Thr-Gln-Gly-D-Ser-D-cOHOrn	Beiderbeck <i>et al.</i> , 1999
<i>Pseudomonas chlororaphis</i>	TR133	R-ChrP-Asp-D-FoHOrn-Lys-(Thr-Ala-Ala-D-FOHOrn-Ala)	Barelmann <i>et al.</i> , 2003
<i>Pseudomonas montelii</i>	Lille 1*	R-ChrP-Asp-Lys-AcOHOrn-Ala-Ser-Ser-Gly-Ser-cOHOrn	Meyer <i>et al.</i> , 2008
<i>Pseudomonas rhodesiae</i>	Lille 25*	R-ChrP-Ser-Lys-FoHOrn-Ser-Ser-Gly-(Lys-FOHOrn-Ser-Ser)	Budzikiewicz <i>et al.</i> , 2004

Tableau 3: Liste des 80 molécules de pyoverdine dont la séquence peptidique est connue. Les acides aminés entre () font partie d'un cycle dans la molécule. ChrP : chromophore de pyoverdine, ChrD : chromophore de dihydropyoverdine, R : chaîne latérale du chromophore. Les noms de souches suivis d'un * sont les souches pour lesquelles l'isomérisation des monomères constituant la chaîne peptidique de la pyoverdine n'a pas été déterminée. Il existe deux autres molécules de pyoverdines qui ont été étudiées, la pyoverdine ML45 produite par *Pseudomonas thivervalensis* dont la structure n'a pas été déterminée et la pyoverdine G17 produite par *Pseudomonas fuscovaginae* dont la composition en acides aminés est connue mais pas la structure primaire.

VI.1.1. Structure générale

La première structure de pyoverdine a été établie par cristallographie au début des années 80 (Teintze *et al.*, 1981). Plus de 60 structures sont aujourd'hui connues grâce aux outils d'analyse tels que la résonance magnétique nucléaire (RMN) et la spectrométrie de masse (Meyer *et al.*, 2008 ; Matthijs *et al.*, 2009) . La structure des molécules de pyoverdines peut être divisée en trois parties distinctes (Budzikiewicz , 2004) (Figure 21) : une chaîne peptidique comportant des acides aminés sous forme L ou D (de 6 à 14 acides aminés en fonction de la souche productrice), un chromophore (ou acide (1S)-5-amino-2,3-dihydro-8,9-dihydroxy-1H-pyrimido-[1,2-a]quinoline-1-carboxylique) et un acide di-carboxylique ou sa forme amide. La molécule de pyoverdine a donc la structure générale suivante :

Chaîne peptidique – chromophore - chaîne latérale

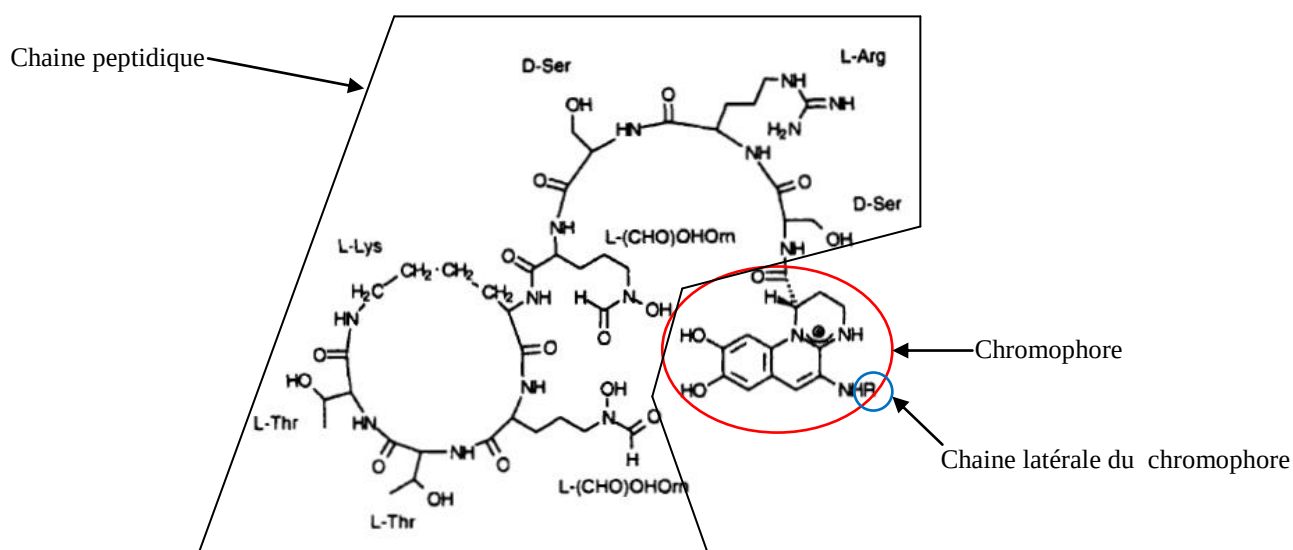


Figure 21: Structure de la pyoverdine produite par *Pseudomonas aeruginosa* PAO1 (Meyer *et al.*, 1997). Cette figure montre que la structure de la molécule de pyoverdine se compose de trois parties : la chaîne peptidique, le chromophore et sa chaîne latérale représenté par le radical R (acide dicarboxylique ou sa forme amide).

VI.1.1.1. La chaîne peptidique

La partie peptidique des pyoverdines est composée de 6 à 14 monomères sous forme L ou D, qui peuvent, entre autres, être acétylés, formylés. Ce peptide est lié au chromophore par une liaison amide entre l'acide aminé à l'extrémité N-terminale et l'acide carboxylique

porté par le C1 du chromophore ou C3 dans le cas des isopyoverdines (Figure 22). La molécule de pyoverdine comporte trois sites destinés à la coordination du fer ferrique (Fe^{3+}) au sein du sidérophore, deux de ces sites se trouvent sur la chaîne peptidique sous la forme de groupements hydroxamates ou hydroxycarboxylates (Budzikiewicz, 2004, Caboche *et al.*, 2010)).

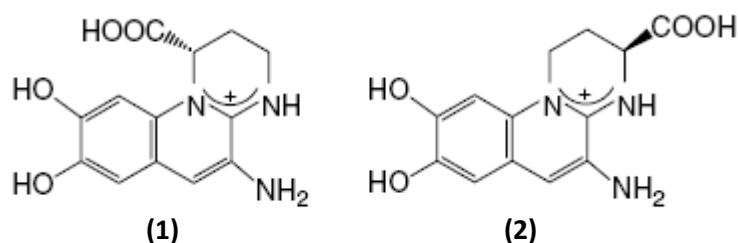


Figure 22: Représentation schématique du chromophore de pyoverdine avec localisation des atomes de carbones sur lesquels se fixent les chaînes peptidiques. Fixation de la chaîne peptidique par le groupement carboxyle porté par le C1 du chromophore pour les pyoverdines **(1)** et par le groupement carboxyle porté par le C3 du chromophore pour les isopyoverdines **(2)** (Budzikiewicz *et al.*, 2006).

VI.1.1.2. Le chromophore

Le chromophore est responsable des propriétés de fluorescence de la molécule lorsqu'elle n'est pas liée au fer et possède le troisième site de coordination du fer par l'intermédiaire d'un groupement catécholate. Lorsque le sidérophore est complexé à une molécule de fer, il prend une forme cyclique et perd sa fluorescence. La production des pyoverdines peut être accompagnée par la production de composés apparentés qui sont considérés comme leurs précurseurs biosynthétiques. Ces composés possèdent la même chaîne peptidique mais diffèrent par la structure de leur chromophore (Budzikiewicz, 2004). Les différentes structures de chromophores sont reprises sur la figure 23. Le chromophore le plus commun chez les pyoverdines est le n°1, le n°2 est le chromophore retrouvé chez les isopyoverdines qui ne diffèrent que par la position de l'acide carboxylique ici porté par le C-3 du chromophore.

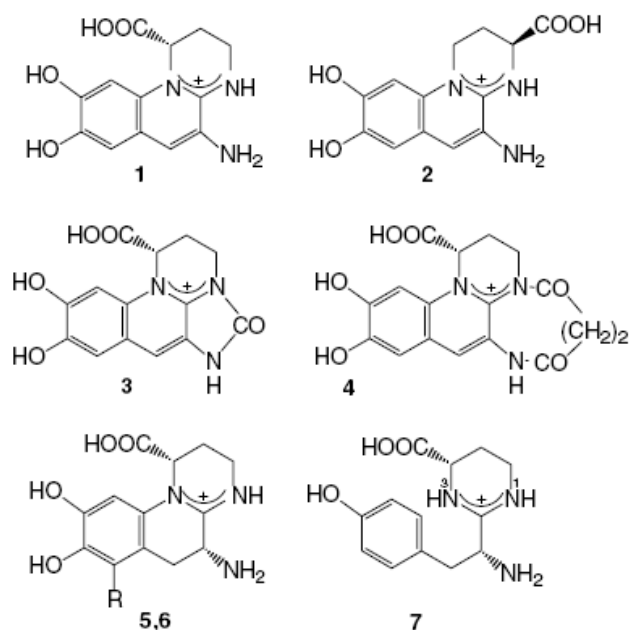


Figure 23: Structures des différents chromophores de sidérophores. 1. Pyoverdine, 2. Isopyoverdine, 3. Azotobactine, 4. Succinopyoverdine, 5. R=H, dihydropyoverdine, 6. R=SO₃H, dihydropyoverdine-7-sulfonique acide, 7. Ferribactine (Budzikiewicz *et al.*, 2006).

VI.1.1.3. La chaîne latérale du chromophore

L'acide di-carboxylique est lié au chromophore via une liaison amide avec le groupement NH₂ porté par le C-5 du chromophore. Cette petite chaîne latérale du chromophore est dérivée du cycle de l'acide citrique (Schäfer *et al.*, 1991). Six chaînes latérales différentes ont été observées liées au chromophore : l'acide glutamique et l'acide α-cétoglutarique, l'acide succinique et la succinamide (forme amide), et l'acide malique et la malamide (forme amide). Une souche de *Pseudomonas* peut produire dans un même milieu plusieurs molécules de pyoverdine, séparables par chromatographie, ne variant que par la nature de cette chaîne latérale (Budzikiewicz *et al.*, 1993).

VI.1.2. Biosynthèse des pyoverdines

Le chromophore (Mossialos *et al.*, 2002) et la chaîne peptidique des pyoverdines (Lehoux *et al.*, 2000) sont tous deux synthétisés par le mécanisme NRPS. La chaîne latérale du chromophore est quant à elle dérivée du cycle de l'acide citrique comme discuté ci-dessus.

Jacques Ravel et Pierre Cornelis (2003) ont étudié l'organisation des synthétases chez différentes souches de *Pseudomonas* produisant des pyoverdines (Figure 24). Sur cette figure,

l'ensemble des gènes intervenant dans la synthèse de la molécule de pyoverdine proprement dite ainsi que ceux intervenant dans le transport ou encore la régulation de ce type de molécules sont représentés.

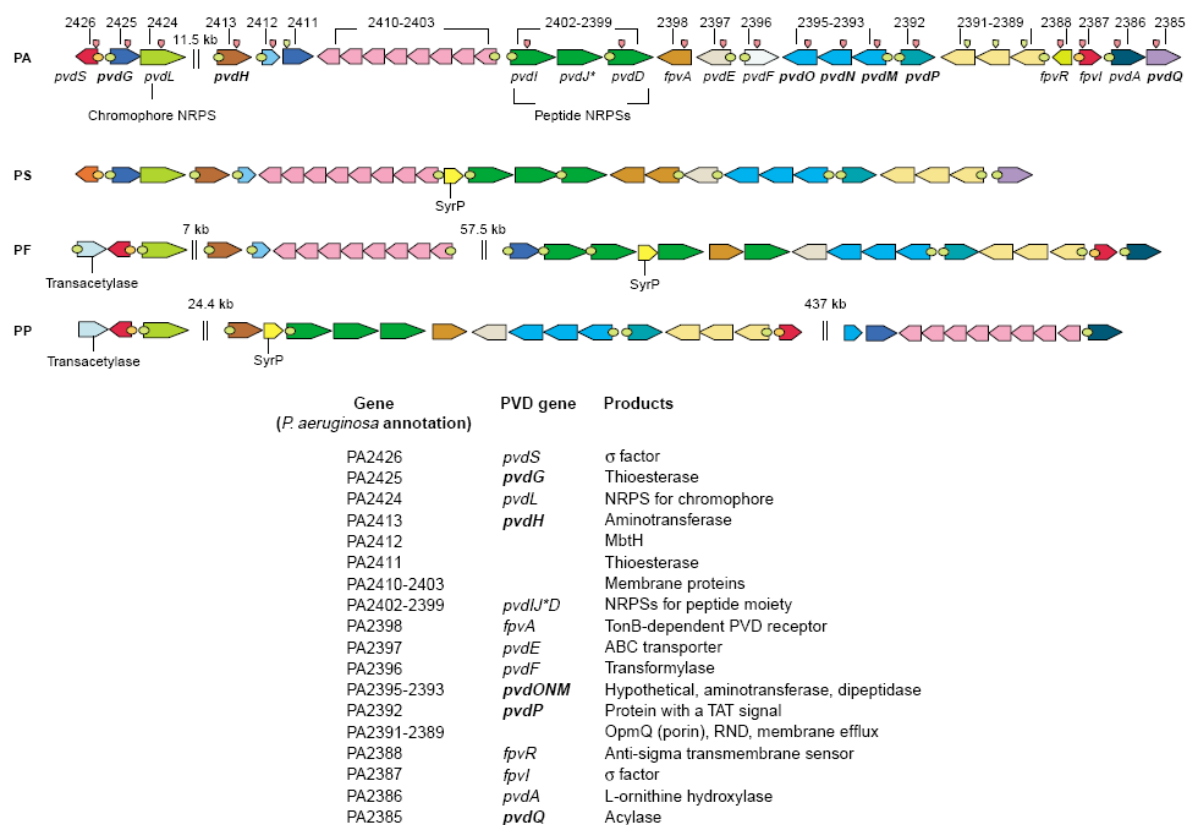


Figure 24: (Ravel et Cornelis, 2003) Organisation des clusters de gènes de la pyoverdine synthétase chez différentes espèces de *Pseudomonas*. PA :*Pseudomonas aeruginosa*, PS :*Pseudomonas syringae*, PF :*Pseudomonas fluorescens*, PP :*Pseudomonas putida*. Dans cette figure, les gènes impliqués dans la synthèse de la molécule de pyoverdine proprement dite sont indiqués pour la chaîne peptidique en vert foncé (*pvdI*, *pvdJ* et *pvdD*) et pour le chromophore en vert pale (*pvdL*).

VI.1.2.1. Biosynthèse de la chaîne peptidique

La chaîne peptidique des pyoverdines est synthétisée par le mécanisme de synthèse non ribosomiale. Les gènes impliqués dans la synthèse de la chaîne peptidique de la pyoverdine produite par *Pseudomonas aeruginosa* PAO1 ont été identifiés et nommés *pvdI*, *pvdJ* et *pvdD* (Figure 24) (Ravel et Cornelis, 2003). Les gènes *pvdI*, *pvdJ* et *pvdD* ont été identifiés comme étant annotés respectivement PA2402, PA2400 et PA2399 dans le génome de la souche *Pseudomonas aeruginosa* PAO1. La synthétase de la chaîne peptidique est composée de 8 modules répartis comme suit : PA2402 code pour 4 modules, PA2400 code

pour 2 modules et PA2399 code pour 2 modules. La chaîne peptidique de la molécule de pyoverdine de cette souche est donc composée de 8 monomères.

VI.1.2.2. Biosynthèse du chromophore

Les ferribactines sont des analogues structuraux des pyoverdines à la différence qu'elles possèdent un chromophore immature, elles sont considérées comme le précurseur des pyoverdines. Différents travaux ont tenté de montrer le mécanisme de synthèse du chromophore des pyoverdines (Taraz *et al.*, 1991 ; Nowak-Thompson et Gould, 1994). En 1999, Stinzi *et al.* ont identifié un opéron de 4 gènes (*pvcA-D*) impliqué dans la maturation du chromophore de la pyoverdine produite par *Pseudomonas aeruginosa* mais Clarke-Pearson et Brady ont montré en 2008 que les enzymes codés par ces gènes n'intervenaient pas dans la maturation du chromophore. La maturation du chromophore, qui est commun à toutes les pyoverdines, implique une série de réactions d'oxydo-réduction d'un produit de condensation d'une D-tyrosine (D-Tyr) et d'un acide L-2,4-diaminobutyrique (L-Dab) (Dorrestein *et al.*, 2003). Le gène *pvdL*, gène conservé au sein des génomes de *Pseudomonas* fluorescents, code pour une synthétase NRPS impliquée dans l'incorporation de la D-tyrosine et de l'acide L-2,4-diaminobutyrique (Mossialos *et al.*, 2002) dans le sidérophore.

VI.1.3. Régulation de la synthèse des pyoverdines

La plupart des micro-organismes, lorsqu'ils se développent dans des conditions en fer soluble limitantes, ils produisent des molécules chélatrices du fer pour satisfaire aux besoins métaboliques auxquels le fer est lié. Lors d'une carence en fer, les pyoverdines sont produites en excès dans le milieu environnant et vont fixer le fer sous sa forme ferrique (Fe^{3+}), la forme la plus représentée. Les complexes sidérophores/fer appelés ferrisidérophores, sont reconnus par des récepteurs sur la membrane externe des micro-organismes, ces récepteurs sont des porines qui vont permettre le transport du ferrisidérophore vers le cytoplasme. Dans le cytoplasme, le fer ferrique est réduit par l'action d'une réductase en fer ferreux (Fe^{2+}), l'affinité du sidérophore pour le fer ferreux est faible donc ce dernier est libéré du complexe sidérophores/fer. Le sidérophore est recyclé et le fer est stocké, et peut être utilisé directement par le micro-organisme pour satisfaire ses besoins métaboliques. Lorsque le fer ferreux est en

concentration importante dans le milieu dans lequel évolue le micro-organisme, il va diffuser à travers la membrane bactérienne et agir comme un répresseur de la synthèse des sidérophores et de leurs récepteurs membranaires.

VI.2. Les lipopeptides cycliques de *Pseudomonas*

Les lipopeptides cycliques (CLPs) sont produits par plusieurs *Pseudomonas*, incluant des pathogènes comme *Pseudomonas syringae*, *Pseudomonas tolaasii*, *Pseudomonas fuscovaginae*, *Pseudomonas corrugata*, et par plusieurs souches classées comme antagonistes comme *Pseudomonas fluorescens* et *Pseudomonas putida* (Nielsen *et al.*, 2002). Les CLPs sont des composés formés d'une chaîne lipidique d'une longueur variant de 6 à 16 atomes de carbone et d'une chaîne peptidique de 8 à 25 acides aminés.

En fonction de la nature et de la longueur de leurs chaînes lipidique et peptidique, les lipopeptides de *Pseudomonas* sont classés en six groupes (Gross et Loper, 2009). Ces lipopeptides ont tous une activité antibiotique et surfactante.

- **les viscosines**

Les viscosines sont des lipopeptides qui sont composés d'un acide gras ($C_{10}HO$ ou $C_{12}HO$) et de 9 monomères. La massetolide A (NOR00333) est un lipopeptide produit par *Pseudomonas fluorescens* qui appartient à ce groupe. Dans Norine, 14 lipopeptides appartiennent à ce groupe.

- **les syringomycines**

Les syringomycines sont des lipopeptides qui sont composés d'un acide gras (3 hydroxy-fatty acyl) et de 9 monomères. La cormycine A (NOR00360) est un lipopeptide produit par *Pseudomonas corrugata* qui appartient à ce groupe. Dans Norine, 11 lipopeptides appartiennent à ce groupe.

- **les amphisines**

Les amphisines sont des lipopeptides qui sont composés d'un acide gras ($C_{10}HO$) et de 11 monomères. L'arthrofactine (NOR00343) est un lipopeptide produit par *Pseudomonas sp.* MIS38 qui appartient à ce groupe. Dans Norine, 5 lipopeptides appartiennent à ce groupe.

- **les putisolvines**

Les putisolvines sont des lipopeptides qui sont composés d'un acide gras ($CH_3(CH_2)_4CO$) et de 12 monomères. La putisolvine I (NOR00361) est un lipopeptide produit par *Pseudomonas putida* PCL1445 qui appartient à ce groupe. Dans Norine, 3 lipopeptides appartiennent à ce groupe.

- **les tolaasines**

Les tolaasines sont des lipopeptides à activité antibiotique, surfactante et de toxine, ces lipopeptides sont composés d'un acide gras (C_8HO ou $C_{10}HO$) et de 18 à 22 monomères. La tolaasine A (NOR00875) est un lipopeptide produit par *Pseudomonas tolaasii* qui appartient à ce groupe. Dans Norine, 7 lipopeptides appartiennent à ce groupe.

- **les syringopeptines**

Les syringopeptines sont des lipopeptides à activité antibiotique et de toxine, ces lipopeptides sont composés d'un acide gras ($CH_3(CH_2)_{2-8}CH(OH)CH_2CO$) et de 22 à 25 monomères. La syringopeptine 25A (NOR00355) est un lipopeptide produit par *Pseudomonas syringae* qui appartient à ce groupe. Dans Norine, 11 lipopeptides appartiennent à ce groupe.

Les CLPs sont synthétisés par le mécanisme NRPS et plusieurs synthétases impliquées dans la synthèse de lipopeptides cycliques ont été décrites dans la littérature. Par exemple, l'organisation modulaire de la synthétase de l'orfamide A (Loper et Gross, 2007) (Figure 25) ou celle de la synthétase de la syringafactine A (Berti *et al.*, 2007) (Figure 26) ont été mises en évidence. Le tableau 4 reprend l'ensemble des CLPs connus produits par les bactéries du genre *Pseudomonas*.

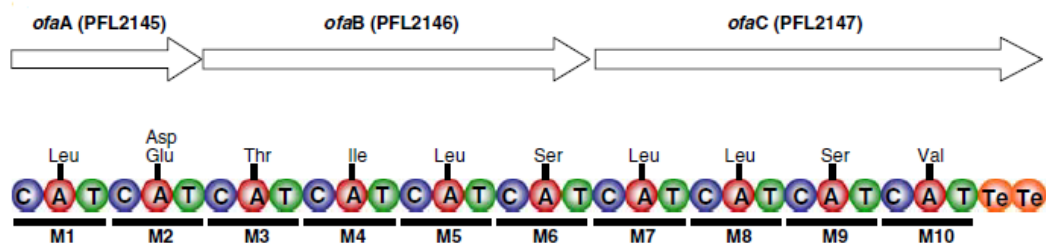


Figure 25: (Loper et Gross, 2007) Représentation de l'opéron de gènes de synthèse de l'orfamide A et de la synthétase codée par ce cluster, avec la prédiction de la composition de l'orfamide A.

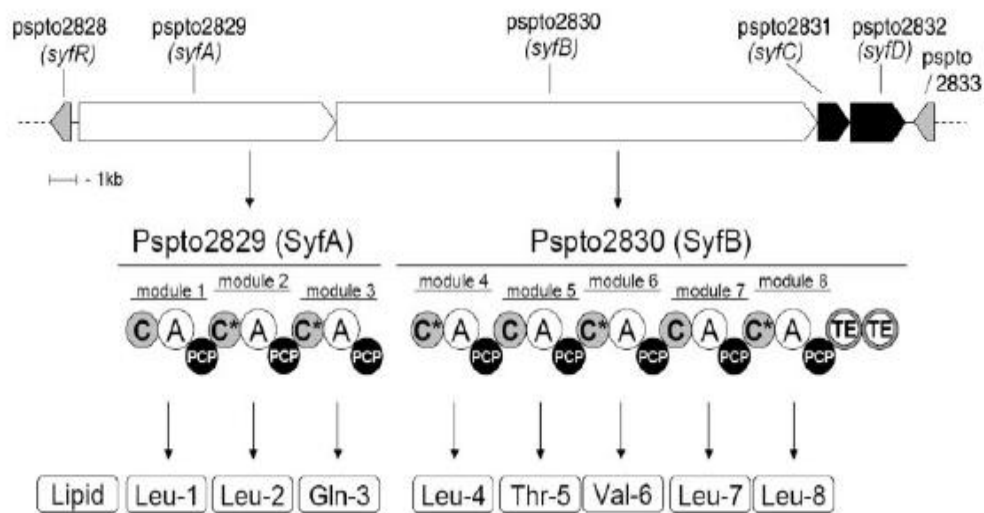


Figure 26: (Berti *et al.*, 2007) Représentation du cluster de gènes de la syringafactine A, de la structure de la synthétase et de la composition de la syringafactine A (NOR01075) produit par cette synthétase .

souches productrices	lipopeptides	séquence peptidique	références
<i>P. sp.</i> DSS73	amphisine	C10:0-OH(3)_D-Leu_D-Asp_D-aThr_D-Leu_D-Leu_D-Ser_Leu_D-Gln_Leu_Ile_Asp	Sorensen <i>et al.</i> , 2001
<i>P. sp.</i> MIS38	arthrofactine	C10:0-OH(3)_D-Leu_D-Asp_D-Thr_D-Leu_D-Leu_D-Ser_Leu_D-Ser_Ile_Ile_Asp	Morikawa <i>et al.</i> , 1993
<i>P. sp.</i> DS41	lokisine	C10:0-OH(3)_D-Leu_D-Asp_D-aThr_D-Leu_D-Leu_D-Ser_Leu_D-Ser_Leu_Ile_Asp	Sorensen <i>et al.</i> , 2002
<i>P. fluorescens</i>	pholipeptine	C10:0-OH(3)_D-Leu_Asp_Thr_D-Leu_D-Leu_D-Ser_D-Leu_D-Ser_D-Leu_Ile_D-Asp	Ui <i>et al.</i> , 1997
<i>P. fluorescens</i> 96.578	tensine	C10:0-OH(3)_D-Leu_D-Asp_D-aThr_D-Leu_D-Leu_D-Ser_-Leu_D-Gln_Leu_Ile_Glu	Henriksen <i>et al.</i> , 2000
<i>P. corrugata</i>	corpeptine A	C10:0-OH(3)_dhAbu_Pro_Ala_Ala_Ala_Val_Val_dhAbu_Hse_Val_aIle_dh-Ala_Ala_Ala_Ala_Val_dhAbu_aThr_Ala_Dab_Ser_Ile	Emanuele <i>et al.</i> , 1998
	corpeptine B	C12:1(5)-OH(3)_dhAbu_Pro_Ala_Ala_Ala_Val_Val_dhAbu_Hse_Val_aIle_dh-Ala_Ala_Ala_Ala_Val_dhAbu_aThr_Ala_Dab_Ser_Ile	
<i>P. corrugata</i>	cormycine A	C16:0-OH(3.4)_Ser_D-Orn_Asn_D-Hse_His_aThr_dhAbu_OH-Asp_4Cl-Thr	Scaloni <i>et al.</i> , 2004
<i>P. fuscovaginae</i>	fuscopeptine A	C8:0-OH(3)_dhAbu_D-Pro_Leu_D-Ala_D-Ala_D-Ala_D-Ala_D-Val_Gly_D-Ala_D-Val_D-Ala-D-Val_dhAbu_D-aThr_Ala_Dab_D-Dab_Phe	Baré <i>et al.</i> , 1999
	fuscopeptine B	C10:0-OH(3)_dhAbu_D-Pro_Leu_D-Ala_D-Ala_D-Ala_D-Ala_D-Val_Gly_D-Ala_D-Val_D-Ala-D-Val_dhAbu_D-aThr_Ala_Dab_D-Dab_Phe	
<i>P. sp.</i>	plusbacine A1	C14:0-OH(3)_D-aThr_D-Ala_3OH-Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	Shoji <i>et al.</i> , 1992
	plusbacine A2	IC15:0-OH(3)_D-aThr_D-Ala_3OH-Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
	plusbacine A3	IC16:0-OH(3)_D-aThr_D-Ala_3OH-Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
	plusbacine A4	C16:0-OH(3)_D-aThr_D-Ala_3OH-Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
	plusbacine B1	C14:0-OH(3)_D-aThr_D-Ala_Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
	plusbacine B2	IC15:0-OH(3)_D-aThr_D-Ala_Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
	plusbacine B3	IC16:0-OH(3)_D-aThr_D-Ala_Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
	plusbacine B4	C16:0-OH(3)_D-aThr_D-Ala_Pro_Arg_D-OH-Asp_D-Ser_3OH-Pro_OH-Asp	
<i>P. putida</i> PCL1445	putisolvine I	C6:0_Leu_Glu_Leu_Ile_Gln_Ser_Val_Ile_Ser_Leu_Val_Ser	Kuiper <i>et al.</i> , 2004
	putisolvine II	C6:0_Leu_Glu_Leu_Ile_Gln_Ser_Val_Ile_Ser_Leu_Leu_Ser	
	putisolvine III	C6:0_Leu_Glu_Leu_Ile_Gln_Ser_Val_Ile_Ser_Leu_Ile_Ser	
<i>P. syringae</i> DC3000	syringafactine A	C10:0-OH(3)_Leu_Leu_Gln_Leu_Thr_Val_Leu_Leu	Berti <i>et al.</i> , 2007
	syringafactine B	C10:0-OH(3)_Leu_Leu_Gln_Leu_Thr_Leu_Leu_Leu	
	syringafactine C	C10:0-OH(3)_Leu_Leu_Gln_Leu_Thr_Ile_Leu_Leu	
	syringafactine D	C12:0-OH(3)_Leu_Leu_Gln_Leu_Thr_Val_Leu_Leu	
	syringafactine E	C12:0-OH(3)_Leu_Leu_Gln_Leu_Thr_Leu_Leu_Leu	
	syringafactine F	C12:0-OH(3)_Leu_Leu_Gln_Leu_Thr_Ile_Leu_Leu	
<i>P. syringae</i>	pseudomycine A	C14:0-OH(3.4)_Ser_D-Dab_Asp_Lys_Dab_aThr_dhAbu_OH-Asp_4Cl-Thr	Ballio <i>et al.</i> , 2004
	pseudomycine B	C14:0-OH(3)_Ser_D-Dab_Asp_Lys_Dab_aThr_dhAbu_OH-Asp_4Cl-Thr	

<i>P. syringae</i>	pseudomycine C	C16:0-OH(3.4)_Ser_D-Dab_Asp_Lys_Dab_aThr_dhAbu_OH-Asp_4Cl-Thr	Segre <i>et al.</i> , 1989
	pseudomycine C2	C16:0-OH(3)_Ser_D-Dab_Asp_Lys_Dab_aThr_dhAbu_OH-Asp_4Cl-Thr	
	syringomycine A1	C10:0-OH(3)_Ser_D-Ser_Dab_D-Dab_Arg_Phe_dhAbu_OH-Asp_4Cl-Thr	
	syringomycine E	C12:0-OH(3)_Ser_D-Ser_Dab_D-Dab_Arg_Phe_dhAbu_OH-Asp_4Cl-Thr	
	syringomycine G	C14:0-OH(3)_Ser_D-Ser_Dab_D-Dab_Arg_Phe_dhAbu_OH-Asp_4Cl-Thr	
<i>P. syringae</i>	syringostatine A	C14:0-OH(3)_Ser_D-Dab_Dab_D-Hse_Orn_aThr_dhAbu_OH-Asp_4Cl-Thr	Sorensen <i>et al.</i> , 1996
	syringostatine B	C14:0-OH(3.4)_Ser_D-Dab_Dab_D-Hse_Orn_aThr_dhAbu_OH-Asp_4Cl-Thr	
<i>P. syringae</i>	syringotoxine	C14:0-OH(3)_Ser_D-Dab_Gly_D-Hse_Orn_aThr_dhAbu_OH-Asp_4Cl-Thr	Ballio <i>et al.</i> , 1990
<i>P. syringae</i>	syringopeptine 25A	C10:0-OH(3)_dhAbu_D-Pro_D-Val_Ala_D-Ala_Val_D-Leu_D-Ala_D-Ala_dhAbu_D-Val_dhAbu_D-Ala_D-Val_D-Ala_D-Ala_dhAbu_D-aThr_D-Ser_D-Ala_Val_Ala_Dab_D-Dab_Phe	Scaloni <i>et al.</i> , 1997
<i>P. syringae</i>	syringopeptine 22	C10:0-OH(3)_dhAbu_D-Pro_D-Val_Val_D-Ala_D-Ala_D-Val_D-Val_dhAbu_D-Ala_D-Val_Ala_D-Ala_dhAbu_D-aThr_D-Ser_D-Ala_dhAbu_Ala_Dab_D-Dab_Tyr	Grgurina <i>et al.</i> , 2002
<i>P. tolaasii</i>	tolaasine A	Pda_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Ile_D-Hse_D-Dab_Lys	Bassarrello <i>et al.</i> , 2004
	tolaasine B	C8:0-OH(3)_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Ile_D-Hse_D-Dab_Lys	
	tolaasine C	C8:0-OH(3)_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Ile_D-Hse_D-Dab_Lys	
	tolaasine D	C8:0-OH(3)_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Leu_D-Hse_D-Dab_Lys	
	tolaasine E	C8:0-OH(3)_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Leu_Gly_D-Dab_Lys	
	tolaasine I	C8:0-OH(3)_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Leu_D-Hse_D-Dab_Lys	Jourdan <i>et al.</i> , 2003
	tolaasine II	C8:0-OH(3)_dhAbu_D-Pro_D-Ser_D-Leu_D-Val_D-Ser_D-Leu_D-Val_Val_D-Gln_Leu_D-Val_dhAbu_D-aThr_Leu_Gly_D-Dab_Lys	
<i>P. fluorescens</i>	massetolide A	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-alle_Leu_D-Ser_Leu_D-Ser_Ile	Gérard <i>et al.</i> , 1997
	massetolide B	C11:0-OH(3)_Leu_D-Glu_D-aThr_D-alle_Leu_D-Ser_Leu_D-Ser_Ile	
	massetolide C	C12:0-OH(3)_Leu_D-Glu_D-aThr_D-alle_Leu_D-Ser_Leu_D-Ser_Ile	
	massetolide D	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-alle_Leu_D-Ser_Leu_D-Ser_Leu	
	massetolide E	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-Val_Leu_D-Ser_Leu_D-Ser_Val	
	massetolide F	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-Val_Leu_D-Ser_Leu_D-Ser_Leu	
	massetolide G	C11:0-OH(3)_Leu_D-Glu_D-aThr_D-Val_Leu_D-Ser_Leu_D-Ser_Ile	
	massetolide H	C12:0-OH(3)_Leu_D-Glu_D-aThr_D-Val_Leu_D-Ser_Leu_D-Ser_Ile	
	massetolide L	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-alle_Leu_D-Ser_Leu_D-Ser_Val	De Bruijn <i>et al.</i> , 2008
<i>P. fluorescens</i>	pseudophomine A	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-Ile_D-Leu_D-Ser_Leu_D-Ser_Ile	Pedras <i>et al.</i> , 2003
	pseudophomine B	C12:0-OH(3)_Leu_D-Glu_D-aThr_D-Ile_D-Leu_D-Ser_Leu_D-Ser_Ile	
<i>P. fluorescens</i>	viscosine	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-Val_Leu_D-Ser_Leu_D-Ser_Ile	Laycock <i>et al.</i> , 1991

<i>P. fluorescens</i> DR54	viscosamide	C10:0-OH(3)_Leu_D-Gln_D-aThr_D-Val_Leu_D-Ser_Leu_D-Ser_Ile	Nielsen <i>et al.</i> , 1999
<i>P. fluorescens</i> BD5	pseudofactine I	X_Gly_Ser_Thr_Leu_Leu_Ser_Leu_Val	Janek <i>et al.</i> , 2010
	pseudofactine II	X_Gly_Ser_Thr_Leu_Leu_Ser_Leu_Leu	
<i>P. fluorescens</i> Pf5	orfamide A	X_Leu_Asp/Glu_Thr_Ile_Leu_Ser_Leu_Leu_Ser_Val	Gross <i>et al.</i> , 2007
<i>P. reactans</i>	WLIP	C10:0-OH(3)_Leu_D-Glu_D-aThr_D-Val_D-Leu_D-Ser_Leu_D-Ser_Ile	Coraiola <i>et al.</i> , 2006

Tableau 4 : Liste des lipopeptides produits par les bactéries du genre *Pseudomonas*.

Les lipopeptides de *Pseudomonas* possèdent différentes activités biologiques telles qu'antibiotique (arthrofactine, Morikawa *et al.* (1993)) et toxine (corpeptine, Emanuele *et al.* (1998)). Ces lipopeptides ont la capacité de diminuer la tension de surface (activité surfactante).

VI. Outils *in silico* d'analyse des NRPs

Le nombre de NRPs décrits étant en pleine expansion, différents auteurs ont mis en place des outils d'analyse des synthétases NRPS responsables de la synthèse de ces peptides. Les différents outils disponibles à ce jour sont décrits ci-dessous, chaque outil apporte un niveau de précision supplémentaire.

- NRPS/PKS (<http://www.nii.res.in/nrps-pks.html>)

NRPS-PKS est un logiciel permettant l'analyse de synthétases NRPS et PKS (Ansari *et al.*, 2004). Cet outil permet à partir de la séquence protéique sous format FASTA d'une protéine, de définir la composition en domaines et de définir la spécificité des domaines A. BLAST est utilisé pour rechercher dans la base de données les domaines similaires aux séquences composant la requête protéique. De cette manière sont identifiés les différents domaines constituant la synthétase. Ensuite pour chaque domaine A identifié, les 10 acides aminés du code Stachelhaus sont extraits afin de déterminer la spécificité de ce domaine par comparaison aux signatures de Stachelhaus présentes dans la base de données et de déterminer l'acide aminé spécifiquement incorporé par ce domaine A.

