HAL open science

Spécificité d'import et de transformation des agrocinopines et dérivés chez les pathogènes du genre Agrobacterium

Grégory Deicsics

▶ To cite this version:

Grégory Deicsics. Spécificité d'import et de transformation des agrocinopines et dérivés chez les pathogènes du genre Agrobacterium. Biochimie [q-bio.BM]. Université Paris-Saclay, 2020. Français. NNT : 2020UPASQ013 . tel-03793798

HAL Id: tel-03793798

https://theses.hal.science/tel-03793798v1

Submitted on 2 Oct 2022

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

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

université PARIS-SACLAY

Spécificité d'import et de transformation des agrocinopines et dérivés chez les pathogènes du genre *Agrobacterium*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°569 : Innovation Thérapeutique du Fondamental à l'Appliqué (ITFA)
Spécialité de doctorat : Biochimie et Biologie structurale
Unité de Recherche : Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.
Référent : Faculté de pharmacie

Thèse présentée et soutenue à Gif-Sur-Yvette, le 21 octobre 2020, par

Grégory DEICSICS

Composition du Jury

Nicolas BAYAN PR (UP-Saclay), I2BC - Orsay (UMR 9198)	Président
Anne POUPON DR, HDR (INRAE), PRC - Tours (UMR 0085)	Rapporteur & Examinatrice
Marc GRAILLE DR, HDR (CNRS), BIOC - Palaiseau (UMR 7654)	Rapporteur & Examinateur
Muriel MASI Maitre de Conférence (UP-Saclay), I2BC – Orsay (UMR 9198)	Examinatrice
Matthieu FONVIELLE CR (INSERM), CRC - Paris (UMRS 1138)	Examinateur
Pr. Herman VAN TILBEURGH PR (UP-Saclay), I2BC - Orsay (UMR 9198)	Directeur de thèse

Thèse de doctorat

NNT : 2020UPASQ013

La réalisation d'une thèse est une épreuve qui n'est facile pour personne.
Elle a été pour moi une étape, qui paradoxalement m'a montré la personne que je veux et ne veux pas devenir.

Remerciements

Dans ces quelques lignes je tiens à remercier les personnes ayant contribué de loin ou de près au bon déroulement de cette thèse.

Dans un premier temps, j'aimerais remercier ***Anne Poupon*** *et* ***Marc Graille*** *pour avoir accepté d'être rapporteurs de cette thèse,* ***Muriel Masi*** *et* ***Matthieu Fonvielle*** *d'être examinateurs ainsi que* ***Nicolas Bayan*** *pour avoir accepté de la présider.*

Je tiens particulièrement à remercier ***Herman Van Tilbeurgh*** *de m'avoir accueilli et encadré dans l'équipe Fonction et Architecture des Assemblages MacroMoléculaires me permettant ainsi de soutenir ma thèse. Merci également de m'avoir aidé et accordé du temps pour la rédaction de cette thèse.*

Un grand, grand merci à ***Sophie Quevillon-Cheruel*** *pour tout … ton accueil, tes encouragements, tes conseils, tes remarques toujours constructives, ton investissement, ta joie de vivre et surtout ta gentillesse. Je ne pensais pas rencontrer un jour une personne si rayonnante et admirable et j'ai été ravi de faire ta connaissance. Durant cette thèse, tu as été présente pour moi, tu m'as aidé à me relever et je t'en serais éternellement reconnaissant. Encore Merci !*

Je tiens à remercier ***tous*** *les membres de* ***l'équipe FAAM*** *de m'avoir aussi bien accueilli et grâce à laquelle je me suis très vite bien senti. J'ai pu découvrir un autre monde scientifique, un monde sein et agréable, où passer du temps est devenu un plaisir. J'ai rencontré des gens fantastiques, dans un milieu de vie d'entraide, d'écoute permettant beaucoup plus de productivité et un bien-être sans égal. Je souhaite à tout le monde d'évoluer dans une équipe semblable.*

Merci à ***Pauline Blin****,* ***Almudena Gonzàlez-Mula****,* ***Marta Torres Bejar****,* ***Léa Lo Bello****,* ***Claire Cargemel****,* ***Sophia Missoury****,* ***Marion Dos Santos Malhao****,* ***Christophe Velours****,* ***Stéphane Plancqueel****,* ***Adrien Fleck*** *pour nos discutions, nos soirées et tous ces bons moments. Merci également à* ***Thanh-Lan Lai*** *et* ***Raynald Cossard*** *pour leur soutien durant cette période difficile. Enfin, une mention particulière à* ***Florian Lamouche****,* ***Quentin Barrière*** *et* ***Quentin Nicoud*** *qui ont été de véritables piliers et amis pour moi pendant cette aventure.*

Durant ces 3 années j'ai été amené à travailler et échanger avec plusieurs personnes et je te tiens particulièrement à remercier ***Almudena Gonzàlez-Mula****,* ***Magalie Aumont-Nicaise, Michel Desmadril*** *pour leur aide, leur expertise et leur gentillesse.*

Pour finir, je tiens à remercier ***ma compagne, ma famille, mes amis et mes proches*** *qui ont toujours cru en moi et m'ont toujours soutenu.*

Table des matières

Liste des abréviations 9

Avant-propos 11

INTRODUCTION 15

a. Les bactéries du genre Agrobacterium 17

i. Caractéristiques 17

ii. Taxonomie 17

iii. Structure du génome 19

b. La galle du collet 20

i. Cycle infectieux du pathogène 20

1. Chimiotactisme 20

2. Attachement aux cellules végétales 23

3. Activation des gènes de virulence 24

4. Transfert de l'ADN-T au génome de la plante 25

5. Expression des gènes de l'ADN-T 28

c. Les opines 29

i. Les différents types d'opines et de plasmides Ti 29

ii. Processus d'utilisation des agrocinopines 32

1. Opéron acc 32

2. Rôles des agrocinopines 33

3. Les Glycérophosphodiesterases 33

a. La phosphodiesterase AccF de l'opéron *acc* 35

iii. Quorum sensing 36

1. Principe 36

2. Mécanisme de communication quorum sensing chez Agrobacterium 37

3. La régulation du quorum sensing chez Agrobacterium 38

4. Propagation de la virulence 39

d. L'import et la transformation des opines 39

i. L'import des opines 39

1. Les transporteurs membranaires d'opine 40

2. Les Periplasmic Binding Proteins (PBP) 41

a. Structure globale des PBPs 41

b. Mécanisme d'interaction 41

c. Classification des différentes PBP 41

Objectifs de la thèse : 45

RESULTATS 47

Chapitre 1 : 49

Etude de la spécificité de la fixation des agrocinopines par la protéine périplasmique PBP AccA chez les pathogènes du genre *Agrobacterium*. 49

a. Introduction 51

b. Résultats 52

i. Clonages, expressions et purifications 52

ii. Synthèse de molécules signal et Transfert conjugatif 55

iii. Etude des interactions protéine:ligands des différentes PBP AccA. 59

1. PBP $AccA_{C58}$ 59

2. PBP $AccA_{Bo542}$ 60

3. PBP $AccA_{S4}$ 63

4. Conclusion 65

iv. Cristallogenèse 65

v. Diffraction aux rayons X et résolution des structures 66

vi. Analyse et comparaison structurale de la protéine AccA des différentes souches et en complexe avec divers ligands. 66

1. Les structures obtenues 66

2. Description structurale des PBP AccA. 67

a. Structure globale des PBP $AccA_{Bo542/S4}$ 67

b. Site de fixation des PBP $AccA_{C58/Bo542/S4}$ 68

i. Analyse des poches de fixation d'$AccA_{C58/Bo542/S4}$ 69

ii. Analyse des interactions protéine:ligands 70

iii. Modélisation de l'Agrocinopine A dans le site de fixation de la PBP $AccA_{Bo542}$ 72

iv. Choix de mutants 74

vii. Test des mutants identifiés 74

1. Analyse $AccA_{BO542}$ mutant P44A 75

2. Analyse $AccA_{BO542}$ mutant P44N 76

3. Analyse $AccA_{BO542}$ mutant F145Y/G440Q 77

viii. Hypothèses concernant la spécificité de la protéine $AccA_{Bo542}$ 78

1. Révision des choix de mutation d'$AccA_{Bo542/C58/S4}$ 78

c. Conclusion 80

Chapitre 2 : 83

Etude structure-fonction de la phosphodiesterase AccF 83

a. ***Introduction*** 85
b. ***Résultats*** 86
i. Clonage, expression et purification 86
ii. Cristallogenèse 90
iii. Diffraction aux rayons X et résolution des structures 90
iv. Analyse et comparaison structurale de la protéine AccF. 91
1. Les structures obtenues 91
2. Description structurale de la protéine. 92
3. Description de l'interaction protéine-ligand 97
v. Le mécanisme enzymatique des glycerophosphodiesterases 100
vi. Choix des mutants 103
vii. Etude de l'activité enzymatique de la protéine $AccF_{C58/Bo542}$ sauvage et de ses mutants. 106
viii. Explication structurale de l'activité allostérique observée chez la protéine $AccF_{Bo542/C58}$ sauvage 108
1. L'enzyme $AccF_{Bo542}$ 108
2. L'enzyme $AccF_{C58}$ 111
ix. Etude des interactions Enzyme:Ligands 112
x. Impact de l'ion sur la stabilité de la protéine 114
c. ***Conclusion*** 115
Discussion et perspectives 117
a. ***Spécificité des PBP $AccA_{Bo542}$ et $AccA_{S4}$*** 120
i. L'apport de connaissance des travaux effectués 121
ii. Perspectives 122
1. Comprendre la spécificité de reconnaissance 122
2. Implications in vivo de la spécificité 124
b. ***Phosphodiesterase AccF d'Agrobacterium C58 et Bo542*** 124
i. L'apport de connaissance des travaux effectués 124
ii. Perspectives 128
Matériels et Méthodes 132
a. ***Synthèse et purification des agrocinopines et dérivés.*** 134
b. ***Bilan des clonages et mutagénèses*** 134
c. ***Expression des protéines*** 135
i. Transformation par choc thermique 135
ii. Tests d'expression 136

iii. Expressions à grande échelle 136
d. Purification des protéines 137
e. Signaux de quorum sensing et transfert conjugatif 138
i. Détermination de l'efficacité du transfert conjugatif 139
ii. Mesure de la production de molécules signal 139
f. Cristallisation et collecte de données 140
i. Cristaux de la protéine AccA 141
ii. Cristaux de la protéine AccF 141
g. Détermination des structures et affinements 141
i. Collecte des données et affinement des différentes structures 141
ii. Structure de la protéine AccA 142
iii. Structure de la protéine AccF 142
h. Détermination de l'activité enzymatique d'AccF 143
i. Caractérisation des interactions protéines:ligands et de la stabilité des protéines 143
i. Mesure de la constante de dissociation par titration de fluorescence 143
ii. Détermination des températures d'inflexion (Tycho) 144
iii. Analyse et détermination des températures de fusion (DSC) 144
iv. Mesure de la constante de dissociation par titrage de calorimétrie isotherme 145
Annexes 149
Tables cristallographiques $AccF_{Bo542}$ 151
..... 152
Tables cristallographiques $AccA_{C58/Bo542/S4}$ 154
Références 157

Liste des abréviations

nM/µM/mM: Nanomolaire/Micromolaire/Millimolaire
Å: Angström
A2P: Arabinose-2-phosphate
ABC: ATP binding cassette
ABA: acide abscissique
ADN: acide désoxyribonucléique
ADN-T: ADN transféré chez A. tumefaciens
AHLs: acyl-homoserine-lactone synthase
AgroA/B/C/D: Agrocinopine A/B/C/D
AIA: acide indole-acétique
AMP: adénosine 5'-monophosphate
ARN: acide ribonucléique
ATP: adénosine triphosphate
Ca: Calcium
Cα: Carbone alpha
Co: Colbat
CFBP : Collection Française de Bactéries associées aux Plantes
DO: Densité Optique
DSC: differencial scanning calorimetry
ECF: Energy-Coupling Factor
G2P: Glucose-2-phosphate
IAM: indole-3-acétamide
IMC: inner membrane complex
Kb: kilobase
kDa: kilodalton
LB: Luria Broth
LB/RB: Left Border/Right Border
IMAC: immobilized metal affinity chromatography
IPTG: isopropyl β-D-1-thiogalactopyranoside
IPT : isopentényl transferase
ITC: isothermal titration calorimetry
MCP: methyl-accepting chemotaxis protein
Mg : Magnésium
Mn : Manganèse
NAHL: N-acylhomosérine lactone
NBD: Nucleotide Binding Domain
NCBI: National Center for Biology Information
Ni: Nickel
OGM : Organisme génétiquement modifié
OMC: outer membrane complex
PBP: Periplasmic binding protein
PDB: Protein Data Bank

pTi: plasmid Tumor inducing
QS: Quorum sinsing
RMSD: Root mean square deviation
SDS-PAGE: Sodium dodecylsulfate-Polyacrylamide Gel Electrophoresis
SAM: S-adenosylmethionine
SBP: solute-binding protein
SEC : Size exclusion chromatography
SSB: Single Stand Binding
SST4 : Système de sécrétion de type 4
T4CP : Type IV coupling protein
Ti: Temperature of inflection
TIM: Triosephosphate Isomerase
Tm: Temperature of melting
TMD: Transmembrane Domaine
TRAP: Tripartite ATP-Independent Periplasmic

Avant-propos

Agrobacterium tumefaciens est une bactérie bien connue pour son utilisation lors de la création d'organismes génétiquement modifiés ou OGM. Cependant, ce microorganisme est en premier lieu une bactérie du sol qui peut potentiellement être pathogène et infecter plus de 600 espèces de plantes différentes. Cette pathogénicité est portée par un plasmide nommé Ti (Tumor inducing). La présence de ce plasmide engendre la formation de tumeurs et un détournement du métabolisme de la plante. Ceci a un impact sur l'esthétique et la croissance des plantes infectées. Il y a donc un intérêt agronomique et économique à s'intéresser au mécanisme d'infection de cette bactérie pathogène.

En détournant le métabolisme de la plante, la bactérie lui fait produire des sources nutritives uniquement utilisables par la bactérie, grâce à un opéron nommé *acc* (agrocinopine catabolism) présent sur le plasmide Ti. Cet opéron est, entre autre, impliqué dans l'import et la transformation de ces sources nutritives nommées opines, et plus particulièrement agrocinopines dans le cadre de cette étude.

L'objectif de cette thèse a été de mieux comprendre le fonctionnement et la spécificité d'interaction protéine-ligand de certaines protéines de l'opéron *acc* chez différentes souches bactériennes du genre *Agrobacterium*.

La partie Introduction de cette thèse regroupe les connaissances actuelles concernant la bactérie pathogène *Agrobacterium tumefaciens* et donne notamment une description génomique et taxonomique des souches utilisées. Elle évoque également la maladie développée par la plante suite à l'infection par le pathogène, l'utilisation des opines qui favorise la croissance bactérienne, et le partage de virulence. Enfin, j'ai réalisé une description des familles de protéines étudiées durant cette thèse.

La partie Résultats regroupe les travaux que j'ai effectués durant ma thèse sur les deux sujets étudiés. Le premier sujet concerne la protéine AccA de l'opéron *acc* responsable de la reconnaissance périplasmique des agrocinopines. Je me suis particulièrement intéressé à la spécificité de liaison d'AccA des souches Bo542 et S4 d'*A. tumefaciens* vis-à-vis des différentes agrocinopines. De précédentes études ayant montré la non spécificité de cette protéine chez la souche modèle C58 d'*A. tumefaciens* pour les différentes agrocinopines, il était intéressant de savoir s'il en était de même pour les souches Bo542 et S4. Le second sujet s'est orienté vers l'étude d'une seconde protéine de l'opéron *acc*, la glycerophosphodiesterase AccF. Cette enzyme est une métalloprotéine responsable du clivage et de la transformation des agrocinopines en vue de leur utilisation. Le but a été de comprendre le mécanisme de son fonctionnement en réalisant une étude structurale et enzymatique.

Pour finir, la discussion et les perspectives portant sur la spécificité de liaison des protéines AccA vis-à-vis des agrocinopines et du mécanisme de la phosphodiesterase AccF sont présentées.

La réalisation de cette thèse a été soutenue par l'Ecole doctorale Innovation thérapeutique : du fondamental à l'appliqué (ED569), dans le cadre d'une bourse du Ministère de l'enseignement supérieur et de la recherche (MESR), entre 2017 et 2020.