- NRPS predictor et NRPS predictor 2

(<http://ab.inf.uni-tuebingen.de/toolbox/index.php?view=domainpred>)

NRPS Predictor (Rausch *et al.*, 2005) est un outil dans lequel la spécificité des domaines A est caractérisée par 34 acides aminés. En effet, les 10 acides aminés du code Stachelhaus se trouvent dans un rayon de 5,5Å autour du site de reconnaissance du substrat

alors que les acides aminés situés dans un rayon de 8Å autour du site de reconnaissance du substrat peuvent intervenir dans ce mécanisme de reconnaissance. 34 acides aminés sont situés dans un rayon de 8Å autour du site de fixation du substrat et ce sont ces 34 acides aminés qui sont utilisés dans l'analyse par cet outil. Chaque acide aminé est représenté par 12 valeurs correspondant à ses propriétés physico-chimiques comme par exemple sa polarité, son volume ou son point isoélectrique, et chaque domaine A est caractérisé par un vecteur contenant 408 valeurs (34 acides aminés * 12). Les domaines A sont regroupés en fonction de la similarité de leur substrat en deux niveaux de clustering. Les « *large clusters* » regroupent les domaines A dont le substrat est similaire d'un point de vue physico-chimique, par exemple les domaines A qui activent les acides aminés qui ont une longue chaîne latérale chargée positivement (Orn, Lys, Arg). Les « *small clusters* » regroupent les domaines A dont le substrat est proche, comme les domaines A activant les acides aminés portant un noyau aromatique non polaire (Phe et Trp). Cet outil permet la détermination de la spécificité de chaque domaine A, à partir d'une séquence protéique, à travers trois résultats : le « *large cluster* », le « *small cluster* » et la prédiction de l'acide aminé activé à partir des 10 acides aminés du code Stachelhaus. Depuis avril 2011, une nouvelle version de cet outil est disponible : NRPS Predictor 2 (Röttig *et al.*, 2011). Cet outil permet une prédiction plus précise de l'acide aminé activé par un domaine A donné grâce à deux nouveaux niveaux de prédiction. Ces deux nouveaux éléments de prédiction sont le « *three cluster* », les substrats y sont regroupés en trois classes en fonctions de leur propriétés physico-chimiques (hydrophobes aliphatiques, hydrophobes aromatiques et hydrophiles), et le « *nearest neighbour* » qui donne le substrat possédant la signature la plus proche de celle extraite par l'analyse de la séquence protéique soumise (basé sur le code Stachelhaus).

- **PKS/NRPS analysis website** (<http://nrps.igs.umaryland.edu/nrps/>)

Cet outil permet la prédiction des domaines constituant la synthétase ainsi que la spécificité des domaines A à partir d'une séquence protéique au format FASTA codant pour une synthétase (Bachmann et Ravel, 2009). Contrairement au logiciel précédent, l'analyse de la spécificité ne s'effectue pas à partir des 10 acides aminés du code Stachelhaus mais à partir de 8 acides aminés réellement impliqués dans la fixation du substrat. En effet, l'étude de plus de 150 domaines A a permis de montrer que l'utilisation de 8 des acides aminés du code Stachelhaus était suffisant pour prédire la spécificité de plus de 80% des domaines A (Challis *et al.*, 2000). La cystéine en position 331 a été éliminée des acides aminés critiques car sa

chaîne latérale est en dehors du site de fixation du substrat ainsi que la lysine en position 517 car elle est conservée au sein de tous les domaines A. Les domaines composant la synthétase sont déterminés par rapport aux motifs conservés pour chaque domaine. Ensuite les 8 acides aminés impliqués dans la liaison au substrat sont extraits grâce à un alignement avec la séquence protéique du domaine A identifié comme sélectionnant la phénylalanine (PheA) dans la synthèse de la gramicidine S (Conti *et al.*, 1997). La séquence extraite est comparée à la base de données afin de déterminer la spécificité du domaine A.

- **NP.searcher** (<http://dna.sherman.lsi.umich.edu/>)

NP.searcher (Li *et al.*, 2009) permet à partir du téléchargement d'un génome entier de prédire les clusters de gènes NRPS/PKS ainsi que la séquence du produit de chacun de ces clusters. Cet outil permet une analyse préliminaire globale du potentiel NRPS d'un génome. NP.Searcher utilise BLAST pour comparer les séquences des signatures extraites avec celles présentes dans la banque de données dans le but de déterminer la spécificité des domaines A et AT.

- **Logiciel NRPS ProBioGEM**

Ce logiciel d'analyse des synthétases NRPS est issu d'une collaboration entre le laboratoire ProBioGEM et le laboratoire d'informatique fondamentale de Lille. Il permet à partir d'une séquence protéique sous format FASTA d'identifier les séquences consensus des différents domaines de la synthétase et de donner l'organisation modulaire de la synthétase codée par la séquence donnée en requête. Outre cette fonctionnalité, ce logiciel permet également de prédire la séquence du produit final de la synthétase étudiée avec une précision moins importante que les autres programmes cités ci-dessus.

- **Norine** (<http://bioinfo.lifl.fr/norine/>)

Norine est la première plateforme exclusivement dédiée aux molécules d'origine non ribosomiale (Caboche *et al.*, 2008), elle est issue de la collaboration entre deux équipes de recherche de l'Université de Lille 1 (le laboratoire ProBioGEM et le laboratoire d'informatique fondamentale de Lille (LIFL)). Elle regroupe aujourd'hui 1122 peptides rassemblés en 205 familles. La banque de données recense les informations telles que la structure du peptide ainsi que l'activité biologique et l'organisme producteur avec les

références bibliographiques associées. Norine propose également différents outils de recherche de peptides comme la recherche par structure qui permet de chercher un peptide par sa structure globale ou par homologie de motifs en fonction des monomères et des motifs d'association de ces monomères qu'il possède (Caboche *et al.*, 2009).

Deux nouveaux outils d'analyse des synthétases NRPS ont été récemment publiés, ils n'ont pas été utilisés dans ce travail.

- **AntiSMASH** (<http://antismash.secondarymetabolites.org/>)

Cet outil (Medema *et al.*, 2011) permet de déterminer des clusters de gènes de biosynthèse de métabolites secondaires ainsi que la prédiction de la spécificité au substrat du domaine AT pour les PKS et du domaine A pour les NRPS. La spécificité du substrat est déterminée par la signature des 24 acides aminés du site actif pour les domaines AT (PKS) et par, d'une part la méthode des signatures protéiques et de l'autre par NRPS Predictor 2, pour les domaines A (NRPS). La compilation des résultats permet la prédiction de la composition du produit de la synthétase.

- **Non ribosomal peptide synthase substrate predictor (NRPSsp)**
(<http://www.nrpssp.com/index.php>)

NRPSsp (Prieto *et al.*, 2011) possède sa propre banque de données qui associe un domaine d'adénylation (A) à un substrat spécifiquement incorporé. Cette banque de données permet de prédire la présence de domaines A dans une séquence sous format FASTA et de prédire la nature du monomère spécifiquement incorporé.

Matériel et Méthodes

II. Matériel et méthodes

II.1. Origine des produits

Les produits chimiques utilisés dans ce travail sont pour la plupart des produits Sigma Aldrich, les produits issus d'autres firmes seront précisés dans le texte.

II.2. Milieux de culture

Les deux milieux de culture suivants sont utilisés pour réaliser des pré-cultures bactériennes à partir de cellules conservées dans du glycérol 50%, V/V par congélation à -80°C.

- Milieu de Luria-Bertani (LB)

Le milieu LB est préparé à partir de 10 g de tryptone, 5 g de d'extrait de levure et de 10 g de NaCl pour un volume final de 1 litre. Le pH est ajusté à 7 par le NaOH puis le milieu est stérilisé 20 min à 121°C. Le milieu solide correspondant est obtenu en ajoutant 1,7% (p/v) d'agar.

- Milieu 869

Le milieu 869 est préparé à partir de 10 g de tryptone, 5g d'extrait de levure, 5g de NaCl, 1g de glucose et 0,345 g de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ pour un volume final de 1 litre. Le pH de ce milieu est amené à pH 7 au moyen de NaOH 40%. Le milieu est ensuite stérilisé 20 min à 121°C. Le milieu solide correspondant est obtenu en ajoutant 1,7% (p/v) d'agar.

La production de pyoverdine est influencée par la composition du milieu, la production de ce sidérophore est favorisée par une teneur élevée en phosphate dans le milieu ainsi que par une carence en fer dans le milieu de culture. Pour la préparation des milieux qui favorise la production de pyoverdine, l'ensemble de la verrerie utilisée a été traitée à l'acide nitrique 10% dans le but de s'affranchir de traces de fer adsorbées sur les parois. Le dipyridyl, agent chélatant, est ajouté au milieu de culture à raison de 200µM de concentration finale. Les milieux de culture suivants ont donc été choisis pour produire la pyoverdine :

- Milieu casamino acid (CAA)

Le milieu CAA est préparé à partir de 5 g de casaminoacides (Biokar Diagnostics), 0,9 g de K_2HPO_4 et de 0,25 g de $MgSO_4$ pour un volume final de 1 litre. Le pH est ajusté à 7 par du NaOH 1M et le milieu est stérilisé 20 min à 121°C.

- Milieu succinate

Le milieu succinate est préparé à partir de 6 g de K_2HPO_4 , 3 g de KH_2PO_4 , 4 g d'acide succinique, 1 g de $(NH_4)_2SO_4$ et de 0,2 g de $MgSO_4$ pour un volume final de 1 litre. Avant stérilisation le pH du milieu est ajusté à 7 par du NaOH 6M.

- Milieu minéral+glucose

Ce milieu est préparé à partir de 0,45 g de $MgSO_4$, 3 g de KH_2PO_4 , 6g de K_2HPO_4 , 0,39 g de $CaCl_2$ et 0,001 g de $CuSO_4$ pour un volume final de 1 litre. Le pH est ajusté à 7 et le milieu est stérilisé 20 min à 121°C. Après autoclavage, 5 g/L de glucose sont ajoutés au milieu après stérilisation de la solution sur filtre stérile 0,22 μm ainsi que 3 g/L d'acide aminé stérilisé de la même manière. Ce milieu est aussi utilisé sans acide aminé, la source d'azote est alors remplacée par 1g/L de $(NH_4)_2SO_4$.

- Milieu King B (King *et al.*, 1954)

Le milieu King B est composé de 20 g de mélange de peptones, 1,5 g de K_2HPO_4 , 1,5 g de $MgSO_4$ et de 14 g d'agar pour 1L de milieu. 37 g du milieu commercial (Laboratoire Conda S.A., Madrid, Espagne) sont mis en suspension dans 1L d'eau milliQ et 10 mL de glycérol sont ajoutés. Ce milieu est stérilisé à l'autoclave à 121°C pendant 15 min. Ce milieu de culture est utilisé pour la conservation des souches de *Pseudomonas* sur boîtes, elles sont ensuite stockées à 4°C.

II.3. Les souches bactériennes utilisées

Les souches bactériennes utilisées dans cette étude sont exclusivement des souches appartenant au genre *Pseudomonas* groupe fluorescent. Elles sont caractérisées par une fluorescence due à la production d'un sidérophore appelé pyoverdine ou pseudobactine.

Dans l'objectif d'effectuer un travail d'analyse bioinformatique sur ces souches, nous avons utilisé uniquement celles dont le génome était complètement séquencé au début du travail, vous trouverez dans le tableau 5 les souches de *Pseudomonas* fluorescents utilisés.

Souches	Taille du	Identifiant	Auteurs	Contacts
<i>Ps. aeruginosa</i> PA7	6,5 Mb	NC_009656	Cornelis <i>et al.</i> , 1989	P. Cornelis
<i>Ps. putida</i> KT2440	6,2 Mb	NC_002947	Matthijs <i>et al.</i> , 2009	P. Cornelis
<i>Ps. entomophila</i> L48	5,9 Mb	NC_008027	Vodovar <i>et al.</i> , 2006	F. Bocquaert
<i>Ps. putida</i> F1	5,9 Mb	NC_009512	Wackett et Gibson, 1988	disponible
<i>Ps. fluorescens</i> SBW25	6,7 Mb	NC_012660	Moon <i>et al.</i> , 2008	P.B. Rainey
<i>Ps. fluorescens</i> Pf5	7,1 Mb	NC_004129	Gross <i>et al.</i> , 2007	H. Gross
<i>Ps. syringae</i> pv. <i>Tomato</i>	6,4 Mb	NC_004578	Feil <i>et al.</i> , 2005	S. Lindow
<i>Ps. syringae</i> B728a	6,1 Mb	NC_007005	Feil <i>et al.</i> , 2005	S. Lindow
<i>Ps. fluorescens</i> Pf0-1	6,4 Mb	NC_007492	Meyer <i>et al.</i> , 2008	S. Matthijs
<i>Ps. Putida</i> W619	5,8 Mb	NC_010501	Weyens <i>et al.</i> , 2010	J.Vangronsveld
<i>Ps. putida</i> GB1	6,1 Mb	NC_010322	Gueszvain et Tebo, 2010	S. Matthijs
<i>Ps. brassicacearum</i> NFM421	6,8 Mb	NC_015379	Ortet <i>et al.</i> , 2011	S. Matthijs

Tableau 5 : Tableau des souches de *Pseudomonas* fluorescents utilisées dans cette étude.

Les souches sont conservées à long terme à -20°C et à -80°C dans du glycérol 50%. Elles sont repiquées régulièrement sur le milieu de conservation King B et conservées à 4°C.

Production de pyoverdines par les souches de *Pseudomonas* fluorescents

Les cultures liquides de *Pseudomonas* fluorescents sont réalisées dans un erlenmeyer de 500 mL avec 100 mL de milieu de culture. La verrerie utilisée est préalablement traitée à l'acide nitrique 10% afin d'éliminer au maximum les traces de fer présentes pour favoriser la production des pyoverdines, et stérilisée par autoclavage à 121°C pendant 20 min. L'inoculation est réalisée soit à partir de 100µL de pré-culture d'une nuit en milieu LB ou à partir de colonies conservées sur boîtes de King B à 4°C. Les cultures sont mises à incuber à 30°C sous 140 tpm pendant 24 à 48h jusqu'à atteindre une stabilité optique à 400 nm, qui permet de mesurer la concentration en pyoverdines à pH7 (Meyer et Abdallah, 1978). La culture est ensuite centrifugée à 13000g pendant 10 min et le surnageant est récupéré puis son pH ajusté à une valeur de 7 afin de limiter la dégradation des pyoverdines.

L'inoculation sur milieu solide est réalisée à partir d'une pré-culture en début de phase exponentielle et les milieux sont mis à incuber 48 h à 30°C.

II.4. Méthodes d'analyse

II.4.1. Méthodes d'analyse biologique

II.4.1.1. Concentration des échantillons

Les surnageants de culture de *Pseudomonas* sont concentrés par passage sur colonne C18 Altech (Grace-Darison, Templemars, France). Les différentes étapes qui vont suivre sont réalisées sous vide. La colonne est conditionnée par le passage de 20 mL de méthanol 100 % puis rincée par 10 mL d'eau ultra pure. 2 mL d'échantillon sont déposés sur la colonne et la colonne est séchée par passage d'air pendant 5 à 10 minutes. Ensuite l'échantillon est élué par le passage de 8 mL de méthanol 50 %. Le méthanol 50 % est évaporé grâce à un évaporateur sous vide (Speed Vac Plus SC110A, Savant, GMI, Ramsey, USA) jusqu'à l'obtention d'un dépôt sec. Ce dépôt sec est ensuite repris dans 200 µL d'eau milliQ pour analyse en spectrométrie de masse.

II.4.1.2. Spectrométrie de masse MALDI-TOF

Les analyses en spectrométrie de masse ont été réalisées à Gembloux AgrobioTech (Université de Liège) avec la collaboration du Professeur Bernard Wathelet de l'Unité de Chimie Biologique.

L'analyse en spectrométrie de masse d'un échantillon est réalisée par une série d'étapes successives. L'échantillon est solubilisé dans une matrice sur une cible, les molécules à étudier sont vaporisées et ionisées dans la source de l'appareil. Les ions ainsi formés sont séparés lors de cette étape en fonction de leur ratio masse/charge (m/z) par un analyseur et viennent terminer leur course sur le détecteur. Le spectre de masse est obtenu par traitement du signal émis par le détecteur.

Pour l'étude des molécules d'intérêt, la source d'ionisation choisie est le MALDI (Matrix Assisted Laser Desorption/Ionization) et l'analyseur est de type TOF (Time Of Flight). Le MALDI-TOF est utilisé en mode réflectron, celui-ci permet de dévier les ions avec un champ électrique doublant ainsi la longueur du chemin de vol de l'ion et augmente de manière significative la résolution de l'instrument. Plus la résolution est élevée, plus les pics seront fins sur le spectre de masse et il sera alors possible d'observer deux molécules de masses moléculaires très proches (précision de l'ordre du millième de dalton).

Les analyses ont été réalisées sur un spectromètre de masse Ultraflex Brüker Daltonics® en mode réflectron. Le potentiel d'accélération a été fixé à 20 kV, de pulse 50 ns avec un analyseur de type TOF.

La matrice utilisée est constituée d'une solution d'acide α -cyano-4-hydroxycinnammique (CHCA) à 10mg/mL dans un mélange 1:2 (v/v) d'acétonitrile et d'acide trifluoroacétique (TFA) 0,1%.

Les échantillons étant sous forme solide, l'équivalent de 1 μ L d'échantillon est prélevé à partir d'une colonie sur boîte de Pétri à l'aide d'un cône de pipette et solubilisé dans 20 μ L de la solution de matrice décrite ci-dessus. Les dépôts effectués sur la cible sont de 1 μ L pour les échantillons.

La solution de standard est préparée en mélangeant 1 μ L de peptide calibration standard (Brüker) à 9 μ L de solution de matrice. Le standard utilisé pour l'analyse de peptides est composé de 7 peptides de masses connues, ces peptides sont repris dans le tableau 6.

peptides	[M+H] ⁺ mono isotopic	[M+H] ⁺ moyen
Angiotensine II	1046,5420	1047,20
Angiotensine I	1296,6853	1297,51
Substance P	1347,7361	1348,66
bombesine	1619,8230	1620,88
ACTH Clip 1-17	2093,0868	2094,46
ACTH Clip 18-39	2465,1990	2466,73
Somatostatine 28	3147,4714	3149,61

Tableau 6: Composition en peptides de la solution de calibration du spectromètre de masse.

Avant analyse, les spots déposés sur la cible sont deshydratés. 150 tirs au laser sur chaque spot sont accumulés afin d'obtenir le spectre de masse final. Le logiciel Flexcontrol (Brüker) a été utilisé pour piloter le spectromètre de masse et le logiciel de traitement des résultats Flex analysis (Brücker) a été utilisé pour analyser les pics détectés et déterminer les masses correspondantes en mode positif.

II.4.2. Méthodes d'analyse bioinformatique

II.4.2.1. Protocole d'analyse bioinformatique des génomes de *Pseudomonas*

Le premier objectif de notre travail a été de caractériser au moyen d'outils bioinformatiques, le potentiel de synthèse de peptides non ribosomiaux des dix-sept souches de *Pseudomonas* dont le génome est complètement séquencé et accessible librement dans les banques de données. Le protocole utilisé pour l'analyse de ces génomes est décrit par la figure 27.

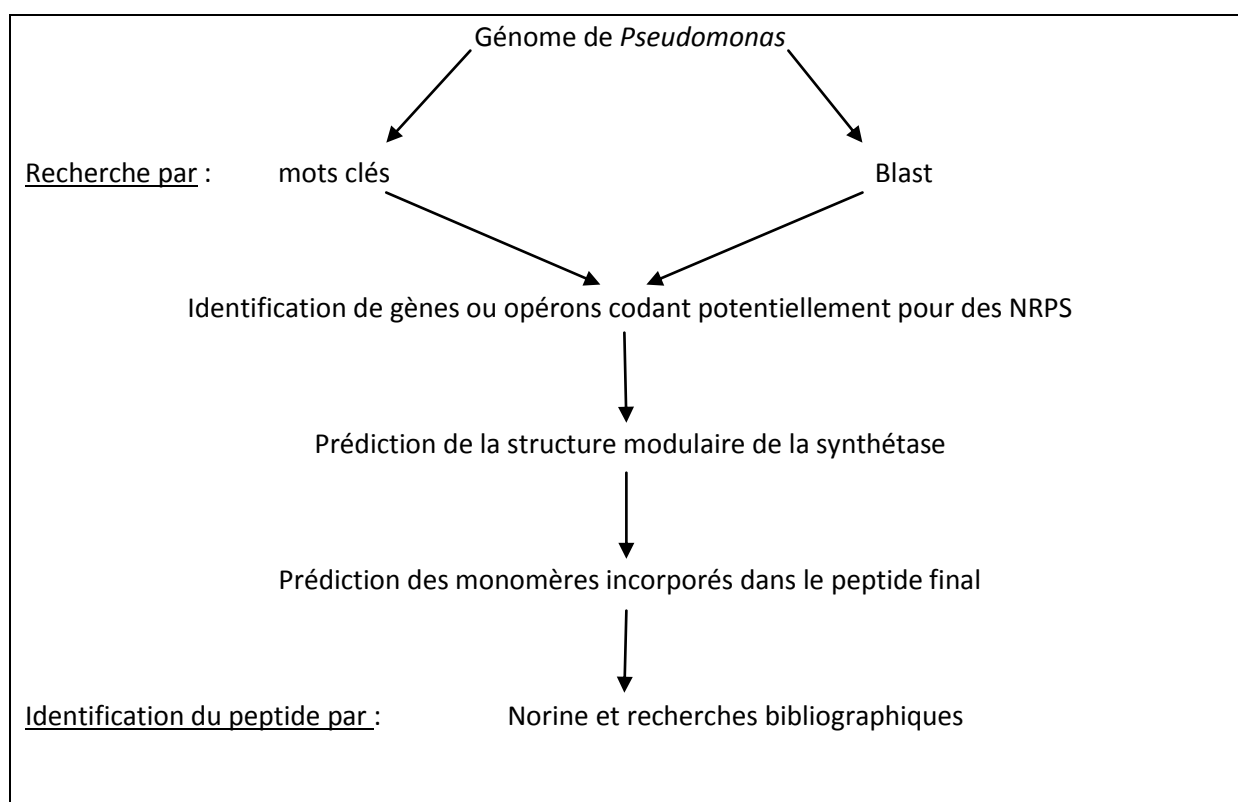


Figure 27: Stratégie générale d'exploration des génomes.

La première étape consiste à rechercher dans les génomes disponibles, des gènes codant potentiellement pour des synthétases NRPS, par deux approches. Soit une recherche en utilisant les mots-clés suivants « NRPS », « adenylation », « synthetase », « synthase », « peptide », « nonribosomal », « lipopeptide » et « siderophore » en se basant sur l'annotation qui a été faite pour ces gènes lors du séquençage du génome. Soit par une recherche par comparaison de séquences (BLAST) afin de déterminer s'il existe des similarités entre des gènes du génome étudié et d'autres gènes connus comme codant pour

des synthétases NRPS. Les séquences utilisées pour la recherche par BLAST ont été principalement la protéine PvdL (AAG05812) codant pour la synthèse du chromophore des molécules de pyoverdine ainsi que la protéine MycB (AAF08796) connue pour son implication dans la synthèse d'un lipopeptide chez *Bacillus subtilis*, la mycosubtiline.

La seconde étape consiste à prédire la structure des synthétases NRPS, c'est-à-dire leur organisation en modules puis en domaines et d'autre part de prédire le monomère potentiellement incorporé par le domaine d'adénylation de chaque module. Cette prédiction est réalisée à partir de la séquence protéique de la synthétase. Trois logiciels disponibles en ligne (NRPS predictor 2, PKS/NRPS analysis website et NRPS/PKS) permettent de réaliser cette analyse. La comparaison des résultats obtenus *via* ces 3 logiciels permet d'avoir une première prédiction de la composition du peptide final. Néanmoins, la plupart du temps, ces logiciels ne prennent pas en compte les domaines secondaires présents dans les modules. Par exemple, le domaine d'épimérisation (domaine E) permet la modification d'un acide aminé de sa forme L vers sa forme D. C'est pourquoi la compilation des résultats des 3 logiciels ainsi qu'une analyse précise de la composition modulaire est nécessaire pour obtenir une prédiction fiable du peptide.

Après avoir prédit la composition du peptide produit par la synthétase, la troisième étape d'analyse consiste en l'identification de ce peptide. Deux approches sont utilisées, d'une part une recherche par comparaison de structure grâce aux outils présents dans Norine qui permettent d'identifier des peptides ayant une structure proche de celle du peptide prédit et d'autre part, une recherche bibliographique afin de savoir si le peptide prédit a déjà été identifié.

II.4.2.2. Outils bioinformatiques d'analyse des synthétases et de prédiction de la composition du peptide final

Les outils d'études des synthétases NRPS cités ci-dessous ont été utilisés sur les années 2010 et 2011 dans le but d'analyser les différents génomes de *Pseudomonas* fluorescents disponibles.

- **NP.searcher** (<http://dna.sherman.lsi.umich.edu/>)

Cet outil est utilisé pour une analyse préliminaire du potentiel de synthèse NRPS au sein d'un génome. NP.Searcher analyse le génome entier donné en requête et nous donne les clusters de gènes responsables de la synthèse d'une NRPS ainsi que la composition du NRP produit. Cette analyse préliminaire permet de déterminer les génomes de *Pseudomonas* présentant un potentiel intéressant de synthèse NRPS.

Suite à cette analyse, les génomes sont étudiés grâce à une recherche par mots clés dans MBGD avec les mots clé cités précédemment à la recherche de synthétase NRPS de pyoverdines et de lipopeptides. Une analyse par BLAST est aussi réalisée sur le NCBI, en utilisant tBLASTn qui va permettre de rechercher un pourcentage d'identité entre une requête protéique (la protéine PvdL (AAG05812) ainsi que la protéine MycB (AAF08796) et un génome sélectionné. Les résultats croisés de la recherche par mots clés et par BLAST a permis de mettre en évidence les gènes codant pour des synthétases NRPS dans les génomes étudiés.

Les synthétases ainsi définies vont être analysées par les outils bioinformatiques suivants :

- **PKS/NRPS analysis website** (<http://nrps.igs.umaryland.edu/nrps/>)
- **NRPS/PKS** (<http://www.nii.res.in/nrps-pks.html>)

Ces deux premiers outils permettent de déterminer la composition de la synthétase étudiée en modules et domaines, ainsi que la prédiction de la séquence du produit de la synthétase en analysant les domaines A la composant permettant l'incorporation d'un acide aminé particulier. L'outil qui suit permet de prédire la composition du produit de la synthétase étudiée.

- **NRSPredictor et NRSPredictor 2**
(<http://nrps.informatik.uni-tuebingen.de/Controller?cmd=SubmitJob>)

Le fonctionnement des quatre outils d'analyse est décrit dans le paragraphe VII. de l'introduction. La compilation et l'interprétation des résultats, donnés sur l'étude d'un gène codant pour une synthétase NRPS, permet à la fois de prédire la composition en domaines de la synthétase ainsi que la composition en monomères du produit de cette dernière. En effet,

chacun de ces outils apporte une information complémentaire quant à la prédiction du produit des synthétases. La prédiction obtenue par l'analyse des gènes codant une synthétase NRPS est donc obtenue par une compilation et une interprétation croisée des résultats obtenus par ces différents outils. Pour l'ensemble de ces outils, la requête est donnée sous format FASTA.

II.4.2.3. Banques de données

- National Center for Biotechnology Information (NCBI)
(<http://www.ncbi.nlm.nih.gov/>)

Le NCBI est un centre national de recherche américain pour l'information en biologie moléculaire fondé en 1988. Le NCBI est spécialisé dans la recherche en bioinformatique et propose des banques de données publiques telles que GenBank, dbSNP et RefSeq qui contiennent des séquences nucléotidiques et/ou protéiques. Il développe aussi des outils pour l'analyse des génomes et fournit des informations biomédicales.

- Microbial Genome Database for Comparative Analysis (MBGD)
(<http://mbgd.genome.ad.jp/>)

MBGD est une banque de données japonaise pour l'analyse de génomes complètement séquencés (Uchiyama *et al.*, 2009). Elle regroupe en avril 2012, 1382 génomes bactériens, 116 génomes d'Archées et 34 génomes d'eucaryotes. L'interface de MBGD permet une analyse des génomes en utilisant une recherche par mots clés se basant sur l'annotation des gènes.

L'utilisation de ces deux banques de données a permis d'extraire les génomes de *Pseudomonas* entièrement séquencés et de les analyser, grâce à MBGD, à la recherche de gènes codant pour des NRPS potentielles par l'intermédiaire d'une recherche par mots clés.

- Norine (<http://bioinfo.lifl.fr/norine/>)

Norine a été utilisée pour la recherche de pyoverdines et de lipopeptides cycliques de *Pseudomonas* identifiés par l'analyse bioinformatique des génomes de *Pseudomonas* séquencés par l'intermédiaire de ces outils de comparaison de structure.

II.4.2.4. Outils de comparaison de séquences

L'alignement multiple permet d'aligner plusieurs séquences protéiques ou nucléiques afin de déterminer les éléments communs entre ces séquences pour mettre en évidence les zones similaires en termes d'acides aminés ou nucléotides. Différents outils sont disponibles en ligne afin d'effectuer ces alignements :

- **ClustalW** (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>)

ClustalW (Thompson *et al.*, 2002) est un outil d'alignement multiple de séquences protéiques ou nucléotidiques. Il permet d'identifier les zones similaires entre plusieurs séquences et propose la possibilité de représenter le résultat de l'alignement sous forme d'un arbre de distance.

- **Weblogo** (<http://weblogo.berkeley.edu/>)

Cet outil est une application disponible en ligne (Crooks *et al.*, 2004) qui génère à partir d'un alignement multiple de séquences des représentations graphiques appelées 'logo'. Chaque logo est composé d'une pile de lettres pour chaque position de l'alignement. La hauteur d'une lettre représente la fréquence à laquelle on retrouve cette lettre dans l'alignement. Pour déterminer la séquence consensus issue de l'alignement il convient de lire la ligne supérieure du logo qui correspond aux lettres les plus représentées à chaque position de l'alignement. La figure 28 montre un exemple de résultat d'alignement obtenu par weblogo.



Figure 28: Exemple de résultat obtenu par Weblogo à partir d'un alignement multiple de séquences protéiques. Les séquences utilisées pour générer ce logo sont des séquences de 34 acides aminés de domaines A de synthétase de pyoverdine impliquées dans la reconnaissance d'un même substrat, la sérine.

- **MUSCLE** (<http://www.drive5.com/muscle/>)

MUSCLE est un outil d'alignement multiple de séquences nucléotidiques et protéiques. Dans ce travail les séquences protéiques, au format FASTA, de différents domaines de synthétases NRPS de pyoverdines ont été utilisées. Le format de sortie choisi pour les résultats d'alignement est de type « Phylip » afin de pouvoir procéder à la construction d'un arbre de distance à partir de ces alignements et les options par défaut sont conservées.

II.4.2.5. Outil d'analyse des résultats de spectrométrie de masse

Les analyses en spectrométrie de masse réalisées dans cette étude, nous ont amené à générer et analyser une quantité importante de résultats. Cette observation nous a conduit à la conception d'un logiciel de calcul de masse moléculaire, dédié au calcul de masse moléculaire de pyoverdines, appelé « **Pyomass** ». Ce logiciel a été développé en collaboration avec le laboratoire d'informatique fondamentale de Lille (LIFL) dans le cadre d'un projet avec des étudiants. Ce logiciel a été développé en JAVA pour une compatibilité multiplateforme. Pyomass propose différentes fonctionnalités qui seront développées au cours de leur utilisation dans la partie résultats (Sections IV.2. et V.2.). Ce logiciel permet d'une part de calculer la masse moléculaire d'une molécule de pyoverdine à partir de la connaissance de la composition de la chaîne peptidique en acides aminés, du chromophore et des six chaînes latérales différentes pouvant être portées par le chromophore, le logiciel va générer le calcul des masses en fonction de chaque chaîne latérale et des ions potentiellement liés à la molécules dans les analyses en spectrométrie de masse (H^+ , K^+ ou Na^+). Une autre fonctionnalité majeure de ce logiciel est de permettre de changer un acide aminé par un autre à une position particulière de la chaîne peptidique et de générer les différents calculs inhérents à ces modifications dans le but de prédire les masses attendues lors de l'enrichissement du milieu de culture en un acide aminé. La dernière fonction de pyomass consiste à partir d'une chaîne peptidique donnée et d'un acide aminé donné en excès dans le milieu de culture, ainsi qu'un intervalle de masses moléculaires, de donner les différentes modifications pouvant avoir affecté la chaîne peptidique, ceci dans le but d'aider à l'analyse des résultats de spectrométrie de masse.

Résultats

III. Analyse des peptides d'origine non-ribosomiale produits par les *Pseudomonas fluorescents*

Les *Pseudomonas* du groupe fluorescent produisent un certain nombre de métabolites secondaires en fonction des conditions environnementales dans lesquelles elles croissent. Un grand nombre de métabolites secondaires produits par des micro-organismes sont connus pour être synthétisés par le mécanisme non-ribosomal. Les deux groupes de peptides non-ribosomiaux (NRPs) principalement produits par ces bactéries sont des lipopeptides cycliques (CLPs) et les pyoverdines. Dans ce chapitre, une étude statistique va être entreprise pour caractériser la taille et les monomères incorporés dans ces deux grands groupes.

III.1. Les pyoverdines

Comme décrit dans l'introduction, les pyoverdines sont constituées de trois parties distinctes : le chromophore, la chaîne latérale du chromophore et la chaîne peptidique. Notre étude statistique sur la taille et la composition en monomères a été effectuée sur 80 molécules de pyoverdines de la banque de données Norine (57 molécules de pyoverdine) et de l'analyse bibliographique. Les structures de ces 80 pyoverdines ont été présentées dans le chapitre « Introduction bibliographique ». Les résultats sont montrés dans les figures 29 et 30.

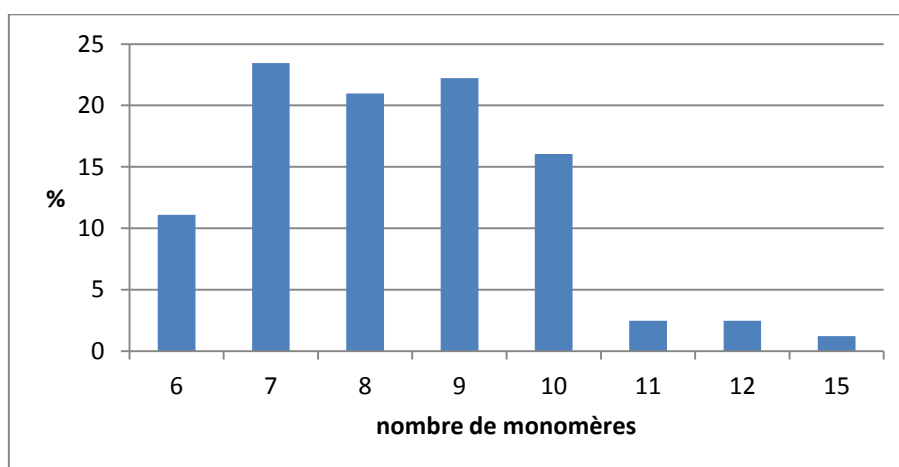


Figure 29: Répartition des molécules de pyoverdines en fonction de la longueur de la chaîne peptidique. Etude réalisée sur 80 molécules de pyoverdines.

La figure 29 montre que la longueur de la chaîne peptidique est comprise entre 6 et 15 monomères, mais seules 5 molécules ont une longueur plus grande que 10.

La diversité des monomères retrouvés dans ces molécules est reportée sur la figure 30. Puisque l'isomérisie des monomères n'est pas connue pour toutes les pyoverdines, l'analyse ne porte que sur les 61 molécules pour lesquelles elle est connue. Parmi les 487 monomères composant ces pyoverdines, il n'y a que 35 monomères différents, dont 6 n'ont qu'une occurrence (D-Asp, aThr, D-OH-But-OH-Orn, OH-Orn, OH-His, Val). Les six monomères les plus représentés sont des acides aminés protéogéniques (L/D-sérine, L-lysine, glycine, L-thréonine et L-alanine) et couvrent plus de 50 % des effectifs. La sérine sous sa forme L représente à elle seule 11,5 % des monomères rencontrés dans la composition des pyoverdines et 9,5 % sous sa forme D. Toutes les pyoverdines, dont l'isomérisie des monomères est connue, contiennent au moins un monomère sous forme D, ils représentent donc près de 30 % des effectifs. Sur les 80 molécules de pyoverdines recensées dans le tableau 2, 79 possèdent au moins une occurrence d'ornithine ou d'un de ces dérivés (cOHOrn, AcOHOrn, D-FoOHOrn, D-AcOHOrn, OHButOrn, D-OHOrn et OHOrn). L'ornithine et ses dérivés représentent donc près de 20% des monomères.

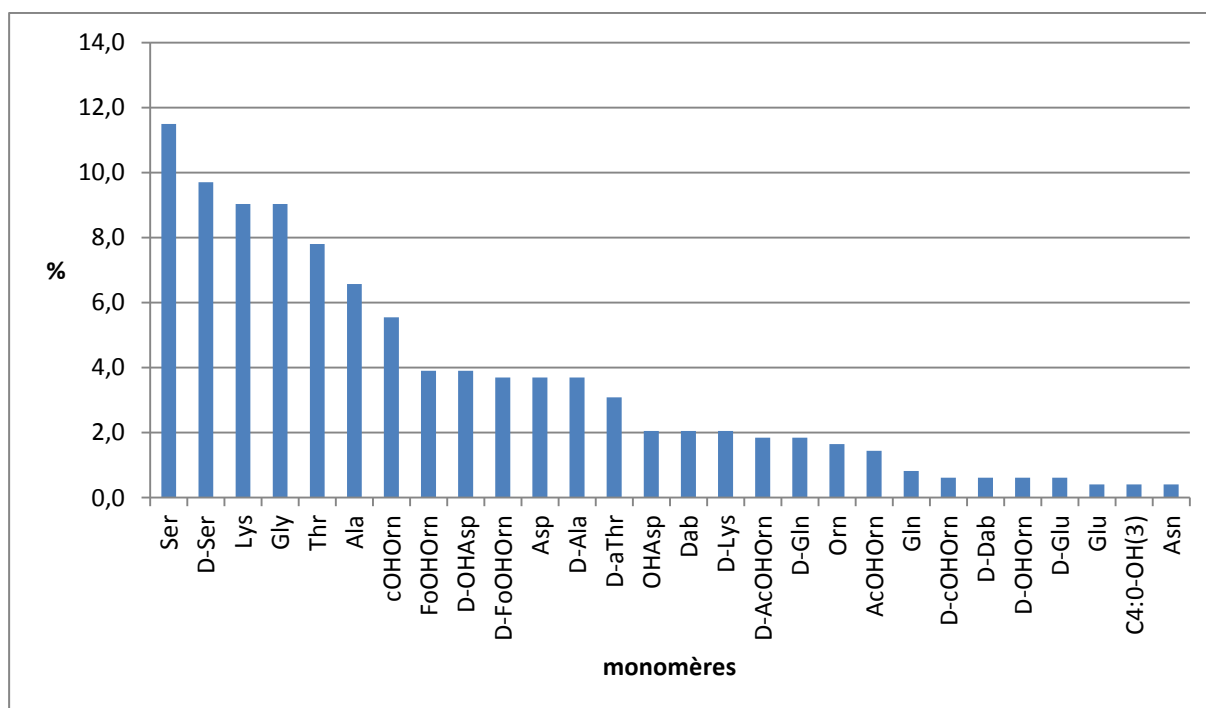
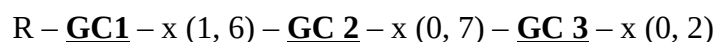


Figure 30: Proportions des monomères retrouvés dans les 61 molécules de pyoverdines dont l'isomérisie des monomères est connue, soit un total de 487 monomères. Les six monomères n'ayant qu'une seule occurrence n'apparaissent pas sur ce graphique (D-Asp, aThr, D-OH-But-OH-Orn, OH-Orn, OH-His et Val).