INTRODUCTION

a. ***Les bactéries du genre Agrobacterium***

La première bactérie du genre *Agrobacterium* a été isolée en 1897 par le botaniste italien Fridiano Cavara. Ses travaux ont ensuite été repris par deux chercheurs américains nommés E.F. Smith et C.O. Townsend qui ont isolé la bactérie d'une galle de marguerite et l'ont nommée *Bacterium tumefaciens* (Kado, 2014). C'est en 1942 que la bactérie sera finalement nommée *Agrobacterium tumefaciens* (Conn, 1942).

i. Caractéristiques

La famille bactérienne des Rhizobiacceae est composée de plusieurs genres qui peuvent avoir pour rôle d'aider ou de perturber le développement de la plante. Certaines bactéries du genre *Rhizobium* sont capables d'établir une relation symbiotique avec les plantes légumineuses via des nodosités afin de leur permettre de fixer l'azote atmosphérique. Cependant, les bactéries du genre *Agrobacterium*, majoritairement saprophytes et non pathogènes, peuvent être des phytopathogènes si elles possèdent des éléments génétiques de virulence. C'est le cas des bactéries de l'espèce *Agrobacterium tumefaciens* qui, grâce au plasmide de virulence nommé Ti pour « Tumor inducing », engendre le développement de la galle du collet. Ce genre bactérien se représente sous forme de bacille, à coloration gram négative, avec une taille de l'ordre du micromètre, et possède des flagelles pour permettre sa mobilité. Enfin ces bactéries ont une température optimale de croissance qui se situe entre 25 et 28°C.

ii. Taxonomie

Les bactéries du genre *Agrobacterium* ont été classées de 3 manières différentes. L'attribution du nom d'espèce en fonction de la pathogénicité a dans un premier temps permis de les différencier (Young, 2008). Ainsi *A. tumefaciens* est responsable du développement d'une maladie nommée la galle du collet, *A. vitis* induit la galle de la vigne aussi appelée broussin de vigne, *A. rubi* engendre la galle du framboisier, *A. rhizogene* provoque des chevelus racinaires ou « hairy roots », *A. larrymoorei* est responsable de la galle du figuier et enfin *A. radiobacter* regroupe les souches non pathogènes. Contrairement à *A. vitis* qui développe la galle chez les plants de vigne, A tumefaciens infecte différent type de plants au niveau du collet qui est une région de la plante localisée à la base de la tige.

Dans un second temps, c'est le classement en biovars, reposant sur une variabilité des propriétés biochimiques et physiologiques des souches, qui a permis une distinction des différentes souches. Le biovar 1 regroupe les souches d'*A. tumefaciens* et d'*A. radiobacter*. Le biovar 2 regroupe les souches d'*Agrobacterium rhizogenes* renommé *Rhizobium rhizogenes*.

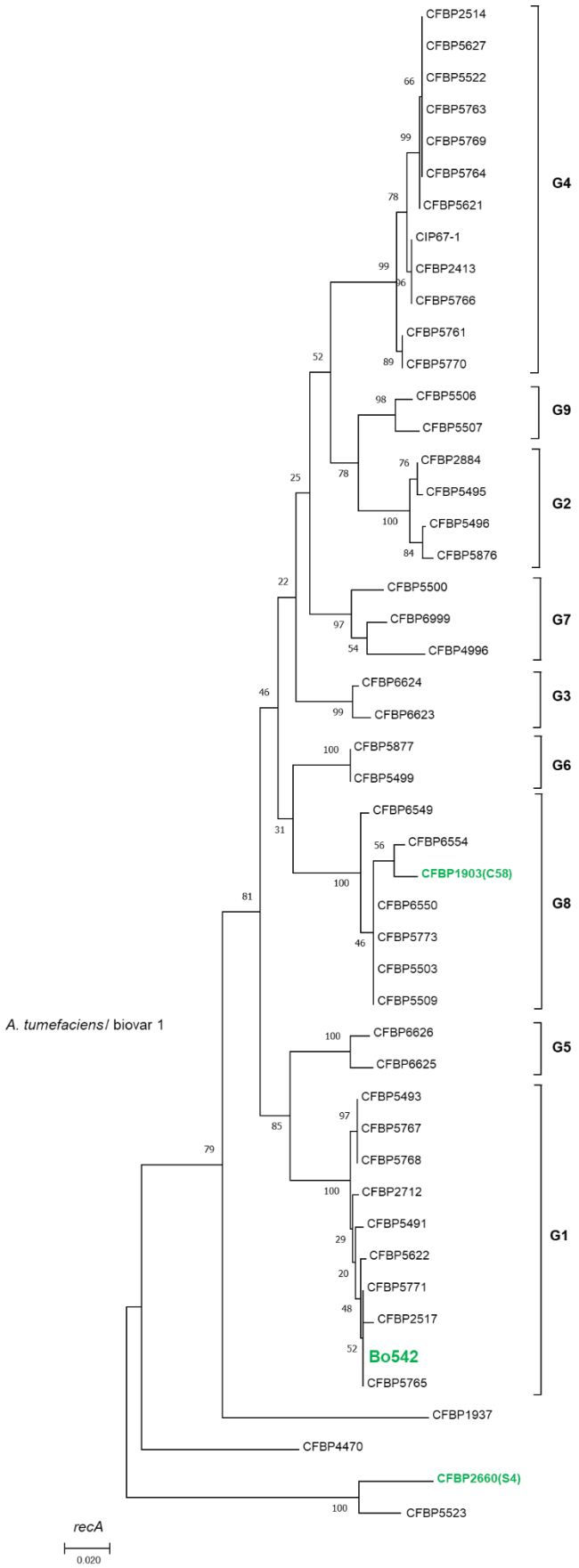


Figure 1: Arbre phylogénique partiel du genre Agrobacterium construit à partir de la comparaison du gène recA. Les souches C58, Bo542 et S4 utilisées et évoquées dans ce manuscrit sont représentées en vert.

Enfin, le biovar 3 ne regroupe que les souches d'*Agrobacterium vitis* renommées « *Allorhizobium vitis* » (Ramírez-Bahena et al., 2014).

Enfin, les souches d'*Agrobacterium* ont été classées par groupes génomiques en se basant sur des analyses de polymorphisme et de relations phylogénétiques. Ces travaux ont permis d'identifier 10 groupes génomiques (G1 à G9 et G13) pour l'espèce d'*Agrobacterium tumefaciens* (Figure 1). De plus, des travaux ont mené à identifier un gène spécifique ainsi qu'à l'identification de fonctions écologiques particulières au sein de chaque groupe génomique. Le groupe G8 a donc été proposé comme espèce à part entière et a été renommé *Agrobacterium fabrum*. Aujourd'hui, il est possible de trouver différentes appellations dans la bibliographie et j'utiliserai les noms d'*Agrobacterium tumefaciens* pour les souches C58 (*A. fabrum*) et Bo542 et d'*Agrobacterium vitis* (*Allorhizobium vitis*) pour la souche S4.

iii. Structure du génome

Il existe une multitude de souches différentes au sein de l'espèce *Agrobacterium tumefaciens*. Ces espèces sont généralement composées de 4 réplicons : un chromosome circulaire d'environ 2.8 Mb (atu0001 à atu2835), un chromosome linéaire d'environ 2 Mb (atu3000 à atu4898) et deux plasmides, à savoir le plasmide At (*Agrobacterium tumefaciens*) d'environ 0.5 Mb et le plasmide Ti (Tumor inducing) d'environ 0.2 Mb. Les plasmides ne sont pas nécessaires à la survie de la bactérie mais leur confère un avantage adaptatif lorsque celle-ci est présente dans le sol dans le cas du plasmide At (Platt, Morton, Barton, Bever, & Fuqua, 2014), et un caractère virulent qui aide à la croissance au sein des tumeurs en ce qui concerne le plasmide Ti.

Le plasmide At ne contient aucun gène de survie mais plutôt des gènes codant pour des systèmes de transport et d'utilisation de certaines molécules nutritives. Ce plasmide confère donc un avantage dans des conditions pauvres en nutriments lorsque la bactérie le possède et est localisée dans le sol, vis-à-vis de celles qui n'en possèdent pas (Platt et al., 2014). Il contient également tous les gènes impliqués dans sa réplication et sa conjugaison et est présent chez la majorité des souches d'*Agrobacterium tumefaciens*, qu'elles soient virulentes ou non (Chen, Chen, Wood, & Nester, 2002; Lassalle et al., 2011). Enfin, même si sa présence semblerait avoir un impact sur la taille des tumeurs, ce plasmide n'est en aucun cas un facteur de virulence et ne permet pas le développement de la galle du collet contrairement au plasmide Ti (Matthysse, Jaeckel, & Jeter, 2008; Morton, Merritt, Bever, & Fuqua, 2013; Nair, Liu, & Binns, 2003; Platt et al., 2014).

Le plasmide Ti n'est pas non plus nécessaire à la survie de la bactérie mais il est tout de même indispensable pour lui octroyer son pouvoir pathogène et provoquer la galle du collet. Il

contient l'ADN-T qui est transféré et intégré au génome de la plante lors de l'infection, ainsi que les gènes *vir* codant pour les protéines permettant son transfert et son intégration. Les gènes impliqués dans l'import et l'utilisation des opines sont également présents sur ce plasmide et sous forme d'opéron. Enfin le plasmide Ti contient les régulons *tra* et *trb*, contrôlés par les gènes *traI* et *traR* du quorum sensing (QS), qui permettent la conjugaison du réplicon (Dessaux & Faure, 2018).

b. ***La galle du collet***

La galle du collet est une maladie contractée par les plantes lorsqu'elles sont en contact direct avec une bactérie pathogène nommée *Agrobacterium tumefaciens*. Son nom « galle du collet » vient de la particularité phénotypique de cette maladie à former des excroissances tumorales ou galles au niveau du site d'infection (Figure 2). Le développement de ces tumeurs est favorisé en présence de blessures qui sont souvent localisées au niveau du collet. Cette région correspond à la région charnière entre la partie racinaire et la pousse de la plante. Ces tumeurs ont pour la première fois étaient observées en 1679 par l'italien Malpighi (Kado, 2014).

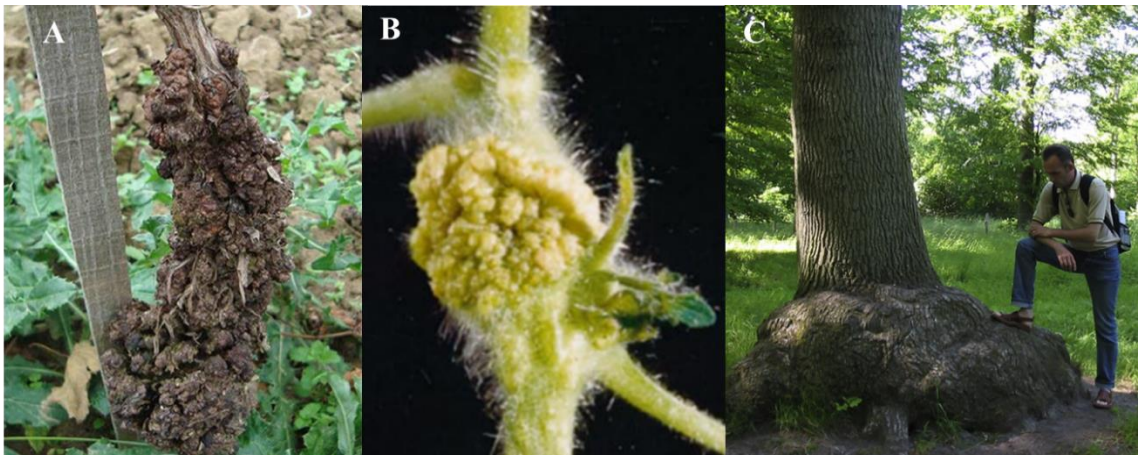


Figure 2 : Excroissance tumorales provoquées par le pathogène *Agrobacterium*. A : Galle sur plant de vigne (Kuzmanović et al.,2012). B: Excroissance tumorale de pied de tomate (Aloni, Wolf, Feigenbaum, Avni, & Klee, 1998) C : Tumeur de grande taille d'un chêne (P ă curar et al., 2011).

i. Cycle infectieux du pathogène

La bactérie potentiellement pathogène est naturellement présente dans le sol et son coté pathogène nécessite la présence d'un plasmide de virulence nommé plasmide Ti (Tumor inducing) pour pouvoir infecter les plantes. De plus, un contact avec une blessure de plante est indispensable pour pouvoir initier le développement de la galle du collet. Cependant, la bactérie peut être éloignée de ce site de blessure et doit donc se déplacer. Le déplacement de la bactérie est rendu possible car *A. tumefaciens* est une bactérie mobile munie de flagelles placés sous le contrôle de molécules chimioattractantes libérées lors de la blessure.

1. Chimiotactisme

Le chimiotactisme se définit par la capacité d'une substance chimique à avoir un effet d'attraction ou de répulsion sur une cellule vivante. Ces molécules contrôlent le mécanisme de mobilité des bactéries via le sens de rotation du flagelle ou via la vitesse de rotation (Figure 3). Lorsque le sens du flagelle est impacté, comme chez *E. coli* par exemple, en l'absence de molécules chimioattractantes, la rotation horaire des flagelles est désordonnée et permet un changement de direction. A l'inverse, la rotation du flagelle antihoraire est ordonnée quand la bactérie est en présence de molécules chimioattractantes et permet un avancement significatif dans cette même direction. Il existe également un système impliquant cette fois la vitesse de rotation des flagelles, notamment chez *Sinorhizobium meliloti* et *A. tumefaciens*. En absence d'attractants chimiques, la vitesse du flagelle est ralentie et permet un changement de direction. Lorsque ces molécules sont présentes, il y a au contraire une ré-augmentation de la vitesse, engendrant une avancée importante dans la direction de ces molécules (Figure 3) (Guo, Huang, & Yang, 2017; Scharf & Schmitt, 2002). Ainsi, lors de la blessure d'une plante, de nombreuses molécules sont libérées, dont des composés phénoliques, des sucres et de l'acide organique qui engendrent une baisse de pH. En présence de ces molécules, la bactérie est attirée au site de blessure.

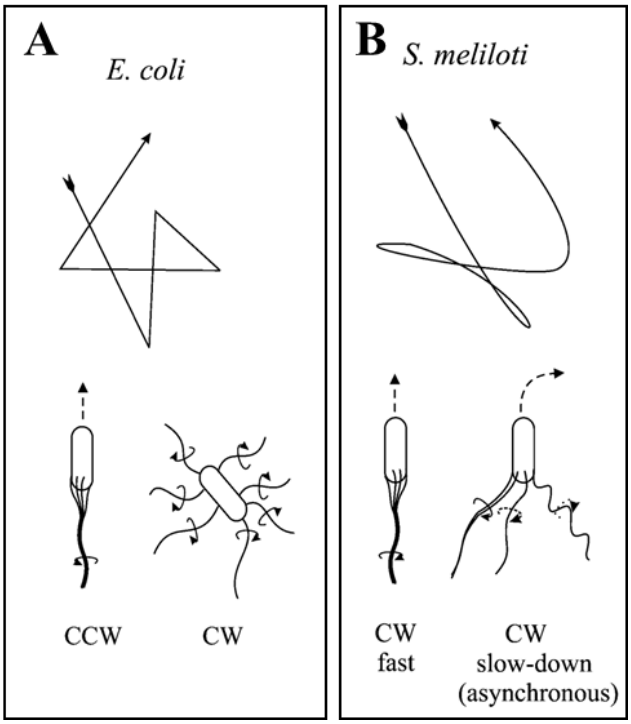
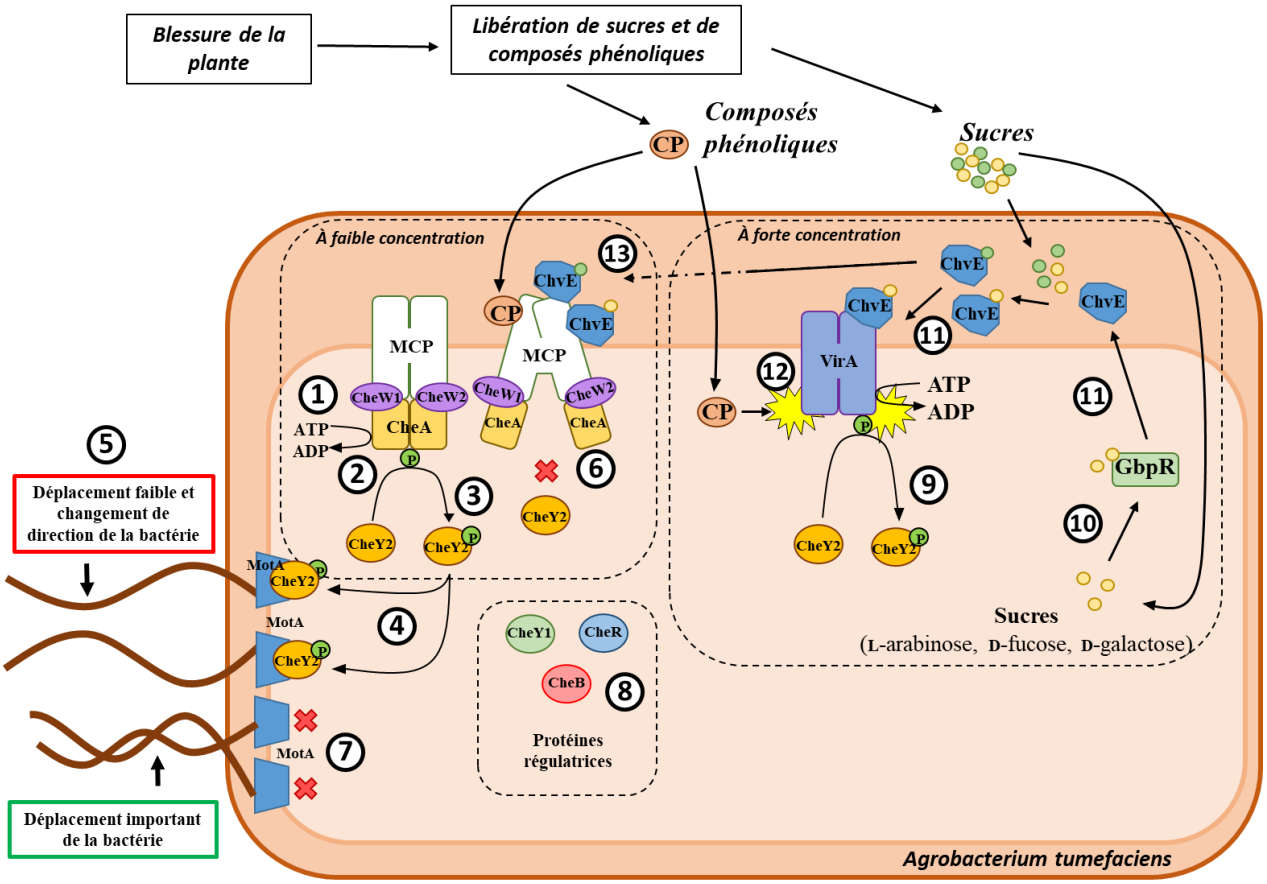


Figure 3: **Représentation visuelle du parcours de la bactérie selon deux mécanismes de chimiotactisme ainsi que du mouvement des flagelles.** A: Représentation du chimiotactisme chez *E.coli* (CCW: CounterClockWise = sens antihoraire = déplacement rectiligne; CW ClockWise = sens horaire = déplacement perturbé/changement de direction brutal). B: Représentation du chimiotactisme chez *S. meliloti* et *A. tumefaciens* (CW fast: ClockWise fast = sens horaire rapide = déplacement plus rectiligne; CW slow-down: ClockWise slow-down = sens horaire ralenti = déplacement perturbé et moins rectiligne/changement de direction lent). (Scharf & Schmitt, 2002)

Chez *A. tumefaciens*, une voie de signalisation permet à la bactérie de se rapprocher du site de blessure de la plante. A faible concentration, les molécules chimioattractantes (composés phénoliques et sucres) sont reconnues par divers chimiorécepteurs transmembranaires nommés MCP (methyl-accepting chemotaxis protein) (Guo et al., 2017). Dans le cas où il y a absence de molécules chimioattractantes, une autokinase (CheA) se lie à MCP et cette liaison est stabilisée par la protéine CheW (Figure 4/①). Une fois liée et stabilisée, CheA est capable de s'autophosphoryler lorsqu'elle est dimérique (Figure 4/②). Lorsque CheA est phosphorylée, le phosphate est transféré à la protéine CheY2 qui se fixe à MotA (une protéine motrice des flagelles) et engendre une désynchronisation des flagelles et un ralentissement de la motricité bactérienne (Figure 4/③④) (Deakin, Sanderson, Goswami, & Shaw, 1997). Ce ralentissement va avoir pour conséquence des changements de direction de la bactérie et une augmentation des chances de s'orienter vers des molécules chimioattractantes (Figure 4/⑤). A l'inverse, en présence de molécules chimioattractantes, une interaction entre ces molécules et MCP induit un changement de conformation de MCP empêchant l'autophosphorylation de CheA (Figure 4/⑥). La conséquence étant une resynchronisation des flagelles qui vont engendrer un déplacement important vers la



source chimioattractante (Figure 4/⑦).

Figure 4 : Schéma de la voie de signalisation du chimiotactisme ainsi que sa voie de régulation en cas de fortes concentrations en sucre, acide organique et composé phénolique (Deakin et al., 1997; Guo et al., 2017; Harighi, 2008, 2009; Wright, Deakin, & Shaw, 1998).

Il existe également 4 protéines impliquées dans la régulation de la phosphorylation de CheA : CheY1 qui agit comme une phosphatase permettant de diminuer la quantité de CheA phosphorylée, CheR et CheB qui sont des methyltransferases spécifiques de MCP. L'activité de CheR permet de methyler et CheB de demethyler MCP, pour favoriser ou non la phosphorylation de CheA (Harighi, 2008, 2009) (Figure 4/⑧).

A forte concentration en chimioattractants, le chimiotactisme bactérien est régulé par une deuxième voie de signalisation qui implique notamment la protéine VirA. VirA est, comme CheA, une autokinase transmembranaire. Elle est active et dimérique à pH 5-6 et en présence de composés phénoliques (Brencic, Winans, & Colonization, 2005; Stachel, Nester, & Zambryski, 1986). Une fois active et dimérique, elle est capable de céder son phosphate à CheY2 (Figure 4/⑨), ce qui entraine une diminution du déplacement de la bactérie lors de fortes concentrations en molécules chimioattractantes (Figure 4/④).

De plus, l'activité de VirA est elle-même régulée par la présence de sucres, en particulier en augmentant sa sensibilité pour les composés phénoliques. En effet, la présence de sucres engendre la levée de la répression transcriptionnelle de ChvE par GbpR (Figure 4/⑩⑪) (Doty, Chang, & Nester, 1993). Lorsque ChvE est exprimée, elle se lie à un sucre pour ensuite interagir avec VirA et augmenter sa sensibilité pour les composés phénoliques (Guo et al., 2017) (Figure 4/⑪⑫).

Enfin, VirA et ChvE sont deux protéines indispensables à la transition entre le chimiotactisme et l'activation des protéines de virulence. VirA peut phosphoryler CheY2 pour diminuer le déplacement bactérien (Figure 4/⑨④), mais peut également phosphoryler VirG indispensable à l'activation des protéines de virulence (Cf sous chapitre 3 : Activation des gènes de virulence, page 26) (Guo et al., 2017; He et al., 2009; Hu, Zhao, DeGrado, & Binns, 2013; Kemner, Liang, & Nester, 1997). ChvE augmente la sensibilité de VirA pour les composés phénoliques (Figure 4/⑫) et favorise donc la diminution du déplacement bactérien (Figure 4/⑨④). Mais ChvE entraine également la diminution du déplacement bactérien en induisant le changement conformationnel de MCP et donc la non-dimérisation de CheA (Figure 4/⑬⑥) (Guo et al., 2017).

2. *Attachement aux cellules végétales*

Suite au déplacement de nombreuses bactéries pathogènes au niveau du site de blessure, celles-ci se retrouvent à proximité des molécules libérées par la blessure de la plante. La détection par la bactérie de certaines molécules induit la sécrétion d'exopolysaccharides engendrant la formation de

biofilms pour permettre aux bactéries de créer une interaction avec la plante et de se retrouver en conditions favorables pour résister aux molécules antibactériennes et inhibitrices de croissance (Melene A. Thompson, Onyeziri, & Fuqua, 2018). Différents exopolysaccharides sont secrétés dont la cellulose, le succinoglycan et le β-1,2-glucan cyclique. La cellulose établie un lien direct entre la bactérie et les cellules de la plante tandis que le succinoglycan a pour effet de stabiliser les interactions plante-bactérie et d'augmenter la viscosité du milieu. Enfin, les molécules dues à la blessure de la plante engendrent un changement osmotique néfaste pour la bactérie. La production de β-1,2-glucan cyclique permet à la bactérie de s'adapter à ce changement osmotique et d'établir une interaction avec la plante et plus particulièrement avec ses lectines (Breedveld & Miller, 1994; Brencic, Winans, et al., 2005). Les gènes impliqués et mis en avant dans la synthèse et l'expression de ces exopylosaccharides sont les gènes *exo* et les gènes *chvA* et *chvB* (Halder, Banerjee, & Bandopadhyay, 2017; Melene A. Thompson et al., 2018). Comme évoqué précédemment, l'expression de ces gènes est permise grâce aux molécules libérées par la plantes lorsque celle-ci est blessée mais surtout grâce à la libération d'acide organique puisqu'il entraine une baisse du pH. Cette diminution du pH (Figure 5/①) engendre une dénaturation du répresseur ExoR (Figure 5/②) et induit l'autophosphorylation de ChvG, une autokinase senseur transmembranaire (Figure 5/③). Le phosphate est ensuite transféré à la protéine ChvI, un activateur transcriptionnel (Figure 5/④) (Li et al., 2002; Wu, Lin, Shaw, & Lai, 2012). Cette dernière régule la synthèse de différentes protéines responsables de la synthèse d'exopolysaccharides et l'expression de la protéine VirG (Figure 5/⑤) (Li et al., 2002).

3. *Activation des gènes de virulence*

L'expression de la protéine VirG, induite par la diminution du pH, joue un rôle important lors de l'infection de la plante par la bactérie. En effet, cette protéine fait partie d'une voie de signalisation mettant en jeu le couple protéique VirA/VirG qui permet d'activer la transcription des gènes de virulence. VirA est une kinase transmembranaire, active à bas pH (5-6) (Figure 5/①), qui doit se lier à un composé phénolique (acetosyringone ou ferulate) cytoplasmique (Figure 5/⑥) pour être capable de s'autophophoryler en présence d'ATP (Brencic, Angert, & Winans, 2005) (Figure 5/⑦). Le phosphate est ensuite transféré à la protéine VirG (Figure 5/⑧) afin de lui permettre d'activer la transcription des gènes de virulence (Figure 5/⑨). De plus, en présence d'une forte concentration en sucre, la répression de la protéine périplasmique ChvE est levée et est exprimée dans le périplasme (Figure 5/⑩) (Doty et al., 1993). Comme évoqué précédemment, cette protéine peut se lier aux sucres et augmenter fortement la sensibilité de la kinase VirA pour les composés phénoliques (Figure 5/⑪⑫). Ainsi cette interaction augmente significativement la vitesse

d'activation des protéines de virulence. (Guo et al., 2017)

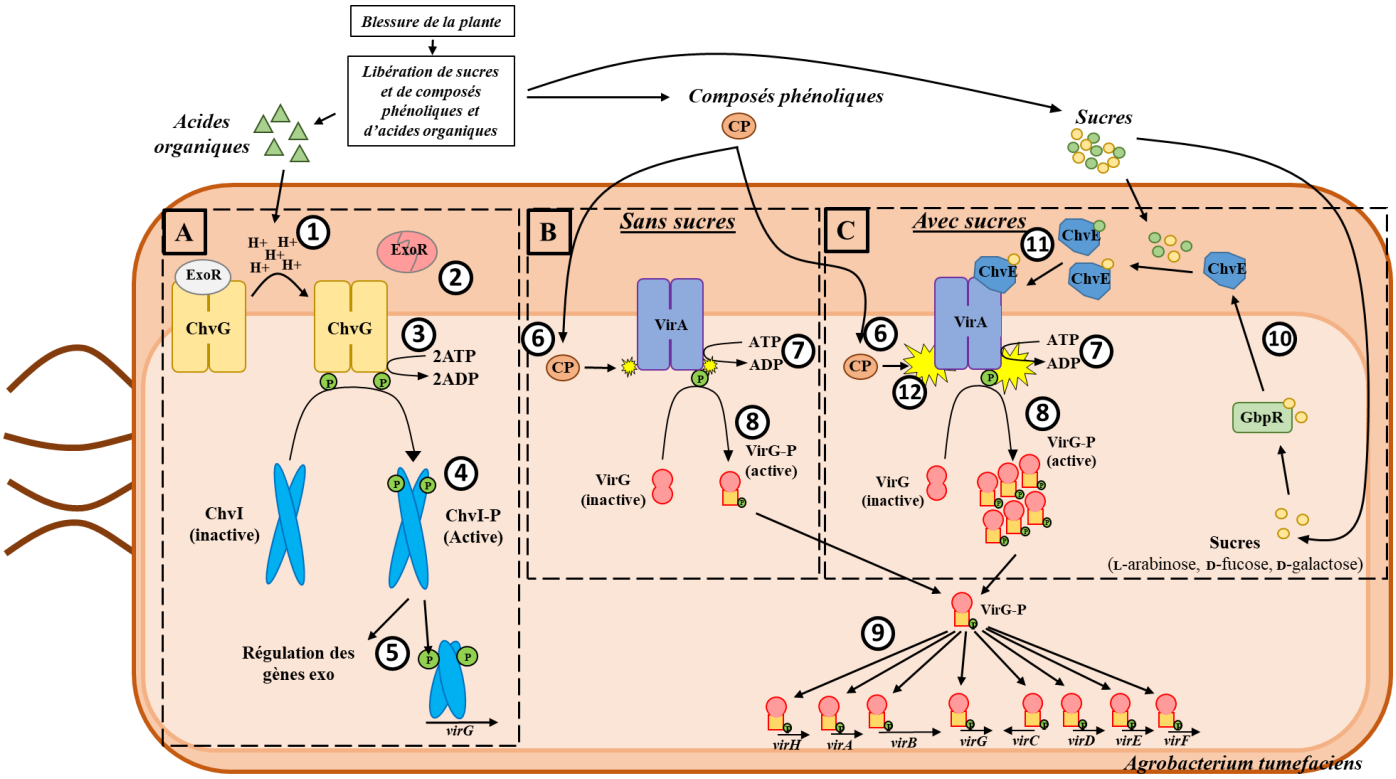


Figure 5 : Schéma représentatif des différentes voies de signalisation permettant l'activation de la transcription des gènes de virulence *vir*. A : Voie de signalisation qui en présence d'acide organique permet la transcription des gènes exo et du gène *virG*. **B** : Voie de signalisation engendrant l'activation de la protéine VirG. **C** : Sensibilité de la VirA pour les composés phénoliques augmentés en présence de sucres.

4. *Transfert de l'ADN-T au génome de la plante*

La galle du collet se caractérise par le développement de tumeurs au site d'infection et par un détournement du métabolisme de la plante. Ceci est rendu possible grâce au transfert d'une portion d'ADN bactérien nommé ADN-T pour « ADN de transfert », située sur le plasmide de virulence Ti. Les protéines de virulence exprimées possèdent un rôle dans la réplication, le transfert et l'intégration de l'ADN-T dans le génome de la plante. L'ADN-T est localisé sur le plasmide Ti et est délimité par deux séquences d'une vingtaine de paires de bases nommées LB et RB pour « Left Border » et « Right Border » (Lacroix & Citovsky, 2019; Peralta & Ream, 1985). Cet ADN est reconnu par le complexe protéique VirD1/D2. VirD1 est une hélicase et sépare les deux brins d'ADN, pendant que VirD2 le clive avec son activité endonucléase. Une fois l'ADN simple brin clivé et libéré, la protéine VirD2 reste fixée covalemment à son extrémité 5' pour former le « brin-T » (Figure 6/①).