Pour leur activité biologique, les pyoverdines comportent généralement trois fonctions impliquées dans la coordination du fer. Ces trois fonctions sont les catécholates, hydroxamates et hydrocarboxylates et sont portées par des monomères entrant dans la composition de la pyoverdine. Le groupement catécholate est porté par le chromophore toujours présent et se présente sous la forme de fonctions hydroxyles libres. Le groupement hydroxamate est apporté par des monomères hydroxylés qui sont majoritairement les dérivés de l'ornithine sous forme hydroxylés (Ac-OH-Orn, Fo-OH-Orn et cOH-Orn). Ces deux groupements, catécholates et hydroxamates, sont retrouvés ensembles dans la majorité des molécules étudiées. Enfin, le groupement hydroxycarboxylate, présent sous la forme d'un hydroxy-aspartate (OH-Asp) n'est retrouvé que dans 44 % des molécules étudiées et y remplace le groupement hydroxamate. Plus de 50% des molécules de pyoverdines dont la structure est connue, ont une chaîne peptidique qui se termine par un groupement chélateur du fer. Grâce à ces observations, il est possible de définir une représentation linéaire de la molécule de pyoverdine comme suit :



Où : R est la chaîne latérale portée par le chromophore

GC 1, 2 et 3 sont les groupements chélateurs (GC1 est représenté par le chromophore)

et x représente n'importe quel acide aminé

En fonction de la nature et de la position des groupements chélateurs 1 et 2, on peut définir 5 classes structurales de pyoverdines :

1. **GC 2 et GC 3** sont des dérivés hydroxylés de l'ornithine. Ils portent un groupement hydroxamate qui intervient dans la coordination du fer.
2. **GC 2 et GC 3** sont des hydroxy-aspartates. Ils portent un groupement hydroxycarboxylate qui intervient dans la coordination du fer.
3. **GC 2** est un dérivé hydroxylé de l'ornithine et **GC 3** un hydroxy-aspartate.
4. **GC 2** est un hydroxy-aspartate et **GC 3** est un dérivé hydroxylé de l'ornithine.
5. **GC 2** est une hydroxy-histidine et **GC 3** est un dérivé de l'ornithine.

Les différentes structures de pyoverdines connues peuvent être classées dans ces classes structurales sur la base de la place des groupements chélateurs dans la molécule ainsi

que selon la nature de ses groupements chélateurs. Le tableau 7 reprend l'ensemble des 5 classes définies ci-dessus.

	GC 1	nb mono	GC 2	nb mono	GC 3	nb mono	% pyoverdines par classe (nb)
Classe 1	R-ChrP	1 à 4	*Orn	0 à 7	*Orn	0 à 2	52,5% (42)
Classe 2		1 à 3	OH-Asp	2 à 3	OH-Asp	1 à 3	3,75% (3)
Classe 3		1 à 2	*Orn	2 à 6	OH-Asp	1 à 2	13,75% (11)
Classe 4		1 à 4	OH-Asp	2 à 6	*Orn	0	28,75% (23)
Classe 5		2	OH-His	2	*Orn	0	1,25% (1)

Tableau 7: Répartition des molécules de pyoverdines connues en fonction des cinq classes structurales décrites ci-dessus. *Orn représente les dérivés hydroxylés de l'ornithine. Le pourcentage se base sur les 80 structures complètes de pyoverdines décrites dans le chapitre I. nb mono = nombre de monomères.

III.2. Les lipopeptides cycliques (CLPs)

Les lipopeptides cycliques sont composés de deux parties distinctes : la chaîne peptidique et un acide gras. Dans ce cas, l'analyse statistique a porté sur 64 molécules de CLPs produit par les *Pseudomonas* issues de Norine et d'une étude bibliographique. Dans cette étude l'acide gras n'est pas pris en compte, mais uniquement les variants de la chaîne peptidique. Seul un monomère de la chaîne change pour les variants d'un même peptide. Les résultats sont présentés dans les figures 31 et 32. La longueur des chaînes peptidiques de CLPs de *Pseudomonas* varie entre 8 et 25 monomères. Plus de 55 % des CLPs possèdent une chaîne peptidique d'une longueur de 8 ou 9 monomères (Figure 31).

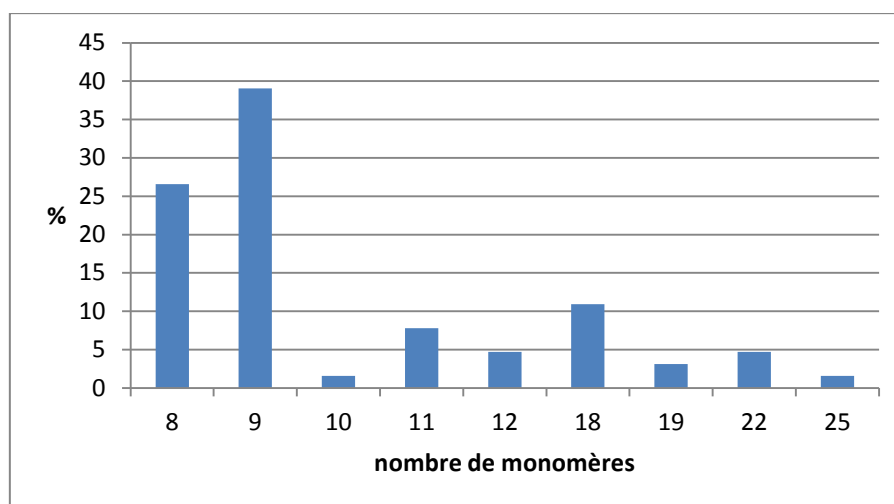


Figure 31: Répartition des molécules de lipopeptides en fonction de la longueur de la chaîne peptidique. Etude réalisée sur les 64 molécules actuellement décrites.

La diversité des 663 monomères présents dans les lipopeptides cycliques de *Pseudomonas* a été analysée, seuls 41 monomères différents sont observés. La figure 32 présente la distribution de ces monomères. La leucine (Leu), présente dans 39 CLPs, est le monomère le plus représenté, avec un pourcentage de plus de 16%. Les syringafactines B et E contiennent six monomères de leucine sur les huit monomères qui composent leurs chaînes peptidiques. Les monomères sous forme D représentent 37,5 % des monomères présents dans la composition de la chaîne peptidique des lipopeptides cycliques. Contrairement à ce qui a été observé chez les pyoverdines, l'ornithine et ses dérivés représentent moins de 3 % de l'ensemble des monomères présents dans les lipopeptides chez *Pseudomonas*. Les monomères hydrophobes représentent 46% des monomères et renforcent ainsi le caractère hydrophobe des lipopeptides.

En 2010, Caboche *et al.* ont décrit la répartition des 30 monomères les plus représentés dans les peptides non ribosomiaux recensés dans Norine, on observe que la répartition de ces monomères est différente de celle observée ici. Néanmoins le monomère le plus abondant reste la leucine. Chez les lipopeptides de *Pseudomonas*, les deux monomères les plus représentés sont la leucine (sous formes L et D) (monomère hydrophobe) et la sérine sous forme D (monomère hydrophile). La présence de monomères hydrophobes et hydrophiles dans la composition des lipopeptides contribue au caractère amphiphile nécessaire à leur activité biologique.

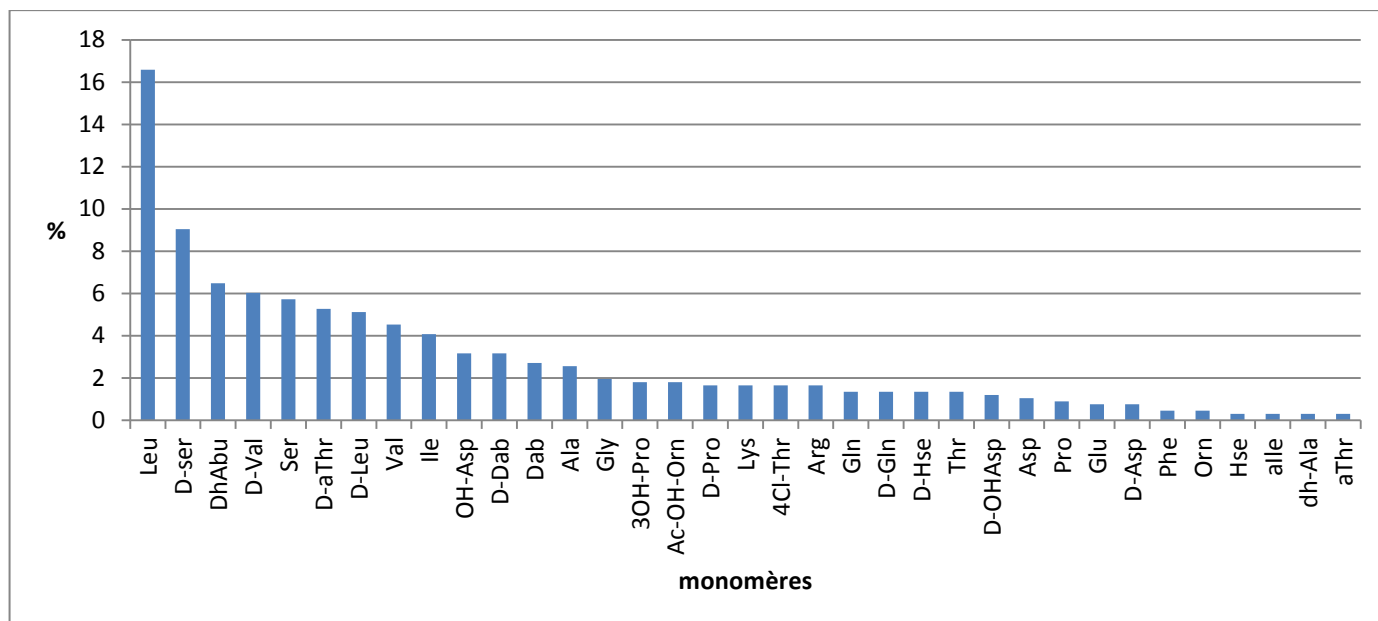


Figure 32: Pourcentage des monomères retrouvés dans les molécules de lipopeptides cycliques. Etude réalisée sur 64 molécules pour un total de 663 monomères. Les monomères n'ayant qu'une seule occurrence n'apparaissent pas dans ce graphique pour en simplifier la lecture (D-Thr, D-Orn, OH-His, Asn, Tyr et His).

III.3. Comparaison de la composition des CLPs et des pyoverdines

Les chaînes peptidiques des pyoverdines et des CLPs de *Pseudomonas* diffèrent par la répartition de leur longueur et par la nature de leurs monomères. Les pyoverdines ont une longueur moyenne plus petite que les lipopeptides. Les pyoverdines ont besoin de 3 monomères chélatant le fer, espacés d'au grand maximum 7 monomères. La figure 30 montre que dans les chaînes peptidiques des CLPs connus il y a une part de monomères sous forme D (37,5 %) légèrement supérieure à celle observée chez les pyoverdines (30%). D'autre part, la proportion de dérivés de l'ornithine est très faible dans les lipopeptides (moins de 3 %) contre près de 25 % dans les molécules de pyoverdines. Comme décrit précédemment l'ornithine et ses dérivés apportent à la molécule de pyoverdine une fonction essentielle à son activité sidérophore qui lui permet de lier le fer, alors que la présence de ces groupements n'est pas indispensable à l'activité surfactante des lipopeptides. Toutefois chez *Pseudomonas*, il existe deux exemples de lipopeptides à activité sidérophore, d'une part la ferrocine qui porte trois sites de fixation du fer par l'intermédiaire de groupements hydroxamates portés par des acétyles-hydroxy-ornithines et d'autre part la corrugatine qui possède également trois sites de fixation du fer portés par deux hydroxy-aspartate et une hydroxy-histidine. La leucine est un monomère très représenté dans la composition des CLPs, et dans celle des NRPs en général,

alors qu'elle n'est pas présente chez les pyoverdines. Plus généralement, les acides aminés hydrophobes ne sont pas présents dans les pyoverdines (seule une valine est présente dans un peptide) ce qui pourrait être favorable à une bonne solubilité en solution aqueuse. De plus, la diversité des monomères est plus importante dans les CLPs. En effet les lipopeptides comptent 41 monomères différents, dont 17 acides aminés protéogéniques. Les pyoverdines comptent 35 monomères, dont 11 acides aminés protéogéniques (L et D). Cette différence peut en partie être due au fait que les CLPs sont en moyenne plus longs que les pyoverdines, avec une faible répétition des monomères.

Dans ce chapitre, nous avons réalisé une analyse des lipopeptides cycliques et des pyoverdines produits par les Pseudomonas. La comparaison de la composition des ces deux types de peptides non-ribosomiaux met en évidence le lien entre la composition en monomères et le mode d'action du peptide, notamment l'activité sidérophore des pyoverdines grâce à l'importante présence de l'ornithine et de ses dérivés dans la composition du peptide. De plus, cette étude a permis de montrer la forte proportion de monomères sous forme D chez les lipopeptides cycliques et les pyoverdines de Pseudomonas. Le chapitre qui va suivre va traiter de la recherche du potentiel de synthèse non-ribosomiale au sein des génomes de Pseudomonas et de la structure des synthétases produisant ces NRPs.

IV. Analyse des synthétases de pyoverdines

VI.1. Introduction

Ce chapitre présente les résultats obtenus lors de l'analyse de génomes de bactéries du genre *Pseudomonas* dont la séquence est disponible dans les banques de données. Dans un premier temps, des outils bio-informatiques nous ont permis d'extraire les synthétases non-ribosomiales des génomes impliqués dans la synthèse de pyoverdines puis de prédire la structure des peptides produits. Dans un deuxième temps, des analyses par spectrométrie de masse des molécules produites par les souches que nous avons pu obtenir ont permis de valider la structure de certaines des molécules prédites. La diversité de ces molécules a également été testée expérimentalement à l'aide d'expériences de feeding. Les résultats obtenus sont présentés par type de molécules et en distinguant les molécules déjà connues des nouvelles.

En 2012, vingt-deux génomes de *Pseudomonas* ont été complètement séquencés (deux de ces génomes *Pseudomonas putida* S16 et *Pseudomonas stutzeri* ATCC17588 n'ont pas été séquencés car leur séquence est disponible depuis peu). La banque de données MBGD (Microbial Genome Database) contient ces différents génomes dont les informations sont reprises dans la tableau 8.

Souches bactériennes	taille du génome	n° accession
<i>Pseudomonas aeruginosa</i> LESB58	6,6 Mb	NC_011770.1
<i>Pseudomonas aeruginosa</i> PA7	6,6 Mb	NC_009656.1
<i>Pseudomonas aeruginosa</i> PAO1	6,3 Mb	NC_002516.2
<i>Pseudomonas aeruginosa</i> UCBPP-PA14	6,5 Mb	NC_008463.1
<i>Pseudomonas mendocina</i> NK-01	5,5 Mb	NC_015410.1
<i>Pseudomonas mendocina</i> ymp	5,1 Mb	NC_009439.1
<i>Pseudomonas brassicacearum</i> NFM421	6,8 Mb	NC_015379.1
<i>Pseudomonas entomophila</i> L48	5,9 Mb	NC_008027.1
<i>Pseudomonas fluorescens</i> Pf-5	7,1 Mb	NC_004129.6
<i>Pseudomonas fluorescens</i> PfO-1	6,4 Mb	NC_007492.2
<i>Pseudomonas fluorescens</i> SBW25	6,7 Mb	NC_012660.1
<i>Pseudomonas putida</i> F1	6,0 Mb	NC_009512.1
<i>Pseudomonas putida</i> GB-1	6,1 Mb	NC_010322.1
<i>Pseudomonas putida</i> KT2440	6,2 Mb	NC_002947.3
<i>Pseudomonas putida</i> W619	5,8 Mb	NC_010501.1
<i>Pseudomonas stutzeri</i> A1501	4,6 Mb	NC_009434.1
<i>Pseudomonas syringae</i> B728a	6,1 Mb	NC_007005.1
<i>Pseudomonas syringae</i> 1448A	5,9 Mb	NC_005773.3
<i>Pseudomonas syringae</i> DC3000	6,4 Mb	NC_004578.1
<i>Pseudomonas fulva</i> 12-X	4,9 Mb	NC_015556.1

Tableau 8 : Liste des génomes de *Pseudomonas* entièrement séquencés issus de la banque de données MBGD. Ce tableau reprend les génomes des différentes souches de *Pseudomonas* séquencés avec leur taille et leur numéro d'accension.

IV.2. Présentation de l'outil bio-informatique Pyomass

Le calcul de la masse des peptides non-ribosomiaux à partir de leur structure monomérique n'est pas pris en charge par les logiciels dédiés aux peptides ribosomiaux puisque ces peptides sont composés d'une grande variété de monomères (526 répertoriés dans Norine) et que des liaisons non-peptidiques peuvent lier certains monomères entre eux. Dans le cas des pyoverdines, la structure monomérique est relativement simple puisque ces molécules sont linéaires (parfois partiellement cycliques), composées d'un chromophore avec une chaîne latérale, lié à une chaîne peptidique et qu'elles ne comptent que 35 monomères différents. Il est donc possible d'automatiser le calcul de leur masse à partir de leur structure monomérique. Un outil de calcul de masse moléculaire dédié à l'étude des pyoverdines a donc été développé dans le cadre d'une collaboration avec le LIFL (Laboratoire d'Informatique Fondamentale de Lille).

Cet outil offre une interface conviviale qui permet de sélectionner les acides aminés composant la chaîne peptidique, préciser s'il existe des cycles dans cette chaîne, ainsi que de préciser le type de chromophore (généralement ChrP ou ChrI). Le résultat en sortie est un tableau reprenant la séquence donnée pour la pyoverdine et les masses de la pyoverdine en fonction des six chaînes latérales (R) pouvant être portées par le chromophore. Cet outil étant lié à l'analyse en masse, le résultat donne aussi les valeurs m/z des ions monochargés $[M+H]^+$, $[M+Na]^+$ et $[M+K]^+$, pour chacune des masses calculées. La figure 33 montre une capture d'écran de Pyomass ainsi qu'une fenêtre de résultats qui illustre le cas d'utilisation ci-dessus. Les autres fonctionnalités de Pyomass seront détaillées au moment de leur utilisation dans la suite de ce chapitre.

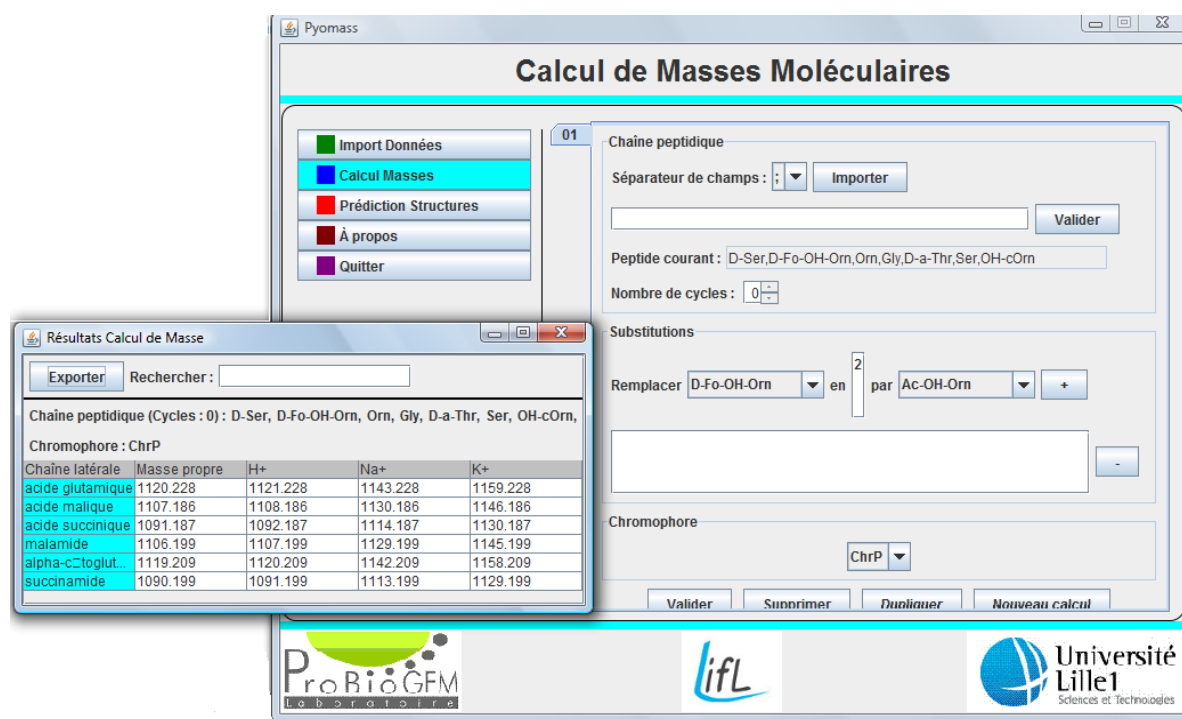


Figure 33: Capture d'écran du logiciel « Pyomass ». L'exemple pris dans cette figure est le calcul de la masse de la pyoverdine produite par *Pseudomonas aeruginosa* PA7. La fenêtre au premier plan est le tableau de résultats reprenant les masses calculées en fonction de la chaîne latérale (R) portée par le chromophore.

IV.3. Caractérisation des gènes qui codent pour des synthétases de pyoverdines connues

Les 20 génomes de *Pseudomonas* séquencés (tableau 8) ont été analysés à la recherche de clusters de gènes de synthèse de pyoverdine selon le protocole du paragraphe II.3.2.1. de la partie « matériel et méthodes ». Cette recherche débute par l'identification à l'aide de BLAST

du gène codant la protéine responsable de la synthèse du chromophore de la pyoverdine (appelé pvdL), puis continue par la recherche des gènes codant la chaîne peptidique de la molécule à l'aide de mots clés dans la banque de données MBGD. La présence d'un potentiel de synthèse d'une pyoverdine a pu être mis en évidence dans seize des vingt génomes étudiés. Parmi les 16 souches produisant une pyoverdine, 7 produisent une pyoverdine dont la structure a déjà été décrite dans la littérature (cf tableau 9). L'isomérisation des acides aminés présents dans ces pyoverdines n'a été décrite que pour trois d'entre elles.

souches bactériennes	séquence de la pyoverdine	références
<i>P. aeruginosa</i> PA7	R-ChrP-D-Ser-D-FoHOrn-Orn-Gly-D-aThr-Ser-cHOrn	Budzikiewicz, 2004
<i>P. aeruginosa</i> PAO1	R-ChrP-D-Ser-Arg-D-Ser-FoHOrn-(Lys-FoHOrn-Thr-Thr)	Briskot <i>et al.</i> , 1989
<i>P. entomophila</i> L48	R-ChrP-Ala-Asn-Dab-OHHis-Gly-Gly-Ala -Thr-Ser-cHOrn	Matthijs <i>et al.</i> , 2010
<i>P. fluorescens</i> Pf-5	R-ChrP-Asp-FoHOrn-Lys-(Thr-Ala-Ala-FoHOrn-Lys)	Paulsen <i>et al.</i> , 2005
<i>P. fluorescens</i> PfO-1	R-ChrP-Ala-AcHOrn-Orn-Ser-Ser-Ser-Arg-OHAsp-Thr	Meyer <i>et al.</i> , 2008
<i>P. fluorescens</i> SBW25	R-ChrP-D-Ser-Lys-Gly-FoHOrn-(Lys-Ser-D-FoHOrn)	Moon <i>et al.</i> , 2008
<i>P. putida</i> KT2440	R-ChrP-Asp-Orn-(OHAsp-Dab)-Gly-Ser-cHOrn	Matthijs <i>et al.</i> , 2009

Tableau 9: Liste des sept souches de *Pseudomonas* dont la structure de la pyoverdine est connue. Les acides aminés entre () sont inclus dans un cycle.

Pyomass a été utilisé pour calculer les masses des 7 pyoverdines du tableau 8, à partir des séquences données dans les articles. Les résultats sont repris dans le tableau 10, la masse indiquée correspond à la masse propre calculée avec la succinamide comme chaîne latérale (R) du chromophore car cette chaîne est la plus fréquente.

souches bactériennes	Masse Pyomass (en Da)	Masse expérimentale (en Da)
<i>P. aeruginosa</i> PA7	1090,199	1090,644
<i>P. aeruginosa</i> PAO1	1332,505	1332,712
<i>P. putida</i> KT2440	1072,156	1072,451
<i>P. fluorescens</i> Pf-5	1286,476	1286,613
<i>P. fluorescens</i> PfO-1	1380,503	1380,787
<i>P. fluorescens</i> SBW25	1159,319	1159,62
<i>P. entomophila</i> L48	1297,434	1313,366

Tableau 10: Masses calculées par Pyomass et masses déterminées expérimentalement par spectrométrie de masse MALDI-TOF pour les pyoverdines connues et reprises dans la littérature. La ligne en gras met en évidence une différence de masses.

Le tableau 10 montre une cohérence entre les masses calculées et les masses obtenues expérimentalement par spectrométrie de masse MALDI-TOF pour les pyoverdines produites par six des sept souches de *Pseudomonas* étudiées. En effet, une incohérence a été observée pour *Pseudomonas entomophila* L48. La séquence publiée pour cette molécule doit présenter une erreur, car une différence de 16 Da est observée entre les deux masses.

La recherche puis l'analyse bio-informatique des 7 synthétases produisant ces pyoverdines a été réalisée en utilisant les outils de prédiction décrits dans le chapitre II (NRPS/PKS, PKS/NRPS et NRPS Predictor). Ces outils permettent de prédire la composition en domaines et modules d'une synthétase ainsi que la composition du peptide produit par celle-ci. La figure 34 montre l'exemple de l'étude de la synthétase de pyoverdine de *Pseudomonas putida* KT2440. De la même manière, la prédiction de la structure des synthétases en modules et en domaines, ainsi que la composition du peptide produit par la synthétase a été réalisée pour les seize souches de *Pseudomonas* possédant le potentiel génique de synthèse d'une pyoverdine.

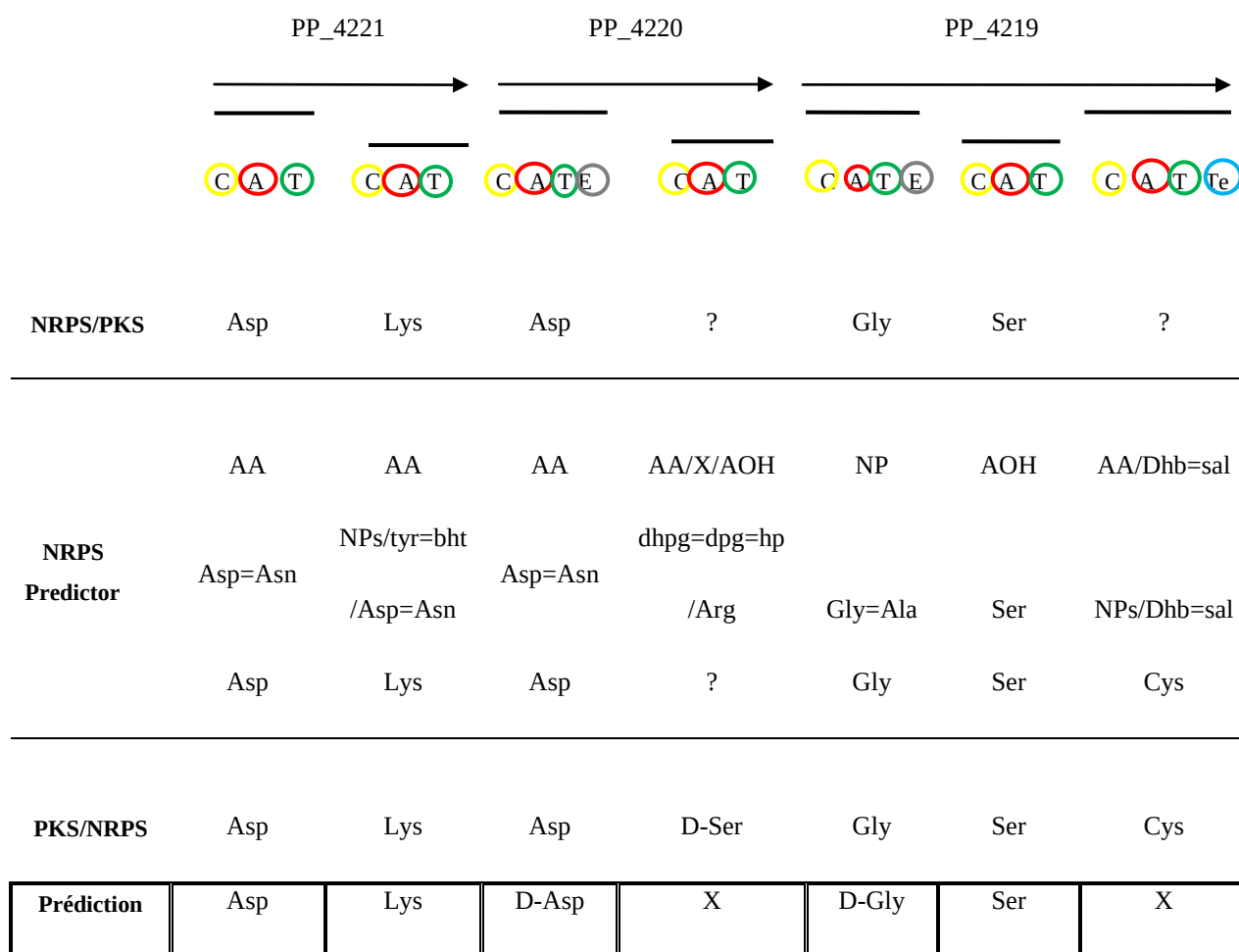


Figure 34: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas putida* KT2440
AOH= ser, thr, ser/thr, Dhpg, Hpg AA= asp, asn, glu, gln, Aad NPs= val, leu, ile, Abu, Iva

NP=gly,ala,val,leu,ile,Abu,Iva et X ou ? quand le monomère est indéterminé. La dernière ligne notée ‘prédiction’ donne la composition prédite, à partir de l’analyse des résultats des différents outils de prédiction, du produit de la synthétase. La prédiction d’un acide aminé incorporé est validée lorsqu’au moins deux des logiciels de prédiction donnent un résultat identique. Les formes D sont identifiées sur la base de la mise en évidence de domaines d’épimérisation.

Le tableau 11 propose une comparaison entre la structure donnée dans la littérature et celle prédite par les outils bioinformatiques.

souches bactériennes	séquences publiées puis prédites de la partie peptidique des pyoverdines									
<i>P. aeruginosa</i> PA7	D-Ser	D-FoHOOrn	Orn	Gly	D-aThr	Ser	cHOOrn			
	D-Ser	D-Orn	X	Gly	D-Thr	Ser	Orn			
<i>P. aeruginosa</i> PAO1	D-Ser	Arg	D-Ser	FoOH Orn	(Lys	FoHOOrn	Thr	Thr)		
	D-Ser	Glu	D-Ser	Orn	Lys	Orn	Thr	Thr		
<i>P. fluorescens</i> Pf-5	Asp	FoHOOrn	Lys	(Thr	Ala	Ala	FoHOOrn	Lys)		
	Asp	D-Orn	Lys	Thr	Ala	D-ala	D-Orn	Lys		
<i>P. fluorescens</i> PfO-1	Ala	AcHOOrn	Orn	Ser	Ser	Ser	Arg	OHAsp	Thr	
	D-Ala /Gly	D-Orn	D-X	Ser	Ser	D-Ser	D-Arg	Asp	Thr	
<i>P. fluorescens</i> SBW25	D-Ser	Lys	Gly	FoOH Orn	Lys	Ser	D-FoHOOrn			
	D-Ser	Lys	Gly	Orn	Lys	D-Orn	Ser			
<i>P. putida</i> KT2440	Asp	Orn	OHAsp	Dab	Gly	Ser	cHOOrn			
	Asp	Lys	D-Asp	X	D-Gly	Ser	X			
<i>P. entomophila</i> L48	Ala	Asn	Dab	OHHis	Gly	Gly	Ala	Thr	Ser	cHOOrn
	D-Ala	X	His	D-Asp	Gly	Gly	Ser	Thr	D-Ser	Orn

Tableau 11: Séquences des pyoverdines produites par les bactéries du groupe des *Pseudomonas* fluorescents (séquences décrites dans la littérature sont au-dessus et les séquences issues de l’analyse bioinformatique des différents génomes sont présentées en-dessous, souche par souche).

Dans ce tableau des prédictions notées ‘X’ signifient que les outils de prédiction donnent des résultats divergents et/ou que certains de ces outils ne proposent pas de prédiction. La présence d’un monomère sous forme D dans l’analyse bioinformatique est prédite par la présence d’un domaine d’épimérisation (E) dans le module incorporant ce monomère. Il est à noter que l’isomérisation des monomères n’a pas été déterminée pour l’ensemble des structures de pyoverdines connues d’où la présence de monomères sous forme D dans la prédiction et non dans la structure issue de la littérature.

Une analyse au cas par cas des résultats du tableau 11 est détaillée ci-dessous :

- *Pseudomonas aeruginosa* PA7

Pour cette prédiction, l'isomérisation des acides aminés est confirmée par la présence de domaines d'épimérisation dans la structure de la synthétase. Les dérivés de l'ornithine (D-FoOHOrn et cOHOrn) sont prédits comme des ornithines « simples » par les outils de prédiction. Le monomère en position 3 dans la chaîne peptidique n'est pas prédit (X) alors qu'il s'agit d'une ornithine. La présence de l'ornithine en position 3 de la chaîne est confirmée par le calcul de la masse dans le tableau 2. Le monomère en position 5 de la chaîne peptidique est prédit comme étant une thréonine sous forme D alors qu'il a été décrit dans la littérature comme une D-allo-thréonine. En effet les outils existants ne permettent pas de différencier les deux formes.

- *Pseudomonas aeruginosa* PAO1

Pour cette prédiction, l'isomérisation des acides aminés est également confirmée par la présence de domaines d'épimérisation dans la structure de la synthétase. La séquence prédite pour la pyoverdine de cette souche montre en position 1 et 3 de la chaîne peptidique des monomères sous forme D tout comme dans la séquence publiée. Le monomère en position 2 est prédit comme étant un glutamate alors qu'il s'agit bien d'une arginine d'après la masse obtenue expérimentalement. Comme pour la pyoverdine précédente les dérivés de l'ornithine sont prédits comme des ornithines « simples ».

- *Pseudomonas fluorescens* Pf5

L'isomérisation des acides aminés composant la chaîne peptidique de la pyoverdine produite par cette souche n'est pas déterminée, mais a pu être prédite à l'aide des analyses bio-informatiques. Les deux formyl-hydroxy-ornithine (FoOHOrn) sont prédites comme étant des ornithines. La présence du cycle dans la chaîne peptidique ne peut pas être prédit par bioinformatique. Les monomères prédits sont dans le même ordre que dans la séquence publiée.

- *Pseudomonas fluorescens* Pf0-1

L'isomérisation des acides aminés composant la chaîne peptidique de la pyoverdine produite par cette souche n'est pas déterminée, mais a pu être prédite à l'aide des analyses bio-informatiques. Les analyses bio-informatiques prédisent pour le premier module un domaine d'épimérisation et le monomère reconnu par le domaine A est « D-Gly/Ala ». Or, la glycine ne peut avoir de forme D étant donné qu'elle ne possède pas de carbone asymétrique, le

monomère en position 1 est donc une D-Alanine. Le monomère en position 3 est prédit comme « D-X » et il s'agit d'une ornithine dans la séquence publiée. La difficulté de prédiction de l'ornithine a déjà été observée pour la pyoverdine de *Pseudomonas aeruginosa* PA7 ci-dessus. L'ornithine prédite en position 2 est en fait une acétyl-hydroxy-ornithine (AcOHOrn) et l'aspartate en position 8 de la chaîne est un hydroxy-aspartate (OH-Asp).

- *Pseudomonas fluorescens* SBW25

L'isomérisation des monomères présents dans la chaîne peptidique de la pyoverdine produite par cette souche a été prédite correctement. Les deux ornithines prédites sont des FoOHOrn. L'ordre des trois derniers acides aminés n'est pas le même dans la prédiction et la séquence publiée car ils sont impliqués dans un cycle. D'une part, l'ordre des monomères dans le cycle dépend du sens dans lequel le cycle est lu, d'autre part, le cycle n'est pas été prédit par les outils d'analyse bio-informatique existants.

- *Pseudomonas putida* KT2440

L'isomérisation des acides aminés composant la chaîne peptidique de la pyoverdine produite par cette souche n'est pas déterminée. Les monomères non prédits et notés 'X' sont le Dab et la cOHOrn. La prédiction de la composition de cette pyoverdine est détaillée sur la figure 34.

- *Pseudomonas entomophila* L48

Dans cette prédiction, 4 monomères prédits ne correspondent pas à la structure publiée (positions 2, 3, 4 et 7 de la chaîne peptidique). Néanmoins, nous avons obtenu par analyse en spectrométrie de masse MALDI-TOF, la même masse que celle obtenue par les auteurs, soit 1313Da.

Pour les six molécules de pyoverdine étudiées ci-dessus, la masse confirme bien la structure publiée. Néanmoins pour la pyoverdine produite par *Pseudomonas entomophila* L48, le tableau 9 montre une incompatibilité entre la structure publiée et la masse observée expérimentalement. Une analyse plus détaillée de la pyoverdine produite par cette souche a donc été effectuée.

Analyse de la structure de la pyoverdine produite par *Pseudomonas entomophila* L48

La pyoverdine produite par *Pseudomonas entomophila* L48 a été étudiée et sa structure a été déterminée par Matthijs *et al.* (2009) par spectrométrie de masse. La structure de la chaîne peptidique de cette molécule de pyoverdine est composée des dix acides aminés suivants :

R_ChrP_Ala _ Asn _ Dab _ OHHis _ Gly _ Gly _ Ala _ Thr _ Ser _ cOHOrn

1 2 3 4 5 6 7 8 9 10

Toutefois notre étude bioinformatique du génome de cette souche a montré que la séquence prédite pour la chaîne peptidique de cette pyoverdine était différente de la séquence publiée. Le monomère en position 7 serait une sérine et non une alanine.

	PSEEN_3234 PSEEN_3232 PSEEN_3231					
PKS/NRPS	Gly	?		?	Asp	?
NRPS Predictor	NP	NP/dhb=Sal		NP	AA	NP
	Gly=Ala	NPs/dhb=Sal		NPs	Asp=Asn	Gly=Ala
	Ala	?		His	Asp	Gly
NRPS/PKS	D-Ala	Cys		His	Asp	Gly
	PSEEN_3230 PSEEN_3229					
PKS/NRPS	?	Ser	Thr	Ser	5hOrn	
NRPS Predictor	NP	AOH	AOH	AOH	AA/dhb=sal	
	Gly=Ala	Ser	Thr=Dht	Ser	NPs/dhb=Sal	
	Gly	Ser	Thr/Dht	Ser	?	
NRPS/PKS	Gly	Ser	Thr	Ser	Lys	

Figure 35: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas entomophila* L48
 AOH= ser, thr, ser/thr, Dhpg, Hpg AA= asp, asn, glu, gln, Aad NPs= val, leu, ile, Abu, Iva
 NP=gly,ala,val,leu,ile,Abu,Iva.

La séquence prédite est donc la suivante :

D-Ala _ X_ His _ D-Asp _ Gly _ Gly _ Ser _ Thr _ D-Ser _ Orn

Un alignement des séquences de domaines A incorporant l'alanine et la sérine, au sein de différentes molécules de pyoverdine dont la structure est bien déterminée, a été réalisé (figure 36). On observe sur cet arbre que la séquence du domaine A incorporant l'alanine en position 7 (noté « PSEEN3230 A2 ») est similaire aux séquences des domaines A incorporant la sérine (par exemple avec le domaine A incorporant la sérine en position 9, noté « PSEEN3229 A2 »). Cette observation conforte les prédictions obtenues avec les différents logiciels de prédiction, l'acide aminé en position 7 serait bien une sérine.

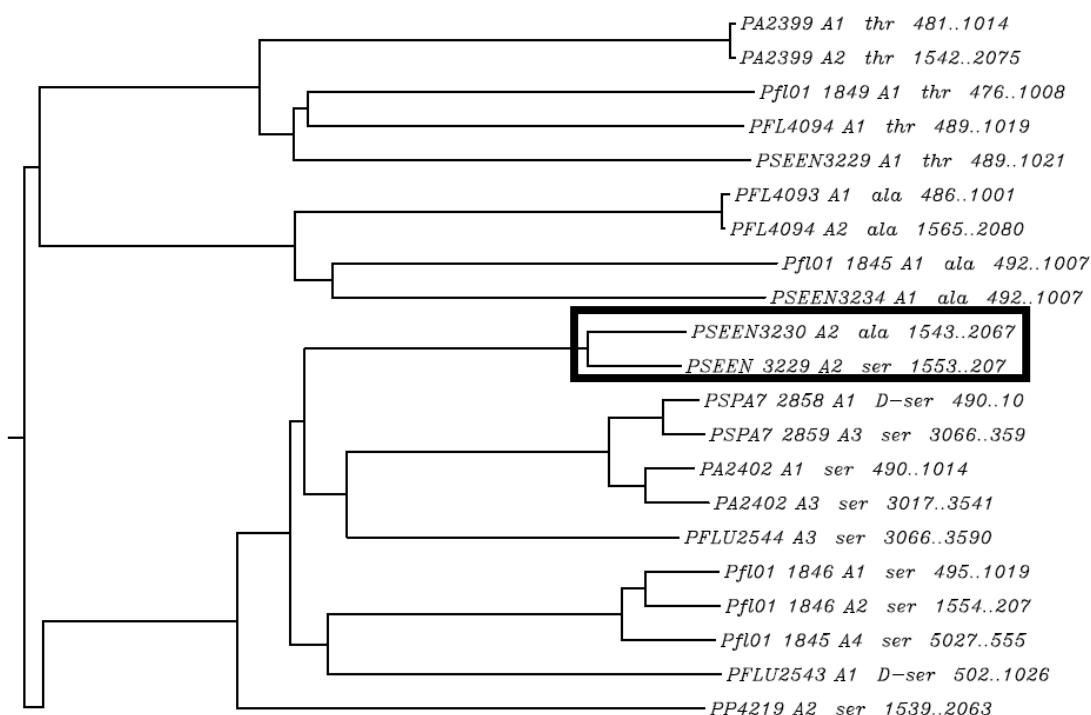


Figure 36: Arbre de distance issu d'un alignement de séquences de domaines A, de synthétases de pyoverdine, (environ 500 aa) incorporant des acides aminés connus .

Afin de confirmer la structure de cette pyoverdine, une analyse LC-MS-MS a été réalisée sur le surnageant de culture de *Pseudomonas entomophila* L48 cultivée dans des conditions de carence en fer. Le spectre MS2 (figure 37) du pic correspondant à la molécule de pyoverdine produite par cette souche confirme que le monomère en position 7 sur la chaîne peptidique est une sérine. Le tableau 12 reprend les résultats du couplage LC-MS-MS pour les fragments b4 à b8.

fragment	séquences	Masse
b4	Succa_ChrP_Ala_Asn_Dab_OH-His	795,4
b4-H2O		777,4
b5	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly	852,4
b5-H2O		834,6
b6	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly_Gly	909,4
b6-H2O		891,4
b7	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly_Gly_ Ser	996,5
b7-H2O		978,5
b8	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly_Gly_ Ser _Thr	1097,8
b8-H2O		1079,5

Tableau 12: Résultats de l'étude LC-MS-MS du spectre correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 .

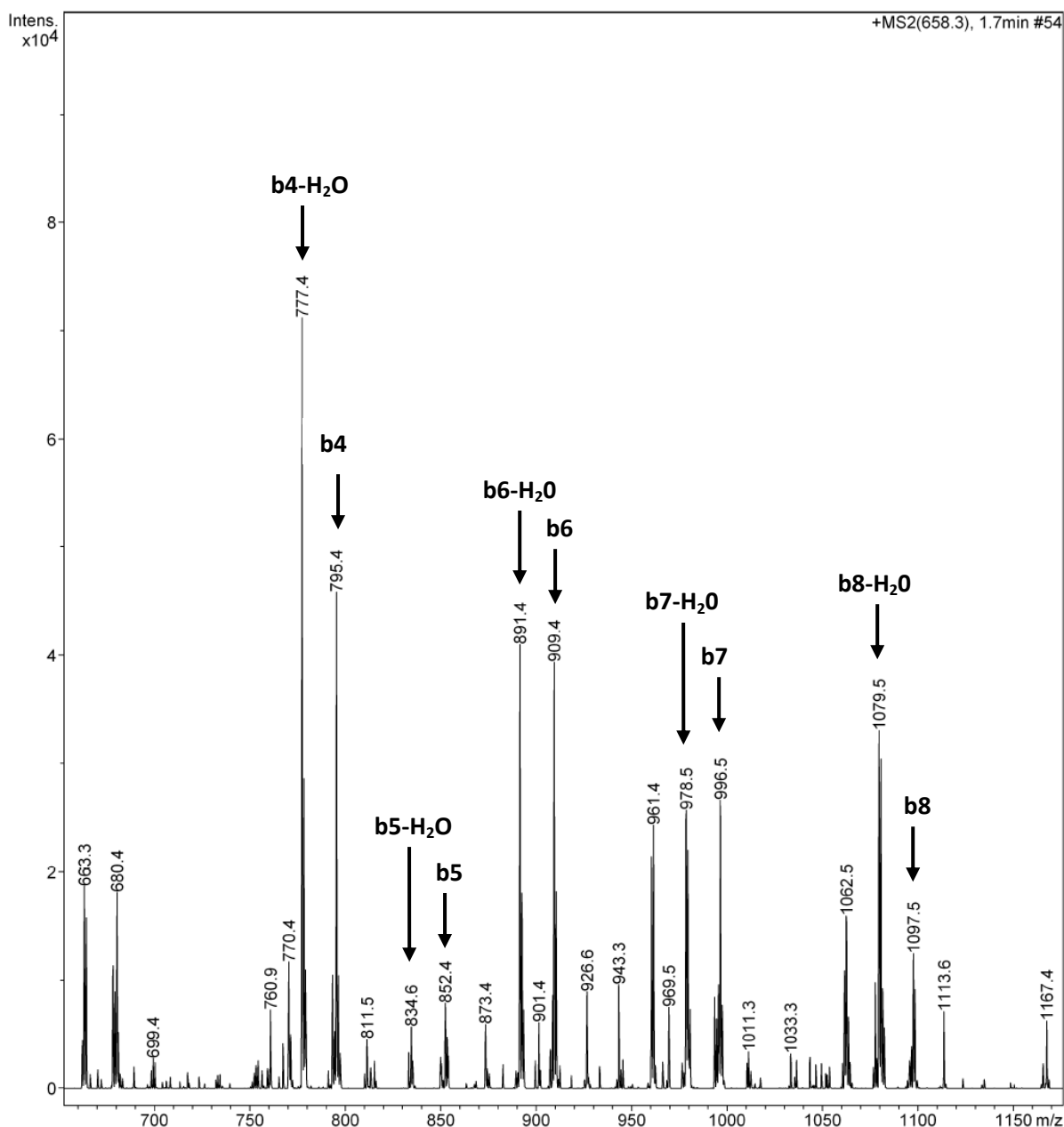


Figure 37: Spectre MS2 du pic correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 ($[M+H]^+ = 1314,366$ Da).

IV.4. Détermination de nouvelles signatures du code NRPS spécifiques aux pyoverdines

Les résultats précédents montrent que certains acides aminés introduits dans les pyoverdines ne peuvent être prédits par les outils actuels de prédiction. Dans ce chapitre, deux approches seront utilisées pour tenter d'améliorer ces prédictions. D'une part nous utiliserons les outils de phylogénie pour s'assurer qu'ils ne permettent pas à eux seuls comme on vient de

le démontrer pour L48 d'obtenir des résultats plus intéressants et ensuite nous analyserons les signatures du code NRPS des domaines d'adénylation dont la spécificité est non prédictible afin de définir de nouvelles signatures.

Une étude phylogénétique a donc été entreprise dans un premier temps sur la séquence entière du domaine A soit environ 500 acides aminés. La figure 38 montre l'arbre obtenu après alignement des séquences de domaines A des synthétases de pyoverdines des sept souches de *Pseudomonas* dont la séquence de la pyoverdine est connue.

Dans cette figure, on observe que les domaines A se regroupent en fonction de la nature de substrat reconnu. On observe également sur cette figure, que les domaines A incorporant la lysine ainsi que l'ornithine et ses dérivés sont regroupés. Cette observation montre la difficulté à distinguer une lysine et une ornithine, mais elle est faite à l'échelle du domaine A entier, soit plus de 500 acides aminés. Dans la suite du travail l'étude des domaines A sera réduite à celle des signatures des 10 acides aminés situés dans un rayon de 5,5Å autour du site de fixation du substrat et intervenant dans la liaison de ce dernier au sein du domaine A. Ces acides aminés sont utilisés par les outils de prédiction des monomères reconnus par les domaines A.

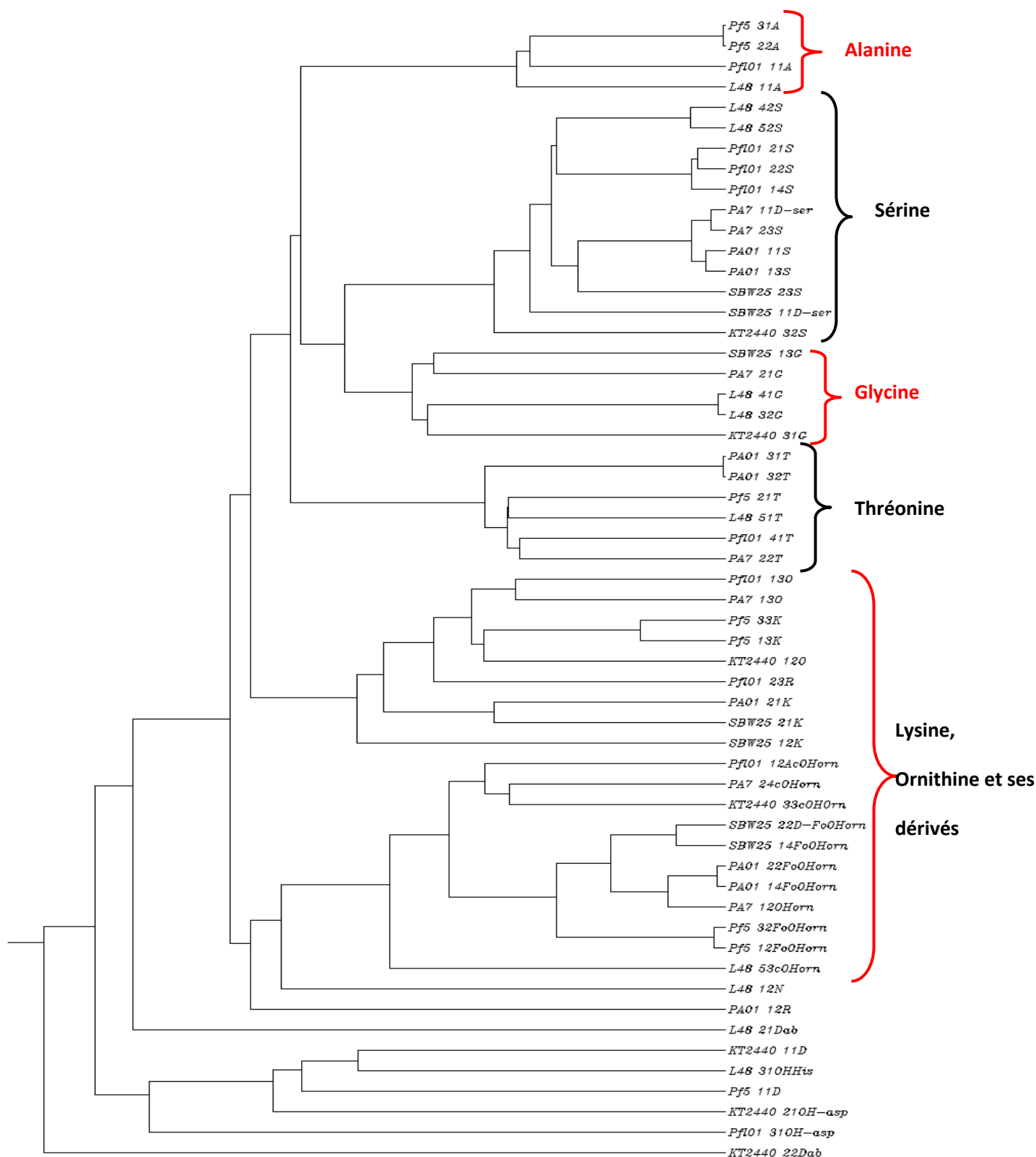


Figure 38: Représentation sous la forme d'un arbre de distances de l'alignement des séquences des domaines A des synthétases de pyoverdines dont la séquence est connue. Les domaines A sont décrits comme suit sur cette figure : nom de la souche de *Pseudomonas*, le premier chiffre correspond à la position de la protéine dans la synthétase, le deuxième chiffre correspond à la position du module dans la protéine, l'acide aminé incorporé par le domaine A.

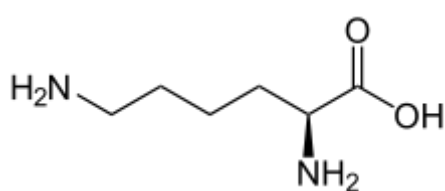
Les outils d'analyse des NRPS (NRPS/PKS, PKS/NRPS et NRSPredictor2) nous ont permis d'extraire les 10 acides aminés impliqués dans la reconnaissance du substrat pour les 61 domaines d'adénylation des 7 synthétases dont les peptides ont été validés expérimentalement. Ainsi, nous avons pu associer des signatures à un monomère spécifique incorporé chez les pyoverdines (Tableau 13). Ces signatures peuvent maintenant être utilisées pour améliorer la prédiction des structures des pyoverdines produites par les autres souches de *Pseudomonas*. En effet, nous allons pouvoir prédire non seulement plus de monomères, mais aussi de façon plus précise.

Signatures								Monomère incorporé	Nb domaine A
L	W	N	N	A	L	-	T	Ala	2
L	W	N	N	A	L	T	Y		2
A	E	D	I	G	A	I	T	Arg	1
S	E	D	V	G	A	V	T		1
W	E	Y	N	A	G	I	T	Asn	1
I	W	E	L	T	A	D	D	Dab	1
S	A	A	I	A	E	V	W		1
I	L	A	I	G	L	I	W	Gly	5
G	E	D	I	I	T	V	V	Lys	1
A	E	D	N	G	T	V	S		2
G	E	D	H	G	T	V	V		2
L	T	K	I	G	H	V	G	OH-Asp	1
L	T	K	V	G	H	V	G		1
G	E	A	-	A	G	V	T	cOHOrn	2
G	E	C	C	G	G	V	T		1
G	E	V	C	G	G	V	T	FoOHOrn	6
G	E	A	Y	F	G	V	T	AcOHOrn	1
G	E	V	C	G	G	V	T	OH-Orn	1
G	E	D	H	G	T	V	T	Orn	2
G	E	D	H	G	T	V	C		1
V	W	H	V	S	L	I	D	Ser	11
F	W	N	I	G	M	V	H	Thr	3
F	W	N	V	G	M	V	H		2

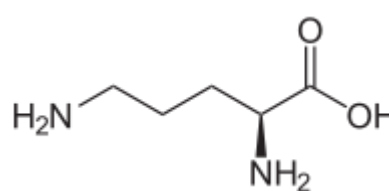
Tableau 13: Liste des signatures extraites à partir des domaines A des sept souches dont la structure de la pyoverdine est connue. La dernière colonne du tableau indique le nombre de domaines A portant chaque signature. Les tirets (-) sont le résultat d'une insertion/délétion lors de l'alignement des séquences de domaines A nécessaire à l'extraction des signatures. Les signatures sont réduites à 8 acides aminés car elles débutent toutes par un aspartate (D) et se terminent par une lysine (K).

Le tableau 13 montre les signatures, extraites des synthétases de pyoverdines, impliquées dans la reconnaissance du substrat du domaine A. On observe des signatures qui permettent de différencier les dérivés de l'ornithine (Ac-OH-Orn, Fo-OH-Orn, cOH-Orn et OH-Orn) entre eux lors des prédictions. Néanmoins les signatures extraites des domaines A incorporant la Fo-OH-Orn et OH-Orn sont identiques, ces deux dérivés de l'ornithine ne

pourront pas être différenciés dans les prédictions. On observe également que l'une des signatures extraites pour l'ornithine est très proche d'une des signatures extraite pour la lysine. En effet, ces deux signatures présentent 9 acides aminés sur 10 en communs (à savoir 'DGEDHGTV**C**K' pour l'ornithine et 'DGEDHGTV**V**K' pour la lysine), la différence se trouve au niveau du neuvième acide aminé de la signature. Nous avons observé une faiblesse lors de l'utilisation dans les logiciels d'analyse des NRPS dans la prédiction de la lysine et de l'ornithine. L'ornithine est souvent prédite comme étant une lysine par ces outils. Les structures de ces deux acides aminés sont très proches (voir ci-dessous), elles ne diffèrent que par la présence d'un groupement CH₂ supplémentaire pour la lysine.



Lysine (C₆H₁₄N₂O₂)



Ornithine (C₅H₁₂N₂O₂)

Les isomères L et D sont déterminés par la présence ou l'absence de domaines d'épimérisation dans les modules incorporant les monomères. Dans le domaine A il n'existe pas de signature permettant de distinguer un isomère sous forme L d'un isomère sous forme D, alors que certains logiciels dédiés aux NRPS proposent des monomères sous forme D en résultat de leurs prédictions du monomère sélectionné par un domaine A. Les logiciels d'analyse des NRPS possèdent des banques de séquences de domaines A associant une signature à un monomère incorporé présent sous forme L ou D dans le produit final de la synthétase. Cette information risque donc d'être parfois erronée.

Les données ainsi obtenues vont permettre d'améliorer les prédictions effectuées par les outils bioinformatiques disponibles notamment d'identifier les monomères prédits « X » par ces outils.

IV.5. Etude des synthétases de pyoverdines et de leurs produits pour les autres souches de *Pseudomonas*

Les synthétases des pyoverdines dont le produit n'est pas connu vont être analysées dans ce chapitre. Le tableau 14 reprend les séquences prédites par les outils dédiés aux NRPS pour chacune des neuf souches de *Pseudomonas* dont la composition de la pyoverdine n'est

pas connue ainsi que les prédictions réalisées en exploitant les nouvelles signatures précédemment obtenues. Par ailleurs, une analyse par spectrométrie de masse sur colonies a été réalisée pour les différentes souches en notre possession.

Pour chaque structure prédite, une recherche bibliographique et dans Norine a été entreprise afin de rechercher toute ressemblance avec des structures déjà décrites.

Souches bactériennes	Séquences prédites								
	D-Ser	X	OHOrn	Gln	D-Gln	D-OHOrn	Gly		
<i>P. aeruginosa</i> LESB58	D-Ser	X	OHOrn	Gln	D-Gln	D-OHOrn	Gly		
	D-Ser	Dab	FoOHOrn	Gln	D-Gln	D-FoOHOrn	Gly		
<i>P. aeruginosa</i> UCBPP-PA14	D-Ser	X	D-Ser	OHOrn	Lys	OHOrn	Thr	Thr	
	D-Ser	Arg	D-Ser	FoOHOrn	Lys	FoOHOrn	Thr	Thr	
<i>P. putida</i> F1	Asp	D-X	D-X	Thr	Gly	D-Ser	Ser	Asp	Thr
	Asp	D-FoOHOrn/ OHOrn	D-Dab	Thr	Gly	Ser	Ser	OHAsp	Thr
<i>P. putida</i> GB-1	Asp	Lys	D-Asp	Ser	Gly/Ala	D-Thr	Lys	X	
	Asp	Lys	D-OHAsp	Ser	Ala	D-Thr	Lys	cOHOrn	
<i>P. brassicacearum</i> NFM421	Gly	D-OHOrn	Gly	Gly	D-Ser	Gly/Ala	Asp	Thr	
	Ala	D-FoOHOrn/ OHOrn	Ala	Gly	Ser	Ala	OHAsp	Thr	
<i>P. syringae</i> pv. <i>Phaseolicola</i> 1448A	Lys	D-Asp	Thr	Thr	Ser	D-Asp	Ser		
	Lys	D-OHAsp	Thr	Thr	Ser	D-OHAsp	Ser		
<i>P. syringae</i> pv. <i>Syringae</i> B728a	X	D-Asp	Thr	Thr	Ser	D-Asp	Ser		
	Lys	D-OHAsp	Thr	Thr	Ser	D-OHAsp	Ser		
<i>P. syringae</i> pv. <i>Tomato</i> DC3000	Lys	D-Asp	Thr	Thr	Ser	D-Asp	Ser		
	Lys	D-OHAsp	Thr	Thr	Ser	D-OHAsp	Ser		
<i>P. putida</i> W619	D-Ser	X	Asp	D-Ser	X	Gly	Thr		
	D-Ser	Asn	OHAsp/Asp	D-Ser	cOHOrn	Gly	Thr		

Tableau 14: Séquences des pyoverdines produites par les bactéries du groupe des *Pseudomonas* fluorescents et dont la structure n'est pas connue (séquences prédites par les logiciels d'analyse des NRPS au-dessus et séquences issues de l'analyse des signatures en-dessous, souche par souche).

Les résultats du tableau 13 vont être étudiés au cas par cas :

- *Pseudomonas aeruginosa* LESB58

L'étude des signatures a permis de déterminer la nature du monomère prédit « X » qui est un Dab (acide 2,4-diaminobutyrique). Les deux monomères prédits OH-Orn par les outils bioinformatiques peuvent, également, être des Fo-OH-Orn. La masse de la pyoverdine prédite grâce à l'étude des signatures est calculée par Pyomass, elle est de 1172,275 Da (avec R= succinamide). La pyoverdine produite par *Pseudomonas aeruginosa* Pa6 a une masse moléculaire de 1172 Da (Gipp *et al.*, 1991) ce qui correspond à la masse de la structure prédite avec les monomères en position 3 et 6 de la chaîne étant des formyl-hydroxy-ornithines (Fo-OH-Orn).

- *Pseudomonas aeruginosa* UCBPP-PA14

L'étude des signatures a permis de déterminer la nature du monomère prédit « X » qui est une arginine. Les deux monomères prédits OH-Orn par les outils bioinformatiques peuvent également être des Fo-OH-Orn. La masse de la pyoverdine prédite grâce à l'étude des signatures est calculée par Pyomass, elle est de 1332,505 Da (avec R= succinamide). La pyoverdine produite par *Pseudomonas aeruginosa* PAO1 a une masse moléculaire de 1332 Da ce qui correspond à la masse de la structure prédite avec les monomères en position 4 et 6 de la chaîne étant des formyl-hydroxy-ornithines (Fo-OH-Orn). Les monomères incorporés dans la pyoverdine de cette souche sont identiques à ceux de la souche PAO1, laissant à penser que ces deux souches produisent la même pyoverdine et possèdent donc la même synthétase.

- *Pseudomonas putida* F1

La structure prédite pour la pyoverdine produite par cette souche contient deux monomères prédit comme 'X' en positions 2 et 3 de la chaîne peptidique. Ils ont été identifiés grâce aux signatures spécifiques des pyoverdines comme étant, respectivement, Fo-OH-Orn ou OH-Orn et Dab. De même l'aspartate en position 8 est identifié comme étant un OH-Asp. La structure de la pyoverdine prédite pour la souche *Pseudomonas putida* F1 est proche la pyoverdine produite par la souche putC (Seinsche *et al.* 1993), la seule différence se trouve au niveau du deuxième monomère de la chaîne peptidique, le monomère prédit est une FoOHOrn au lieu de OHbutOHOrn (monomère non traité par les signatures que nous avons établies).

- *Pseudomonas putida* GB-1

Deux corrections ont pu être effectuées grâce à l'étude des signatures spécifiques aux pyoverdines. La première est la prédiction d'un OH-Asp à la place d'un aspartate en position 3 de la chaîne peptidique et la seconde est l'identification du monomère prédit 'X' en position terminale de la chaîne peptidique qui est une cyclo-hydroxy-Ornithine (cOH-Orn). La pyoverdine prédite pour cette souche est proche la pyoverdine produite par la souche 90-51 (Sultana *et al.* 2000b), la seule différence se trouve en position 5 de la chaîne peptidique, le monomère prédit est une alanine au lieu d'une glycine présente dans la pyoverdine produite par la souche 90-51.

- *Pseudomonas brassicacearum* NFM421

La prédiction à partir des signatures spécifiques aux pyoverdines a permis de déterminer la présence d'alanine en positions 1, 3 et 6 de la chaîne peptidique au lieu de glycine prédite par les outils bioinformatiques. De plus l'aspartate prédit en position 7 de la chaîne a été prédit hydroxy-aspartate. La pyoverdine prédite pour la souche *Pseudomonas brassicacearum* NFM421 est proche de la pyoverdine ATCC39167 (Uria Fernandez *et al.* 2003b), les différences se trouvent au niveau du monomère en position 1 de la chaîne peptidique, le monomère prédit est une alanine au lieu d'une sérine et au niveau du monomère en position 2 où la prédiction est une FoOHOrn ou une OHOrn au lieu d'un AcOHOrn (ces 3 monomères possèdent un groupement hydroxamate). La spectrométrie de masse nous donne une masse de 1133 Da pour la pyoverdine de *Pseudomonas brassicacearum* NFM421 qui correspond exactement à la masse de la pyoverdine produite par ATCC39167 ayant la séquence peptidique suivante : D-Ser-D-AcOHOrn-Ala-Gly-(D-Ser-Ala-OHAsp-Thr). L'étude des signatures du code NRPS pour les monomères en position 1 et 2 de la chaîne peptidique montre bien la différence observée entre la structure prédite et la séquence de la pyoverdine de ATCC39167. Néanmoins, la masse est identique à celle de la structure publiée pour la souche ATCC39167. Ces informations réfutent les prédictions réalisées et tendent à indiquer que les monomères en position 1 et 2 seraient respectivement une sérine et une acétyl-hydroxy-ornithine.

Les deux pyoverdines suivantes sont étudiées plus en détail.

- *Pseudomonas syringae* pv. *Phaseolicola* 1448A, *Pseudomonas syringae* pv. *Syringae* B728a et *Pseudomonas syringae* pv. *Tomato* DC3000

Ces trois souches semblent produire la même pyoverdine, elles sont donc étudiées ensemble. Les monomères prédits comme 'Aspartate' sont en réalité des 'hydroxy-Aspartate' (OH-Asp) selon des signatures spécifiques des pyoverdines. Une seule souche, *Pseudomonas syringae* pv. *Syringae* B728a, possède un monomère prédit 'X' qui a été identifié comme étant une lysine grâce aux signatures spécifiques des pyoverdines. Cette prédiction correspond bien au monomère incorporé par les deux autres souches de *Pseudomonas syringae*.

Le cluster de la synthétase de la pyoverdine produite par *Pseudomonas syringae* pv. *syringae* B728a a été étudié en détails. Il est composé de 5 gènes localisés sur le brin sens aux loci Psyr_1945, Psyr_1957, Psyr_1958, Psyr_1959 et Psyr_1960. Le gène situé au locus

Psyr_1945 est annoté « peptide synthase » et les gènes situés aux loci (Psyr_1957, Psyr_1958, Psyr_1959 et Psyr_1960) sont annotés « amino acid adenylation ». Entre les gènes Psyr_1945 et Psyr_1957, se trouvent des gènes codant des protéines membranaires ainsi que des enzymes appartenant au cluster de la pyoverdine synthétase dont la présence est nécessaire à l'activité du sidérophore. Les séquences protéiques des produits de ces gènes ont été analysées par différents logiciels de prédiction (Figure 39).

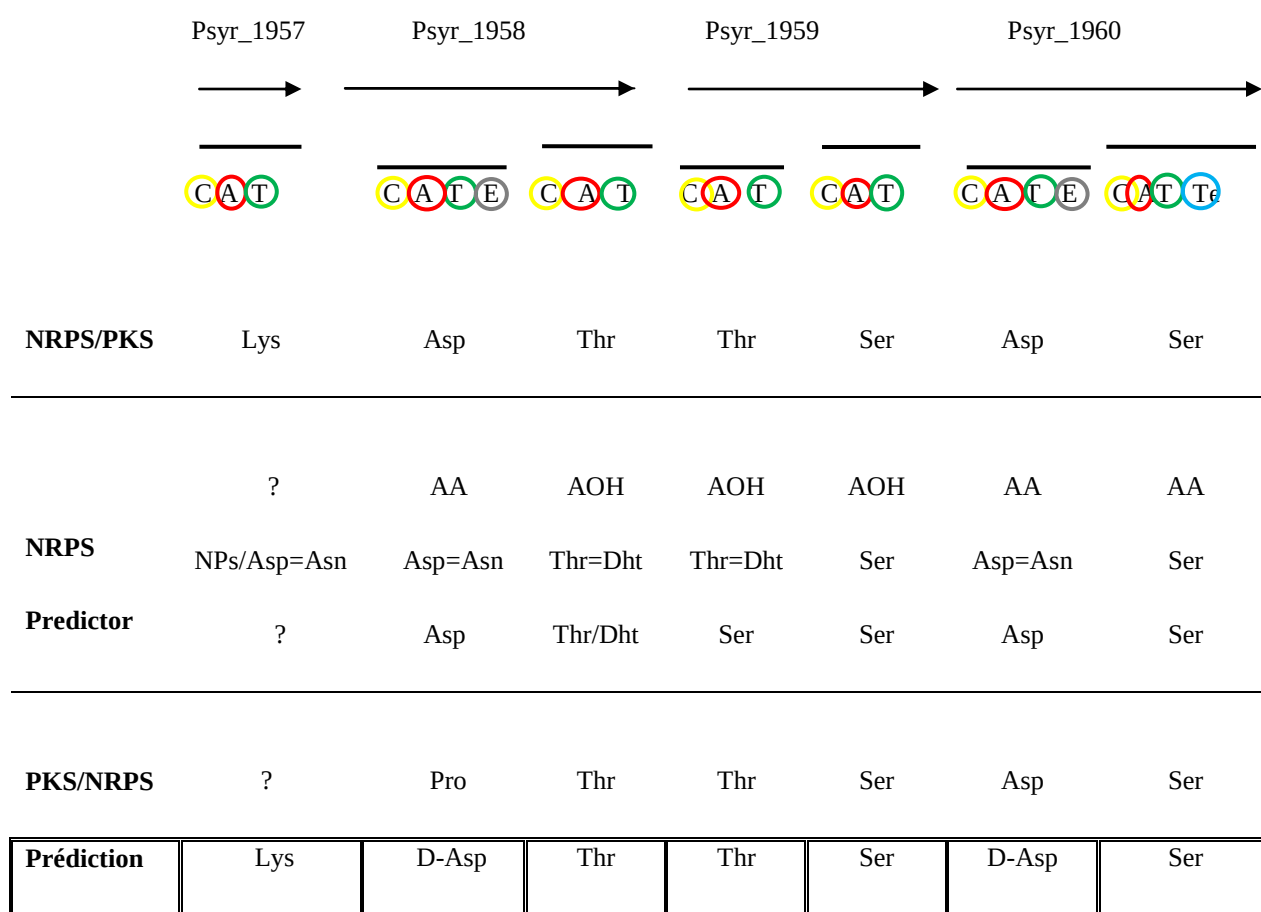


Figure 39:Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas syringae* AOH= Ser,Thr,Ser/Thr,Dhpg,Hpg AA= Asp,Asn,Glu,Gln,Aad NPs= Val,Leu,Ile,Abu, Cette figure reprend les prédictions obtenues par trois logiciels ainsi que l'interprétation de ces résultats dans la case 'prédiction'.

Selon l'interprétation des prédictions obtenues par les différents logiciels, la séquence consensus du peptide produit est « Lys_D-Asp_Thr_Thr_Ser_D-Asp_Ser » et nous l'avons corrigée en « Lys_**D-OH-Asp**_Thr_Thr_Ser_**D-OH-Asp**_Ser » grâce aux signatures spécifiques des pyoverdines. Cette séquence est identique à celle d'une pyoverdine déjà décrite, produite par *Pseudomonas syringae* 19310 (Jülich *et al*, 2001) et par *Pseudomonas syringae* LMG1247 (Bultreys *et al*, 2004) (NOR00199). Le gène Psyr_1945 correspond au

gène pvdL (PA_2424) retrouvé chez *Pseudomonas aeruginosa* PA01 qui est responsable de la synthèse du chromophore de la molécule de pyoverdine (Mossialos *et al.*, 2002). En effet, les protéines codées par ces deux gènes présentent 72% de similarité. L'ensemble de ces informations confirme bien la production d'un sidérophore de type pyoverdine par cette souche. Le même cluster de gènes a été retrouvé chez *Pseudomonas syringae phaseolicola* 1448A et *Pseudomonas syringae* DC3000. L'ensemble de ces informations tend à montrer que les différentes souches de *Pseudomonas syringae* produisent la même molécule de pyoverdine.

- *Pseudomonas putida* W619

Le cluster de la synthétase est composé de 4 gènes dont 2 localisés sur le brin anti-sens aux loci PputW619_3546, PputW619_3545 et 2 sur le brin sens aux loci PputW619_3542 et PputW619_3541. Ces quatre gènes sont annotés « amino acid adenylation containing protein ». Le produit du gène PputW619_3562 présente 72% de similarité avec le produit du gène PvdL de *Pseudomonas aeruginosa* PA01 (PA_2424), ce gène est responsable de la synthèse du chromophore de la pyoverdine. Les séquences protéiques correspondant aux gènes précédents ont été extraites et analysées par les logiciels de prédiction pour déterminer la structure de la synthétase de cette pyoverdine ainsi que la composition potentielle du peptide final. (Figure 40).

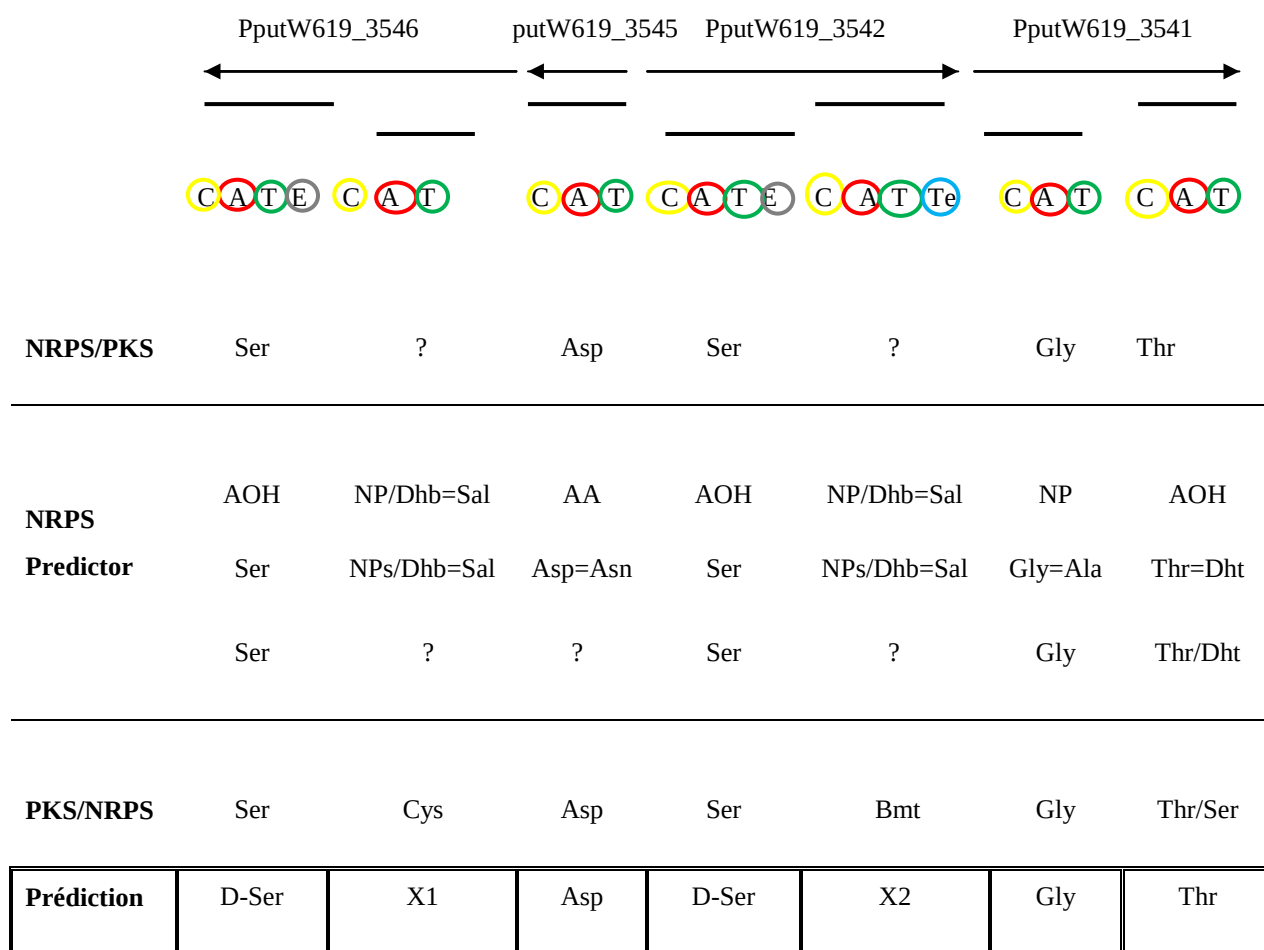


Figure 40: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas putida* W619
 AOH= Ser, Thr, Ser/Thr, Dhpg, Hpg AA= Asp, Asn, Glu, Gln, Aad NPs= Val, Leu, Ile, Abu, Iva NP=Gly, Ala, Val, Leu, Ile, Abu,Iva X=monomère non prédit.