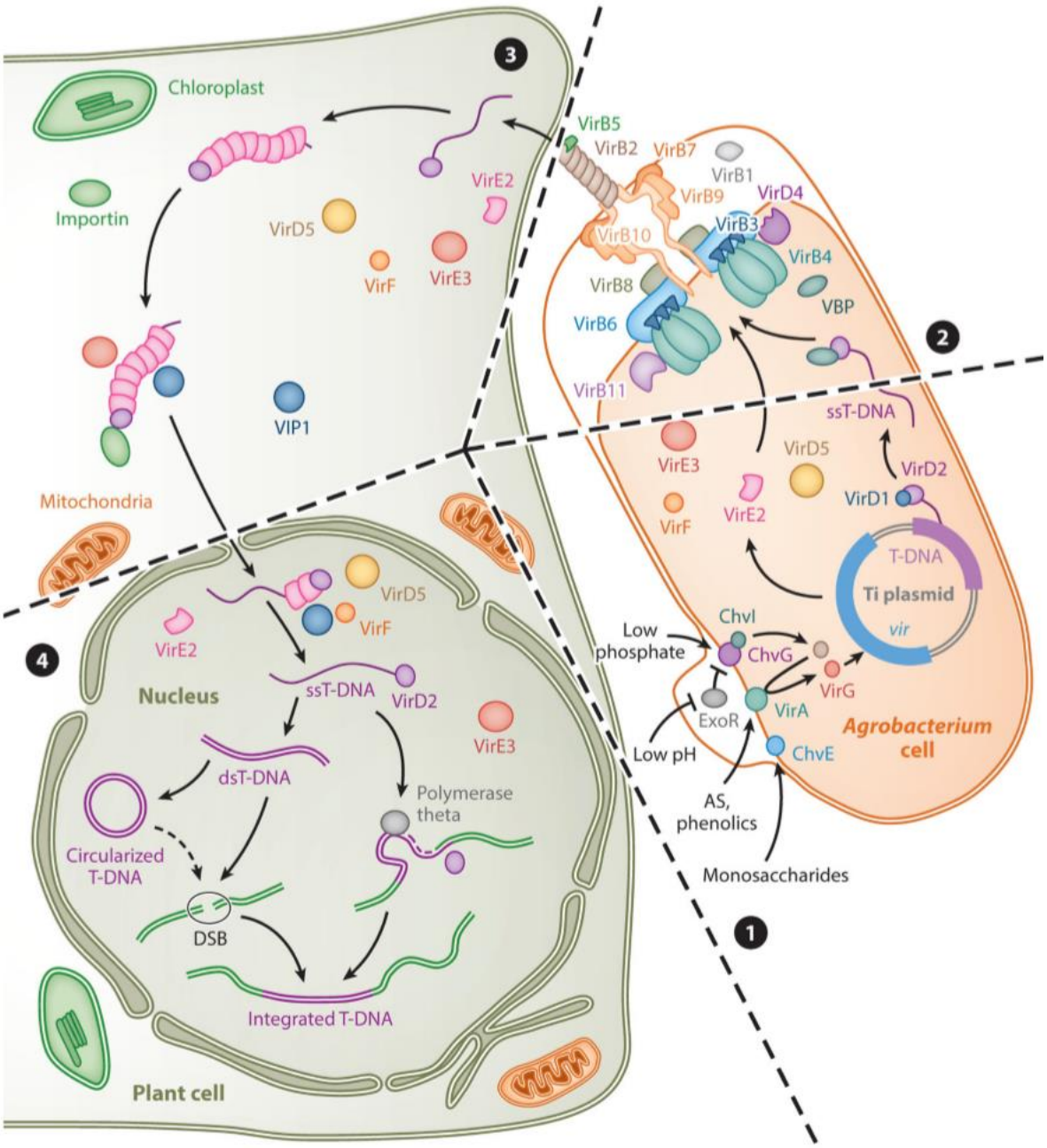


Figure 6 : Représentation schématique des différentes étapes de l'activation des gènes de virulence au transfert de l'ADN-T dans le génome de la plante. ① : Activation des protéines de virulence, réplication de l'ADN-T et adressage au transporteur. ② : Reconnaissance et transfert de l'ADN-T ainsi que les protéines VirE2, VirE3, VirF et VirD5. ③ : Protection de l'ADN-T et adressage au noyau. ④ : Liberation de l'ADN-T, convertion en ADN double brin et intégration au génome par le biais de DSB (Lacroix & Citovsky, 2019).

Cet ADN, et plus particulièrement VirD2 qui y est fixé à son extrémité, interagit avec la protéine VirD4 en vue de son transfert vers la cellule végétale par le biais d'un système de sécrétion de type IV (SST4) (Figure 6/②). Ce système de transport met en jeu 12 protéines (VirB1 à VirB11 et VirD4) réparties en 4 groupes. VirD4 fait partie du groupe T4CP (Type IV coupling protein), et possède une activité ATPasique permettant l'interaction avec le complexe-T ainsi que l'initiation de son transfert vers la cellule végétale. L'IMC (inner membrane complex) est formé des protéines

VirB3, B4, B6, B8 et B11. VirB4 et B11 ont une activité ATPasique fournissant l'énergie nécessaire à la formation du complexe transmembranaire multiprotéique et au transfert du brin-T. Les protéines VirB3, B6 et B8 forment un complexe transmembranaire permettant de traverser la membrane interne. L'OMC (outer membrane complex) est constituée des protéines VirB7, B9 et VirB10, un complexe capable de traverser le périplasme. Ce dernier est lié au pilus composé de VirB2, utile à la formation du canal, et de VirB5 dont la fonction est d'établir une interaction entre le pilus et la cellule cible. Pour finir, la protéine VirB1 est une transglycosidase qui dégrade le peptidoglycane pour permettre la formation du complexe T4SS (Christie, Whitaker, & González-Rivera, 2014; Zupan, Hackworth, Aguilar, Ward, & Zambryski, 2007) (Figures 6/②).

Lors du transfert du brin-T, les protéines VirD5, VirE2 et VirF sont également transférées dans la cellule végétale grâce à la présence d'une séquence consensus en C-terminal (Vergunst et al., 2005) (Figure 6/③). La protéine VirE2 est une SSB-like (Single Stand Binding protein) qui a pour but de protéger le brin-T de l'ADN des dégradations éventuelles par les nucléases de la cellule. En se liant à celui-ci, le complexe formé est nommé complexe-T. L'import du complexe-T dans le noyau (Figure 6/④) est facilité par des importines se liant aux domaines NLS (nuclear localization signal) situé sur les protéines VirE2 et VirD2. De plus VirE2 peut également interagir avec VIP1 (VirE2-Interacting Protein 1) pour aider à l'adressage du complexe-T dans le noyau. Les protéines exprimées et transférées VirD5, VirE3 et VirF ont un rôle dans la préparation de l'ADN-T avant son intégration dans le génome de la plante (Figure 6/③). Une fois dans le noyau, VirF est stabilisé par VirD5 et interagit avec les protéines VIP1/VirE2. Cette reconnaissance engendre la formation d'un nouveau complexe avec les protéines Skp1-Cul1-F-box (SCF), complexe de type ubiquitine ligase. Ceci provoque l'ubiquitination des protéines VIP1 et VirE2 et leur adressage au protéasome qui va assurer leur dégradation. L'ADN se retrouve alors sous la forme simple brin dans le noyau.

Enfin, l'ADN-T est intégré dans le génome de la plante, de façon aléatoire mais les conditions d'intégration sont encore mal connues (S. I. Kim, Veena, & Gelvin, 2007; Lacroix & Citovsky, 2013) (Figure 6/④). Il est en effet possible que l'ADN-T soit intégré en tant que simple brin dans le génome, ou transformé en double brin linéaire ou circulaire avant d'être intégré au niveau d'une région dégradée de l'ADN génomique. Même si ces différentes hypothèses ont été émises et que le mécanisme de cette étape est encore peu compris, de récentes études portent à croire qu'avant l'intégration de l'ADN-T, celui-ci est dans un premier temps converti en ADN double brin pour ensuite être intégré par recombinaison non homologue dans des zones d'ADN génomique double brin dégradée ou DSB (Double-Strand Break) (Figure 6/④) (Lacroix &

Citovsky, 2019).

5. *Expression des gènes de l'ADN-T*

Le but du transfert de l'ADN-T au génome de la plante est de faciliter le développement des bactéries du genre *Agrobacterium* au sein de l'hôte. Afin d'améliorer les conditions de croissance de la bactérie et le transport de nutriments, l'ADN-T contient des gènes qui permettent la synthèse de phytohormones engendrant une perturbation de l'équilibre hormonal. Cette perturbation a un impact sur la prolifération, l'élongation et la différenciation cellulaires chez la plante, conduisant à une multiplication anarchique des cellules transformées et donc à la formation de tumeurs. Les gènes de l'ADN-T directement impliqués dans la synthèse de phytohormones, sont les gènes *iaaM*, *iaaH* et *ipt* (Veselov et al., 2003; Zambryski, Tempe, & Schell, 1989). IaaM est une tryptophane mono-oxygénase qui transforme le tryptophane en indole-3-acétamide (IAM) par décarboxylation oxydative. L'indole-3-acetamide est ensuite hydrolysé en 3-indole-acetique (AIA) par la protéine IaaH. L'AIA est une substance naturelle de la famille des auxines qui favorise la croissance radiale et la formation de cals. Ensuite, l'IPT est une isopentényl transferase qui catalyse la synthèse d'isopentényladénosine-5'-monophosphate à partir de dimethylallyl-pyrophosphate et d'adénosine-5'-monophosphate (AMP) (Figure 7).

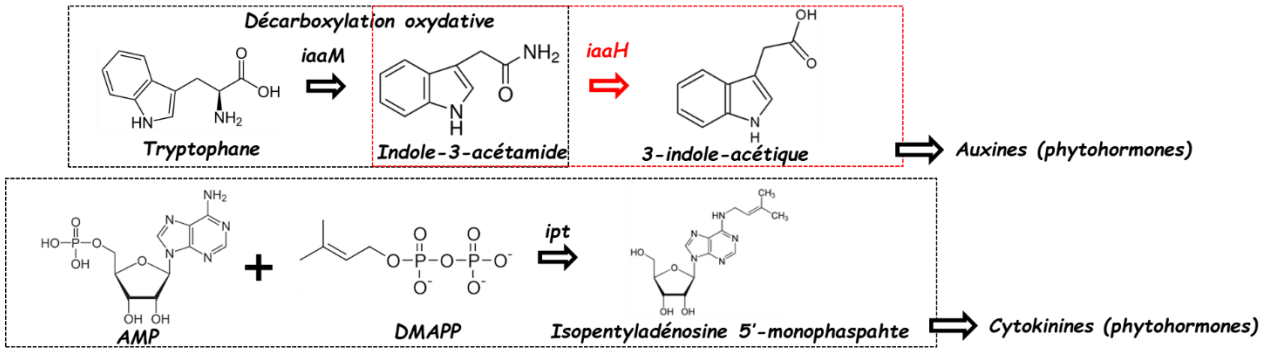


Figure 7 : Représentation des réactions engendrées par la présence des enzyme IaaM, IaaH et IPT amenant à la dérégulation hormonale due à la formation de phytohormones (Veselov et al., 2003; Zambryski et al., 1989).

La molécule produite est une cytokinine qui engendre une stimulation de la division cellulaire et du métabolisme des cellules ainsi que de la différenciation vasculaire. De plus, la forte concentration en auxine et la différenciation vasculaire stimule l'activité de l'acide 1 aminocyclopropane-1-carboxilique (ACC) synthase. Cette enzyme permet la synthèse de l'ACC qui est ensuite dégradé en éthylène par l'ACC oxydase. L'impact de l'éthylène à plus forte concentration est multiple. En effet, il permet à la fois d'améliorer la vascularisation par

restructuration du xylème (Wächter et al., 2003), et de stimuler la synthèse d'acide abscissique (ABA) qui réduit la transpiration au sein des tumeurs et évite leur déshydratation. Enfin, il favorise le transport d'eau et de nutriments depuis les racines jusqu'aux tumeurs en régulant l'expression des gènes impliqués dans l'osmorégulation (Efetova et al., 2007).

En plus de certains gènes engendrant la formation de tumeurs, certains sont quant à eux responsables de la synthèse de molécules carbonées nommées opines utilisées spécifiquement par les bactéries. Ces gènes sont localisés sur l'ADN-T et transférés dans le génome de la plante. Par conséquent, c'est la plante qui synthétise les opines utilisées par la bactérie. Il existe différents types d'opines synthases et également différents types d'opines produites. Leur présence engendre un avantage spécifique aux bactéries pouvant les utiliser.

c. ***Les opines***

Comme évoqué précédemment, les opines sont produites par la plantes dans les tumeurs, sont utilisées comme nutriments par la bactérie, en étant une source de carbone et/ou d'azote, ou comme molécules signal. Certaines opines peuvent également être qualifiées d'opines conjugatives et initier le transfert conjugatif de plasmides, dont le plasmide de virulence, et ainsi participer à la propagation de la virulence (Beck von Bodman, Hayman, & Farrand, 1992).

i. Les différents types d'opines et de plasmides Ti

Il existe aujourd'hui plus d'une vingtaine d'opines différentes produites dans les tumeurs par les plantes infectées par les bactéries du genre *Agrobacterium tumefaciens.* Ces opines peuvent être classées en deux groupes différents.

Le premier groupe comprend les opines pouvant être issues de la condensation entre un acide aminé et un sucre ou un céto-acide (Figure 8/Orange). Ce groupe comprend les opines de la famille des octopines, des chrysopines, des succinamopines, des nopalines, des mannityl-opines, des leucinopines et des cucumopines. Le deuxième groupe comprend des opines particulières correspondant à des phosphodiesters de sucres (Figure 8/Vert).

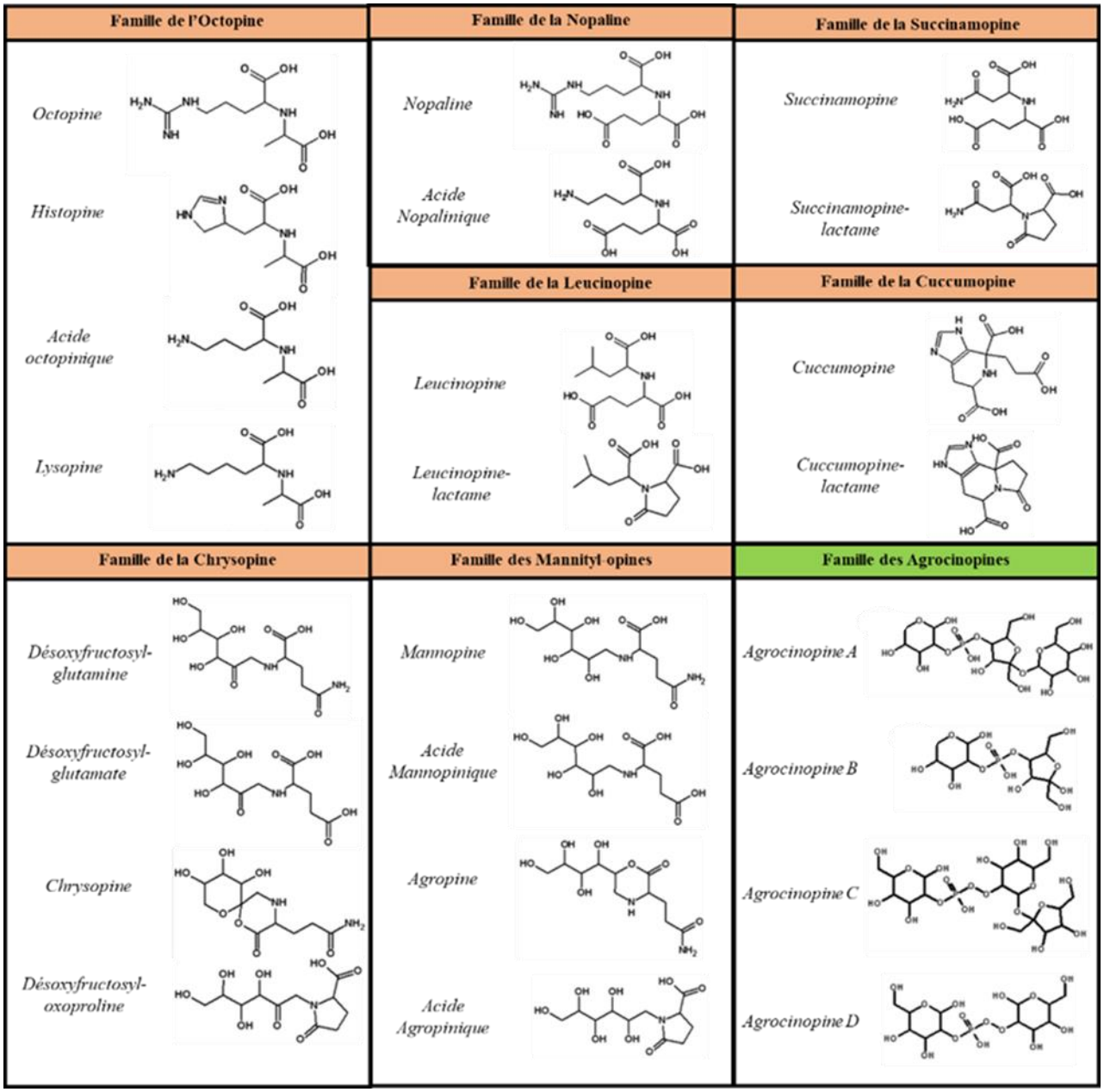


Figure 8 : Représentation des structures chimiques des différentes opines induites par les bactéries pathogènes du genre Agrobacterium. Adapté de (Dessaux & Farrand, 1998)

Toutes les souches du genre *Agrobacterium tumefaciens* n'induisent et n'utilisent pas les mêmes opines. Seules les bactéries possédant les gènes codant pour des opines synthases spécifiques peuvent les utiliser. Par exemple, la souche C58 d'*A. tumefaciens* possède un plasmide de type Nopaline (Tableau 1) et l'ADN-T contient les gènes de la nopaline synthase et de l'agrocinopine synthase qui vont permettre de produire des opines. Une fois produites par la plante, ces opines sont ensuite utilisées par la bactérie grâce aux opérons spécifiques de ces opines. Ces plasmides ont donc été classés en fonction des opines produites par la plante après le transfert de l'ADN-T (Tableau 1).