Les outils dédiés aux NRPS, complétés par l'utilisation de nos signatures spécifiques des pyoverdines nous ont permis de prédire la composition suivante pour le peptide produit : « D-Ser, Asn, OH-Asp, D-Ser, cOH-Orn, Gly, Thr » (avec X1 prédit comme Asn et X2 comme cOH-Orn). La composition de cette pyoverdine correspond, à un monomère près, à celle produite par pour *Pseudomonas fluorescens* BTP2 (NOR00165) :

D-Ser _ Val _ OHAsp _ Gly _ Thr _ D-Ser _ cOHOrn

Par contre, l'ordre des monomères n'est pas le même dans les deux peptides.

Cette séquence ne correspond à aucune structure présente dans Norine donc l'analyse de cette synthétase a permis de mettre en évidence une molécule de pyoverdine potentiellement nouvelle. La structure de la synthétase prédite montre une originalité dans le fait que le domaine de la thioestérase (Te) ne se situe pas à la fin de la synthétase. L'ordre des gènes n'est pas identique à l'ordre d'intervention des protéines dans la synthétase. Ceci peut

s'expliquer par un remaniement au niveau des gènes même si un mauvais assemblage lors du séquençage du génome de cette souche ne peut être exclu.

La composition en monomères présente donc 7 monomères sur 8 en commun entre les deux molécules. Le domaine de thioestérase (Te) ne se trouve pas en fin de synthétase, suite à l'observation des monomères incorporés, on observe qu'en inversant les modules codés par les gènes PputW619_3542 et 3541, la séquence des pyoverdines de ces deux souches est très proche hormis le monomère 'X1'.

Afin de déterminer si les structures de ces pyoverdines sont identiques, une analyse en spectrométrie de masse MALDI-TOF du surnageant de culture de *Pseudomonas putida* W619 a été réalisée. Cette analyse a mis en évidence deux masses m/z 1049.464 Da et 1065.474 Da correspondant à la molécule de pyoverdine portant soit la succinamide, soit la malamide comme chaîne latérale du chromophore. En 2001, Ongena *et al.* ont déterminé la structure de la pyoverdine (NOR00165) produite par *Pseudomonas fluorescens* BTP2 ayant une masse nominale m/z 1049 Da correspondant à la forme $[M+H]^+$ de la molécule portant la succinamide comme chaîne latérale. Les deux masses sont identiques, donc ces deux molécules de pyoverdine sont très probablement identiques. Nous pouvons donc corriger la structure prédite pour W619 en remplaçant l'asparagine par une valine. il reste néanmoins une originalité dans la structure dans la synthétase de la pyoverdine de W619 qui ne peut être vérifiée chez *Pseudomonas fluorescens* BTP2 car le génome n'est pas séquencé.

L'étude d'un arbre phylogénétique des domaines A a aussi été utilisé afin de déterminer la nature des monomères prédits X après l'étude bioinformatique par les logiciels d'analyse des NRPS. Des travaux ont déjà montré que les domaines A clusterisent en fonction du monomère qu'ils incorporent et activent.

Le tableau 15 reprend les différents domaines A pour lesquels le monomère incorporé n'a pas pu être déterminé. L'application de l'étude phylogénétique, réalisée sur les domaines A des synthétases de pyoverdine de structure connue, va être utilisé en incorporant les différentes séquences de domaines A ayant pour substrat un monomère noté 'X' afin de vérifier si l'on détermine les mêmes monomères que lors de l'étude des signatures des 10 acides aminés.

souches	Structure pyoverdine déduite de l'étude bioinformatique								
<i>P. aeruginosa</i> LESB58	D-Ser	X1	OHOrn	Gln	D-Gln	D-OHOrn	Gly		
<i>P. aeruginosa</i> UCBPP-PA14	D-Ser	X2	D-Ser	OHOrn	Lys	OHOrn	Thr	Thr	
<i>P. putida</i> F1	Asp	D-X3	D-X4	Thr	Gly	D-Ser	Ser	Asp	Thr
<i>P. putida</i> GB-1	Asp	Lys	D-Asp	Ser	Gly/Ala	D-Thr	Lys	X5	
<i>P. putida</i> W619	D-Ser	X6	Asp	D-Ser	X7	Gly	Thr		
<i>P. aeruginosa</i> LESB58	D-Ser	X1	OHOrn	Gln	D-Gln	D-OHOrn	Gly		
<i>P. aeruginosa</i> UCBPP-PA14	D-Ser	X2	D-Ser	OHOrn	Lys	OHOrn	Thr	Thr	
<i>P. syringae</i> pv. <i>Syringae</i> B728a	X8	D-Asp	Thr	Thr	Ser	D-Asp	Ser		

Tableau 15: Ce tableau indique les références des différents monomères notés 'Xn' qui vont être utilisé dans l'analyse phylogénétique des domaines A.

Les différents monomères 'Xn' du tableau 15 vont être identifiés en fonction de leur position dans l'arbre des domaines A dont le substrat est déterminé. La figure 41 représente le résultat sous la forme d'un arbre de distance de l'alignement des séquences de domaines A dont le substrat est connu avec les séquences de domaines A dont le substrat est noté 'Xn'. Les 8 monomères notés X1 à X8 ont pu être déterminés également de cette manière, les résultats sont représentés dans le tableau 16.

Xn	Monomères correspondant issu de l'arbre de distance des domaines A
X3	*Orn
X2	Arg (R)
X6	Asn (N)
X5	cOHOrn
X7	cOHOrn
X1	Dab
X4	Dab
X8	Lys (K)

Tableau 16: Identification des monomères prédits 'X' par les outils d'analyse des synthétases disponibles par étude de la séquence du domaine A entier en comparant avec des séquences de domaines A dont le substrat est déjà déterminé.

Les monomères déterminés de cette manière correspondent parfaitement à ceux déterminés par l'étude des signatures des 10 acides aminés. Néanmoins, une exception existe pour le monomère X6 identifié par les deux méthodes comme étant une asparagine (Asn), ce monomère serait en réalité une valine comme décrit dans l'exemple 2 du paragraphe IV.2.1.

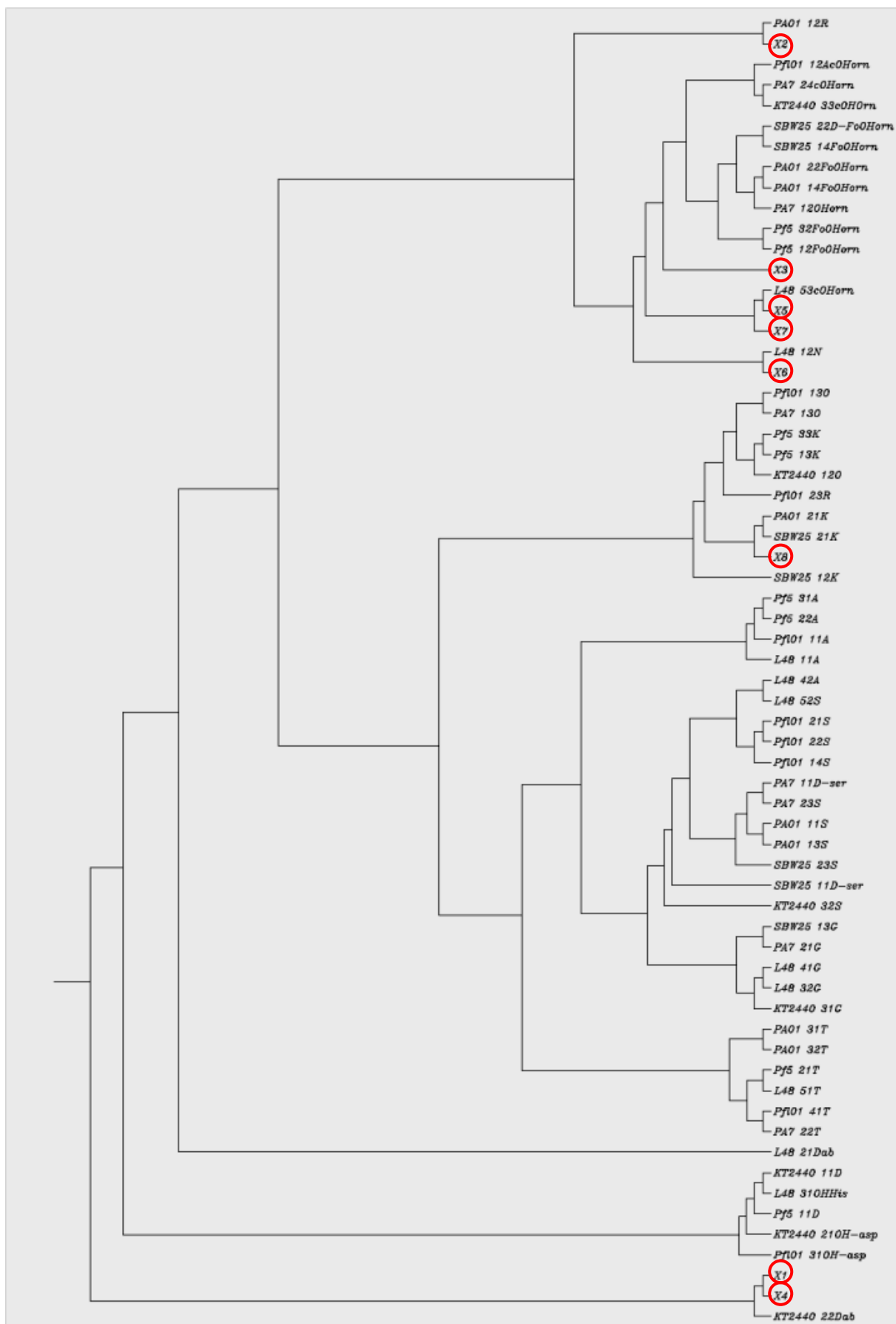


Figure 41: Représentation sous la forme d'un arbre de distance de l'alignement des séquences des domaines A des synthétases de pyoverdines dont la séquence est connue avec les domaines A dont l'analyse bioinformatique n'a pas permis la détermination du substrat (Xn). Les domaines A sont décrits comme suit sur cette figure : nom de la souche de *Pseudomonas*/le premier chiffre correspond à la position de la protéine dans la synthétase/ le deuxième chiffre correspond à la position du module dans la protéine/acide aminé incorporé par le domaine A.

On observe que l'étude du domaine A entier permet de déterminer la nature du monomère qu'il est capable d'incorporer. Néanmoins, on a observé que l'étude des signatures des 10 acides aminés est plus précise pour la distinction entre les différents dérivés de l'ornithine.

Le tableau 17 reprend les résultats discutés ci-dessus avec les séquences prédites grâce à la comparaison des signatures et les séquences de pyoverdine identiques ou les plus proches retrouvés dans la littérature.

souches bactériennes	séquence issue de l'analyse des signatures de Stachelhaus	séquence identique issue de la littérature	pyoverdine
<i>P. aeruginosa</i> LESB58	D-Ser-Dab-FoOHOrn/OHOrn-Gln-D-Gln-D-FoOHOrn/OHOrn-Gly	(D-Ser-Dab)-FoOHOrn-Gln-D-Gln-D-FoOHOrn-Gly	Pa6
<i>P. aeruginosa</i> UCBPP-PA14	D-Ser-Arg-D-Ser-FoOHOrn/OHOrn-Lys-FoOHOrn/OHOrn-Thr-Thr	D-Ser-Arg-D-Ser-FoOHOrn-(Lys-FoOHOrn-Thr-Thr)	PAO1
<i>P. putida</i> F1	Asp-D-FoOHOrn/OHOrn-D-Dab-Thr-Gly-D-Ser-Ser-OHAsp-Thr	Asp-D-OHbutOHOrn-D-Dab-Thr-Gly-D-Ser-Ser-OHAsp-Thr	putC
<i>P. putida</i> GB-1	Asp-Lys-D-OHAsp-Ser-Ala-D-Thr-Lys-cOHOrn	Asp-εLys-D-OHAsp-Ser-Gly-D-αThr-Lys-cOHOrn	90-51
<i>P. putida</i> W619	D-Ser-Asn-OHAsp/Asp-D-Ser-cOHOrn-Gly-Thr	D-Ser-Val-OHAsp-Gly-Thr-D-Ser-cOHOrn	BTP2
<i>P. syringae</i> pv. <i>Phaseolicola</i> 1448A	Lys-D-OHAsp-Thr-Thr-Ser-D-OHAsp-Ser	εLys-D-OHAsp-Thr-(Thr-Ser-D-OHAsp-Ser)	Syr19310
<i>P. syringae</i> pv. <i>Syringae</i> B728a	Lys-D-OHAsp-Thr-Thr-Ser-D-OHAsp-Ser		
<i>P. syringae</i> pv. <i>tomato</i> DC3000	Lys-D-OHAsp-Thr-Thr-Ser-D-OHAsp-Ser		
<i>P. brassicacearum</i> NFM421	Ala-D-FoOHOrn/OHOrn-Ala-Gly-D-Ser-Ala-OHAsp-Thr	D-Ser-D-AcOHOrn-Ala-Gly-(D-Ser-Ala-OHAsp-Thr)	ATCC39167

Tableau 17: Liste des séquences prédites pour les pyoverdines à partir de l'étude des signatures des domaines A en regard de la séquence la plus proche issue de la littérature.

Différents outils d'analyse des synthétases NRPS sont déjà disponibles. Dans cette partie, nous avons appliqué ces logiciels ainsi qu'une étude phylogénétique à l'analyse des synthétases de pyoverdines. Cette étude a permis de montrer que l'analyse phylogénétique des domaines A permet d'identifier l'acide aminé incorporé. La prédiction du produit des synthétases de pyoverdines est améliorée par l'étude des signatures du code NRPS.

V. Amplification de la biodiversité au sein des pyoverdines par feeding

Différents travaux (Baumgart *et al.*, 1991 ; Bonmatin *et al.*, 1995 ; Van Dörhen *et al.*, 1997 ; Kurmayer *et al.*, 2001) ont montré que certains modules de synthétase peptidiques non-ribosomiales pouvaient être permissifs pour l'incorporation d'un acide aminé. En effet, certains modules seraient capables d'incorporer et d'activer au moins deux acides aminés différents. A l'inverse, les modules dits spécifiques ne peuvent incorporer et activer qu'un seul acide aminé quelle que soit les conditions de culture. Dans cette partie du travail, les onze souches de *Pseudomonas* fluorescents en notre possession et dont le génome a été séquencé, ont été cultivées sur des milieux de culture enrichis en un acide aminé particulier (feeding) afin de vérifier s'il est possible d'obtenir de nouveaux variants de pyoverdines et de définir des signatures de domaines d'adénylation permissifs. L'incorporation de cet acide aminé à la place de l'acide aminé naturellement incorporé par un module sera suivie par spectrométrie de masse MALDI-TOF.

V.1. Stratégie d'analyse

Chacune des souches de *Pseudomonas* disponibles est mise en culture sur milieu solide dans différentes conditions de feeding. Chaque souche est cultivée en présence des vingt acides aminés protéogéniques et de l'ornithine comme source d'azote, et sur un milieu témoin où la source d'azote est le sulfate d'ammonium. Chaque culture est réalisée dans des conditions de carence en fer nécessaire à la production de sidérophores par les *Pseudomonas*. L'ensemble des conditions testées et les résultats obtenus sont représentés dans le tableau 18.

	A	R	N	D	C	E	Q	G	H	I	L	K	M	O	F	P	S	T	W	Y	V
PA7	3	3	3	3	2	3	3	3	3	3	3	3	2	3	3	1	4	3	3	3	2
L48	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	3	3	3
Pf5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
KT2440	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
B728a	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	3	3	3
F1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
SBW25	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
W619	3	3	3	3	1	3	3	3	3	2	3	3	2	3	3	3	3	3	2	3	3
PfO-1	3	3	3	3	1	3	3	3	3	3	2	3	2	3	2	2	3	2	2	2	3
GB-1	2	3	3	3	1	3	3	3	3	3	3	3	2	3	3	3	3	2	2	2	3
NFM421	3	3	3	3	1	3	3	3	3	3	3	2	2	2	2	3	3	2	2	2	3

Tableau 18: Conditions de feeding testées en fonction des souches de *Pseudomonas* disponibles. Le code utilisé pour la description des résultats du feeding est le suivant : 1, la souche étudiée n'a pas poussé dans les conditions testées, 2, la souche n'a pas produit de pyoverdine dans les conditions testées, 3, la pyoverdine produite a une

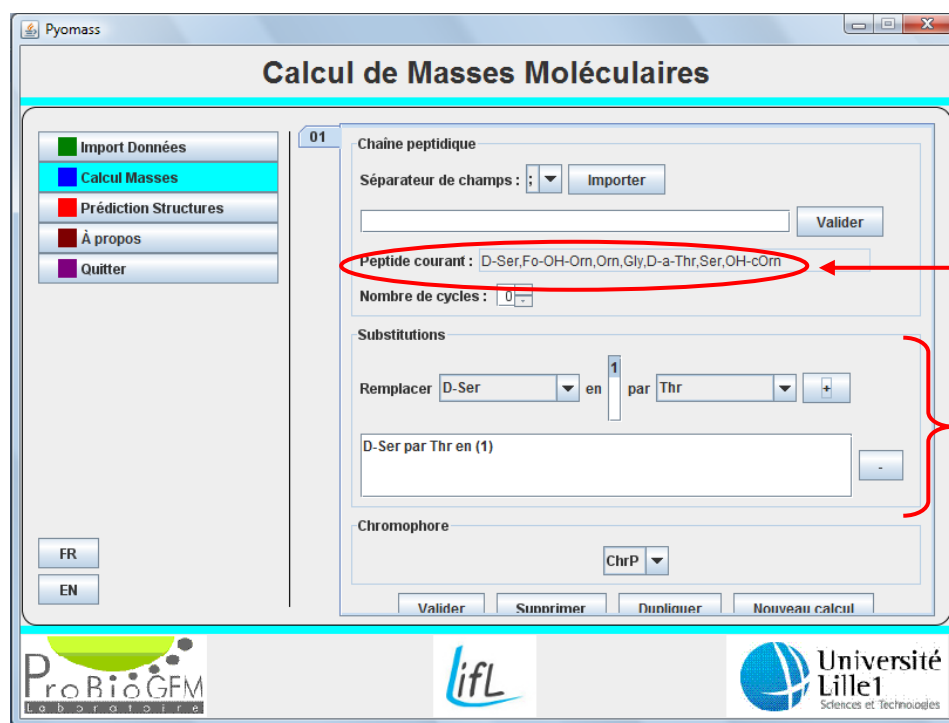
composition identique à la pyoverdine connue pour la souche étudiée et 4, une substitution a été observée dans la composition de la pyoverdine suite aux expériences de feeding.

Dans ce tableau sont reprises les conditions testées pour le feeding, et n'apparaissent sur fond jaune que les cas où l'on observe la production d'une molécule de pyoverdine présentant une masse différente de celle ordinairement produite par la souche étudiée. Néanmoins pour certaines des autres conditions testées, les souches n'ont pas réussi à croître ou n'ont pas produit de pyoverdines ce qui pourrait, dans ce dernier cas, s'expliquer par la présence de fer contaminant qui inhibe la production de sidérophore par la souche de *Pseudomonas*.

V.2. Application de Pyomass à l'étude des résultats de masse du feeding

Nous possédons les masses des peptides produits dans les conditions normales de cultures, ainsi que leur séquence peptidique. Nous mesurons les masses des produits présents sur la boîte de culture au niveau des colonies bactériennes après le feeding. Une masse proche de la masse originale du peptide est potentiellement représentative de l'incorporation de l'acide aminé fourni en excès. La recherche du monomère qui a été remplacé par feeding demande de nombreux calculs, vu les différentes chaînes latérales possibles, qui peuvent être facilement automatisés. Ils ont été intégrés dans Pyomass.

A partir d'une séquence peptidique saisie grâce à l'interface de Pyomass, un acide aminé mis en excès dans le milieu de culture peut être sélectionné (figure 42) puis les calculs automatiques sont lancés pour prédire la substitution d'un acide aminé de la chaîne par l'acide aminé apporté en excès. Le résultat de ces calculs est donné sous la forme d'un tableau (figure 43) qui reprend d'une part les masses propres de la molécule initiale donnée en requête en fonction des différentes chaînes latérales possibles du chromophore ainsi que les ions $(M+H)^+$, $(M+Na)^+$ et $(M+K)^+$ et d'autre part les mêmes calculs avec chacun des monomères du peptide remplacé par l'acide aminé en excès. Les résultats peuvent ensuite être exportés au format Excel.



Séquence requête

Partie dédiée au feeding, on peut y choisir les substitutions avec leur position sur la chaîne peptidique

Figure 42: Capture d'écran du logiciel Pyomass présentant l'interface de calcul qui permet de calculer la masse d'une pyoverdine avec des substitutions dans la chaîne peptidique.

Résultats Calcul de Masse

Exporter Rechercher :

Chaîne peptidique (Cycles : 0) : D-Ser, Fo-OH-Orn, Orn, Gly, D-a-Thr, Ser, OH-cOrn,
Chromophore : ChrP

Chaîne latérale	Masse propre	H+	Na+	K+
acide glutamiqu...	1120.228	1121.228	1143.228	1159.228
acide malique	1107.186	1108.186	1130.186	1146.186
acide succinique	1091.187	1092.187	1114.187	1130.187
malamide	1106.199	1107.199	1129.199	1145.199
alpha-c-toglut...	1119.209	1120.209	1142.209	1158.209
succinamide	1090.199	1091.199	1113.199	1129.199

Chaîne peptidique (Cycles : 0) : **Thr**, Fo-OH-Orn, Orn, Gly, D-a-Thr, Ser, OH-cOrn,
Chromophore : ChrP

Chaîne latérale	Masse propre	H+	Na+	K+
acide glutamiqu...	1134.255	1135.255	1157.255	1173.255
acide malique	1121.213	1122.213	1144.213	1160.213
acide succinique	1105.214	1106.214	1128.214	1144.214
malamide	1120.226	1121.226	1143.226	1159.226
alpha-c-toglut...	1133.236	1134.236	1156.236	1172.236
succinamide	1104.226	1105.226	1127.226	1143.226

Figure 43: Capture d'écran du tableau de résultat de Pyomass qui confronte dans cet exemple les masses de la molécule entrée comme requête avec la même molécule à laquelle on a substitué le premier acide aminé (D-Ser → Thr)

V.3. Analyse des résultats obtenus par spectrométrie de masse MALDI-TOF

Parmi les souches de *Pseudomonas* utilisées pour cette étude, seules trois (*P. aeruginosa* PA7, *P. entomophila* L48 et *P. syringae* B728a) montrent des substitutions au niveau de la chaîne peptidique des pyoverdines qu'elles produisent lors du feeding. Les figures 44, 46 et 48 présentent les spectres de masse MALDI-TOF des molécules de pyoverdines produites par ces trois souches.

V.3.1. *Ps. aeruginosa* PA7

La molécule de pyoverdine produite par cette souche possède la structure suivante :



L'ajout de la sérine dans le milieu de culture de cette souche induit une modification de la composition de la chaîne peptidique de la pyoverdine produite (figure 44). Sur le spectre B de cette figure nous observons un m/z de 1077 Da correspondant à l'ion $[M+H]^+$ d'une molécule de masse de 1076 Da. Selon le résultat du logiciel Pyomass (figure 45), cette masse correspond à la substitution de l'allo-thréonine en position 5 de la chaîne par une sérine. Néanmoins la différence de 14 Da, par rapport à la masse 1090 Da de la molécule de pyoverdine possédant la structure ci-dessus, peut aussi être expliquée par d'autres substitutions comme celle de l'ornithine en position 3 de la chaîne peptidique par un acide 2,4 diaminobutyrique (Dab). La substitution privilégiée reste tout de même le remplacement de la D-allo-thréonine (D-aThr) en position 5 par une sérine, étant donné que la sérine est ajoutée en excès dans le milieu de culture pour le feeding.

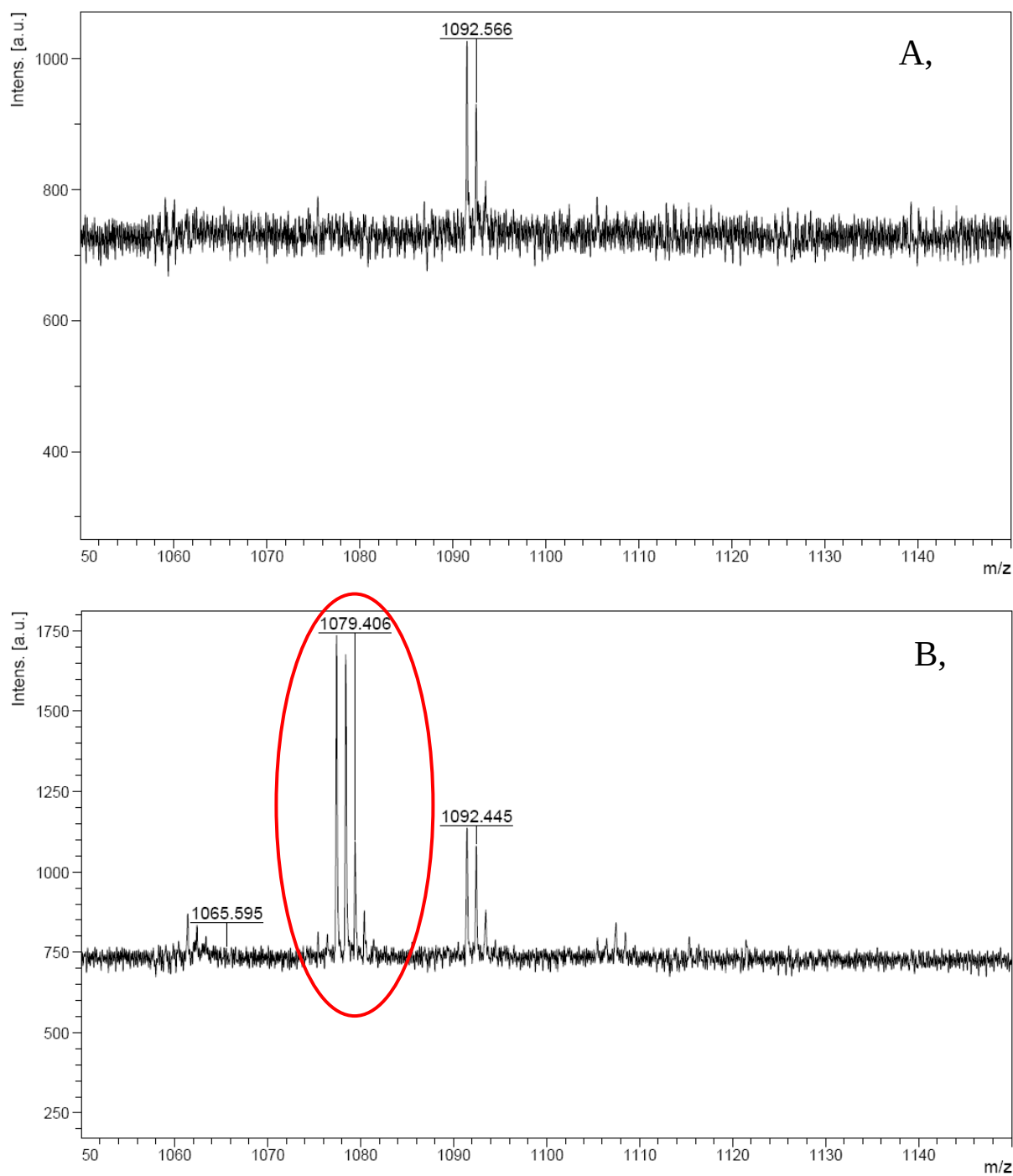


Figure 44: Spectre de masse MALDI-TOF en mode réflectron d'une culture de *Pseudomonas aeruginosa* PA7 produisant une molécule de pyoverdine témoin (A), en présence de sérine dans le milieu de culture (B).

Résultats Calcul de Masse

Exporter Rechercher :

Chaîne peptidique (Cycles : 0) : Ser, Fo-OH-Orn, Orn, Gly, a-Thr, Ser, OH-cOrn,
Chromophore : ChrP

Chaîne latérale	Masse propre	H+	Na+	K+
acide glutami...	1120.228	1121.228	1143.228	1159.228
acide malique	1107.186	1108.186	1130.186	1146.186
acide succiniq...	1091.187	1092.187	1114.187	1130.187
malamide	1106.199	1107.199	1129.199	1145.199
alpha-c-toglu...	1119.209	1120.209	1142.209	1158.209
succinamide	1090.199	1091.199	1113.199	1129.199

Chaîne peptidique (Cycles : 0) : Ser, Fo-OH-Orn, Orn, Gly, **Ser**, Ser, OH-cOrn,
Chromophore : ChrP

Chaîne latérale	Masse propre	H+	Na+	K+
acide glutami...	1106.202	1107.202	1129.202	1145.202
acide malique	1093.16	1094.16	1116.16	1132.16
acide succiniq...	1077.16	1078.16	1100.16	1116.16
malamide	1092.172	1093.172	1115.172	1131.172
alpha-c-toglu...	1105.182	1106.182	1128.182	1144.182
succinamide	1076.172	1077.172	1099.172	1115.172

Figure 45: Fenêtre de résultat du logiciel pyomass. Le tableau supérieur reprend le calcul des masses moléculaires de la pyoverdine de *P. aeruginosa* PA7 en fonction de la chaîne latérale portée par le chromophore et le tableau inférieur reprend le même calcul mais avec la substitution de aThr en position 5 par une sérine.

V.3.2. *Ps. syringae* B728a

La molécule de pyoverdine produite par cette souche possède la structure suivante :



La masse de cette molécule est de 1122 Da. Les expériences de feeding avec cette souche ont montré que la composition de la chaîne peptidique de la pyoverdine était modifiée lors de l'ajout dans le milieu de culture de thréonine ou de sérine.

Lors de l'ajout de thréonine dans le milieu de culture nous observons un m/z de 1134 Da (Figure 46) correspondant à l'ion $[M+K]^+$ d'une molécule de masse de 1095 Da, cette masse correspond à la substitution de la lysine en position 1 par une thréonine avec la succinamide comme chaîne latérale du chromophore. Néanmoins, cette hypothèse est peu

probable du fait de l'absence de l'ion $[M+H]^+$ sur le spectre. Toutefois, cette substitution est la seule possibilité d'obtention de la masse observée dans les conditions testées.

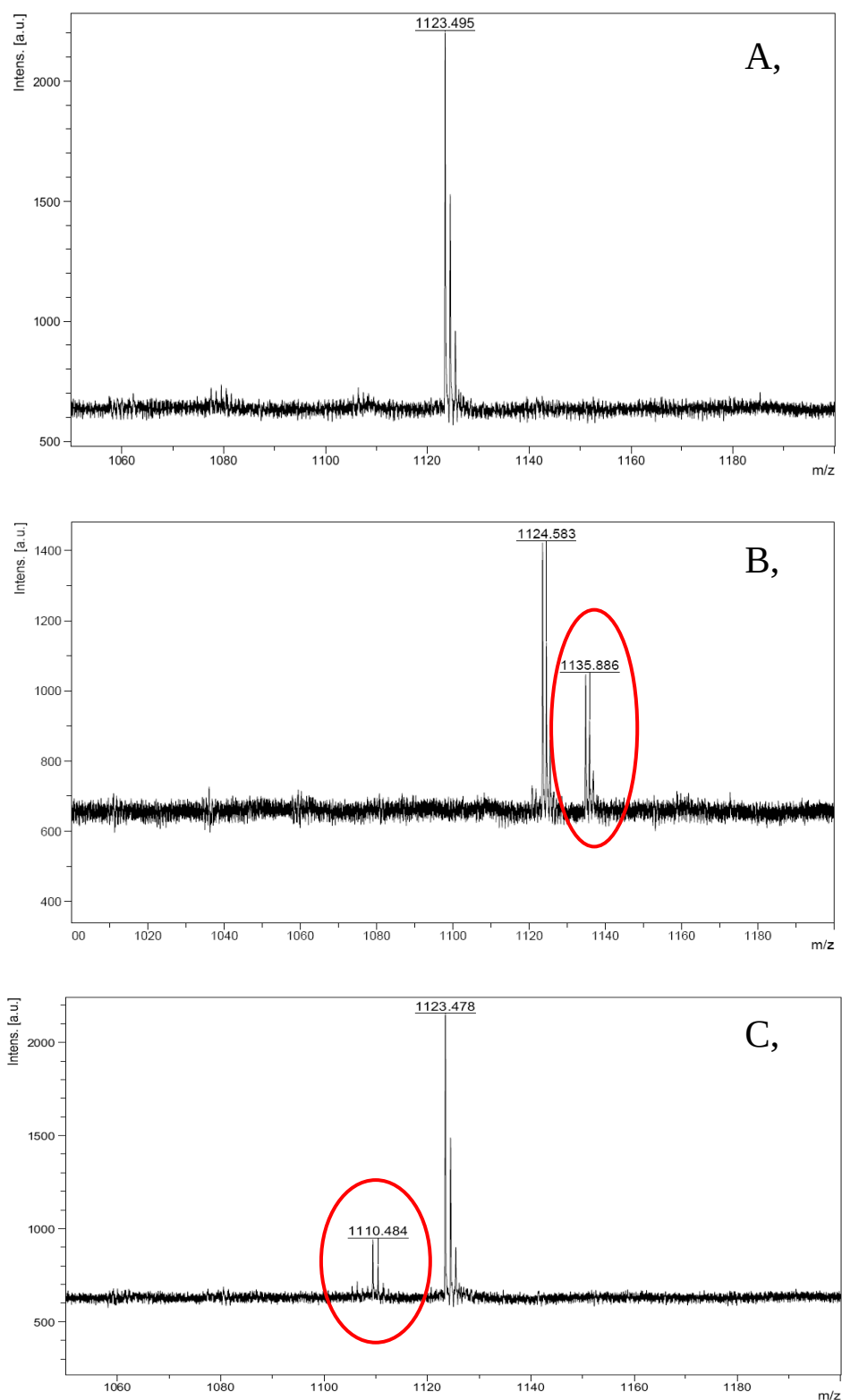


Figure 46: Spectre de masse MALDI-TOF en mode réflectron d'une culture de *Pseudomonas syringae* B728a produisant une molécule de pyoverdine témoin (A), en présence de thréonine dans le milieu de culture (B) et en présence de sérine (C).

Avec l'ajout de la sérine dans le milieu, apparaît un m/z de 1109 Da correspondant à l'ion $[M+H]^+$ d'une molécule de masse 1108 Da qui concorde avec une substitution de la thréonine en position 3 ou 4 de la chaîne peptidique par une sérine avec la succinamide comme chaîne latérale du chromophore. La figure 47 reprend les calculs par Pyomass de la masse de la pyoverdine produite par cette souche en fonction de l'acide aminé ajouté en excès dans le milieu de culture lors de l'expérience, c'est-à-dire la sérine. Cette différence de 14 Da entre la masse de la pyoverdine produite par cette souche et la masse de la pyoverdine obtenue après le feeding peut être expliquée par la substitution de la lysine en position 1 par une ornithine ou encore une substitution de l'aspartate en position 2 ou 6 de la chaîne peptidique par une thréonine. Néanmoins, au vu des conditions de feeding utilisées ici, c'est-à-dire la sérine en excès dans le milieu de culture, l'hypothèse de la substitution de la thréonine en position 3 ou 4 par une sérine reste privilégiée.

1

2

3

4

Résultats Calcul de Masse					
Exporter		Rechercher :			
Chaîne peptidique (Cycles : 1) : Lys, OH-Asp, Thr, Thr, Ser, OH-Asp, Ser,					
Chromophore : ChrP					
Chaîne latér...	Masse propre	H+	Na+	K+	
acide glutam...	1152.196	1153.196	1175.196	1191.196	
acide malique	1139.154	1140.154	1162.154	1178.154	
acide succini...	1123.155	1124.155	1146.155	1162.155	
malamide	1138.167	1139.167	1161.167	1177.167	
alpha-c-togl...	1151.177	1152.177	1174.177	1190.177	
succinamide	1122.167	1123.167	1145.167	1161.167	
Chaîne peptidique (Cycles : 1) : Thr, OH-Asp, Thr, Thr, Ser, OH-Asp, Ser,					
Chromophore : ChrP					
Chaîne latér...	Masse propre	H+	Na+	K+	
acide glutam...	1125.128	1126.128	1148.128	1164.128	
acide malique	1112.086	1113.086	1135.086	1151.086	
acide succini...	1096.086	1097.086	1119.086	1135.086	
malamide	1111.098	1112.098	1134.098	1150.098	
alpha-c-togl...	1124.108	1125.108	1147.108	1163.108	
succinamide	1095.098	1096.098	1118.098	1134.098	
Chaîne peptidique (Cycles : 1) : Lys, OH-Asp, Ser, Thr, Ser, OH-Asp, Ser,					
Chromophore : ChrP					
Chaîne latér...	Masse propre	H+	Na+	K+	
acide glutam...	1138.169	1139.169	1161.169	1177.169	
acide malique	1125.128	1126.128	1148.128	1164.128	
acide succini...	1109.128	1110.128	1132.128	1148.128	
malamide	1124.14	1125.14	1147.14	1163.14	
alpha-c-togl...	1137.15	1138.15	1160.15	1176.15	
succinamide	1108.14	1109.14	1131.14	1147.14	
Chaîne peptidique (Cycles : 1) : Lys, OH-Asp, Thr, Ser, Ser, OH-Asp, Ser,					
Chromophore : ChrP					
Chaîne latér...	Masse propre	H+	Na+	K+	
acide glutam...	1138.169	1139.169	1161.169	1177.169	
acide malique	1125.128	1126.128	1148.128	1164.128	
acide succini...	1109.128	1110.128	1132.128	1148.128	
malamide	1124.14	1125.14	1147.14	1163.14	
alpha-c-togl...	1137.15	1138.15	1160.15	1176.15	
succinamide	1108.14	1109.14	1131.14	1147.14	

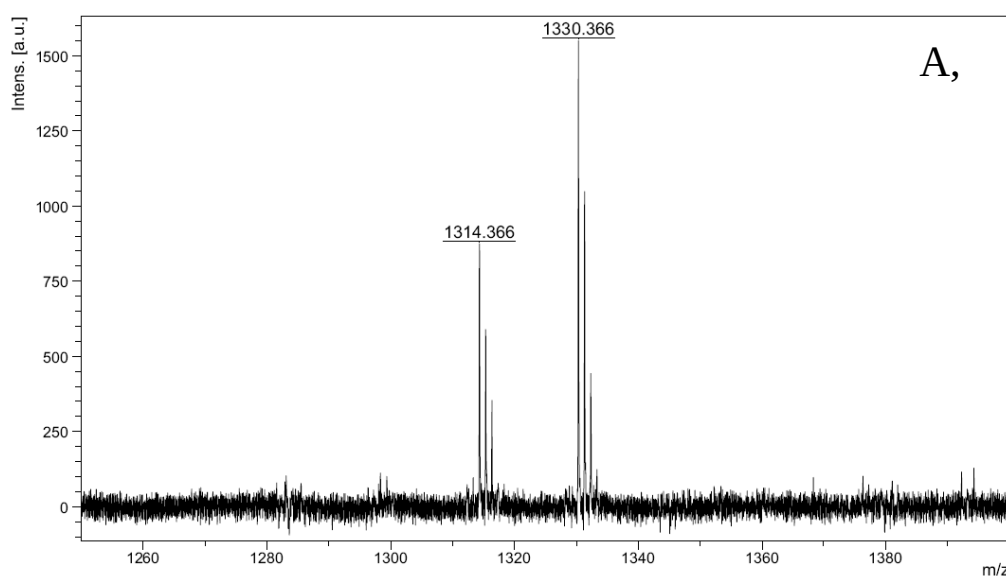
Figure 47: Fenêtre de résultat du logiciel pyomass. Le tableau 1 reprend le calcul des masses moléculaires de la pyoverdine de *P. syringae* B728a en fonction de la chaîne latérale portée par le chromophore . Le tableau 2 reprend le même calcul mais avec la substitution de Lys en position 1 par une thréonine. Et les tableaux 3 et 4 reprennent le calcul avec la substitution de la thréonine en position 3et 4 respectivement par une sérine.

V.3.3. *Ps. entomophila* L48

La molécule de pyoverdine produite par cette souche possède la structure suivante :



La masse de cette molécule de pyoverdine est de 1313 Da avec comme chaîne latérale du chromophore la succinamide. L'ajout de la sérine dans le milieu de culture de *Ps. entomophila* L48 induit une modification de la masse moléculaire de la pyoverdine, sur la figure 48 un massif de pics est présent autour de 1300 Da alors que nous n'observons pas ces pics sur le spectre témoin. Sur le spectre B de cette figure nous observons une m/z de 1300 Da correspondant à l'ion moléculaire $[M+H]^+$ d'une molécule de masse de 1299 Da. Selon le résultat du logiciel Pyomass (figure 49), cette masse correspond à la substitution de la thréonine en position 8 de la chaîne par une sérine, avec la succinamide comme chaîne latérale du chromophore. Cette différence de 14 Da entre les masses observées avec et sans ajout de sérine dans le milieu de culture peut aussi être expliquée par une substitution de l'alanine en position 1 de la chaîne par une glycine. Néanmoins la substitution de la thréonine par une sérine est plus logique au vu des conditions de feeding utilisées ici.



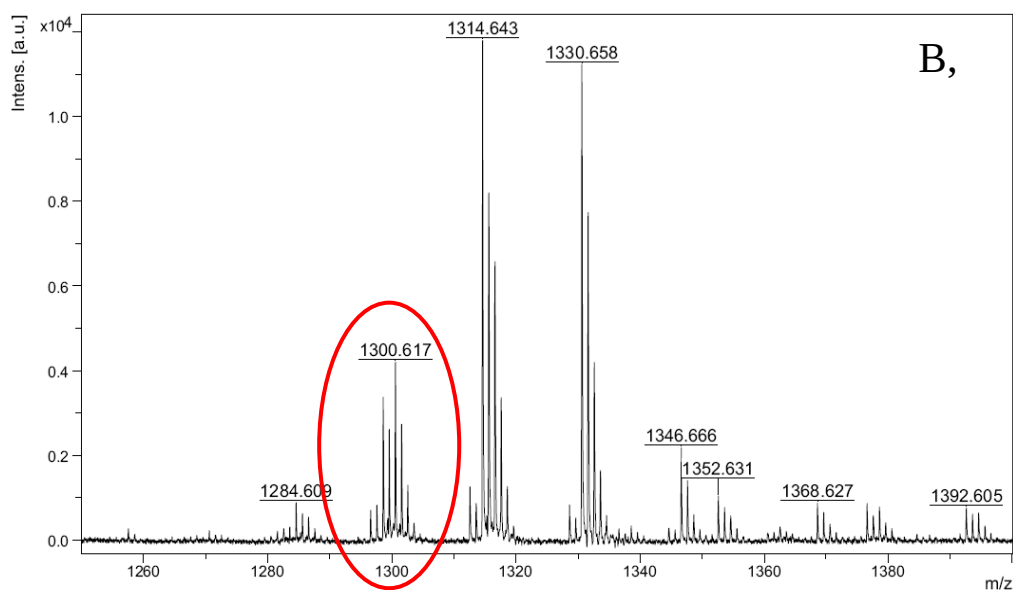


Figure 48: Spectre de masse MALDI-TOF en mode réflectron d'une culture de *Pseudomonas entomophila* L48 produisant une molécule de pyoverdine témoin (A), en présence de sérine dans le milieu de culture (B).

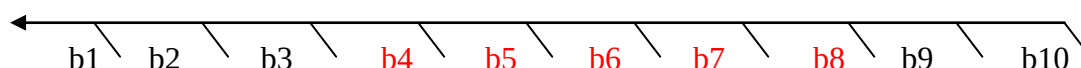
Pour vérifier que la substitution de la thréonine en position 8 par une sérine est bien la raison de l'obtention de cette nouvelle masse, une étude LC-MS-MS été réalisée.

Résultats Calcul de Masse				
Exporter Rechercher :				
Chaîne peptidique (Cycles : 0) : Ala, Asn, Dab, OH-His, Gly, Gly, Ser, a-Thr, Ser, OH-cOrn, Chromophore : ChrP				
Chaîne latérale	Masse propre	H+	Na+	K+
acide glutamique	1343.463	1344.463	1366.463	1382.463
acide malique	1330.421	1331.421	1353.421	1369.421
acide succinique	1314.422	1315.422	1337.422	1353.422
malamide	1329.434	1330.434	1352.434	1368.434
alpha-c-toglutar...	1342.444	1343.444	1365.444	1381.444
succinamide	1313.434	1314.434	1336.434	1352.434
Chaîne peptidique (Cycles : 0) : Ala, Asn, Dab, OH-His, Gly, Gly, Ser, Ser , Ser, OH-cOrn, Chromophore : ChrP				
Chaîne latérale	Masse propre	H+	Na+	K+
acide glutamique	1329.436	1330.436	1352.436	1368.436
acide malique	1316.394	1317.394	1339.394	1355.394
acide succinique	1300.395	1301.395	1323.395	1339.395
malamide	1315.407	1316.407	1338.407	1354.407
alpha-c-toglutar...	1328.417	1329.417	1351.417	1367.417
succinamide	1299.407	1300.407	1322.407	1338.407

Figure 49: Fenêtre de résultat du logiciel pyomass. Le tableau supérieur reprend le calcul des masses moléculaires de la pyoverdine de *P. entomophila* L48 en fonction de la chaîne latérale portée par le chromophore et le tableau inférieur reprend le même calcul mais avec la substitution de aThr en position 8 par une sérine.

V.4. Analyse MS2 du cas de *Pseudomonas entomophila* L48 cultivée en présence de sérine dans le milieu de culture

La culture de *Pseudomonas entomophila* L48 en présence de sérine montre une modification de la masse moléculaire de la pyoverdine qui, selon les prédictions, est due à l'incorporation de la sérine à la place de la thréonine en position 8 de la chaîne peptidique. Une analyse LC-MS-MS (figure 50) d'un surnageant de culture de cette souche cultivée en présence de sérine a permis de déterminer les différents acides aminés constituant la chaîne peptidique de la pyoverdine. Cette analyse confirme la substitution de la thréonine en position 7 par un résidu de sérine. En effet, l'interprétation du tableau 19 montre que de b8 à b4, la séquence peptidique est sérine/sérine/glycine/glycine.



Ala _ Asn _ Dab _ OHHis _ Gly _ Gly _ Ser _ **Ser** _ Ser _ cOHOrn

fragment	séquences	Masse (Da)
b4	Succa_ChrP_Ala_Asn_Dab_OH-His	796,3
b4-H2O		778,3
b5	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly	853,3
b5-H2O		835,5
b6	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly_Gly	910,4
b6-H2O		892,4
b7	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly_Gly_Ser	997,4
b7-H2O		979,4
b8	Succa_ChrP_Ala_Asn_Dab_OH-His_Gly_Gly_Ser_Ser	1084,8
b8-H2O		1066,5

Tableau 19: Résultats de l'étude LC-MS-MS du spectre correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 en présence de sérine dans le milieu de culture.

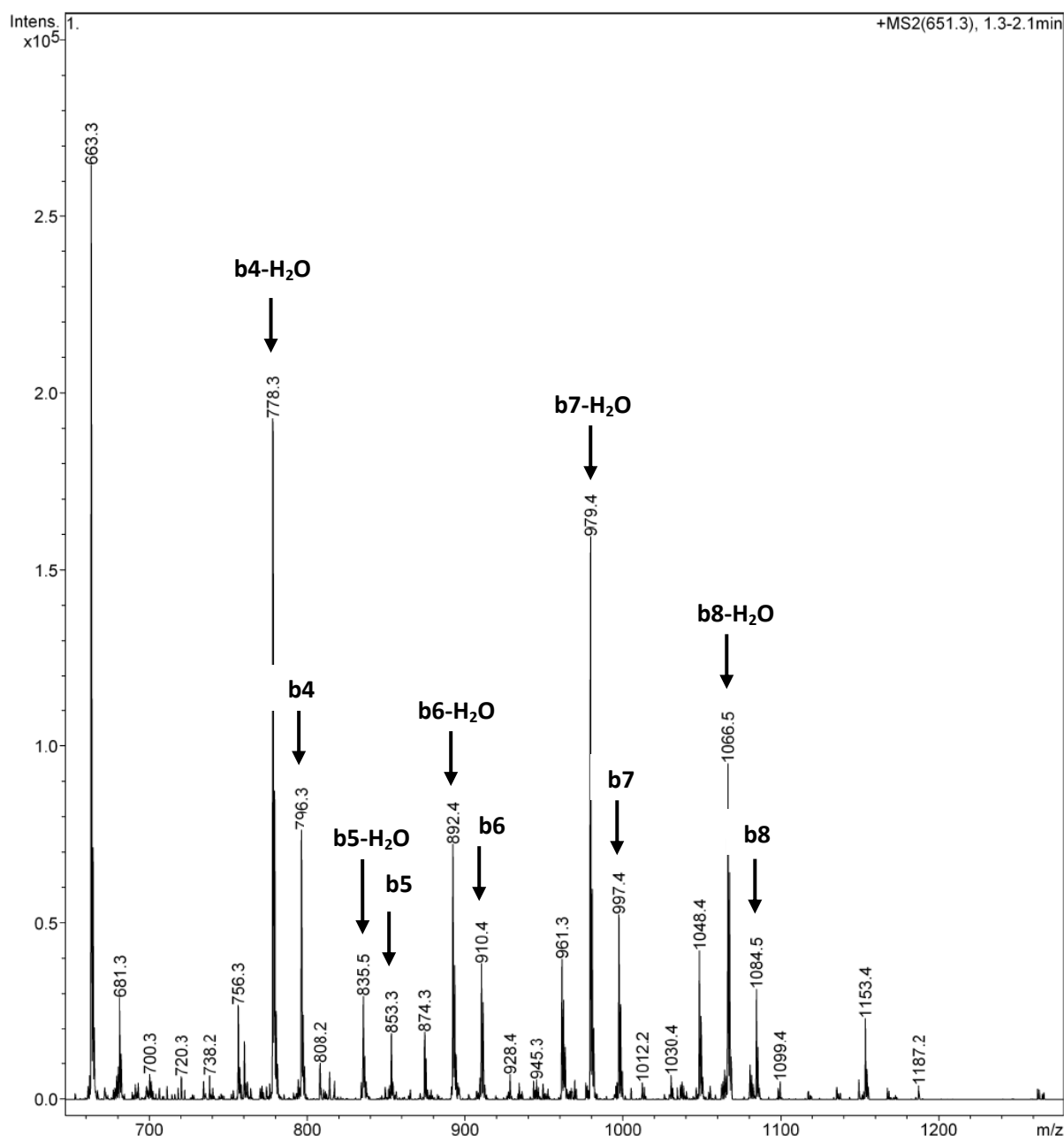


Figure 50: Spectre MS2 du pic correspondant à la molécule de pyoverdine produite par *Pseudomonas entomophila* L48 cultivée en présence de sérine ($[M+H]^+ = 1300,617$ Da), avec les masses correspondantes aux différents fragments du tableau 11.

L'analyse en LC-MS-MS a permis de confirmer la présence d'un domaine A permissif dans la structure de la synthétase de la pyoverdine produite par *Pseudomonas entomophila* L48. Au vu de ce résultat, l'étude des autres cas de substitutions observées en analyses MALDI-TOF est envisagée dans le but de confirmer l'existence d'autres domaines A permissifs dans les synthétases de pyoverdine.

Dans cette partie du travail, nous avons pu mettre en évidence des modifications au niveau de la chaîne peptidique des molécules de pyoverdine lors d'expérience de feeding avec différents acides aminés. Ces expériences de feeding ont permis d'une part de montrer que les substitutions dans la chaîne peptidique des pyoverdines étaient possibles dans certaines conditions et d'autre part de mettre en évidence des domaines A tolérants quant à la reconnaissance d'un acide aminé particulier lors de la synthèse non-ribosomiale.

VI. Analyse des autres NRPS de *Pseudomonas*

Cette partie va traiter de l'étude des synthétases NRPS de lipopeptides et autres peptides retrouvés dans les 20 génomes de *Pseudomonas* complètement séquencés en 2012. En fin de partie, un tableau général reprendra les différents NRPs identifiés dans les génomes étudiés.

VI.1. Sidérophore produit par une synthétase hybride NRPS/PKS : la yersiniabactine

Un cluster de synthétase composé de deux gènes a été identifié sur le brin sens aux loci PSPPH_2897 et PSPPH_2899 chez *Pseudomonas syringae phaseolicola 1448A* et sur le brin anti-sens aux loci PSPTO_2602 et PSPTO_2600 chez *Pseudomonas syringae DC3000*. Les gènes PSPTO_2602 et PSPPH_2897 sont annotés « yersiniabactin non ribosomal peptide synthetase », et les gènes PSPTO_2600 et PSPPH_2899 « irp1 (iron regulatory protein1) yersiniabactin polyketide/non ribosomal peptide synthetase ». Les séquences protéiques correspondant aux gènes cités ci-dessus ont été extraites et utilisées pour la prédiction de la structure de la synthétase et de la composition de la molécule produite (Figure 51). Cette synthétase est de produire une peptide hybride NRPS/PKS.

	PSPTO_2602 / PSPPH_2897	PSPTO_2600 / PSPPH_2899
	T Cy A ? T Cy T	Ks At ? Kr T Cy ? T Te
NRPS/PKS	Cys	
	Cys-like	
NRPS Predictor	Cys-like	
	Cys	
PKS/NRPS	Cys	

Figure 51: Prédiction du peptide produit par la synthétase de pyoverdine de *Pseudomonas syringae phaseolicola 1448A* et *syringae DC3000*.

Cette synthétase a été identifiée sur la base de l'annotation du génome et de sa composition en domaines comme étant responsable de la synthèse d'un sidérophore hybride NRPS/PKS : la yersiniabactine (Ybt). Ce sidérophore est produit par *Yersinia pestis* lors d'une carence en fer dans le milieu environnant. La synthétase de la Ybt est composée de quatre protéines. En comparant la structure prédite de la synthétase (Fig.51) et la structure publiée, les deux gènes de la figure peuvent être identifiés comme étant responsables de la synthèse des protéines HMWP1 et 2 de la synthétase de la yersiniabactine. Les deux autres protéines impliquées dans la synthèse d'un sidérophore de type yersiniabactine (YbtE et YbtU) ont été identifiées par BLAST N (en conservant les paramètres par défaut) en utilisant la séquence protéique codée par les gènes YbtE et YbtU de *Yersinia pestis* (tableau 20). Les domaines, non identifiés dans la figure 4.4. et notés « ? », présents dans la structure de la synthétase prédite ont été identifiés comme étant des domaines Mt (méthyltransférase).

Ybt	gènes	% de similarité avec Ybt (<i>Yersinia pestis</i>)	Annotation
YbtE	PSPPH_2902	57	Ybt synthetase, salicylate ligase component irp5
	PSPTO_2597	57	
YbtU	PSPPH_2900	50	Ybt synthetase, thiazolinyl reductase component
	PSPTO_2598	50	

Tableau 20: Tableau d'identification des gènes YbtE et YbtU chez *Pseudomonas syringae* DC3000 (gènes PSPTO) et *P.phaseolicola* 1448A (gènes PSPPH) par TblastN à partir de la séquence peptidique correspondant aux gènes YbtE et YbtU de *Yersinia pestis*.

VI.2. Synthétases de lipopeptides

VI.2.1. Synthétase de la syringafactine A

Un cluster composé de 2 gènes codant des NRPS a été identifié aux loci Psyr_2576 et Psyr_2577 chez *Pseudomonas syringae* pv. *Syringae* B728a. Ces deux gènes sont localisés sur le brin sens et annotés « amino acid adenylation ». Ce cluster est également retrouvé chez *Pseudomonas syringae* DC3000 aux loci PSPTO_2829 et PSPTO_2830 sur le brin sens, ils sont annotés respectivement « non ribosomal peptide synthase syfA » et « non ribosomal peptide synthase syfB ». Les séquences protéiques de ces deux gènes ont été extraites et analysées en utilisant les logiciels de prédiction (Figure 52).

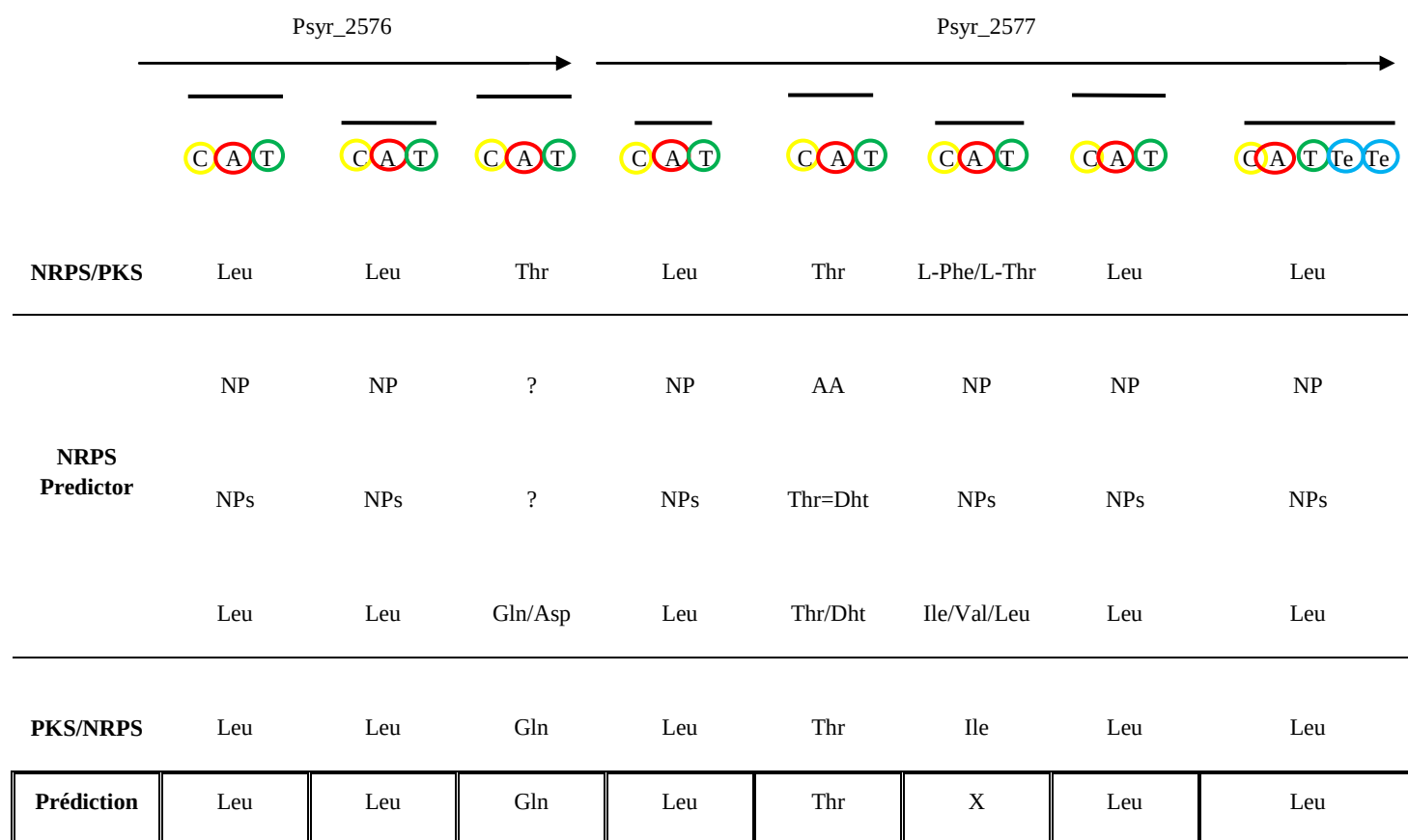


Figure 52: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par le cluster de la syringafactine A chez *Pseudomonas syringae* pv. *Syringae* B728a. AA= asp,asn,glu,gln,Aad NPs= val,leu,ile,Abu,Iva NP=gly,ala,val,leu,ile,Abu,Iva ?= pas de prédiction.

D'après la prédiction, la séquence potentielle du peptide produit par ce cluster serait la suivante « Leu_Leu_Gln_Leu_Thr_X_Leu_Leu ». La synthétase se termine par deux domaines de thioestérase (domaine Te) en tandem, ce qui est spécifique des synthétases de lipopeptides cycliques de *Pseudomonas* (Roongsawang *et al.*, 2011), de plus le premier module de la synthétase débute par un domaine de condensation (domaine C) qui pourrait intervenir dans la fixation d'un acide gras au peptide. Ce peptide est composé de 8 monomères et les observations précédentes permettent de prédire que ce peptide est un lipopeptide. La structure prédite présente 7 monomères sur 8 en commun avec celle d'un lipopeptide présent dans Norine, la syringafactine A (Leu-Leu-Gln-Leu-Thr-Val-Thr-Thr) (NOR01075) produite par *Pseudomonas syringae* DC3000 (Berti *et al.*, 2007). Le monomère en position 6 de la chaîne peptidique est difficilement prédictible car les outils de prédictions ne donnent pas de résultats compatibles. Ce peptide pourrait être une nouvelle forme de syringafactine, possédant une isoleucine à la place d'une valine en position 6. La prédiction est strictement identique pour les produits des synthétases de *Pseudomonas syringae* B728a et DC3000, hormis pour le module 6 de la synthétase.

VI.2.2. Synthétase de la syringomycine

Un cluster composé de deux gènes a été identifié aux loci Psyr_2608 et Psyr_2611 chez *Pseudomonas syringae* pv. *syringae* B728a. Ces deux gènes sont localisés sur le brin anti-sens et sont tous deux annotés « amino acid adenylation ». Les séquences protéiques des produits des gènes de ce cluster ont été extraites et analysées par les logiciels de prédiction (Figure 53).

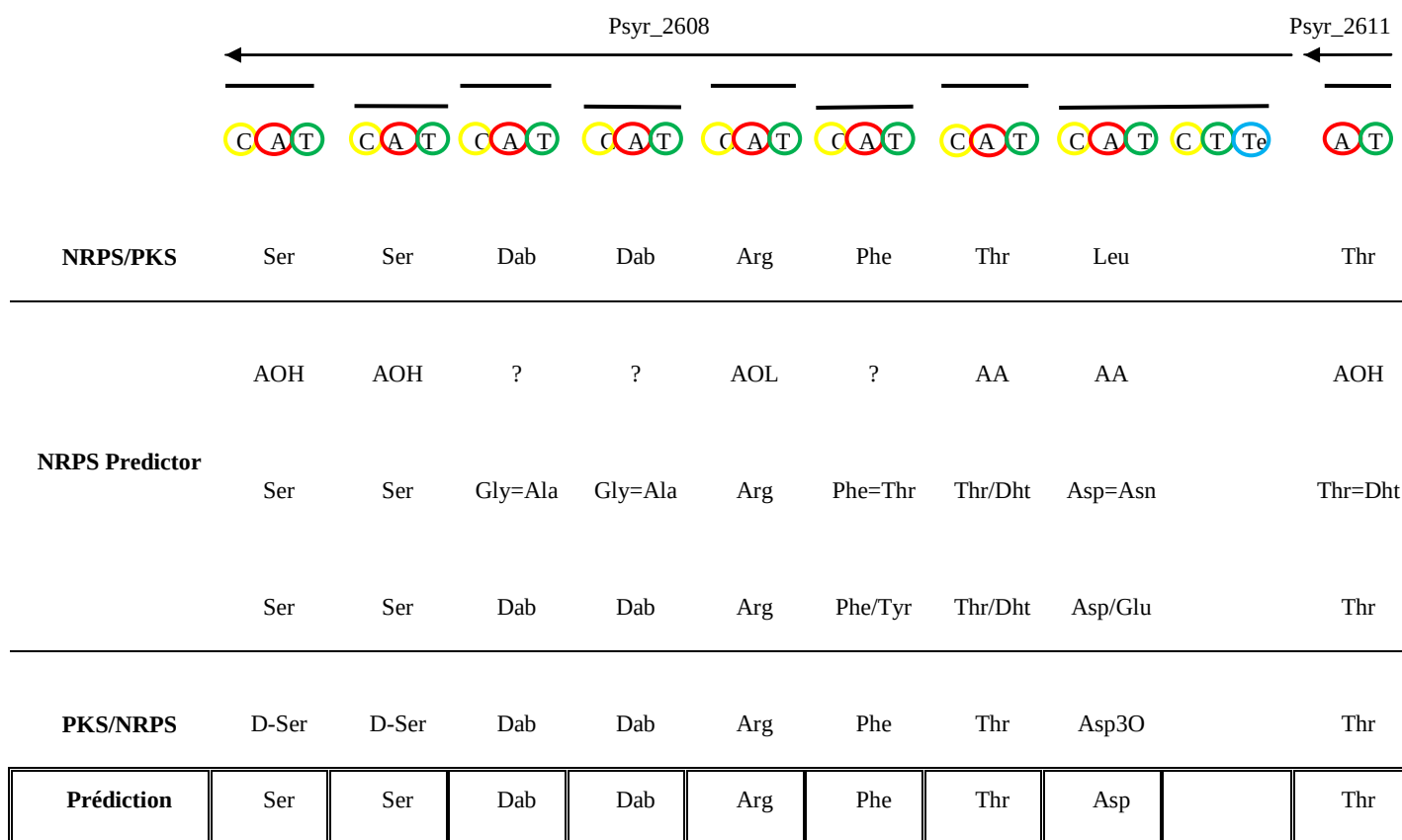


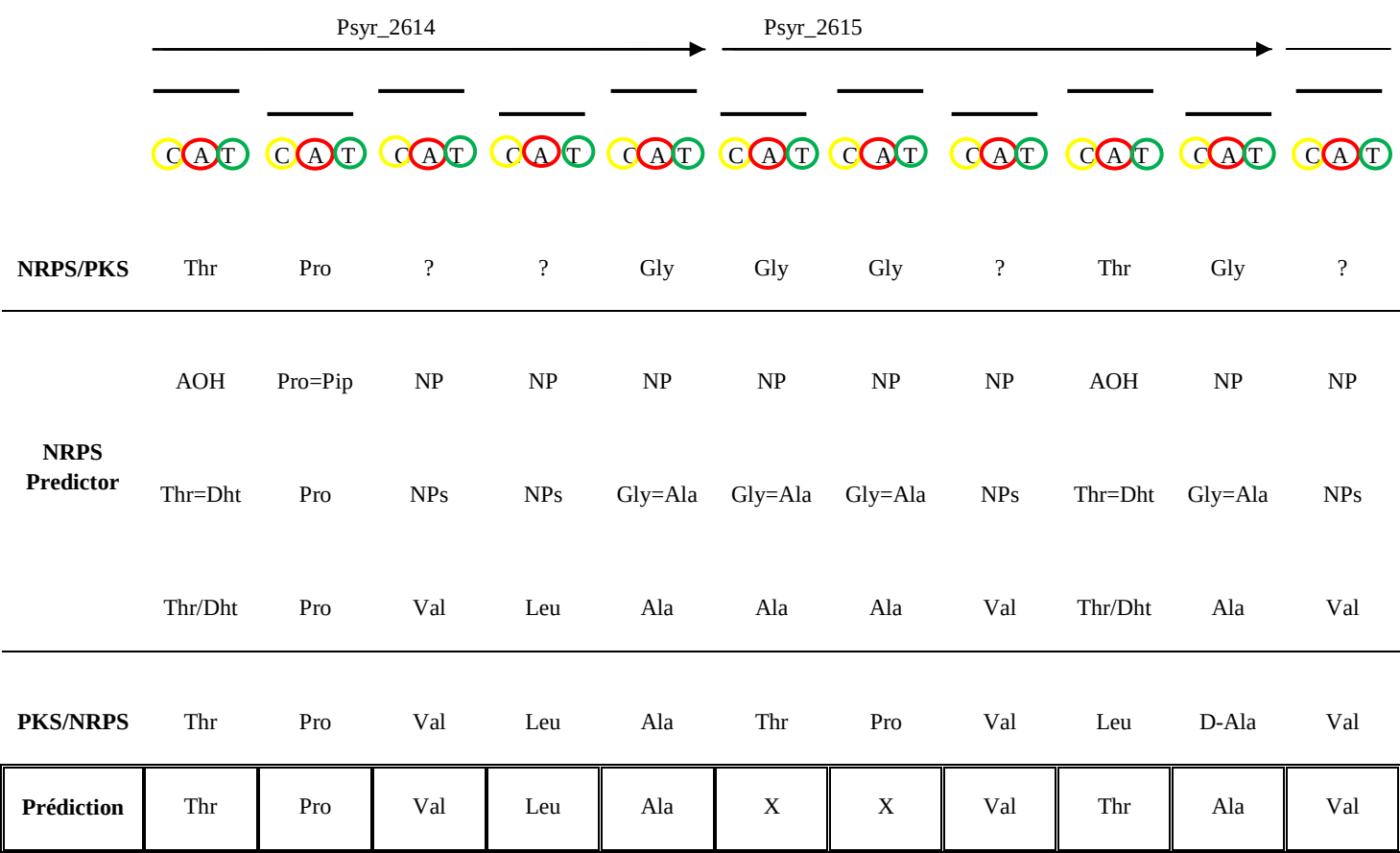
Figure 53: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par le cluster de la syringomycine synthétase de la souche *Pseudomonas syringae* B728a. AOH= Ser, Thr, Ser/Thr, Dhpg, Hpg AOL=Orn, Lys, Arg AA= Asp,Asn,Glu,Gln,Aad.

Le peptide produit par ce cluster est composé de 9 monomères avec la séquence consensus suivante « Ser_Ser_Dab_Dab_Arg_Phe_Thr_Asp_Thr ». Les gènes Psyr_2608 et Psyr_2611 codent des protéines qui contiennent respectivement 8 et 1 module(s). Le dernier module de Psyr_2608 contenant le domaine de la thioestérase ne comporte pas, selon les prédictions, de domaine d'adénylation, de plus le domaine Te ne se situe pas dans le dernier module de la synthétase qui est codé par le gène Psyr_2611. Le module 8 de cette synthétase (CATCTTe) est particulier, en effet il est composé de deux domaines C, deux domaines T, un domaine A et un domaine Te. La structure prédite de ce peptide présente 8 monomères sur 9 en commun avec celle d'un peptide présent dans Norine, la syringomycine E (NOR00356) qui est

un lipopeptide partiellement cyclique à activité antibiotique. La structure de la synthétase de la syringomycine est connue et similaire à la structure de la synthétase prédite. La structure du peptide NOR00356 est « Ser-D-Ser-Dab-D-Dab-Arg-Phe-dhAbu-OH-Asp-4Cl-Thr ». Deux monomères sont sous forme D bien qu'il n'y ait pas de domaine d'épimérisation dans la synthétase. La modification de ces monomères peut être due à l'intervention de racémases après la synthèse du peptide ou la présence de domaines C/E au sein des modules incorporant ces monomères.

VI.2.3. Synthétase de la syringopeptine

Un autre cluster NRPS composé de trois gènes a été identifié aux loci Psyr_2614, Psyr_2615 et Psyr_2617 dans le génome de *Pseudomonas syringae* B728a, sur le brin sens. Ils sont annotés « amino acid adenylation ». Les séquences protéiques des produits des gènes de ce cluster ont été extraites et analysées par les logiciels de prédiction (Figure 54).



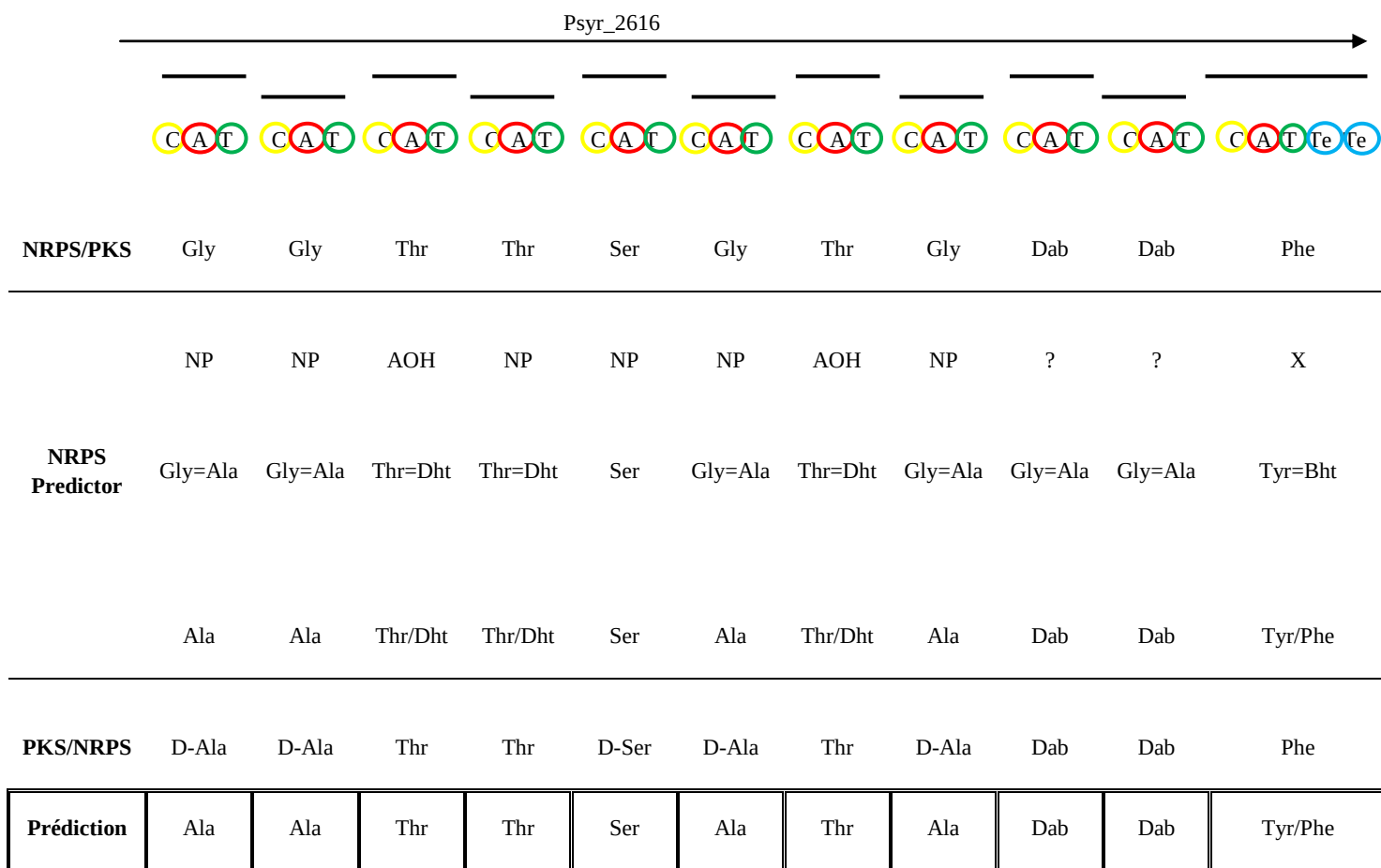


Figure 54 : Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par le cluster de la syringopeptine. AOH= Ser, Thr, Ser/Thr, Dhpg, Hpg NPs= Val, Leu, Ile, Abu, Iva NP=Gly,Ala,Val,Leu,Ile,Abu,Iva.

Le peptide produit par ce cluster est composé de 22 monomères avec la séquence prédite suivante

« Thr_Pro_Val_Leu_Ala_X_X_Val_Thr_Ala_Val_Ala_Ala_Thr_Thr_Ser_Ala_Thr_Ala_Dab_Dab_Tyr/Phe ».

Les produits des gènes Psysr_2614, Psysr_2615 et Psysr_2616 contiennent respectivement 5, 5 et 12 modules CAT.

La synthétase produite par ce cluster se termine par deux domaines Te en tandem et le premier module débute par un domaine C starter, comme vu précédemment pour l'analyse du cluster 2, ces caractéristiques sont spécifiques des lipopeptides cycliques de *Pseudomonas* (CLP).

Les lipopeptides cycliques produits par les *Pseudomonas* sont classés en six groupes (Roongsawang *et al.*, 2011). Ces groupes diffèrent par la longueur de la chaîne peptidique

ainsi que par la nature de l'acide gras. D'après la longueur du peptide produit par ce cluster, le lipopeptide prédit semble appartenir au groupe des tolaasines.

La structure prédite pour ce peptide présente 15 monomères sur 22 (en sachant que la nature des monomères en position 6 et 7 du peptide prédit est indéterminée et que l'on ne tient pas compte de l'isomérisation) en commun avec celle d'un peptide présent dans Norine, la syringopeptine 22 PhvA (NOR00354) qui est un lipopeptide partiellement cyclique à activité antibiotique. Ce peptide comporte 7 monomères sous forme L et 15 monomères sous forme D (Tableau 21) (Scholz-Schroeder *et al.*, 2003). De la même façon que pour la synthétase de la syringomycine, il n'y a pas de domaines d'épimérisation dans la structure de la synthétase, ni de racémases externes recherchées par BLAST dans le génome de cette souche en utilisant comme requête la séquence d'une racémase (ZP_03396269) de *Pseudomonas syringae* tomato T1.