Type de plasmide	Opines synthétisées	Plasmide des souches utilisées (possèdant l'opéron *acc* necessaire à l'utilisation des agrocinopines)
Octopine	Octopine, acide octopinique, lysopine, histopine, mannopine, acide mannopinique et agropinique, sulfonopine	
Nopaline	Nopaline, acide nopalinique, **Agrocinopine A+B**	**pTiC58**
Agropine	Agropine, mannopine, acide mannopinique et agropinique, **Agrocinopine C+D**, leucinopine, leucinopine lactame, L,L succinamopine	**pTiBo542**
Succinamopine	D,L succinamopine, succinamopine lactame, succinamopine	
Lippia	Agrocinopine C+D	
Chrysopine/succinamopine	Chrysopine, composé de type Amadori, L,L succinamopine, L,L leucinopine	
Chrysopine/nopaline	Chrysopine, composé de type Amadori, nopaline	
Grapevine III	Vitopine, rideopine	**pTiS4**

Tableau 1 : Classification des plasmides Ti adaptée de (Dessaux & Farrand, 1998). En vert, les trois types de plasmides des souches utilisées dans cette étude.

Les souches bactériennes étudiées durant ma thèse possèdent des plasmides de type nopaline pour la souche C58 (pTiC58), comprenant les gènes impliqués dans la synthèse de la nopaline, de l'acide nopalinique et des agrocinopines A et B. La souche Bo542 contient un plasmide Ti de type agropine (pTiBo542) responsable de la synthèse des mannityl-opines, des agrocinopines C et D, de la leucinopine et de la leucinopine-lactame. Enfin, la souche S4 possède un plasmide de type vitopine (pTiS4) permettant la synthèse de vitopine. Ce type de plasmide est spécifique aux bactéries de la famille *A. vitis* (Burr 2018).

Les souches C58 et Bo542 d'*A. tumefaciens* sont donc capables de faire produire à la plante des agrocinopines. Sur leur plasmide Ti sont également présents les gènes permettant l'utilisation des opines par la bactérie et plus particulièrement l'opéron *acc* pour les agrocinopines. La souche S4 d'*A. vitis* possède également cet opéron.

Les différentes agrocinopines varient entre elles selon leur ose phosphorylé, arabinose-2-phoshate pour les agrocinopines A et B et glucose-2-phosphate pour les agrocinopines C et D. Elles se différencient par la présence du sucre lié au phosphate, sucrose pour l'agrocinopine A et fructose pour l'agrocinopine B. En ce qui concerne l'agrocinopine C et D, il s'agit respectivement d'un sucrose et d'un glucose (Figure 9).

Souche	C58		Bo542	
Agrocinopine synthétisée				
Molécule « précurseur »	**Agrocinopine A** Arabinose-2-phosphate-fructose-glucose	**Agrocinopine B** Arabinose-2-phosphate-fructose	**Agrocinopine C** Glucose-2-phosphate-glucose-fructose	**Agrocinopine D** Glucose-2-phosphate-glucose

Figure 9 : Représentation des structures chimiques des agrocinopines A, B, C et D pouvant être synthétisées grâce aux gènes *acs* des souches C58 et Bo542 du genre *Agrobacterium*. En violet est représenté le motif arabinose-2-phosphate, en orange le fructose, en vert le glucose et en bleu de glucose-2-phosphate.

ii. Processus d'utilisation des agrocinopines

1. *Opéron acc*

Les agrocinopines sont produites par la plante grâce à l'agrocinopine synthase (*acs*) (Figure 10/②) et excrétées dans l'apoplasme au niveau des tumeurs. Par la suite, les agrocinopines se retrouvent dans le périplasme de la bactérie par diffusion à l'aide de porines non spécifiques (Figure 10/③). Les agrocinopines sont des molécules spécifiques uniquement utilisables par des bactéries possédant un opéron nommé *acc* (Figure 10/①) et généralement localisé sur le plasmide Ti. Cet opéron contient 8 protéines réparties en 3 groupes. Le premier groupe chargé du transport des agrocinopines est formé de 5 protéines. La première, AccA, est une PBP (periplasmic binding protein) qui reconnait les agrocinopines principalement grâce à leur motif pyranose-2-phosphate, pour ensuite l'adresser à un transporteur ABC de type II (AccBCDE) formé des 4 autres protéines (Figure 10/④) (El Sahili et al., 2015). Les protéines AccB et AccC fixent et hydrolysent l'ATP tandis que les protéines AccD et AccE sont des protéines transmembranaires qui forment un canal (H. Kim & Farrand, 1997). Le deuxième groupe est composé des enzymes liées au catabolisme des agrocinopines et notamment l'enzyme AccF qui catabolise le clivage des agrocinopines lorsqu'elles sont présentes dans le cytoplasme (Figure 10/⑤). La phosphatase AccG peut ensuite transformer le pyranose-2-phosphate en pyranose et phosphate (Figure 10/⑥). Enfin, le troisième groupe est constitué d'un répresseur transcriptionnel nommé AccR impliqué dans la régulation de l'expression de nombreux opérons dont les opérons *acc*, *arc*, *rep* et *rct* (Figure 10/⑧). L'opéron *arc* contient plusieurs gènes dont *traR* responsable de la détection des signaux de « Quorum Sensing ». Les opérons *rep* et *rct* sont chargés de la réplication des plasmides Ti et At, respectivement.

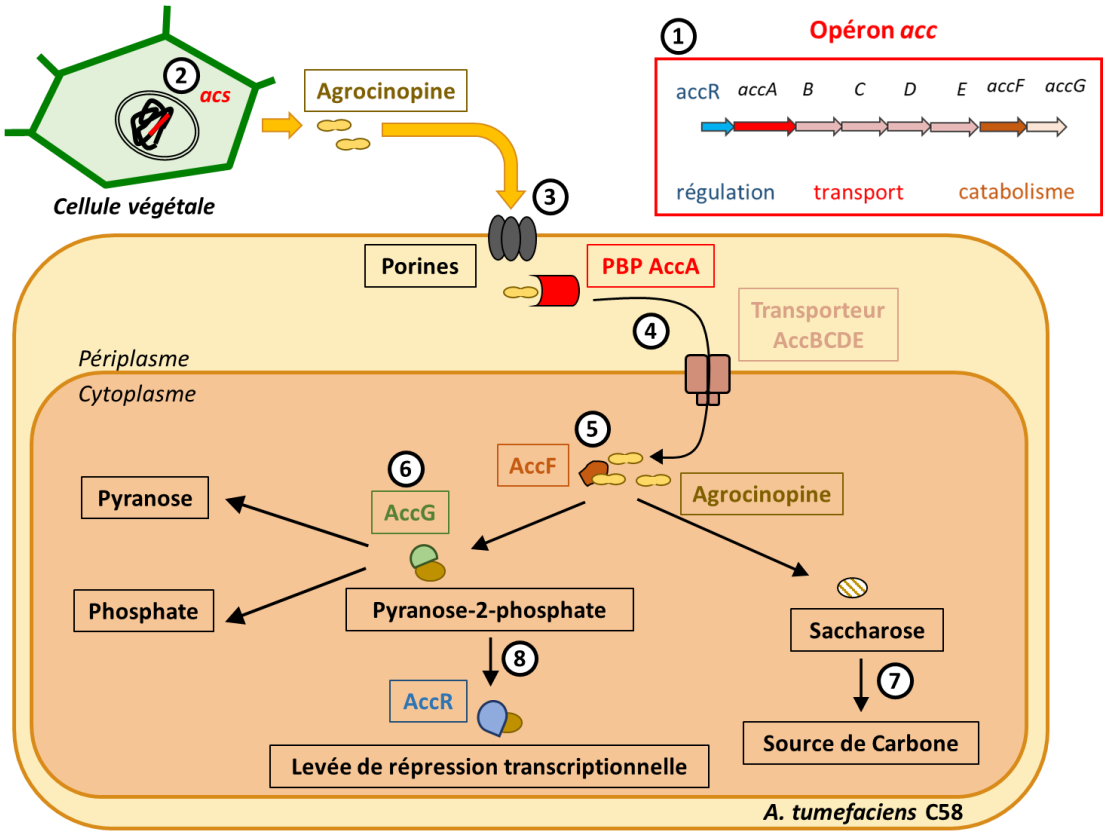


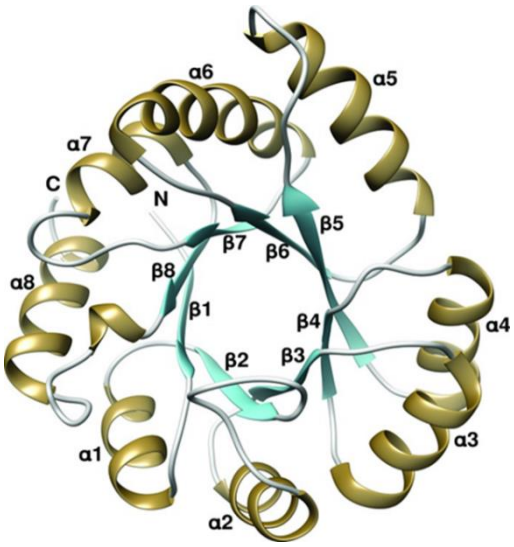
Figure 10 : Représentation de l'opéron *acc* nécessaire à l'utilisation des agrocinopines. Schéma représentant la localisation bactérienne des différentes protéines de cet opéron.

2. *Rôles des agrocinopines*

Une fois internalisées, les agrocinopines peuvent avoir différents rôles au sein de la cellule. L'hydrolyse de l'agrocinopine est catalysée par les enzymes AccF et AccG et engendre la formation de sucres qui servent de source de carbone à la bactérie (Figure 10/⑦). Cependant, lorsque l'agrocinopine est transformée en pyranose-2-phosphate (P2P), cette molécule peut interagir avec le répresseur transcriptionnel AccR. La formation du complexe AccR-P2P engendre la levée de la répression de certaines protéines impliquées dans la communication bactérienne appelée « Quorum Sensing ».

3. *Les Glycérophosphodiesterases*

Les glycérophosphodiesterases sont connues pour être des enzymes impliquées dans la transformation d'un glycérophosphodiester et d'eau en alcool et glycérol-3-phosphate. Cette famille d'enzymes est impliquée dans diverses fonctions biologiques et notamment le métabolisme des phospholipides, les modifications du cytosquelette, la différenciation des neurones moteurs et la synthèse de facteurs de virulence (Corda et al., 2014; Denloye, Dalal, & Klemba, 2012; Jackson et al., 2007; Rao, Bonanno, Burley, & Swaminathan, 2006; Shi, Liu, An, & Liang, 2008)



Les protéines de la famille des glycérophosphodiesterases ont la particularité structurale d'adopter un repliement triosephosphate isomérase (tim) barrel (Corda et al., 2014). Ce repliement extrêmement conservé est présent dans approximativement 10% des enzymes et est constitué d'un enchainement de 8 brins β formant un feuillet entouré d'hélices α : $(\beta/\alpha)_8$ (Figure 11) (Goldman, Beatty, & Landweber, 2016).

Figure 11 : Représentation en ruban d'un repliement caractéristique en triosephosphate isomérase (tim) barrel (Perveen, Rashid, & Papageorgiou, 2016).

Les glycérophosphodiesterases connues ont pour fonction de catalyser le clivage d'une liaison phosphodiester grâce à une molécule d'eau activée (Shi et al., 2008). Le mécanisme réactionnel de ces enzymes est toujours incertain à ce jour. Cependant une étude de 2008 a suggéré un mécanisme réactionnel en deux étapes impliquant deux histidines, un ion divalent et une molécule d'eau (Shi et al., 2008) (Figure 12 A et B). Cette hypothèse est appuyée par la conservation de ces histidines dans différentes séquences de glycérophosphodiesterases (Figure 12 C).

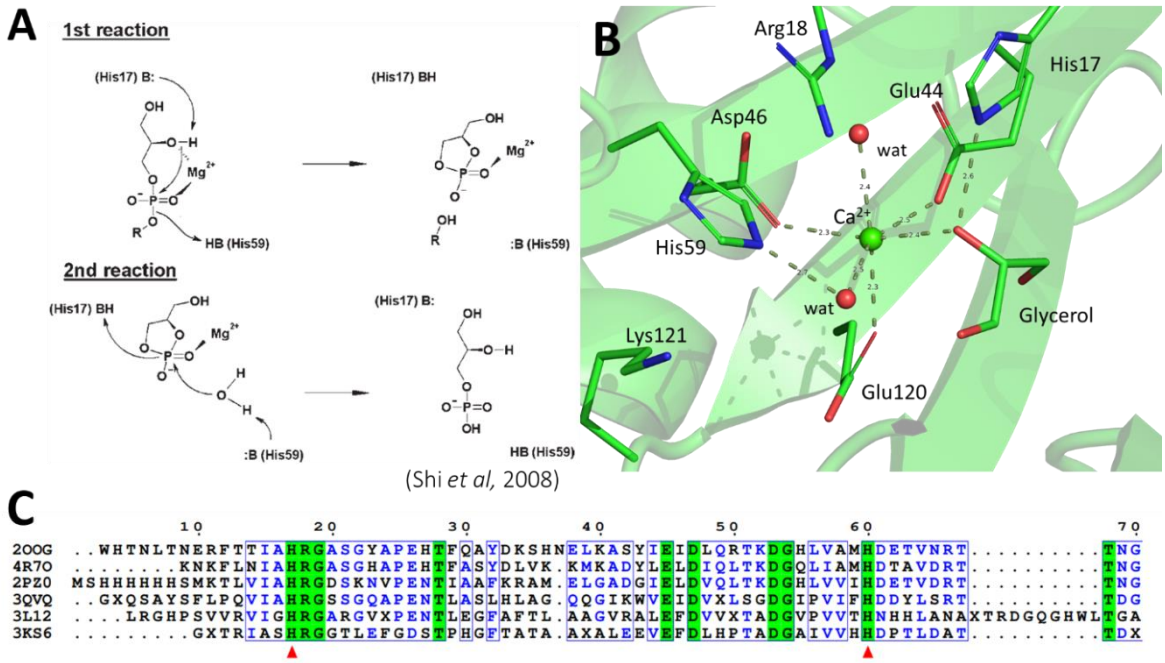


Figure 12 : Représentation du mécanisme réactionnel proposé (**A**), une visualisation du site catalytique de la glycérophosphodiesterase de *Thermoanaerobacter tengcongensis* en complexe avec du glycérol (**B**) (**code PDB: 2PZ0**) ainsi qu'un alignement de séquences comparant les 70 premiers résidus de différentes glycérophosphodiesterases. **A :**

Ce mécanisme réactionnel propose l'implication des histidines 17 et 59 en tant qu'acide et base respectivement, coordonnées par un ion manganèse dans la catalyse du glycérol-1-phosphate. **B :** Structure en 3 dimensions du site catalytique en présence du produit de réaction (glycérol) et d'un ion calcium (code PDB: 2PZ0). **C :** Alignement de séquences de différentes glycérophosphodiesterases mettant en avant (triangles rouges) la conservation des histidines 17 et 60. Le surlignage vert met en avant les résidus identiques, les résidus en bleu sont les résidus similaires et enfin les portions de séquence similaires encadrés en bleu. (Codes PDB : 2OOG (*Staphylococcus aureus*), 4R7O (*Bacillus anthraci*), 2PZ0 (*Thermoanaerobacter tengcongensis)*, 3QVQ (*Oleispira antarctica*), 3L12 (*Silicibacter pomeroyi*) et 3KS6 (*Agrobacterium tumefaciens*)

Lorsque l'on compare le substrat de la glycérophosphodiesterase de *Thermoanaerobacter tengcongensis* avec le substrat agrocinopine D d'AccF d'*A. tumefaciens* Bo542, on remarque quelques similitudes, laissant à penser que ces deux protéines pourraient avoir un mécanisme catalytique semblable (Figure 13).