	1	2	3	4	5	6	7	8	9	10	11
1	Thr	Pro	Val	Leu	Ala	X	X	Val	Thr	Ala	Val
2	DhAbu	D-Pro	D-Val	Val	D-Ala	D-Ala	D-Val	D-Val	DhAbu	D-Ala	D-Val

	12	13	14	15	16	17	18	19	20	21	22
1	Ala	Ala	Thr	Thr	Ser	Ala	Thr	Ala	Dab	Dab	Tyr
2	Ala	D-Ala	D-DhAbu	αThr	Ser	Ala	D-DhAbu	D-Ala	D-Dab	D-Dab	D-Tyr

Tableau 21 : Alignement de la structure prédite du peptide final (1) avec la structure de la syringopeptine décrite par Scholz-Schroeder (2). Le DhAbu est un dérivé de la thréonine.

Pour les deux lipopeptides vus précédemment, la syringomycine E et la syringopeptine, nous avons observé la présence de monomères sous forme D bien qu'il n'y ait pas de domaines d'épimérisation dans la structure prédite de la synthétase. La présence de formes L ou D a été décrite chez les lipopeptides comme influencé par la présence de domaine C à double fonction condensation/épimérisation. L'alignement des séquences des domaines C de la syringopeptine A et de la syringomycine E (figure 55) permet de différencier les domaines C conventionnels et

les domaines C à activité condensation/épimérisation (domaines C/E). Néanmoins, les domaines sypA_C20, 21 et 22 sont classés selon cette figure parmi les domaines C conventionnels alors que les acides aminés incorporés par les modules contenant ces domaines C sont sous forme D. La synthétase ne comporte pas de domaines E, ces trois domaines C seraient donc des domaines C/E pour permettre l'obtention d'acides aminés sous forme D dans le peptide. Les acides aminés aux positions 20, 21 et 22 seraient sous forme L contrairement à la structure décrite et donc la molécule serait différente. On observe également les deux domaines C starter qui se regroupent dans l'analyse.

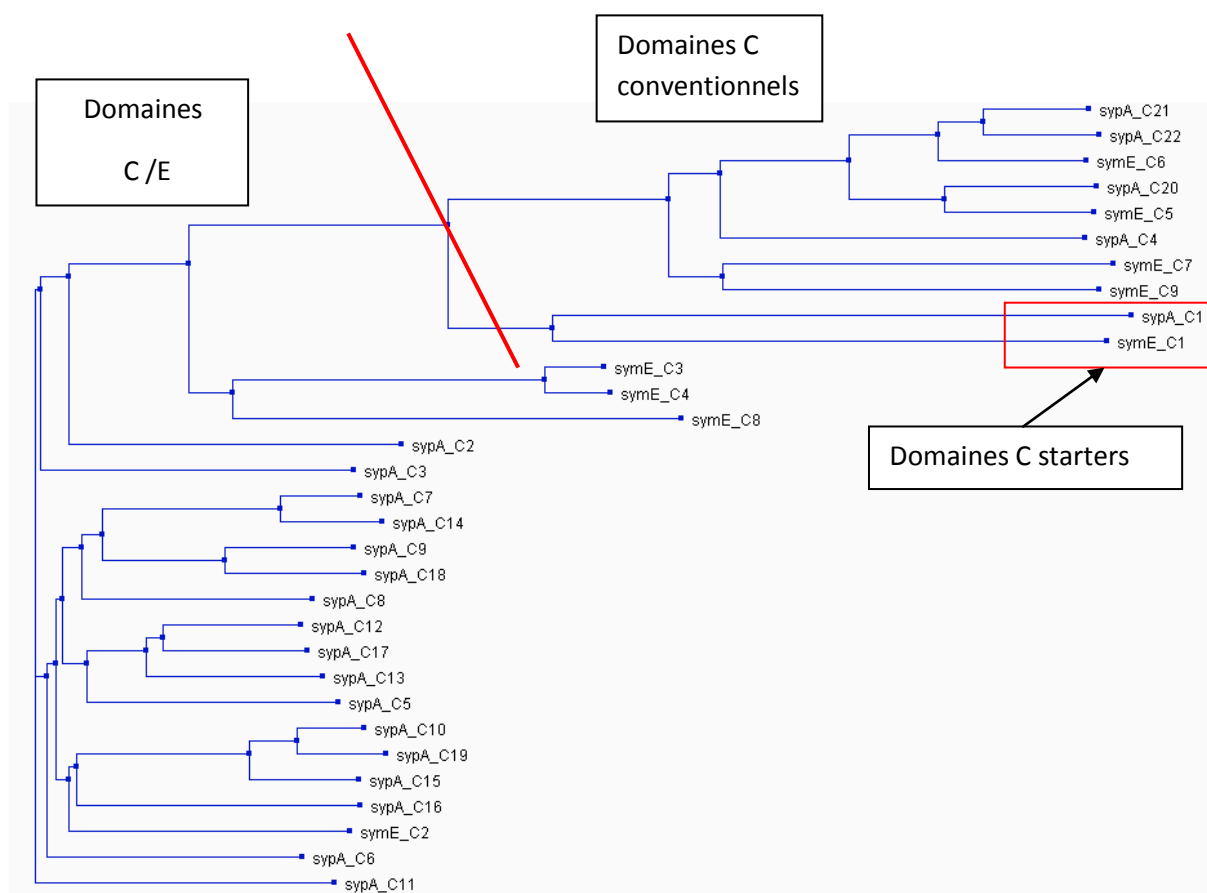


Figure 55: Analyse phylogénétique des domaines C de la syringopeptine A (sypA_Cx) et de la syringomycine E (symE_Cx), deux lipopeptides produits par *Pseudomonas syringae* B728a. « x » représente le numéro du domaine C dans la synthétase.

VI.2.4. Synthétase hybride NRPS/PKS

Une synthétase identifiée comme produisant un hybride NRPS/PKS a été retrouvée dans quatre des génomes analysés (*Pseudomonas entomophila* L48, *Pseudomonas fluorescens* Pf01 et *Pseudomonas syringae* 1448A et B728a). On observe dans cette synthétase deux modules d'initiation et deux thioestérases, il pourrait en fait s'agir de deux synthétases successives responsables de la synthèse de deux peptides de respectivement 4 et 2 monomères.

On peut également poser l'hypothèse que ce cluster est responsable de la synthèse d'un peptide ayant la structure prédite : Arg_Pro_Cys (Thz)_X_Ile_Pro. La synthétase décrite ci-dessous comporte deux modules d'initiation et deux modules de terminaison. La synthèse d'un tel peptide par deux synthétases successives serait réalisée en deux étapes. La première synthétase produit un quadripeptide qui est repris par la deuxième synthétase à laquelle elle ajoute son dipeptide. Aucun peptide présent dans Norine ne présente la même structure. Néanmoins, la cératospongamide (NOR00578), un peptide cyclique anti-inflammatoire isolé à partir d'une algue vivant en symbiose avec une éponge (Tan *et al.*, 2000), présente une structure cyclique proche : [Pro-Thz_Phe_Pro_MeOx-Ile_Phe]. Dans l'article de Vodovar *et al.* en 2006, ce cluster est décrit comme impliqué dans la biosynthèse d'un lipopeptide, néanmoins ceci n'est pas cohérent avec les analyses effectuées. En effet, le module 1 de cette synthétase ne contient pas le domaine de condensation nécessaire à l'incorporation du lipide au peptide. De plus cette synthétase comporte un domaine de cyclisation (Cy) qui forme un groupement thiazole (Thz) par hétérocyclisation d'une cystéine. La présence de ces domaines est généralement observée dans les synthétases de sidérophores. La cyclisation des lipopeptides est en général effectuée par un domaine de thioestérase.

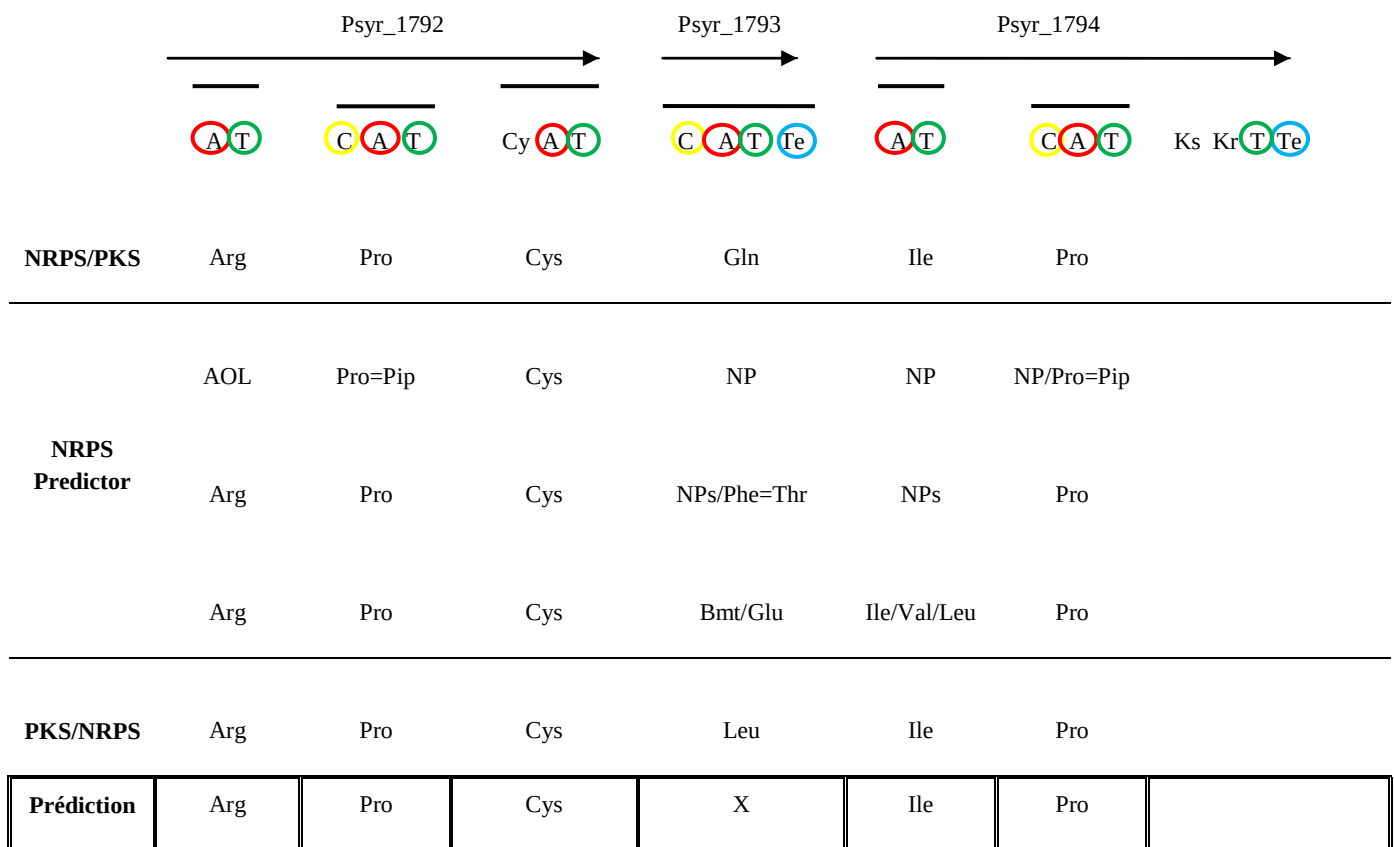


Figure 56: Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par ce cluster . AOL=Orn, Lys, Arg NPs= Val,Leu,Ile,Abu,Iva NP=Gly,Ala,Val,Leu,Ile,Abu,Iva.

VI.2.5. Synthétase de lipopeptide chez *Pseudomonas fluorescens Pf5*

Un cluster de gènes codant des NRPS et composé de 3 gènes a été identifié aux loci PFL_2145, PFL_2146 et PFL_2147 chez *Pseudomonas fluorescens Pf5*, respectivement annotés « ofaA non ribosomal peptide synthetase » « arthrofactin synthetase B » et « non ribosomal peptide synthetase ». Les séquences protéiques correspondant aux gènes de ce cluster ont été extraites et analysées pour déterminer la structure de la synthétase codée par ce cluster ainsi que le peptide qu'il synthétise (Figure 57).

	PFL_2145		PFL_2146				PFL_2147			
NRPS/PKS	Leu	Gln	Thr	?	Leu	Ser	Leu	Leu	Ser	?
NRPS Predictor	NP	?	AOH	NP	NP	AOH	NP	NP	AOH	NP
	NPs	?	Thr=Dht	NPs	NPs	Ser	NPs	NPs	Ser	NPs
	Leu	Asp/Glu	Thr/Dht	Ile/Leu	Leu	Ser	Leu	Leu	Ser	Val
PKS/NRPS	Leu	Gln	Thr	Leu	Leu	Ser	Leu	Leu	Ser	Val
Prédiction	Leu	Gln	Thr	Leu	Leu	Ser	Leu	Leu	Ser	Val

Figure 57:Prédiction de la structure modulaire de la synthétase et de la composition du peptide produit par ce cluster . AOL=Orn, Lys, Arg NPs= Val,Leu,Ile,Abu,Iva NP=Gly,Ala,Val,Leu,Ile,Abu,Iva.

Le gène PFL_2146 est annoté « arthrofactin synthetase B », néanmoins la composition du produit de la synthétase comporte 10 monomères alors que l’arthrofactine est composée de 11 monomères et d’un acide gras. Le cluster de gènes (PFL_2145/PFL_2146/PFL_2147) a été classé comme produisant un lipopeptide cyclique (Raijmakers et al.,2006). Les annotations automatiques de génomes ne sont pas fiables à 100%, il est donc nécessaire de refaire une analyse complète pour chaque NRPS potentielles. La structure du peptide prédite sur la figure 55 présente 8 monomères sur 10 en commun avec la structure de l’orfamide A. Loper et Gross ont décrit en 2007 la biosynthèse de ce lipopeptide antibiotique chez *Pseudomonas fluorescens* Pf5.

VI.3. Analyse des peptides inconnus

L’analyse bioinformatique des génomes de *Pseudomonas* complètement séquencés a permis de retrouver le potentiel de synthèse de divers peptides d’origine non-ribosomiale déjà connus, mais également de mettre en évidence de nouveaux peptides. Cette partie va traiter des synthétases NRPs non connues ainsi que de leurs produits. Cinq clusters de synthèse de NRPs non identifiés sont repris ici.

- **Cluster 1**

<i>Ps. fluorescens 01</i>	Pfl01_2211	Pfl01_2212	Pfl01_2213
Structure synthétase	CAT-CAT	CAT-CAT-CAT-CAT-CAT	CAT-CAT-CAT-CATTeTe
Peptide prédit	Leu_Gln	Gln_Leu/Ile_Leu_Gln_Ser	Leu_Leu_Ser_Leu/Ile

Cette synthétase potentiellement produite par *Ps. fluorescens 01* comporte 3 protéines codées par trois gènes qui se suivent, et se compose de 11 modules. La synthétase débute par un domaine de condensation et se termine par deux domaines de thioestérase en tandem, ces caractéristiques sont représentatives des synthétases de lipopeptides cycliques de *Pseudomonas*. Six des dix monomères prédits pour le produit de cette synthétase sont des leucines, dans le chapitre III il a été noté la forte proportion de monomères « leucine » dans les molécules de lipopeptides produites par les *Pseudomonas*. Dans Norine, trois peptides possèdent 7 monomères sur 11 en commun avec le peptide étudié, il s'agit de la putisolvine II (12 monomères)(NOR00362), la syringafactine B (8 monomères)(NOR01076) et la syringafactine E (8 monomères)(NOR01079). Ces trois peptides sont des lipopeptides antibiotiques et surfactants produits par des *Pseudomonas*. L'ensemble de ces informations tend à confirmer l'hypothèse de la production d'un lipopeptide par ce cluster.

- **Cluster 2**

<i>Ps. syringae DC3000</i>	PSPTO_4518	PSPTO_4519
Structure synthétase	CAT	CAT-CAT -CATTe
Peptide prédit	Ser	Gly_X_Ser

Cette synthétase potentiellement produite par *Ps. syringae DC3000* comporte 2 protéines codées par deux gènes qui se suivent, elle se compose de quatre modules. Une recherche de similarité réalisée avec les outils de comparaison de structure de la banque de données Norine a permis d'identifier 11 peptides qui possèdent 3 monomères sur 4 en commun avec le peptide étudié (isopyoverdine 90-44 (NOR00183), pseudobactine 36167-I et 36167-II (respectivement NOR00906 et NOR00907), pseudobactine A225-I et A225-II (respectivement NOR00908 et NOR00909), pyoverdine 12 (NOR00192), pyoverdine 1.2 dihydro (NOR00904), pyoverdine 1547 (NOR00186), pyoverdine 18-I (NOR00195),

pyoverdine 2798 (NOR00178) et pyoverdine BTP16 (NOR00206)). Ces 11 peptides sont tous des pyoverdines donc des chromopeptides à activité sidérophore. Cependant aucun autre gène impliqué dans la synthèse de ce sidérophore n'a été détecté dans les gènes environnants. De plus les 3 monomères prédits ne sont pas des agents chélateurs du fer. Les informations recueillies ici sont donc trop fragmentaires pour tirer une conclusion sur la nature et le rôle de ce peptide.

- **Cluster 3**

<i>Ps. syringae</i> DC3000	PSPTO_4699
Structure synthétase	CAT-CAT-CAT-CATTe
Peptide prédit	Leu_Ser_X_Ser

Cette synthétase se compose de quatre modules. Une recherche de similarité réalisée avec les outils de comparaison de structure de la banque de données NORINE a permis d'identifier 3 peptides qui possèdent 3 monomères sur 4 en commun avec le peptide étudié. Ces peptides sont les putisolvines I, II et III (respectivement NOR00361, NOR00362 et NOR00363), qui sont des lipopeptides à activité antibiotique et surfactant.

- **Cluster 4**

<i>Ps. syringae</i> B728a	Psyr_3722
Structure synthétase	CAT-CAT-CAT-CAT-CAT-CATTe
Peptide prédit	Leu_Ser_X_Val_Ala_Ser

Cette synthétase se compose de six modules. Une recherche de similarité réalisée avec les outils de comparaison de structures de la banque de données Norine a permis d'identifier 16 peptides qui possèdent 3 monomères sur 6 en commun avec le peptide étudié. Ces peptides peuvent être classés en deux groupes en fonction de leur nature. Le premier groupe est composé de lipopeptides : les putisolvines I, II et III (respectivement NOR00361, NOR00362 et NOR00363) et les syringopeptines 22phvA, 22phvB, 508A, 508B, SC1 et SC2 (respectivement NOR00354, NOR00970, NOR00873, NOR00874, NOR00353 et NOR00871) qui sont des lipopeptides à activité antibiotique et toxique. Le deuxième groupe

de peptides présentant 3 monomères sur 4 en commun avec le peptide étudié sont des peptaibols : hypomurocines B1, B2, B3b, B4 et B5 (respectivement NOR01008, NOR1009, NOR01011, NOR01012 et NOR01013) et les harzianines HC I et HC III (respectivement NOR00998 et NOR00989). Les hypomurocines et les harzianines sont des peptides à activité antibiotique. Cependant, les peptaibols sont synthétisés par des Fungi, il est donc plus probable que ce peptide soit un lipopeptide.

- **Cluster 5**

<i>Ps. brassicacearum</i> <i>NFM421</i>	PSBR_2394	PSBR_2392	PSBR_2391	PSBR_2390	PSBR_2389
Structure synthétase	?TC	AT	CAT-CAT-CAT-CATE	CAT-CAT-C	ATTe
Peptide prédit	X	His	Phe_Gln_Ser_D-Ser	Asp_X	Asp

Cette synthétase comporte 5 protéines et se compose de 8 modules. L'étude de similarité de structure entreprise avec les outils de Norine a permis de déterminer 33 peptides qui possèdent 3 monomères sur 9 en commun avec le peptide étudié. Cette similarité est trop faible pour conclure sur une activité potentielle de ce peptide.

Dans cette partie, le potentiel de synthèse non ribosomiale des vingt génomes de Pseudomonas totalement séquencés en 2011 a été étudié. Cette étude a montré 45 synthétases (tableau 22) produisant des sidérophores et des biosurfactants de nature lipopeptidique. Les synthétases retrouvées dans ces génomes peuvent être des synthétases NRPS mais aussi des synthétases hybrides NRPS/PKS.

souche	cluster de gènes produisant un peptide NRPS (nombre de modules)	peptide	Sidérophore	lipopeptides
Ps. aeruginosa PAO1	PA_2424/ PA_2402 (4)/ PA2400 (2)/ PA_2399 (2)	pyoverdine	•	
	PA_4231-PA_4224	pyocheline	•	
Ps. aeruginosa PA7	PSPA7_2829/ PSPA7_2858 (3)/ PSPA7_2859 (4)	pyoverdine	•	
	PSPA7_0869-PSPA7_0876	pyocheline	•	
Ps. aeruginosa LESB58	PLES_28711/ PLES_28951 (3)/ PLES_28961 (1)/ PLES_28971 (3)	pyoverdine	•	
	PLES_06961-PLES_07031	pyocheline	•	
Ps. aeruginosa PA14	PA14_33280/ PA14_33610 (4)/ PA14_33630 (2)/ PA14_33650 (2)	pyoverdine	•	
	PA14_09210-PA14_09320	pyocheline	•	
Ps. putida W619	PputW619_3562/ PputW619_3546 (2)/ PputW619_3545 (1)/ PputW619_3542 (2)/PputW619_3541 (2)	pyoverdine	•	
Ps. putida F1	Pput_1624/ Pput_1674 (3)/ Pput_1675 (1)/ Pput_1676 (2)/ Pput_1677 (1)/ Pput_1680 (1)	pyoverdine	•	
Ps. putida KT2440	PP_2243/ PP_4221 (2)/ PP_4220 (2)/ PP_4219 (3)	pyoverdine	•	
Ps. putida GB1	PputGB1_3809/ PputGB1_4086 (2)/ PputGB1_4085 (2)/ PputGB1_4084 (1)/ PputGB1_4083 (3)	pyoverdine	•	
Ps. fluorescens Pfl01	Pfl01_3940/ Pfl01_1845 (4)/ Pfl01_1846 (3) / Pfl01_1847 (1)/ Pfl01_1849 (1)	pyoverdine	•	
	Pfl01_2211 (2)/ Pfl01_2212 (5)/ Pfl01_2213 (4)	ND		
	Pfl01_2265 (3)/ Pfl01_2266 (1)/ Pfl01_2267 (2)	ND		
Ps. fluorescens Pf5	PFL_4189/ PFL_4095 (3)/ PFL_4094 (2)/ PFL_4093 (3)	pyoverdine	•	
	PFL_2996/ PFL_2997	yersiniabactine	•	
	PFL_2145 (2)/ PFL_2146 (4)/ PFL_2147 (4)	Orfamide A		•
Ps. fluorescens SBW25	PFLU_6137/ PFLU_2543 (3)/ PFLU_2544 (3)	pyoverdine	•	
	PFLU_3220/ PFLU_3222 (1)/PFLU_3223 (4)/ PFLU_3224 (2)/ PFLU_3225 (1)	ornicorrugatine	•	
	PFLU_4007 (2)/ PFLU_4007 (2)/ PFLU_2552 (4)/ PFLU_2553 (3)	viscosine		•
Ps. syringae DC3000	PSPTO_2135/ PSPTO_2147 (1)/ PSPTO_2148 (2)/ PSPTO_2149 (2)/ PSPTO_2150 (2)	pyoverdine	•	
	PSPTO_2600/ PSPTO_2602	yersiniabactine	•	
	PSPTO_2829 (3)/ PSPTO_2830 (5)	syringofactine		•
	PSPTO_4518 (1)/ PSPTO_4519 (3)	ND		
	PSPTO_4699 (4)	ND		
Ps. syringae 1448A	PSPPH_1911/ PSPPH_1923 (1)/ PSPPH_1924 (2)/ PSPPH_1925 (2)/ PSPPH_1926 (2)	pyoverdine	•	
	PSPPH_2897/ PSPPH_2899	yersiniabactine	•	
	PSPPH_1749 (3)/ PSPPH_1750 (1)/ PSPPH_1751 (2)	ND		
Ps. Syringae B728a	Psyr_1957 (1)/Psyr_1958 (2)/Psyr_1959 (2)/Psyr_1960 (2)	pyoverdine	•	
	Psyr_2582 -Psyr_2589	achromobactine	•	
	Psyr_2608 (9)/Psyr_2611 (1)	Syringomycine E		•
	Psyr_2614 (5)/Psyr_2615 (5)/Psyr_2616 (12)	Syringopeptine 22 PhvA		•

	Psyr_2576 (3)/Psyr_2577 (5)	syringofactine		•
	Psyr_1704 (1)/Psyr_1705 (2)	Syringoline A		•
	Psyr_3722 (6)	ND		
	Psyr_1792 (3)/Psyr_1793 (1)/Psyr_1794 (2)	ND		
Ps. mendocina ymp	Pmen_1902 (3)/ Pmen_1901 (4)	(Awaya et Dubois, 2008)	•	
Ps. mendocina NK-01	MDS_1791(3)/MDS_1792(4)		•	
Ps. brassicacearum NFM 421	PSEBR_a3922(1)/PSEBR_a3921(2)/PSEBR_a3920/PSEBR_a3919(3)/PSEBR_a3918(1)/PSEBR_a3916(1)	pyoverdine	•	
	PSEBR_a2394/PSEBR_a2392(1)/PSEBR_a2391(4)/PSEBR_a2390(2)/PSEBR_a2389(1)	ND		
Ps. entomophila L48	PSEEN_1815/ PSEEN_3234 (2)/ PSEEN_3232 (1)/ PSEEN_3231 (2) / PSEEN_3230 (2)/ PSEEN_3229 (3)	pyoverdine	•	
	PSEEN_3045 (8)/ PSEEN_3044 (4)/PSEEN_3332 (2)	entolysine		•
	PSEEN_2716 (3)/ PSEEN_2717 (1)/ PSEEN_2718 (2)	ND		
	PSEEN_2154 (3)/ PSEEN_2153 (1)/ PSEEN_2150 (1)/ PSEEN_2149 (1)/ PSEEN_2139	ND		
Ps. fulva 12-X				
Ps. stutzeri A1501				

Tableau 22 : Analyse globale du potentiel de synthèse NRPS des génomes du genre *Pseudomonas*. ND : non déterminé. Pour les deux souches de *Pseudomonas mendocina*, Awada *et al.* (2008) ont identifié un cluster de gènes impliqué dans la synthèse d'un sidérophore.

VII. Discussion

VII.1. Bilan des synthétases de *Pseudomonas* fluorescents et de leurs produits

Dans un premier temps, nous nous sommes attachés à réaliser quelques études statistiques sur les NRPs produits par des *Pseudomonas* fluorescents. A partir des 80 pyoverdines connues, nous avons établi pour la première fois une représentation schématique de l'ensemble de ces molécules.

Les travaux présentés ensuite ont permis à partir de l'analyse des génomes de *Pseudomonas* complètement séquencés en 2012, de prédire la structure des synthétases NRPS ainsi que des synthétases hybrides NRPS/PKS codées par ces génomes. Les NRPs produits par ces synthétases sont principalement des lipopeptides et des sidérophores, néanmoins une troisième catégorie de NRPs non déterminés est présente (tableau 22).

L'étude des génomes a permis de prédire la structure des pyoverdines produites par les différentes souches de *Pseudomonas*, pour les quatre souches de *Pseudomonas aeruginosa* dont le génome est complètement séquencé (*Ps. aeruginosa* PAO1, PA7, PA14 et LESB58), la répartition des structures prédites de pyoverdines correspond aux trois sidérovares de *Pseudomonas aeruginosa* déjà décrits (Budzikiewicz, 2004). Une observation équivalente a été faite pour les structures prédites des pyoverdines de *Pseudomonas syringae* (B728a, DC3000 et 1448A), les trois pyoverdines ont exactement la même composition. Pour ces deux espèces de *Pseudomonas*, la diversité chez les pyoverdines est limitée en comparaison des pyoverdines produites par les *Pseudomonas fluorescens* ou *putida*.

Ces travaux ont également permis de corriger la structure publiée de la pyoverdine de *Pseudomonas entomophila* L48 et de décrire les synthétases de plusieurs molécules connues.

Un autre sidérophore, produit par une synthétase hybride NRPS/ PKS a été identifié sur base de l'organisation de la synthétase. Il s'agit de la yersiniabactine potentiellement produite par *Ps. fluorescens* Pf5, *Ps. syringae* DC3000 et 1448A. Un peptide produit par une synthétase hybride NRPS/PKS a été retrouvé dans quatre des génomes analysés (*Ps. entomophila* L48, *Ps. fluorescens* Pf01 et *Ps. syringae* 1448A et B728a). Ce peptide, issu d'une synthèse hybride NRPS/PKS est conservé entre plusieurs souches. Sa production avait déjà été décrite chez *Pseudomonas syringae* (Bultreys et al., 2006).

Différents lipopeptides ont aussi été identifiés tels que l'orfamide et la viscosine (NOR00331), deux lipopeptides partiellement cycliques à activité antibiotique. Les deux synthétases correspondantes ont leur premier module qui débute par un domaine de condensation. D'autres synthétases de lipopeptides produisent potentiellement des molécules proches de la syringopeptine ou de la syringomycine. Il pourrait s'agir de variants ou de molécules qui diffèrent par leur profil en acides aminés de type L et D.

VII.2. Caractérisation de mécanismes de biosynthèse particuliers

L'analyse de ces nombreuses synthétases a permis de mettre en évidence différents mécanismes de biosynthèse potentiellement originaux.

Les synthétases NRPS se terminent par un domaine de thioestérase (Te) qui marque la fin de la synthèse du NRP. Un cas intéressant est à observer dans la synthétase de la pyoverdine produite par *Pseudomonas putida* W619, deux modules qui interviennent dans la synthèse de la chaîne peptidique de la pyoverdine se trouvent dans la synthétase après le module se terminant par un domaine de thioestérase. Cette observation peut être interprétée comme la mise en évidence d'un mode original de synthèse pour cette pyoverdine ou alors comme une erreur lors du séquençage du génome de cette souche et de l'assemblage des fragments d'ADN. Néanmoins, cette observation a déjà été décrite dans la synthétase de la syringomycine E, un lipopeptide partiellement cyclique produit par *Pseudomonas syringae*. Un gène codant pour un module « AT » intervenant dans la synthétase se trouve après le module contenant le domaine de thioestérase (Guenzi *et al.*, 1998). Ces deux exemples sont les seuls exemples de non-colinéarité dans la synthétase des NRPs chez *Pseudomonas*.

Un autre exemple intéressant est celui des gènes voisins codants pour deux synthétases et retrouvés dans quatre des génomes analysés (*Pseudomonas entomophila* L48, *Pseudomonas fluorescens* PF01 et *Pseudomonas syringae* 1448A et B728a). Ces deux synthétases successives pourraient être responsables de la synthèse de deux peptides de respectivement 2 et 4 monomères ou d'un hexapeptide similaire à la cératospingamide. Dans ce dernier cas, le peptide relargué par la première synthétase devrait être repris par la seconde ce qui correspond à un mode de synthèse entièrement original.

La présence de monomères sous forme D est due à la présence de domaines d'épimérisation (E) ou de domaines possédant une double activité C/E. En effet, pour les

différentes souches de *Pseudomonas* possédant le potentiel génique de synthèse de lipopeptides et de pyoverdines (*Ps. syringae* B728a, *Ps. entomophila* L48, *Ps. fluorescens* Pf5 et SBW25), on retrouve dans les structures de synthétases prédites des domaines E pour les synthétases de pyoverdines et des domaines C/E pour les synthétases de lipopeptides. Cette observation montre que la présence de domaines C/E n'est pas espèce dépendante étant donné que l'on ne trouve pas de domaines C/E dans les synthétases de pyoverdines d'une espèce possédant ce type de domaine dans ses synthétases de lipopeptides. Cette différence dans le mode d'action entraînant la présence de monomères sous forme D dans le produit final des synthétases NRPS de lipopeptides et de sidérophores, notamment de pyoverdines, chez *Pseudomonas* pourrait suggérer que ces deux NRPs ont des origines distinctes.

Afin de tester la spécificité des domaines A de la pyoverdine synthétase dans l'activation et l'incorporation d'un acide aminé particulier, des expériences d'enrichissement du milieu de culture en un acide aminé (feeding) ont été réalisées. Ces expériences ont montré que dans certaines des conditions testées, des modifications avaient eu lieu dans la composition de la chaîne peptidique de la pyoverdine. Ces observations montrent une tolérance dans la reconnaissance de l'acide aminé par le domaine A. Des observations similaires ont déjà été faites pour la surfactine (Baumgart *et al.*, 1991 ; Bonmatin *et al.*, 1995), la microcystine (Kurmayer *et al.*, 2001) et la cyclosporine (Van Döhren *et al.*, 1997). Mais ce sont les premiers exemples de substitutions de monomères observés lors de la synthèse des pyoverdines. Ce travail a été effectué en se concentrant sur les modifications affectant la composition de la chaîne peptidique des pyoverdines lors de ces expériences de feeding, le même travail sur les lipopeptides de *Pseudomonas* pourrait permettre l'identification d'autres domaines A tolérants.

VII.3. Développement des outils bioinformatiques

Ce travail a permis le développement d'un logiciel de calcul de masse moléculaire de pyoverdine (appelé « Pyomass ») qui permet de traiter rapidement les résultats de spectrométrie de masse obtenus lors des analyses des pyoverdines produites par les cellules sur milieu solide et lors des expériences de « feeding ».

Nous avons également identifié un certain nombre de signatures originales des domaines d'adénylation impliqués dans la synthèse des pyoverdines. L'utilisation de ces signatures permet d'améliorer la prédiction des structures des pyoverdines produites.

Enfin nous avons montré l'intérêt des études phylogéniques pour caractériser la spécificité des domaines d'adénylation. Ces outils pourraient être utilisés à l'avenir pour compléter les outils de prédiction existants.

VII.4. Perspectives

Ces travaux ouvrent de nombreuses perspectives. Les premières sont relatives à la caractérisation des structures des différents peptides potentiellement produits par les différentes souches étudiées et l'analyse de leurs propriétés par des méthodes biochimiques.

Les différents mécanismes de biosynthèse originaux méritent d'être décryptés par des analyses biochimiques poussées.

Les outils bioinformatiques mis au point dans ce travail doivent être intégrés dans la plateforme d'outils développés par ProBioGEM et le LIFL autour des peptides non-ribosomiaux. Le but étant à terme de disposer d'un entrepot d'outils d'analyse des peptides non ribosomiaux complet permettant une analyse globale à partir d'une séquence codant potentiellement pour une NRPS.

Ce travail révèle également que des outils de prédiction spécifiques à l'espèce voire au type de molécules méritent d'être mis au point dans l'optique d'améliorer les outils existants.

Références bibliographiques

- Amann, C., Taraz, K., Budzikiewicz, H., and Meyer, J.M.(2000). The siderophores of *Pseudomonas fluorescens* 18.1 and the importance of cyclopeptidic substructures for the recognition at the cell surface. *Z. Naturforsch.*, **55c**: 671-680.
- Ansari, Z.A., Yadav, G., Gokhale, R.S., and Mohanty, D. (2004). NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthetases. *Nucl. Acids Res.*, **32**: W405-W413.
- Anthoni, U., Christophersen, C., Nielsen, P. H., Gram, L., and Petersen, B. O. (1995). Pseudomonine, an isoxazolidone with siderophoric activity from *Pseudomonas fluorescens* AH2 isolated from lake Victorian nile perch. *J. Nat. Prod.*, **58**: 1786-1789.
- Awada, J.D., and Dubois, J.L. (2008). Identification, isolation, and analysis of a gene cluster involved in iron acquisition by *Pseudomonas mendocina ymp*. *Biometals*, **21**: 353-366.
- Bachmann, B.O., and Ravel, J. (2009). In silico prediction of microbial secondary metabolic pathways from DNA sequence data. *Meth. Enzymol.*, **458**: 181-217.
- Balibar, C. J., Vaillancourt, F. H., and Walsh, C. T. (2005). Generation of D amino acids residues in assembly of arthrofactin by dual condensation/epimerization domains. *Chem. Biol.*, **12**: 1189-1200.
- Ballio, A., Bossa, F., Collina, A., Gallo, M., Iacobellis, N.S., Paci, M., Pucci, P., Scaloni, A., Segre, A., and Simmaco, M.(1990). Structure of syringotoxin, a bioactive metabolite of *Pseudomonas syringae* pv. *syringae*. *FEBS Lett.*, **269**: 377-380.
- Ballio, A., Bossa, F., Di Giorgio, D., Ferranti, P., Paci, M., Pucci, P., Scaloni, A., Segre, A.L., and Strobel, G.A.(2004). Novel bioactive lipodepsipeptides from *Pseudomonas syringae*: the pseudomycins. *FEBS Lett.*, **355**: 96-100.
- Baré, S., Coiro, V.M., Scaloni, A., Di Nola, A., Paci, M., Segre, A.L., and Ballio, A. (1999). Conformations in solution of the fuscopeptins ;Phytotoxic metabolites of *Pseudomonas fuscovaginae*. *Eur. J. Biochem.*, **266**: 484-492.
- Barelmann, I., Fernandez, D.U., Budzikiewicz, H., and Meyer, J.M.(2003). The pyoverdine from *Pseudomonas chlororaphis* D-TR133 showing mutual acceptance with the pyoverdine of *Pseudomonas fluorescens* CHA0. *Biometals*, **16**: 263-270.
- Barelmann, I., Taraz, K., Budzikiewicz, H., Geoffroy, V.A., and Meyer, J.M. (2002). The structures of pyoverdins from two *Pseudomonas fluorescens* strains accepted mutually by their respective producers. *Z. Naturforsch.*, **57c**: 9-16.
- Bassarello, C., Lazzaroni, S., Bifulco, G., Lo Cantore, P., Iacobellis, N.S., Riccio, R., Gomez Paloma, L., and Evidente, A. (2004). Tolaasins A—E, five new lipodepsipeptides produced by *Pseudomonas tolaasii*. *J. Nat. Prod.*, **67** : 811-816.
- Baumgart, F., Kluge, B., Ullrich, C., Vater, J., and Ziessow, D. (1991). Identification of amino acid substitutions in the lipopeptide surfactin using 2D NMR spectroscopy. *Biochem. Biophys. Res. Com.*, **177**: 998-1005.
- Beiderbeck, H., Risse, D., Budzikiewicz, H., and Taraz, K. (1999a). A new pyoverdin from *Pseudomonas aureofaciens*. *Z. Naturforsch.*, **54c**: 1-5.

- Beiderbeck, H., Taraz, K., and Meyer, J.M. (1999b). Revised structures of the pyoverdins from *Pseudomonas putida* CFBP 2461 and from *Pseudomonas fluorescens* CFBP 2302. *Biometals*, **12**: 331-338.
- Berti, A.D., Greve, N.J., Christensen, Q.H., and Thomas M.G. (2007). Identification of a biosynthetic gene cluster and the six associated lipopeptides involved in swarming motility of *Pseudomonas syringae* pv. *tomato* DC3000. *J. Bacteriol.*, **189**: 6312-6323.
- Billstein, S., and Eagle, G. (2011). Overview of oritavancin. *European infectious disease*, **5**: 98-101.
- Bonmatin, J.M., Labé, H., Grangemard, I., Peypoux, F., Ptak, R., and Michel, G. (1995). Production, isolation and characterization of [Leu4]- and [Ile4] surfactins from *Bacillus subtilis*. *Lett. Pept. Sci.*, **2**: 41-47.
- Brahim, S., Bettaieb, A., and Kenani, A. (2008). Deglycosylated bleaomycin triggers apoptosis in laryngeal carcinoma cells in a caspase and reactive oxygen species independent manner. *J. Oral. Pathol. Med.*, **37**: 352-357.
- Briskott, G., Taraz, K., and Budzikiewicz, H. (1986). Siderophore vom pyoverdin-typ aus *Pseudomonas aeruginosa*. *Z. Naturforsch.*, **41c**: 497.
- Briskott, G., Taraz, K., and Budzikiewicz, H.(1989). Pyoverdine-type siderophores from *Pseudomonas aeruginosa*. *Liebigs Ann.Chem.*, 375-384.
- Budzikiewicz, H. (2004). Siderophores of the Pseudomonadaceae *sensu stricto* (fluorescent and non-fluorescent *Pseudomonas spp.*). *Prog. Chem. Org. Nat. Prod.*, **87**: 81-237.
- Budzikiewicz, H. (1993). Secondary metabolites from fluorescent Pseudomonads. *FEMS Microbiol. Rev.*, **104**: 209.
- Budzikiewicz, H., Fernandez, D. U., Fuchs, R., Michalke, R., Taraz, K., and Ruangviriyachai, C. (1999). Pyoverdines with a Lys ϵ -amino link in the peptide chain. *Z. Naturforsch.*, **54c**: 1021-1026.
- Budzikiewicz, H., Kilz, S., Taraz, K., and Meyer, J. M. (1997). Identical pyoverdines from *Pseudomonas fluorescens* 9AW and from *Pseudomonas putida* 9BW. *Z. Naturforsch.*, **52c**: 721-728.
- Budzikiewicz, H., Schäfer, M., Fernandez, D. U., and Meyer, J. M. (2006). Structure proposal for a new pyoverdin from *Pseudomonas sp.* PS 6.10. *Z. Naturforsch.*, **61c**: 815-820.
- Budzikiewicz, H., Schäfer, M., Uria Fernandez, D., Matthijs, S., and Cornelis, P. (2006). Characterization of the chromophores of pyoverdins and related siderophores by electrospray tandem mass spectrometry. *Biometals*, **20**: 135-144.
- Budzikiewicz, H., Schröder, H., and Taraz, K. (1992). Zur biogenese der Pseudomonas-siderophore: der Nachweis analoger strukturen eines pyoverdin-desferribactin-paares. *Z. Naturforsch.*, **47c**: 26-32.
- Bultreys, A., Gheysen, I., Wathelet, B., Schäfer, M., and Budzikiewicz, H. (2004). The pyoverdin of *Pseudomonas syringae* and *Pseudomonas cichorii*. *Z. Naturforsch.*, **59c**: 613-618.

- Bultreys, A., Gheysen, I., and de Hoffmann, E. (2006). Yersiniabactin production by *Pseudomonas syringae* and *Escherichia coli*, and description of a second yersiniabactin locus evolutionary group. *Appl. Env. Microbiol.*, **72**: 3814-3825.
- Byford, M. F., Baldwin, J. E., Shiau, C. Y., and Schofield, C. J.(1997). The mechanism of ACV synthetase. *Chem. Rev.*, **7**: 2631-2650.
- Caboche, S., Leclère, V., Pupin, M., Kuchero, G., and Jacques, P. (2010). Diversity of monomers in nonribosomal peptides : towards the prediction of origin and biological activity. *J. Bacteriol.*, **192**(19): 5143.
- Caboche, S., Pupin, M., Leclere, V., Fontaine, A., Jacques, P., and Kuchero, G.(2008). NORINE : a database of nonribosomal peptides. *Nucl. Acids Res.*, **36**: D326-331.
- Caboche, S., Pupin, M., Leclère, V., Fontaine, A., Jacques, P., and Kuchero, G. (2008). NORINE : a database of nonribosomal peptides. *Nucl. Acids Res.*, **36**: D326-331.
- Caboche, S., Pupin, M., Leclere, V., Jacques, P., and Kuchero, G. (2009). Structural pattern matching of nonribosomal peptides. *BMC Struct. Biol.*, **9**: 15.
- Clarke-Pearson, M.F., and Brady, S.F. (2008). Paerumumarin, a new metabolite produced by the pvc gene cluster from *Pseudomonas aeruginosa*. *J. Bacteriol.*, **190**: 6927-6930.
- Clugston, S. L., Sieber, S. A., Marahiel, M. A., and Walsh, C. T. (2003). Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidin synthetase is a DCL catalyst. *Biochem*, **42**: 12095-12104.
- Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.*, **16**: 4174-4183.
- Coraiola, M., Lo Cantore, P., Lazzaroni, S., Evidente, A., Iacobellis, N.S., and Della Serra, M. (2006). WLIP and tolaasin I, lipodepsipeptides from *Pseudomonas reactans* and *Pseudomonas tolaasii*, permeabilise model membranes. *Biochim. Biophys. Acta*, **11**: 1713-1722.
- Cornelis, P., and Matthijs, S. (2002). Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. *Env. Microbiol.*, **4**: 787-798.
- Cornelis, P., Honhadel, D., and Meyer, J.M.(1989). Evidence for different pyoverdine-mediated iron uptake systems among *Pseudomonas aeruginosa* strains. *Infection and immunity*, 3491-3497.
- Cox, C. D., Rinehart, K. L., Moore, M. L., and Cook, J. C. (1981). Pyochelin: novel structure of an iron-chelating growth promoter for *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA*, **78**: 4256-4260.
- Crooks, G.E., Hong, G., Chandonia, J.M., and Brenner, S.E. (2004). Weblogo: a sequence logo generator. *Genome Res.*, **14**: 1188-1190.
- De Bruijn, I., De Kock, M.J., De Waard, P., Van Beek, T.A., and Raaijmakers, J.M. (2008). Massetolides A biosynthesis in *Pseudomonas fluorescens*. *J. Bacteriol.*, **190**: 2777-2789.

- Demange, P., Bateman, A., Mertz, C., Dell, A., Piémont, Y., and Abdallah, M. (1990). Structure of pyoverdins Pt, siderophores of *Pseudomonas tolaasii* NCPPB 2192, and pyoverdins Pf, siderophores of *Pseudomonas fluorescens* CCM 2798. Identification of an unusual natural amino acid. *Biochem.*, **29**: 11041-11051.
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayart, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., and Gascuel, O.(2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucl. Acids Res.*, **36**: W465-W469.
- Dorrestein, PC., Poole, K., and Begley, T.P. (2003). Formation of the chromophore of the pyoverdine siderophore by an oxidative cascade. *Org. Lett.*, **5**: 2215-2217.
- Emanuele, M.C., Scaloni, A., Lavermicocca, P., Jacobellis, N.S., Camoni, L., Di Giorhio, D., Pucci, P., Paci, M., Segre, A.L., and Ballio, A. (1998). Corpeptins, new bioactive lipoundecapeptides from cultures of *Pseudomonas corrugata*. *FEBS Lett.*, **433**: 317-320.
- Feil, H., Feil, W.S., Chain, P., Larimer, F., Dibartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M., Goltsman, E., Thiel, J., Malfatti, S., Loper, J.E., Lapidus, A., Detter, J.C., Land, M., Richardson, P.M., Kyrpides, N.C., Ivanova, N., and Lindow, S.E. (2005). Comparaison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. *Proc. Natl. Acad. Sci. USA*, **102**: 11064-11069.
- Geisen, K., Taraz, K., and Budzikiewicz, H.(1992). Neue Siderophore des Pyoverdin-typs aus *Pseudomonas fluorescens*. *Monatsh. Chem.*, **123** : 151-178.
- Georgias, H., Taraz, K., Budzikiewicz, H., Geoffroy, V., and Meyer, J.M.(1999). The structure of the pyoverdin from *Pseudomonas fluorescens* 1.3. Structural and biological relationships of pyoverdins from different strains. *Z. Naturforsch.*, **54c**: 301-308.
- Gérard, J., Loyd, R., Barsby, T., Haden, P., Kelly, M.T., and Andersen, R.J. (1997). Massetolides A-H, antimycobacterial cyclic depsipeptides produced by two pseudomonads isolated from marine habitats. *J. Nat. Prod.*, **60** : 223-229.
- Gipp, S., Hahn, J., Taraz, K., and Budzikiewicz, H. (1991). Zwei pyoverdine aus *Pseudomonas aeruginosa* R. *Z. Naturforsch.*, **46c**: 534-541.
- Green , C. J.(1981). Immunosuppression with cyclosporine A: a review. *Diagn. Histopathol.*, **4**: 157-174.
- Grgurina, I., and Mariotti, F. (2002). Biosynthetic origin of syringomycin and syringopeptin 22, toxic secondary metabolites of the phytopathogenic bacterium *Pseudomonas syringae* pv. *syringae*. *FEBS Lett.*, **462**: 151-154.
- Gross, H., and Loper, J.E. (2009). Genomics of secondary metabolite production by *Pseudomonas* spp. *Nat. Prod. Rep.*, **26**: 1408-1446.
- Grünewald, J., and Marahiel, M. A. (2006). Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides. *Microbiol. Mol. Biol. Rev.*, **70**: 121-146.
- Guenzi, E., Galli, G., Grgurina, I., Gross, D.C., and Grandi, G. (1998). Characterization of the syringomycin synthetase gene cluster. *J. Biol. Chem.*, **273**: 32857-32863.
- Gwose, I., and Taraz, K. (1992). Pyoverdine aus *Pseudomonas putida*. *Z. Naturforsch.*, **47c**: 487-502.

Henriksen, A., Anthoni, U., Nielsen, T.H., Sorensen, J., Christophersen, C., and Gajhede, M. (2000). Cyclic lipoundecapeptide tensin from *Pseudomonas fluorescens* strain 96.578. *Acta Crystallogr. C*, **56**: 113-115.

Hohlneicher, U., Hartmann, R., Taraz, K., and Budzikiewicz, H. (1995). Pyoverdin, ferribactin, azotobactin – a new triade of siderophores from *Pseudomonas chlororaphis* ATCC 9446 and its relation to *Pseudomonas fluorescens* ATCC 13525. *Z. Naturforsch.*, **50c**: 337-344.

Hori, K., Yamamoto, Y., Minetoki, T., Kurotsu, T., Kanda, M., Miura, S., Okamura, K., Furuyama, J., and Saito, J. (1989). Modular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene. *J. Biochem.*, **106**: 639-645.

Hubbard, B. K., and Walsh, C. T. (2003). Vancomycin assembly: nature's way. *Angew Chem. Int.*, **42**: 730-765.

Hutchinson, C. R. (2003). Polyketide and non-ribosomal peptide synthases: falling together by coming apart. *Proc. Natl. Acad. Sci. USA*, **100**: 3010-2.

Jacques, P., Ongena, M., Gwose, I., Seinsche, D., Schröder, H., Delfosse, P., Thonard, P., Taraz, K., and Budzikiewicz, H. (1995). Structure and characterization of isopyoverdin from *Pseudomonas putida* BTP1 and its relation to the biogenetic pathway leading to pyoverdins. *Z. Naturforsch.*, **50c**: 622-629.

Jacques, P. (2011). Surfactin and other Lipopeptides from *Bacillus* spp. In: *Biosurfactants*, Microbiology Monographs 20, G. Soberon-Chavez (ed.), Springer Verlag Berlin Heidelberg, Germany. p. 57-91.

Janek, T., Lukaszewicz, M., Rezanka, T., and Krasowska, A. (2010). Isolation and characterization of two new lipopeptide biosurfactants produced by *Pseudomonas fluorescens* BD5 isolated from water from the Arctic Archipelago of Svalbard. *Bioresour. Technol.*, **101**: 6118-6123;

Jourdan, F., Lazzaroni, S., Mendez, B.L., Lo Cantore, P., De Julio, M., Amadeo, P., Iacobellis, Evidente, A., and Motta, A. (2003). A left handed alpha-helix containing both L- and D- amino acids: the solution structure of the antimicrobial lipodepsipeptide tolaasin. *Proteins*, **52**: 534-543.

Jülich, M., Taraz, K., Budzikiewicz, H., Geoffroy, V., Meyer, J.M., and Gardan, L. (2001). The structure of the pyoverdin isolated from various *Pseudomonas syringae* pathovars. *Z. Naturforsch.*, **56c**: 687-694.

Jülich, M., Taraz, K., Budzikiewicz, H., Geoffroy, V., Meyer, J.M., and Gardan, L. (2001). The structure of the pyoverdin isolated from *Pseudomonas syringae* pathovars. *Z. Naturforsch.*, **56c**: 687.

Khalil-Rivzi, S., Toth, S.I., Van der Helm, D., Vidavski, I., and Gross, M.L. (1997). Structures and characteristics of novel siderophores from plant-deleterious *Pseudomonas fluorescens* A225 and *Pseudomonas putida* ATCC39167. *Biochem.*, **36**: 4163.

King, J.V., Ward, M.K., and Raney, D.E. (1954). Two simple media for the demonstration of pyocyanin and fluorescein. *J. Lab. Clin. Med.*, **44**: 301-307.

- Kuiper, I., Lagendijk, E.L., Pickford, R., Derrick, J.P., Lamers, G.E., Thomas-Oates, J.E., Lugtenberg, B.J., and Bloemberg, G.V. (2004). Characterization of two *Pseudomonas putida* lipopeptide biosurfactants, putisolvin I and II, which inhibit biofilm formation and break down existing biofilms. *Mol. Microbiol.* **51**: 97-113.
- Kurmayer, R., Dittman, E., Fastner, J., and Chorus, I. (2001). Diversity of microcystin genes within a population of the toxic Cyanobacterium *Microcystis* spp. In lake Wannsee (Berlin, Germany). *Microb. Ecol.*, **43**: 107-118.
- Laycock, M.V., Hildebrand, P.D., Thibault, P., Walter, J.A., and Wright (1991). Viscosin, a potent peptidolipid biosurfactant and phytopathogenic mediator produced by a pectolytic strain of *Pseudomonas fluorescens*. *J. agricult. food Chem.*, **39**: 483-489.
- Lehoux, D.E., Sanschagrin, F., and Levesque, R.C. (2000). Genomics of the 35-kb *pvd* locus and analysis of novel *pvdIJK* genes implicated in pyoverdine biosynthesis in *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.*, **190**: 141-146.
- Li, M., Mu Hung, P., Zajkowski, J., Garneau-Tsodikova, S., and Sherman, D.H.(2009). Automated genome mining for natural products. *BMC bioinformatics*, **10**: 185.
- Lipmann, F., Gevers, W., Kleinhauf, H., and Roskoski, R. J. (1971). Polypeptide synthesis on protein templates: the enzymatic synthesis of gramicidin S and tyrocidin. *Adv. Enzymol. Relat. Areas Mol. Biol.*, **35**: 1-34.
- Loper, J.E., and Gross, H. (2007). Genomics analysis of antifungal metabolite production by *Pseudomonas fluorescens* Pf5. *Eur. J. Plant Pathol.*, **119**: 265-278.
- Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.*, **97**: 2651-2674.
- Geszvain, K., and Tebo, B.M. (2010). Identification of a two-component regulatory pathway essential for Mn(II) oxidation in *Pseudomonas putida* GB-1. *Appl. Environ. Microbiol.*, **76**: 1224-1231.
- Matthijs, S., Laus, G., Meyer, J. M., Abbaspour-Tehrani, K., Schäfer, M., Budzikiewicz, H., and Cornelis, P. (2009). Siderophore-mediated iron acquisition in the entomopathogenic bacterium *Pseudomonas entomophila* L48 and its close relative *Pseudomonas putida* KT2440. *Biomaterials*. **21**: 259-271.
- Matthijs, S., Tehrani, K. A., Laus, G., Jackson, R. W., Cooper, R. M., and Cornelis, P. (2007). Thioquinolobactin, a *Pseudomonas* siderophore with antifungal and anti-*Pythium* activity. *Environ. Microbiol.*, **9**: 425-434.
- Medema, M.H., Blin, K., Cimermancic, P., De Jager, V., Zakrzewski, P., Fischbach, M.A., Weber, T., Takano, E., and Breitling, R. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucl. Acids Res.*, **39**: W339-W346.
- Meyer, J. M., Gruffaz, C., Raharinosy, V., Bezverbnaya, I., Schäfer, M., and Budzikiewicz, H. (2008). Siderotyping of fluorescent *Pseudomonas*: molecular mass determination by mass spectrometry as a powerful pyoverdine siderotyping method. *Biomaterials*, **21**: 259-271.

- Meyer, J.M., Stintzi, A., De Vos, D., Cornelis, P., Tappe, R., Taraz, K., and Budzikiewicz, H. (1997). Use of siderophores to type *Pseudomonas aeruginosa* pyoverdine system. *Microbiol.*, **143**: 35-43.
- Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002). Yersiniabactin synthetase: A four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem. Biol.*, **9**: 333-344.
- Miller, D. A., Walsh, C. T., and Luo, L. (2001). C-methyltransferase and cyclization domain activity at the intraprotein PK/NRP switch point of yersiniabactin synthetase. *J. Am. Chem. Soc.*, **123**: 8434-8435.
- Mohn, G., Taraz, K., and Bubzikiewicz, H. (1990). New pyoverdine-type siderophores from *Pseudomonas fluorescens*. *Z. Naturforsch.*, **45b**: 1437-1450.
- Moon, C.D., Zhang, X.X., Matthijs, S., Schäfer, M., Budzikiewicz, H., and Rainey, P.B. (2008). Genomic, genetic and structural analysis of pyoverdine-mediated iron acquisition in the plant growth-promoting bacterium *Pseudomonas fluorescens* SBW25. *BMC Microbiol.* **8**:7.
- Mootz, H. D., and Marahiel, M. A. (1997). The tyrocidin biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.*, **179**: 6843-6850.
- Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2002). Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *Chem. Biochem.*, **3**: 490-504.
- Morikawa, M., Daido, H., Takao, T., Murata, S., Shimonishi, Y., and Imanaka, T. (1993). A new lipopeptide biosurfactant produced by *Arthrobacter* sp. strain MIS38. *J. Bacteriol.*, **175**: 6459-6466.
- Mossialos, D., Ochsner, U., Baysse, C., Chablain, C., Pirnay, J.P., Koedam, N., Budzikiewicz, H., Uria Fernandez, D., Schäfer, M., Ravel, J., and Cornelis, P. (2002). Identification of new, conserved, non ribosomal peptide synthetases from fluorescent *Pseudomonads* involved in the biosynthesis of the siderophore pyoverdine. *Mol. Microbiol.*, **45**: 1673.
- Neilands, J. B. (1995). Siderophores: structure and function of microbial iron transport compounds. *J. Biol. Chem.*, **270**: 26723-6.
- Nettles, J. H., Li, H., Cornett, B., Krahn, J. M., Snyder, J. P., and Downing, K. H. (2004). The binding mode of epothilone A on tubulin by electron crystallography. *Science*, **305**: 866-305.
- Nielsen, T.H., Sorensen, D., Tobiasen, C., Andersen, J.B., Christophersen, C., Givskov, M., and Sorensen, J. (2002). Antibiotic and biosurfactant properties of cyclic lipopeptides produced by fluorescent *Pseudomonas* spp. from the sugar beet rhizosphere. *Appl. Environ. Microbiol.*, **68**: 3416-3423.
- Nielsen, T.H., Christophersen, C., Anthoni, U., and Sorensen, J. (1999). Viscosamide, a new cyclic depsipeptide with surfactant and antifungal properties produced by *Pseudomonas fluorescens* DR54. *J. Appl. Microbiol.*, **87**: 80-90.
- Novak-Thompson, B., and Gould, S.J. (1994). Biosynthesis of the Pseudobactin chromophore from tyrosine. *Tetrahedron*, **50**: 9865-9872.

- O'Brien, I. G., Cox, G. B., and Gibson, F.(1970). Biologically active compounds containing 2,3-dihydroxybenzoic acid and serine formed by *Escherichia coli*. *Biochim. Biophys. Acta*, **201**: 453–60.
- Ongena, M., Jacques, P., van Vyncht, G., Charlier, P., de Pauw, E., Thonart, P., and Budzikiewicz, H. (1997). Structural analysis of two pyoverdins by electrospray and FAB mass spectrometry. *J. Mass Spectrom. Soc. Jpn.*, **46**.
- Ongena, M., Jacques, P., Thonart, P., Gwose, I., Fernandez, D. U., Schäfer, M., and Budzikiewicz, H. (2001). The pyoverdine of *Pseudomonas fluorescens* BTP2, a novel structural type. *Tetrahedron Lett.*, **42**: 5849-5851.
- Ortet, P., Barakat, M., Lalaouna, D., Fochesato, S., Barbe, V., Vacherie, B., Santaella, C., Heulin, T., and Achouak, W.(2011). Complete genome sequence of a beneficial plant root-associated bacterium, *Pseudomonas brassicacearum*. *J. Bacteriol.*, **193**: 3146-3147.
- Paulsen, I.T., Press, C.M., Ravel, J., Kobayashi, D.Y., Myers, G.S.A., Mavrodi, D.V., Deboy, R.T., Seshadri, R., Ren, Q., Madupu, R., Dodson, R.J., Durkin, A.S., Brinkac, L.M., Daugherty, S.C., Sullivan, S.A., Rosovitz, M.J., Gwinn, M.L., Zhou, L., Schneider, D.J., Cartinhour, S.W., Nelson, W.C., Weidman, Watkins, K., Tran, K., Khouri, H., Pierson, E.A., Pierson, L.S., Thomashow, L.S., and Loper, J.E. (2005). Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf5. *Nature Biotechnology*, **23**: 873-878.
- Pedras, M.S., Ismail, N., Quail, J.W., and Boyetchko, S.M. (2003). Structure, chemistry, and biological activity of pseudophomins A and B, new cyclic lipodepsipeptides isolated from the biocontrol bacterium *Pseudomonas fluorescens*. *Phytochem.*, **62**: 1105-1114.
- Persmark, M., Frejd, T., and Matthiasson, B. (1990). Purification, characterization, and structure of pseudobactin 589A, a siderophore from a plant growth promoting *Pseudomonas*. *Biochem.*, **29**: 7348-7356.
- Peypoux, F., Bonmatin, J. M., and Wallach, J. M. (1999). Recent trends in the biochemistry of surfactin. *J. Appl. Microbiol. Biotechnol.*, **51**: 553-563.
- Pfeifer, B. A., Wang, C. C. C., Walsh, C. T., and Kroll, C. (2003). Biosynthesis of Yersiniabactin, a Complex Polyketide-Nonribosomal Peptide, Using *Escherichia coli* as a Heterologous Host. *Appl. Environ. Microbiol.*, **69**: 6698-6702.
- Poppe, K., Taraz, K., and Budzikiewicz, H.(1987). Pyoverdine type from *Pseudomonas fluorescens*. *Tetrahedron*, **43**: 2261-2272.
- Prieto, C., Garcia-Estrada, C., Lorenzana, D., and Martin, J.M.(2011). NRPSp: Non-Ribosomal Peptide Synthase substrate predictor. *Bioinformatics*.
- Raijmakers, J. M., De Bruijn, I., and De Kock, M. J. (2006). Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: diversity, activity, biosynthesis, and regulation. *Mol. Plant Microbe Interact.*, **19**: 699-710.
- Rausch, C., Hoof, I., Weber, T., Wohlleben, W., and Huson, D. H. (2007). Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol. Biol.*, **7**: 78.

- Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D. H. (2005). Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucl. Acids Res.*, **33**: 5799-5808.
- Ravel, J., and Cornelis, P. (2003). Genomics of pyoverdine-mediated iron uptake in pseudomonads. *Trends Microbiol.*, **11**: 195-200.
- Riise, D., Beiderbeck, H., Taraz, K., Budzikiewicz, H., and Gustine, D. (1998). Corrugatin, a lipopeptides siderophore from *Pseudomonas corrugata*. *Z. Naturforsch.*, **53**: 295-304.
- Roongsawang, N., Hase, K., Haruki, M., Imanaka, T., Morikawa, M., and Kanaya, S. (2003). Cloning and characterization of the gene cluster encoding arthrofactin synthetase from *Pseudomonas* sp. MIS38. *Chem. Biol.*, **10**: 869-880.
- Roongsawang, N., Washio, K., and Morikawa, M. (2011). Diversity of nonribosomal peptide synthetases involved in the biosynthesis of lipopeptide biosurfactants. *Int. J. Mol. Sci.*, **12**: 141-172.
- Röttig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011). NRPSpredictor2- a web server for predicting NRPS adenylation domain specificity. *Nucl. Acids Res.*, 1-6.
- Rouhiaden, L., Paulin, L., Suomalainen, S., Hyytiäinen, H., Buikema, W., Haselkorn, R., and Sivonen, K. (2000). Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90. *Mol. Microbiol.*, **37**: 156-167.
- Ruangviriyachai, C., Barelmann, I., Fuchs, R., and Budzikiewicz, H. (2000). An exceptionally large pyoverdine from *Pseudomonas* strain collected in Thailand. *Z. Naturforsch.*, **55c**: 323-327.
- Ruangviriyachai, C., Fernandez, D. U., Fuchs, R., Meyer, J. M., and Budzikiewicz, H. (2001). A new pyoverdine from *Pseudomonas aeruginosa* R'. *Z. Naturforsch.*, **56c**: 933-938.
- Ruangviriyachai, C., Uria-Fernandez, D., Schäfer, M., and Budzikiewicz, H. (2004). Structure proposal for a new pyoverdine from a Thai *Pseudomonas putida* strain. *Spectroscopy*, **18**: 453-458.
- Salah-el-din, A. L. M., Kyslic, P., Stephan, D., and Abdallah, M. A. (1997). Bacterial iron transport: structure elucidation by FAB-MS and by 2 D NMR (^1H , ^{13}C , ^{15}N) of pyoverdine G4R, a peptidic siderophore produced by a nitrogen-fixing strain of *Pseudomonas putida*. *Tetrahedron*, **53**: 12539-12552.
- Scaloni, A., Camoni, L., Di Giorgio, D., Scortichini, M., Cozzolino, R., and Ballio, A. (1997). A new syringopeptin produced by *Pseudomonas syringae* pv. *syringae* strain isolated from diseased twigs of laurel. *Physiol. Mol. Plant*, **51**: 259-264.
- Scaloni, A., Dalla Serra, M., Amodeo, P., Mannina, L., Vitale, R.M., Segre, A.L., Cruciani, O., Lodovichetti, F., Greco, M.L., Fiore, A., Gallo, M., Ambrosio, C., Coraiola, M., Menestrina, G., Graniti, A., and Fogliano, V. (2004). Structure, conformation and biological activity of a novel lipodepsipeptide from *Pseudomonas corrugata*: cormycin A. *Biochem. J.*, **384**: 25-36.
- Schäfer, M., Fuchs, R., Budzikiewicz, H., Springer, A., Meyer, J.M., and Linscheid, M. (2006). Structure elucidation of cyclic pyoverdins and examination of rearrangement reactions

in MS/MS experiments by determination of exact product ion masses. *J. Mass. Spectrom.*, **41**: 1162-1170.

Schäfer, M., Taraz, K., and Budzikiewicz, H. (1991). Zur genese der amidisch an den chromophor von pyoverdinen gebundenen dicarbonsäuren. *Z. Naturforsch.*, **46c**: 398.

Schen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr. Opin. Chem. Biol.*, **7**: 285-295.

Schlegel, K., Fuchs, R., Schäfer, M., Taraz, K., Budzikiewicz, H., Geoffroy, V., and Meyer, J.M. (2001). The pyoverdins of *Pseudomonas* sp. 96-312 and 96-318. *Z. Naturforsch.*, **56c**: 680-686.

Scholz-Schroeder, B.K., Soule, J.D., and Gross, D.C. (2003). The sypA, sypS, and sypC synthetase genes encode twenty-two modules involved in the nonribosomal peptide synthesis of syringopeptin by *Pseudomonas syringae* pv. *syringae* B301D. *Mol. Plant Microbiol. Interact.*, **16**: 271-280.

Schwarzer, D., Finking, R., and Marahiel, M. A. (2003). Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.*, **20**: 275-287.

Segre, A.L., Bachmann, R.C., Ballio, A., Bossa, F., Grgurina, I., Iacobellis, N.S., Marino, G., Pucci, P., Simmaco, M., and Takemoto, J.Y. (1989). The structure of syringomycins A1, E and G. *FEBS Lett.*, **255**: 27-31.

Seinsche, D., Taraz, K., Budzikiewicz, H., and Gondol, D. (1993). Neue pyoverdinsiderophore aus *Pseudomonas putida* C. *J. Prakt. Chem.*, **335**: 157-168.

Shoji, J., Hinoo, H., Katayama, T., Nakaqawa, Y., Ikenishi, Y., Iwatani, K., and Yoshida, T. (1992). Structures of new peptide antibiotics, plusbacins A1-A4 and B1-B4. *J. Antibiot.*, **45**: 824-831.

Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G., and Müller, R. (2000). The myxochelin iron transport regulon of the myxobacterium *Stigmatella aurantiaca* Sg a15. *Eur. J. Biochem.*, **267**: 6476.

Sorensen, D., Nielsen, T.H., Christophersen, C., Sorensen, J., and Gajhede, M. (2001). Cyclic lipoundecapeptide amphisin from *Pseudomonas* sp. strain DSS73. *Acta Crystallogr. C*, **57**: 1123-1124.

Sorensen, J., Nielsen, T.H., Sorensen, D., and Christophersen, C. (2002). Cyclic lipoundecapeptide lokisin from *Pseudomonas* sp. strain DSS41. *Tetrahedron Lett.*, **43**: 4421-4423.

Sorensen, K.N., Kim, K.H., and Takemoto, J.Y., (1996). n vitro antifungal and fungicidal activities and erythrocyte toxicities of cyclic lipodepsinonapeptides produced by *Pseudomonas syringae* pv. *syringae*. *Antimicrob. Agents Chemother.*, **40**: 2710-2713.

Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.*, **6**: 493-505.

Stinzi, A., Johnson, Z., Stonehouse, M., Oschner, U., Meyer, J.M., Vasil, M.L., and Poole, K. (1999). The pvc gene cluster of *Pseudomonas aeruginosa*: role in synthesis of the pyoverdine chromophore and regulation by PtxR and PvdS. *J. Bacteriol.*, **181**: 4118-4124.

- Sultana, R., Füchs, R., Schmickler, H., Schlegel, K., Budzikiewicz, H., Siddiqui, B. S., Geoffroy, V., and Meyer, J. M. (2000a). A pyoverdinin from *Pseudomonas* sp CFML 95-275. *Z. Naturforsch.*, **55c**: 857-865.
- Sultana, R., Siddiqui, B. S., Taraz, K., Budzikiewicz, H., and Meyer, J. M. (2001a). An isopyoverdinin from *Pseudomonas putida* CFML 90-44. *Z. Naturforsch.*, **56c**: 303-307.
- Sultana, R., Siddiqui, B. S., Taraz, K., Budzikiewicz, H., and Meyer, J. M. (2001b). An isopyoverdinin from *Pseudomonas putida* CFML 90-33. *Tetrahedron*, **57**: 1019-1023.
- Sultana, R., Siddiqui, B. S., Taraz, K., Budzikiewicz, H., and Meyer, J. M. (2000b). A pyoverdinin from *Pseudomonas putida* CFML 90-51 with a Lys ϵ -amino link in the peptide chain. *Biomaterials*, **13**: 147-152.
- Tappe, R., Taraz, K., Budzikiewicz, H., Meyer, J.M., and Lefevre, J.F. (1993). Structure elucidation of a pyoverdinin produced by *Pseudomonas aeruginosa* ATCC 27853. *J. Prakt. Chem.*, **335**: 83-87.
- Teintze, M., Hossain, M.B., Barnes, C.L., Leong, J., and Van der Helm, D. (1981). Structure of ferric pseudobactin, a siderophore from a plant growth promoting *Pseudomonas*. *Biochem.*, **20**: 6446-6457.
- Teintze, M., Hossain, M.B., Barnes, C.L., Leong, J., and Van der Helm, D. (1981). Structure of ferric pseudobactin, a siderophore from a plant growth promoting *Pseudomonas*. *Biochem.*, **20**: 6446-6457.
- Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2002). UNIT 2.3 multiple sequence alignment using clustalW and clustalX. *Current protocols in Bioinformatics*, 1-22.
- Uchiyama, I., Higuchi, T., and Kawai, M.(2009). MBGD update 2010: toward a comprehensive resource for exploring microbial genome diversity. *Nucl. Acids Res.*, **35**: D361-D365.
- Ui, H., Miyake, T., Linuma, H., Imoto, M., Naganawa, H., Hattori, S., Hamada, M., Takeuchi, T., Umezawa, S., and Umezawa, K. (1997). Pholipeptin, a novel cyclic lipoundecapeptide from *Pseudomonas fluorescens*. *J. Org. Chem.*, **62**: 103-108.
- Uria-Fernandez, D., Fuchs, R., Schäfer, M., Budzikiewicz, H., and Meyer, J. M.(2003a). The pyoverdinin of *Pseudomonas fluorescent* G173, a novel structural type accompanied by unexpected natural derivatives of the corresponding ferribactin. *Z. Naturforsch.*, **58c**: 1-10.
- Uria-Fernandez, D., Geoffroy, V., Schäfer, M., Meyer, J. M., and Budzikiewicz, H., (2003b). Structure revision of pyoverdins produced by plant-growth promoting and plant-deleterious *Pseudomonas* species. *Monatsh. Chem.*, **134**: 1421-1431.
- Van Dörhen, H., Keller, U., Vater, J., and Zocher, R. (1997). Multi-functional peptide synthetases. *Chem. Rev.*, **97**: 2675-2706.
- Vodovar, N., Vallenet, D., Cruveiller, S., Rouy, Z., Barbe, V., Acosta, C., Cattolico, L., Jubin, C., Lajus, A., Seuren, B., Vacherie, B., Wincker, P., Weissenbach, J., Lemaitre, B., Médique, C., and Bocard, F. (2006). Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium *Pseudomonas entomophila*. *Nat. Biotechnol.*, **24**: 673-679.

- Voss, J., Taraz, K., and Budzikiewicz, H. (1999). A pyoverdin from the Antartica strain 51 W of *Pseudomonas fluorescens*. *Z. Naturforsch.*, **54c**: 156-162.
- Vossen, W., and Taraz, K. (1999). Structure of the pyoverdine PVD 2908 – a new pyoverdin from *Pseudomonas sp.* 2908. *Biometals*, **12**: 323-329.
- Wackett, L.P., and Gibson, D.T. (1988). Degradation of trichloroethylene by toluene dioxygenase in whole-cell studies with *Pseudomonas putida* F1. *Appl. Environ. Microbiol.*, **54**: 1703-1708.
- Weber, M., Taraz, K., Budzikiewicz, H., Geoffroy, V., and Meyer, J.M.(2000). The structure of a pyoverdine from *Pseudomonas sp.* CFML 96.188 and its relation to other pyoverdines with a cyclic C-terminus. *Biometals*, **13**: 301-309.
- Weyens, N., Truyens, S., Dupae, J., Newman, L., Taghavi, S., Van der Lelie, D., Carleer, R., and Vangronsveld, J. (2010). Potential of the TCE-degrading endophyte *Pseudomonas putida* W619-TCE to improve plant growth and reduce TCE phytotoxicity and evapotranspiration in poplar cuttings. *Environ. Pollut.*, **158**: 2915-2919.
- Wong-Lung-Sang, S., Bernardini, J.J., Hennard, C., Kyslic, P., Dell, A., and Abdallah, M.A. (1996). Bacterial siderophores: structure elucidation, 2D ^1H and ^{13}C NMR assignments of pyoverdins produced by *Pseudomonas fluorescens* CHA0. *Tetrahedron Lett.*, **37**: 3329-3332.

Résumé

La synthèse peptidique non ribosomiale est réalisée par de grands complexes multienzymatiques, appelés synthétases, qui permettent la synthèse de peptide par une voie indépendante de l'intervention des ribosomes. Les peptides produits par cette voie de synthèse (NRPs) sont largement étudiés et divers outils d'analyse bioinformatique qui permettent la prédiction de la structure d'une synthétase ainsi que de la composition de son produit, sont disponibles. Ce type de synthèse peptidique a été décrit chez plusieurs microorganismes. Notre choix de modèle d'étude s'est porté sur les *Pseudomonas* fluorescents producteurs de deux types de NRPs, les lipopeptides cycliques (CLPs) et des sidérophores dont le plus représenté est la pyoverdine. L'étude de ces NRPs a été entreprise par des études expérimentales et bioinformatiques. Ces travaux ont permis de montrer le potentiel de synthèse non ribosomiale par une étude bioinformatique des 20 génomes de *Pseudomonas* disponibles. L'étude des synthétases de pyoverdine et l'extraction des signatures des domaines d'adénylation (A) ont permis d'améliorer la prédiction issue des outils d'analyse disponibles. Des expériences de feeding suivies par spectrométrie de masse MALDI-TOF ont permis de mettre en évidence des domaines A permissifs au sein des synthétases de pyoverdines.

Abstract

The non-ribosomal peptide synthesis is performed by large multienzymatic complexes, called synthetases, which allow the synthesis of a peptide by a way independent of the intervention of the ribosomes. The peptides produced by this synthetic way (NRPS) are widely studied and various bioinformatics analysis tools that allow the prediction of the structure of the synthetase and the composition of his product, are available. This kind of peptide synthesis has been described in several microorganisms. Our choice focus on the fluorescent pseudomonads producing two types of NRPS, the cyclic lipopeptides (CLPS) and siderophores, the most represented is the pyoverdin. The study of these NRPS was performed by experimental and bioinformatics analysis. This work has demonstrated the potential of non-ribosomal synthesis by a bioinformatics study of the 20 available genomes of *Pseudomonas*. Study of pyoverdine synthetases and extracting signatures of adenylation domains (A) have allowed to improving the prediction of the available tools. Feeding experiments followed by MALDI-TOF helped to highlight permissive A domains in pyoverdins synthetases.