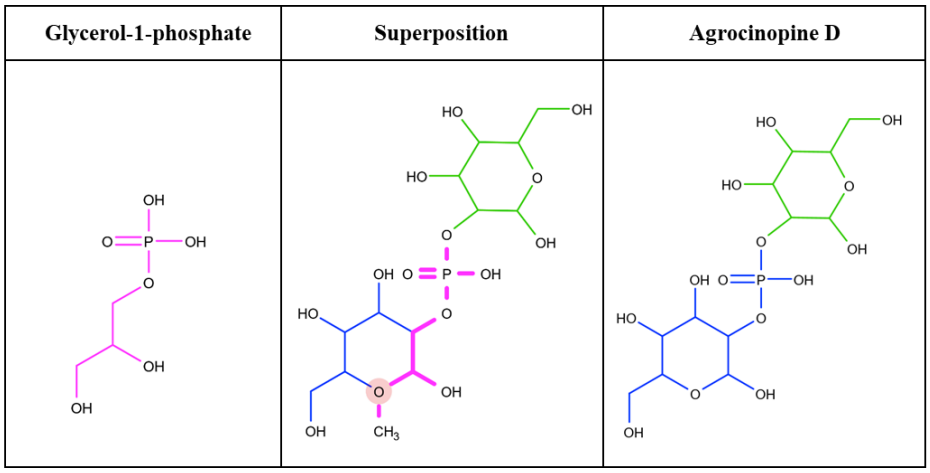


Figure 13 : Superposition du glycérol-1-phosphate (substrat de la glycerophosphodiesterase de *Thermoanaerobacter tengcongensis*) avec l'agrocinopine D (substrat de la phosphodiesterase AccF d'*A.tumefaciens* Bo542), permettant de mettre en évidence la similitude de ces deux substrats.

La réalisation d'alignements et d'analyses de séquences a par la suite permis de montrer que la protéine AccF était une phosphodiesterase dont le mécanisme catalytique pourrait correspondre à celui de la glycérophosphodiesterase de *Thermoanaerobacter tengcongensis*. Ces travaux de recherche qui ont été réalisés durant ma thèse ont été plus amplement développés et décrits dans le chapitre 2 de ce manuscrit page 85.

a. La phosphodiesterase AccF de l'opéron *acc*

Les agrocinopines entières ne sont pas utilisées telles-quelles par les bactéries et nécessitent une maturation pour servir de source de carbone ou d'effecteurs au répresseur transcriptionnel AccR (El Sahili et al., 2015). Les agrocinopines présentes dans le cytoplasme sont donc reconnues par une phosphodiesterase nommée AccF (Figure 10/⑤ page 35). Protéine cytoplasmique d'environ 30 kDa, cette enzyme est chargée de transformer les agrocinopines en pyranose-2-phosphate et en sucrose (H. Kim & Farrand, 1998). La transformation des agrocinopines joue un rôle primordial dans le partage de la virulence, dans la communication bactérienne et dans le développement des

bactéries utilisant les agrocinopines.

iii. Quorum sensing

1. Principe

Le quorum sensing est un moyen de communication intercellulaire. Il régule la synthèse de molécules ainsi que la transcription de certains gènes, en fonction de la concentration de la population bactérienne. Ce type de communication est utilisé pour l'adaptation ou la survie des bactéries en environnements et en milieux complexes. Ainsi il est possible pour une même communauté bactérienne de communiquer et de synchroniser la transcription de certains gènes.

Ce mécanisme est basé sur la production et la perception de molécules « signal » nommées auto-inducteurs. Chez les bactéries à gram négatif les auto-inducteurs sont des acyl-homoserine lactone (AHL). Mis en évidence pour la première fois en 1970 par Kenneth NEALSON chez *Vibrio fisheri*, le couple LuxI/LuxR est reconnu comme jouant un rôle important dans la synthèse et la perception des molécules auto-inductrices (Haygood & Nealson, 1985; Turovskiy, Kashtanov, Paskhover, & Chikindas, 2007). LuxI est une acyl-homoserine-lactone synthase (AHLs) et LuxR un activateur transcriptionnel qui régule la transcription des gènes (Fuqua, Winans, & Greenberg, 1994). Selon les genres ou espèces bactériennes, les auto-inducteurs produits sont différents, le but étant que les bactéries puissent reconnaitre spécifiquement les signaux de leur propre espèce. Les AHL sont composées d'un noyau lactone conservé et d'une chaine carbonée variable (Figure 12) (Welsh & Blackwell, 2016).

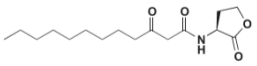
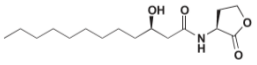
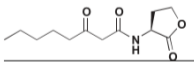
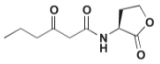
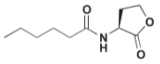
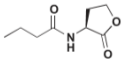
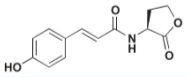
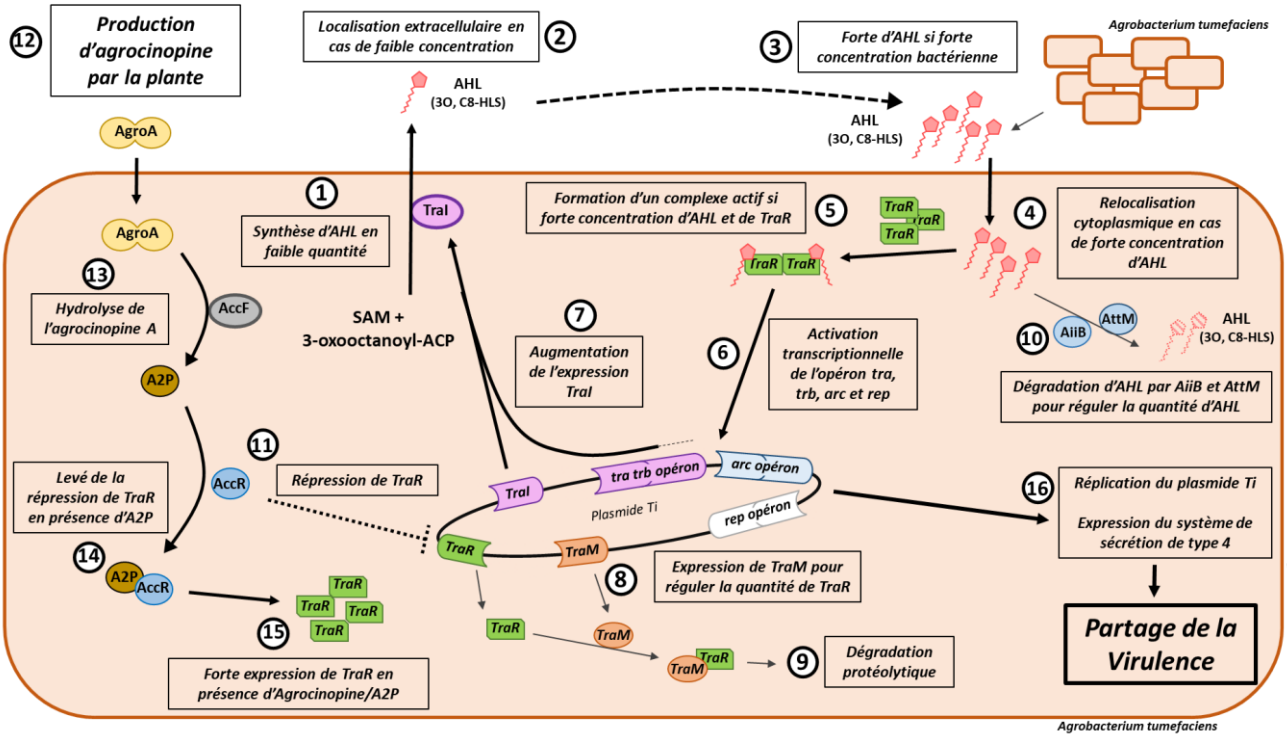
AHL	Bacteria	AHL synthase/receptor
	*Pseudomonas aeruginosa* *Pseudomonas putida*	LasI/LasR PpuI/PpuR
	*Acinetobacter baumannii*	AbaI/AbaR
	*Agrobacterium tumefaciens*	TraI/TraR
	*Vibrio fischeri* *Pectobacterium carotovora*	LuxI/LuxR ExpI/ExpR
	*Chromobacterium violaceum* *Burkholderia cenocepacia*	CviI/CviR CepI/CepR
	*Pseudomonas aeruginosa* *Serratia marcescens*	RhlI/RhlR SpnI/SpnR
	*Rhodopseudomonas palustris*	RpaI/RpaR

Figure 14 : Représentation chimique des différents types d'AHL utilisés pour le quorum sensing chez les bactéries Gram négatives. (Welsh & Blackwell, 2016).

2. *Mécanisme de communication quorum sensing chez Agrobacterium*

Afin de communiquer avec les autres cellules, *Agrobacterium* secrète des AHL nommées OC8HSL. La formation de ces AHL est permise grâce à une acyl-homoserine-lactone synthase LuxI-like nommée TraI (White & Winans, 2007). Cette enzyme catalyse une liaison amide entre la S-adenosylmethionine (SAM) et l'3-oxo-octanoyl-ACP. Après lactonisation de la SAM, une molécule d'3-oxo-C8-HSL (OC8HSL) est formée (Figure 15/①) (Val & Cronan, 1998).

Les AHL sont sécrétées dans le milieu extracellulaire et sont synthétisées en faible quantité non perceptible par les bactéries (Figure 15/②). Cependant si la densité bactérienne atteint un



certain seuil, la concentration de ces molécules devient plus élevée et elles deviennent cette fois perceptibles (Figure 15/③④). Le récepteur et activateur TraR reconnait alors l'AHL et forme un homodimère actif qui se fixe sur les séquences promotrices des gènes cibles via la présence de boites *tra* (Figure 15/⑤⑥⑦) (Zhu & Winans, 1998). Cependant tout ce mécanisme nécessite d'être régulé.

Figure 15 : Représentation simplifiée et schématique du mécanisme de quorum sensing chez le pathogène *Agrobacterium tumefaciens*. L'emplacement et le sens des gènes et opérons présents sur le plasmide Ti n'ont pas été respectés.

3. *La régulation du quorum sensing chez Agrobacterium*

Le quorum sensing est un système de communication finement régulé au cours des différentes étapes de la communication. TraR est la protéine qui interagit avec l'acyl-homoserine lactone (AHL) pour former un dimère actif qui active la transcription de nombreux gènes (Figure 15/⑥). Si l'expression de ces gènes est réalisée au mauvais moment du cycle bactérien, ceci peut avoir un impact sur le fitness de la bactérie, voire même un impact létal. Il est donc extrêmement important que ce mécanisme soit bien régulé afin d'être réalisé au bon moment.

TraR, étant au cœur de ce système de communication, fait partie des éléments les plus ciblés pour la régulation. Pour rappel, le quorum sensing est un mode de communication qui déclenche la transcription de certains gènes ou opérons à partir d'une certaine concentration bactérienne. Ceci permettant de réaliser certaines actions de façon synchronisée et au bon moment. La présence d'antagonistes permet d'éviter son déclenchement lorsque les concentrations de TraR ou d'AHL dans le cytoplasme sont faibles. Ainsi TraM, une protéine anti-activatrice, forme un complexe avec TraR, l'empêchant de se lier aux régions promotrices de l'ADN et déclenchant sa dégradation (Figure 15/⑧⑨) (Luo, Qin, & Farrand, 2000; Zhu & Winans, 1998). A faible concentration en AHL extracellulaire et donc dans des conditions de faible concentration bactérienne, la probabilité qu'une AHL se localise dans le cytoplasme est faible. Cette probabilité croît en fonction de la densité bactérienne. Cependant, il est possible qu'une faible quantité d'AHL se retrouve dans le cytoplasme en risquant d'initier la transcription de certains gènes. Si c'est le cas, elle est alors dégradée par deux lactonases nommées AiiB et AttM (Figure 15/10).

Ces deux mécanismes de régulation sont négligeables et inefficaces en cas de forte concentration en TraR et AHL. La forte concentration en AHL se traduit par une forte concentration bactérienne, alors qu'une forte concentration de TraR implique une levée de répression. En effet, la transcription de *traR* est placée sous le contrôle du répresseur transcriptionnel nommé AccR vu

précédemment et faisant parti de l'opéron *acc* (Figure 15/⑪). La levée de cette répression est permise en présence d'agrocinopine qui est internalisée puis maturée en arabinose-2-phosphate (A2P) ou glucose-2-phosphate (G2P). Ce pyranose-2-phosphate se lie par la suite à AccR et lève la répression transcriptionnelle de *traR*. (Figure 15/⑫⑬⑭⑮)

Ces mécanismes de régulation permettent donc au quorum sensing de se déclencher en présence d'agrocinopine et d'une forte concentration en AHL, nécessitant donc une localisation bactérienne dans la tumeur de la plante et une densité bactérienne élevée. Suite à cela, les gènes cibles et plus particulièrement les opérons *arc*, *rep* et *tra* et *trb* sont exprimés.

4. *Propagation de la virulence*

Les opérons et protéines qui sont exprimés suite à l'activation transcriptionnelle servent essentiellement au partage de la virulence et plus particulièrement à la réplication et au transfert du plasmide Ti. L'opéron *arc* (agrocinopine regulation of the conjugation) possède 5 gènes dont l'activateur transcriptionnel TraR. Les rôles des autres protéines de cet opéron n'ont pas encore été clairement identifiés. L'opéron *rep* (replication) est composé de 3 gènes et code pour des protéines impliquées dans la réplication et la stabilisation de la copie du plasmide Ti (Pappas, 2008).

Les gènes codant pour les protéines *tra* sont répartis en deux opérons : *traAFBH* et *traCDG*. Les protéines TraA, TraC et TraD forment le relaxosome nécessaire au transfert de la copie du plasmide Ti. TraF est requise pour la formation de la piline alors que TraG est une protéine de couplage permettant d'établir une liaison entre le relaxosome et le SST4. Le SST4, indispensable au transfert du plasmide Ti, est formé de onze protéines allant des gènes *trbB* à *trbI*. (Cho & Winans, 2007) (Figure 15/⑯).

d. *L'import et la transformation des opines*

Le sujet principal de mes travaux de thèse concerne la spécificité d'import et la maturation des agrocinopines et dérivés chez le pathogène *Agrobacterium tumefaciens*. Durant cette étude, deux types de protéines ont été mis en avant : Les PBP (periplasmic binding protein) et les phosphodiesterases. C'est pour cette raison que les deux parties suivantes seront consacrées à la description de la famille de ces protéines et plus généralement à l'import et la transformation des opines.

i. L'import des opines

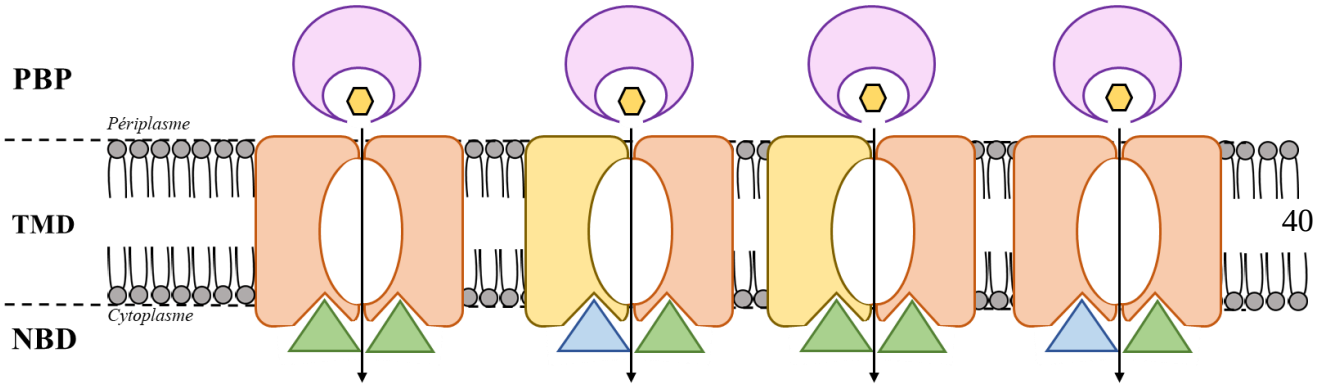
Une fois synthétisées, les opines sont utilisées par la bactérie pathogène. Le mécanisme de leur reconnaissance et de leur internalisation met en jeu plusieurs protéines dont un transporteur membranaire.

1. *Les transporteurs membranaires d'opine*

Chez les bactéries, les transporteurs membranaires peuvent avoir un rôle important, notamment pour l'import de nutriments, la pathogénicité et la virulence. Le type de transport membranaire le plus utilisé est le transporteur de type ABC (ATP Binding Casette). Il est présent dans les 3 règnes du vivant et permet l'import ou l'export de nombreuses molécules. Les substrats importés par les transporteurs ABC peuvent être de natures diverses, à savoir des sucres, des métaux, des acides aminés, des peptides, des antibiotiques et d'autres molécules (Berntsson, Smits, Schmitt, Slotboom, & Poolman, 2010).

Chez *Agrobacterium*, les opines peuvent servir de source de carbone ou être impliquées dans le partage de la virulence. Leur import dans le cytoplasme est permis par un transporteur membranaire de type ABC. Ce type de transport est subdivisé en 3 types. Le type III aussi appelé ECF pour Energy-Coupling Factor se distingue structuralement ainsi qu'au niveau mécanistique par rapport aux types I et II. Il ne sera pas mentionné dans ce chapitre (Wilkens 2015). Les deux autres types (I et II) possèdent un domaine TMD (transmembrane domain) permettant de créer un canal membranaire, et un domaine NBD (nucleotide-binding domain) qui fixe et hydrolyse l'ATP. La différence entre ces deux types I et II de transporteur ABC se trouve au niveau structural. Les sous unités du domaine TMD sont composées de 5 à 6 hélices α pour le type I alors que le type II peut en avoir de 10 à 12. Le domaine NBD quant à lui peut être différent mais reste globalement très conservé structuralement entre différentes souches.

Ces domaines peuvent être des homodimères ou des hétérodimères. TMD et NBD dimérisés forment le complexe TMD/NBD (Figure 16). Pour importer des molécules dans le cytoplasme, ce complexe protéique recrute une PBP localisée dans le périplasme qui reconnaît spécifiquement le substrat et l'adresse au transporteur ABC pour déclencher son internalisation. De nombreuses études ont montré que la spécificité du substrat est déterminée par la liaison PBP-substrat et par la



formation du complexe PBP-substrat-ABC transporteur.

Figure 16 : Représentation schématique de différents complexes TMD/NBD en homo/heterodimère.

2. *Les Periplasmic Binding Proteins (PBP)*

Les PBP (periplasmic binding protein) ou SBP (solute-binding protein) font partie d'une famille de protéines possédant une taille variant entre 25 à 70 kDa. Elles sont présentes sous forme libre dans le périplasme chez les bactéries à Gram négatif et sont enchâssées dans la membrane plasmique chez les bactéries à Gram positif (Sutcliffe & Russell, 1995). Leur rôle est de reconnaitre spécifiquement des molécules solubles dans le périplasme et de les adresser à un transporteur membranaire, permettant leur translocation dans le cytoplasme.

a. Structure globale des PBPs

Les premières données structurales de PBP ont fait l'objet d'un article concernant l'ABP (arabinose-binding protein) d'*E. coli*, dont la structure a été résolue en 1976 (Phillips et al, 1976). Depuis cette découverte, des centaines de structures de PBP sont désormais disponibles dans la base de données structurales PDB (www.rcsb.org). Ces protéines sont constituées de deux domaines globulaires reliés par une région charnière. Ces deux lobes sont eux-mêmes composés d'un feuillet β entouré d'hélices α (Quiocho & Ledvina, 1996). Les protéines sont dynamiques et peuvent subir un changement conformationnel important grâce à cette région charnière.

b. Mécanisme d'interaction

Les PBPs ont la particularité de se retrouver sous deux formes : la forme ouverte (apo) ou la forme fermée (holo). Lorsque la PBP interagit avec son ligand, elle se referme pour augmenter la surface d'interaction. Les deux domaines sont alors rapprochés grâce à une rotation permise par la région charnière. Cette rotation varie en fonction des différentes classes de PBP et peut aller de quelques degrés à 60° comme par exemple pour la protéine LivJ liant des acides aminés (Trakhanov et al., 2005). L'interaction de la PBP avec le transporteur membranaire déclenchera ensuite le relargage du substrat et sa translocation vers le cytoplasme. L'interaction de la PBP avec le transporteur déclenche en effet un changement de conformation qui entraine la libération du ligand dans le domaine transmembranaire du transporteur.

c. Classification des différentes PBP

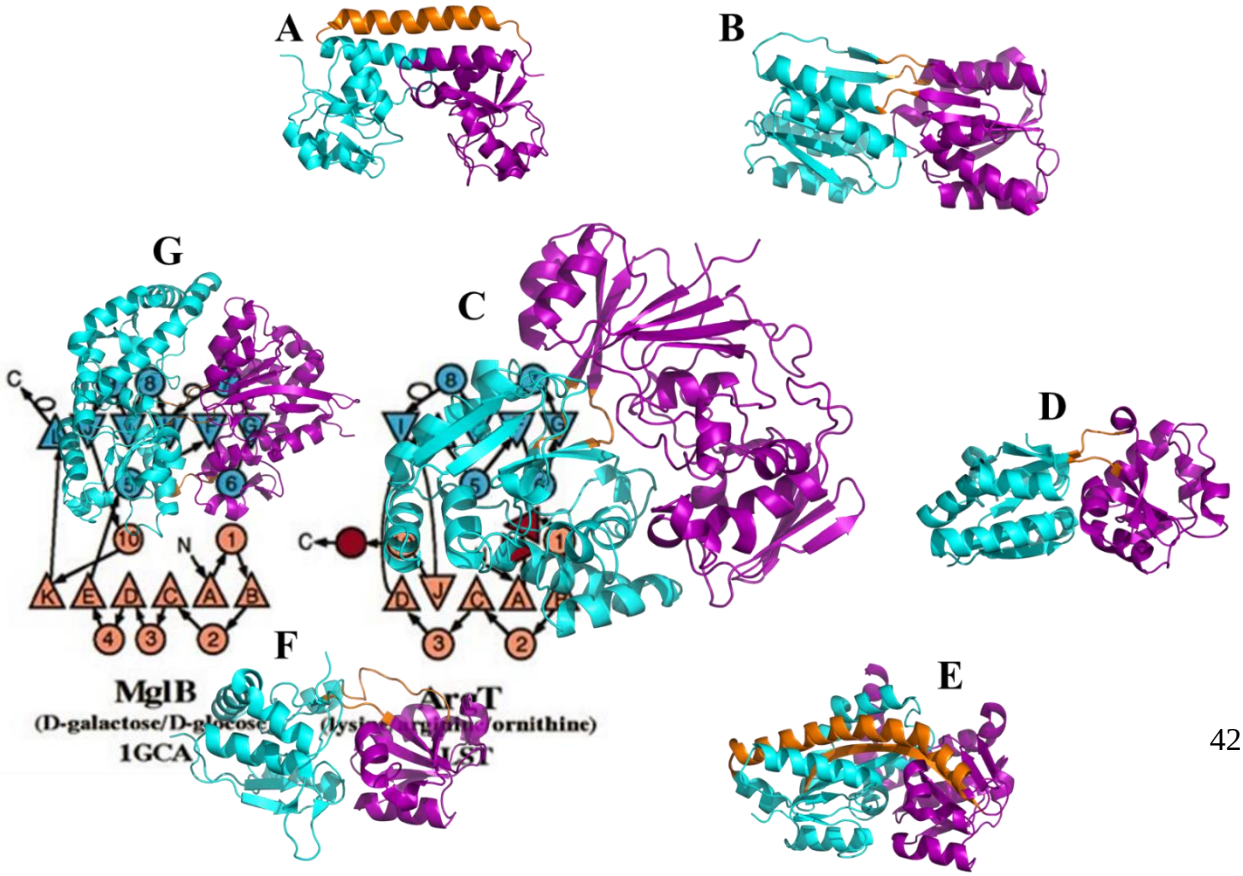
Plusieurs classifications des PBP ont été réalisées ces dernières années. La première, basée sur la topologie des deux domaines en mettant en avant l'ordre des brins β, a identifié deux classes. Pour

la classe I, les brins β sont organisés dans l'ordre $\beta_2\beta_1\beta_3\beta_4\beta_5$ tandis que pour la classe II l'organisation des brins est $\beta_2\beta_1\beta_3\beta n\beta_4$ (*n* représentant le brin après l'alternance de domaine) (Fukami-Kobayashi, Tateno, & Nishikawa, 1999) (Figure 17). En 2003, la classe III fut additionnée et représente les PBP possédant une hélice α dans la région charnière (Karpowich, Huang, Smith, & Hunt, 2003).

Figure 17 : Représentation topologique des deux classes de PBP classifiées en 1999 (Fukami-Kobayashi et al., 1999)

Avec le temps, de nouvelles structures tridimensionnelles de PBP ont été obtenues et des différences structurales plus précises au niveau de la charnière entre les deux lobes ont été remarquées. En 2010, l'analyse de ces différences sur 114 structures a conduit à une reclassification des PBP en 6 groupes allant de A à F (Berntsson et al., 2010) (Figure 18). Enfin, en 2016, l'analyse de 501 structures a permis l'identification d'un nouveau groupe nommé G. (Fukamizo, Kitaoku, & Suginta, 2019)

Les PBP du groupe A ont la particularité de posséder une région charnière composée d'une



hélice α (Figure 18A). La présence de cette hélice α engendre une rigidité de la structure et donc un faible changement conformationnel entre les formes ouverte et fermée. Par exemple, la PBP BtuF qui permet l'import de vitamine B12 chez *E. coli* n'engendre une ouverture entre les deux lobes que de 4°. Cette classe de PBP peut avoir une taille qui varie de 30 à 36 kDa et ces PBP peuvent fixer et déclencher l'internalisation de métaux, de sidérophores, d'hèmes, de la vitamine B12, de molybdate ou de tungstate.

Figure 18 : **Classification des PBP** adaptée de (Berntsson et al., 2010; El Sahili et al., 2015; Fukamizo et al., 2019) Lobe 1 en cyan, lobe 2 en violet et charnière en Orange. **Code PDB** : Classe A : BtuF (PDB code: 1N2Z), Classe B : RBP (PDB code: 1DRJ), Classe C : AccA (PDB code: 4ZEB), Classe D : ModA (PDB code: 4XXU), Classe E : UehA (PDB code: 3FXB), classe F : HisJ (PDB code: 1HSL) et classe G : AlgQ (PDB code: 1Y3N).

Le groupe B possède des protéines d'une taille d'environ 37 à 47 kDa et leur région charnière caractéristique est constituée de 3 chaines polypeptidiques servant de linker entre les domaines (Figure 18 B). Cette classe de PBP est plus flexible et reconnait des substrats tels que des monosaccharides, des acides aminés et des glutamates.

La classe C contient les PBP étudiées durant ma thèse. En 2010, cette classe fut décrite comme étant composée de protéines ayant deux domaines reliés par un troisième domaine charnière. Des études plus récentes ont montré que les PBP de cette classe sont finalement constituées de deux lobes avec une région charnière composée de 2 petits segments polypeptidiques flexibles (El Sahili et al., 2015; Fukamizo et al., 2019). Cette flexibilité peut engendrer une ouverture d'environ 12° entre les deux lobes comme par exemple pour la PBP AccA d'*Agrobacterium tumefaciens*. Les protéines de cette classe peuvent également interagir et permettre d'internaliser des métaux et des peptides (Fukamizo et al., 2019).

La classe D est caractérisée par une région charnière constituée de deux petites chaines polypeptidiques de 4 à 5 résidus. Elle se différencie des autres classes par la présence d'un motif « loop-helix-loop » à l'interface des deux lobes. Ces PBP d'environ 30 à 40 kDa peuvent reconnaitre et fixer différents substrats comme des sucres (Higgins, Abbott, Boulanger, & Boraston, 2009), des polyamines (Erika V. Soriano et al., 2008), ou des métaux (Shouldice et al., 2004).

Les domaines de la classe E sont reliés par deux longues chaines polypeptidiques et une longue hélice α. Cette famille de PBP est constituée de protéines d'une taille d'environ 34 à 40 kDa. Les protéines de cette classe interagissent uniquement avec des transporteurs de type TRAP (Tripartite ATP-Independent Periplasmic). Ces transporteurs contrairement au ABC transporteur n'utilise pas d'ATP mais un gradient électrochimique de protons ou de sodium pour permettre le transport. Ces PBP peuvent transporter des acides aminés (Rucktooa et al., 2007), des dipeptides et

des acides organiques (Gonin et al., 2007).

La classe F regroupe des PBP plus petites avec une taille d'environ 27 à 37 kDa. Elles possèdent des domaines très flexibles grâce à leur région charnière constituée de deux longues chaines polypeptidiques d'environ 8 à 10 résidus contre 4-5 habituellement. Cette charnière est relativement flexible puisque dans certains cas elle peut permettre une ouverture d'environ 43° pour la protéine NocT par exemple (Lang 2014). Cette flexibilité facilite l'interaction avec les molécules et peut donc reconnaitre et interagir avec une plus large variété de substrats et notamment avec des acides aminés (Oh et al., 1994), des dipeptides (Williams et al., 2004), de la thiamine, pyrimidine, ou sulfonates.

Enfin la classe G a été mentionnée pour la première fois en 2016 (Scheepers, Lycklama a Nijeholt, & Poolman, 2016). Les quelques protéines qui la composent ont une taille d'environ 60 kDa et n'ont pas de repliement ou de similarité avec les autres classes. Il s'agit donc d'une nouvelle classe de PBP fixant l'alginate (Mishima, Momma, Hashimoto, Mikami, & Murata, 2003; Momma, Mishima, Hashimoto, Mikami, & Murata, 2005).

Objectifs de la thèse :

L'opéron *acc*, et donc la capacité d'utiliser des agrocinopines, joue un rôle primordial dans le partage de la virulence, la communication bactérienne et dans le développement de certaines bactéries pathogènes du genre *Agrobacterium*. La connaissance d'un mécanisme reliant des fonctions aussi importantes est très intéressante.

Il existe différentes souches du genre *Agrobacterium* qui possèdent les gènes nécessaires à l'utilisation des différentes agrocinopines, regroupés dans l'opéron *acc* (Figures 10/① page 35).

La souche la plus étudiée est la souche modèle C58 d'*A. tumefaciens* qui induit la synthèse d'agrocinopines A et B (constituées d'arabinose-2-phosphate lié à un sucre) dans la tumeur de la plante pour ensuite les utiliser (Figure 9). De récentes études ont montré que cette souche peut importer par le biais de la PBP AccA (Figure 10 page 35) également les agrocinopines C et D (constituées de glucose-2-phosphate lié à un sucre) et n'est donc pas spécifique de ses propres agrocinopines. Ces travaux ont mis en évidence l'importance du motif pyranose-2-phosphate (arabinose-2-phophate ou glucose-2-phosphate) pour l'import de molécules par la PBP AccA (El Sahili et al., 2015).

L'objectif de la première partie de mes travaux a été dans une premier temps d'étudier la spécificité d'interaction des PBP AccA vis-à-vis des agrocinopines et dérivés (arabinose-2-phoshate et glucose-2-phosphate), chez les souches bactériennes Bo542 et S4 du genre *Agrobacterium* (Figure 10/④ page 35). J'ai ainsi réalisé des mesures d'interaction entre les protéines et leurs différents ligands. Les résultats obtenus ont mis en évidence une différence de spécificité d'interaction entre les PBP des souches C58, Bo542 et S4, ce qui m'a amené dans un second temps

à étudier structuralement ces différentes protéines pour identifier les résidus impliqués dans l'interaction protéine-ligand et expliquer au niveau atomique leurs spécificités.

La deuxième partie de mes travaux s'est orientée vers l'enzyme AccF de l'opéron *acc* responsable de la transformation des agrocinopines en pyranose-2-phosphate et en sucre (Figure 10/⑤ page 35). Protéine encore très peu étudiée chez les bactéries du genre *Agrobacterium*, c'est une enzyme clef dans le mécanisme d'utilisation des agrocinopines comme source de carbone ou comme molécule effectrice dans le mécanisme de communication bactérien « quorum sensing ». J'ai réalisé l'étude structurale et enzymatique de cette protéine afin de la caractériser et de mieux comprendre son mécanisme réactionnel.

RESULTATS

Chapitre 1 :

Etude de la spécificité de la fixation des agrocinopines par la protéine périplasmique PBP AccA chez les pathogènes du genre *Agrobacterium*.

a. Introduction

Dans le cadre de mon doctorat, j'ai étudié AccA de la souche Bo542 d'*A. tumefaciens*, ce qui m'a permis de déterminer et d'expliquer sa spécificité envers les agrocinopines. Deux souches d'*A. tumefaciens* (Bo542 et C58) sont capables de faire produire à la plante des agrocinopines (A et B pour la souche C58 et C et D pour la souche Bo542). Cependant, de précédentes recherches réalisées par le laboratoire avaient montré la non spécificité de AccA chez la souche C58 d'*A. tumefaciens* (El Sahili et al., 2015) : la souche C58 est capable d'utiliser les 4 agrocinopines A, B, C ou D comme seule source de carbone. Elle est également capable d'initier la synthèse de signaux de quorum sensing et de partager son plasmide Ti en réalisant du transfert conjugatif en présence des quatre agrocinopines.

Le travail que j'ai réalisé durant cette étude a donc été d'analyser la spécificité d'interaction de la protéine $AccA_{Bo542}$ vis à vis des agrocinopines. J'ai également utilisé la PBP $AccA_{S4}$ d'*A. vitis* afin d'avoir un élément de comparaison supplémentaire lors de mes travaux. Des méthodes telles que l'ITC (isothermal titration calorimetry), la DSC (differencial scanning calorimetry) ou la fluorescence m'ont permis de mesurer l'interaction entre la protéine et ces ligands, et de mettre en évidence la spécificité de la PBP $AccA_{Bo542}$ pour les agrocinopines C et D. Ensuite mon but a été d'expliquer cette spécificité au niveau structural grâce aux différents modèles que j'ai obtenu de la protéine en complexe avec les agrocinopines ou dérivés. Les analyses et conclusions mentionnées dans la suite de cette étude font référence aux différentes agrocinopines (A, B, C et D). Cependant, nos collaborateurs chimistes (Mohammed AMHAR et Yves QUENEAU) n'ont pu synthétiser que les agrocinopines A et D ainsi que leurs dérivés G2P et A2P. Les interprétations concernant les agrocinopines B et C ne sont donc pas justifiées expérimentalement mais sont des hypothèses appuyées grâce aux similitudes structurales qu'ont les agrocinopines entre elles (Figure 19).

Nos collaborateurs ont malgré tout tenté de synthétiser l'agrocinopine C. Cependant, même si la molécule synthétisée possède le même poids moléculaire, la structure cristallographique que j'ai obtenue en complexe avec $AccA_{C58}$ a permis de montrer qu'elle est différente d'un point de vue structural. Je ne montre pas cette structure dans ce manuscrit. Le composé synthétisé se différencie de la vraie agrocinopine C par la position de la liaison osidique entre le fructose et le glucose (Figure 19). L'agrocinopine C lie le fructose et le glucose en position 2 du glucose. Or dans la structure du complexe $AccA_{C58}$-molécule synthétisée, le fructose est lié en position 4 du glucose. Je n'ai donc pas pu utiliser les résultats en lien avec cette molécule synthétisée pour interpréter mes données d'interaction, qui ont malgré tout été réalisées.

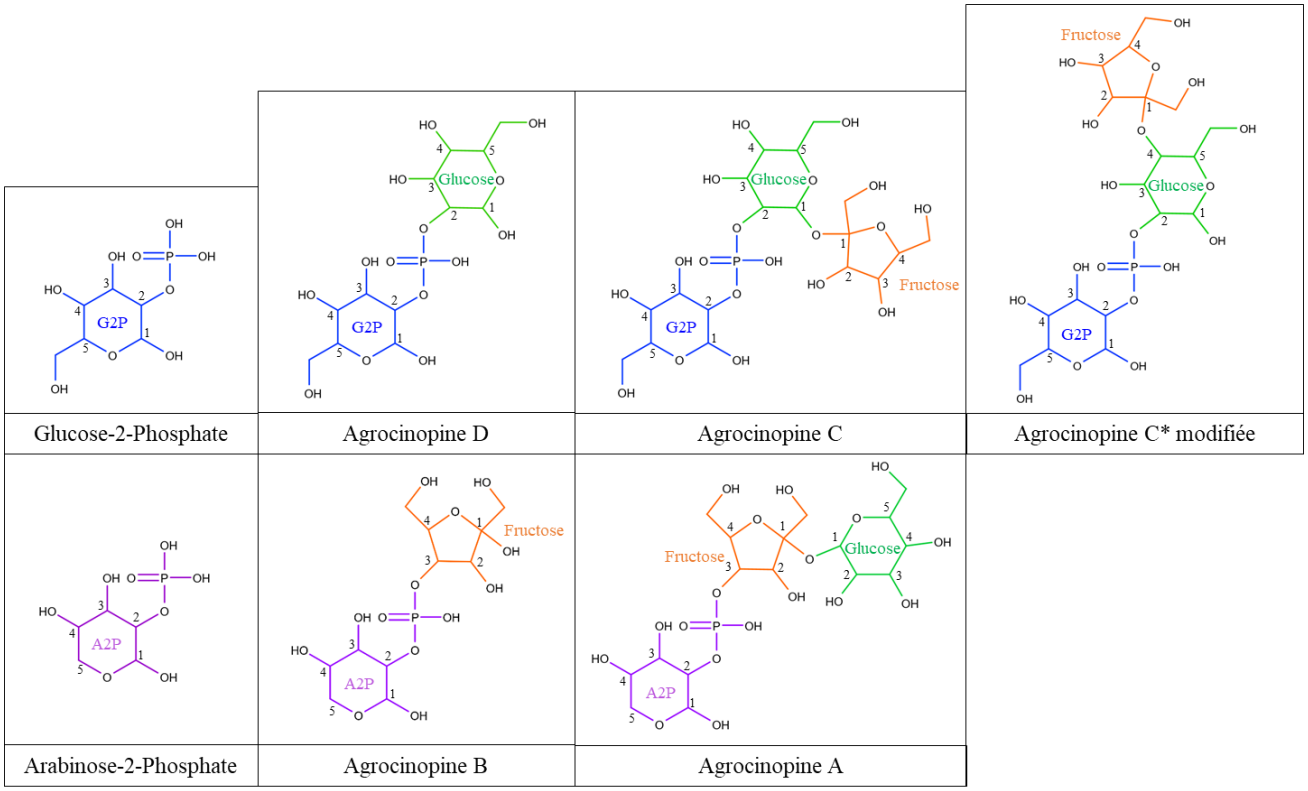


Figure 19: Représentation des différentes agrocinopines et dérivés. En bleu est représenté le groupement glucose-2-phosphate (G2P), en violet le groupement arabinose-2-phosphate, en vert les groupement glucose et en orange le groupement fructose. Les numéros représentent les carbones des cycles des différents sucres. La différence entre l'agrocinopine C et l'agrocinopine C* modifiée se situe au niveau de la position de la liaison entre le fructose et le glucose : Position 1 du glucose pour l'agrocinopine C et position 4 pour l'agrocinopine C* modifiée.

b. ***Résultats***

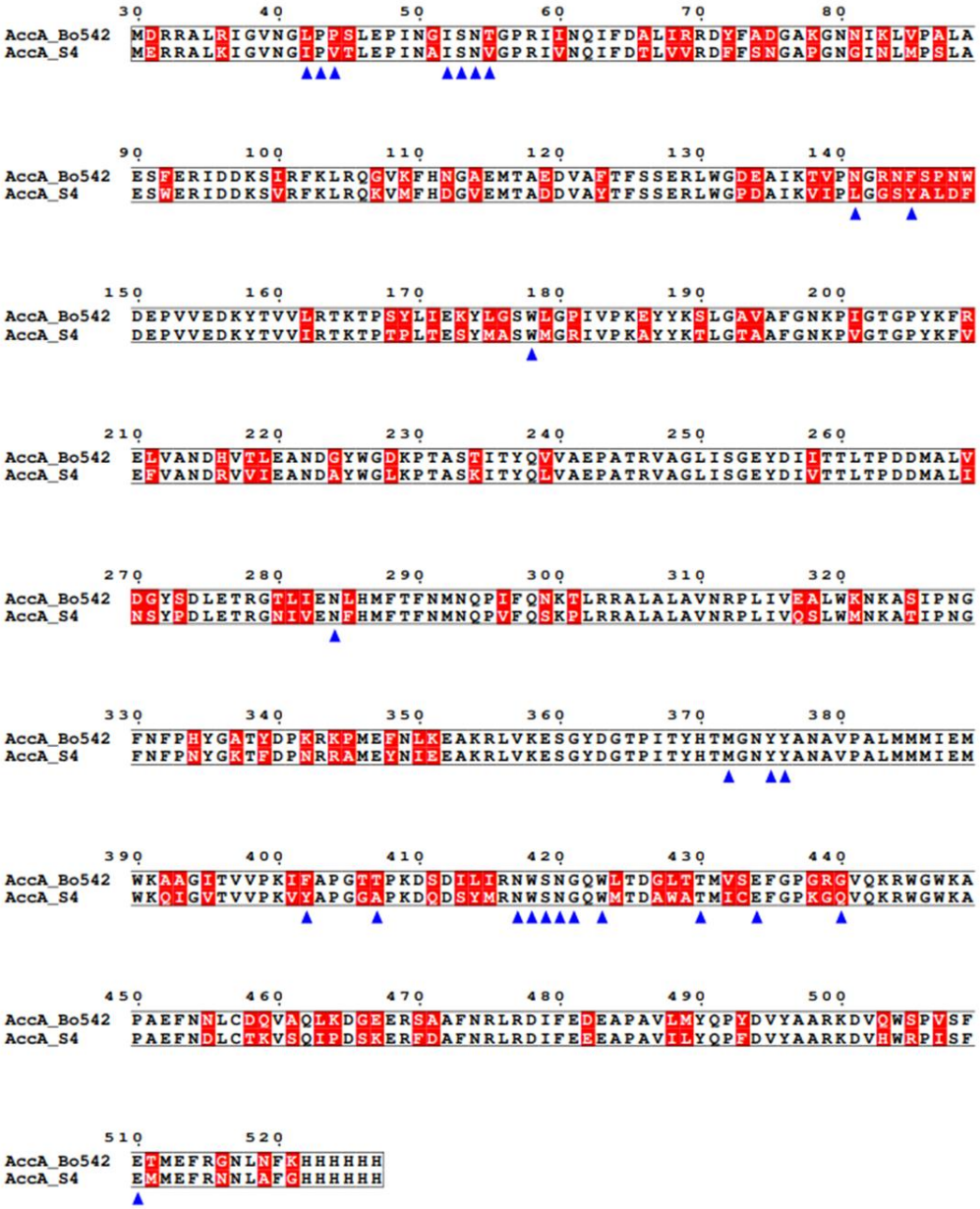
Tous les matériels et méthodes utilisés dans ce chapitre se trouvent à la page 132.

i. Clonages, expressions et purifications

Les gènes codant pour les protéines $AccA_{Bo542}$ d'*A. tumefaciens* (sauvage et les mutants P44A P44N et F145Y/G440Q) et $AccA_{S4}$ d'*A. tumefaciens* (sauvage) ont été synthétisés avec une étiquette codant pour 6 histidines et clonés dans un vecteur de surexpression bactérien par l'entreprise Genescript (Matériel et Méthodes). Les mutations des résidus de chaque mutant utilisé ont été choisies aux vues de leur localisation dans le site de fixation. Ces mutants ont été choisis et utilisés après une première étude structurale d'$AccA_{C58}$ (code pdb : 4ZEB), d'$AccA_{Bo542}$ sauvage et d'$AccA_{S4}$.

La PBP AccA de la souche Bo542 est une protéine composée de 521 résidus pour un poids moléculaire de 58,2 kDa. Lors de la synthèse du gène, la séquence signal permettant l'adressage de la protéine au périplasme a été supprimée et une étiquette six histidines a été ajoutée. La protéine

recombinante possède donc un poids moléculaire de 56.2 kDa et 498 résidus. Il en est de même pour la PBP AccA de la souche S4 (poids moléculaire final de la construction His-tagguée : 59.6 kDa). La seule différence notable entre les deux protéines, est la variation de certains résidus



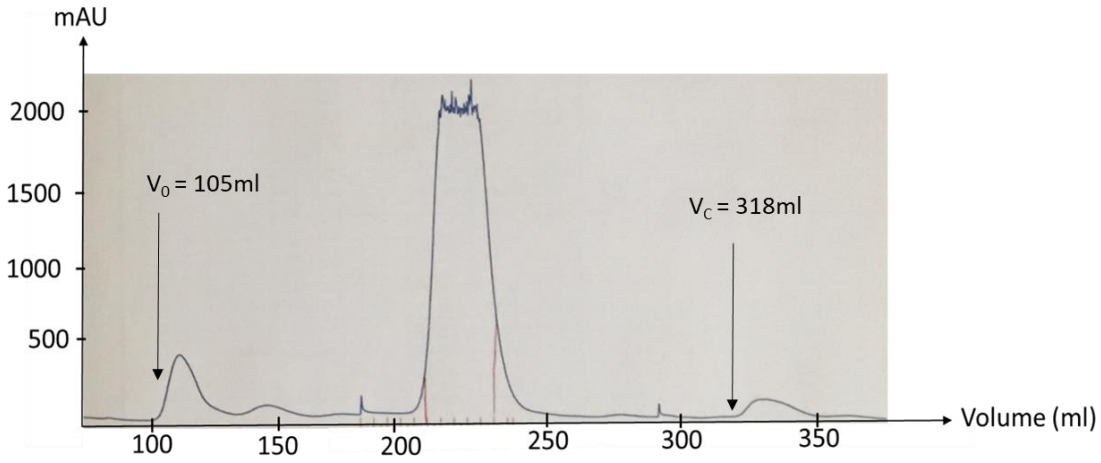
(Figure 20). Ces deux protéines possèdent un pourcentage d'identité de séquence de 74%. La localisation de ces variations sera étudiée par la suite.

Figure 20: Alignement des séquences protéiques d'AccA$_{Bo542}$ et AccA$_{S4}$. Les résidus différents entre les protéines sont encadrés en rouge, et les flèches bleues représentent les résidus situés dans la poche de fixation des deux protéines.

Les tests d'expression des deux protéines AccA$_{Bo542/S4}$ ont été réalisés avant mon arrivée au laboratoire. Les bactéries BL21 (DE3) et Rosetta (DE3) plysS ont été retenues pour la production d'AccA$_{Bo542}$ et AccA$_{S4}$, respectivement. Les conditions d'expression à grande échelle des

différentes protéines, en vue de leur purification, sont référencées dans le Tableau 14 en annexe et dans la partie Matériel et Méthodes page 132.

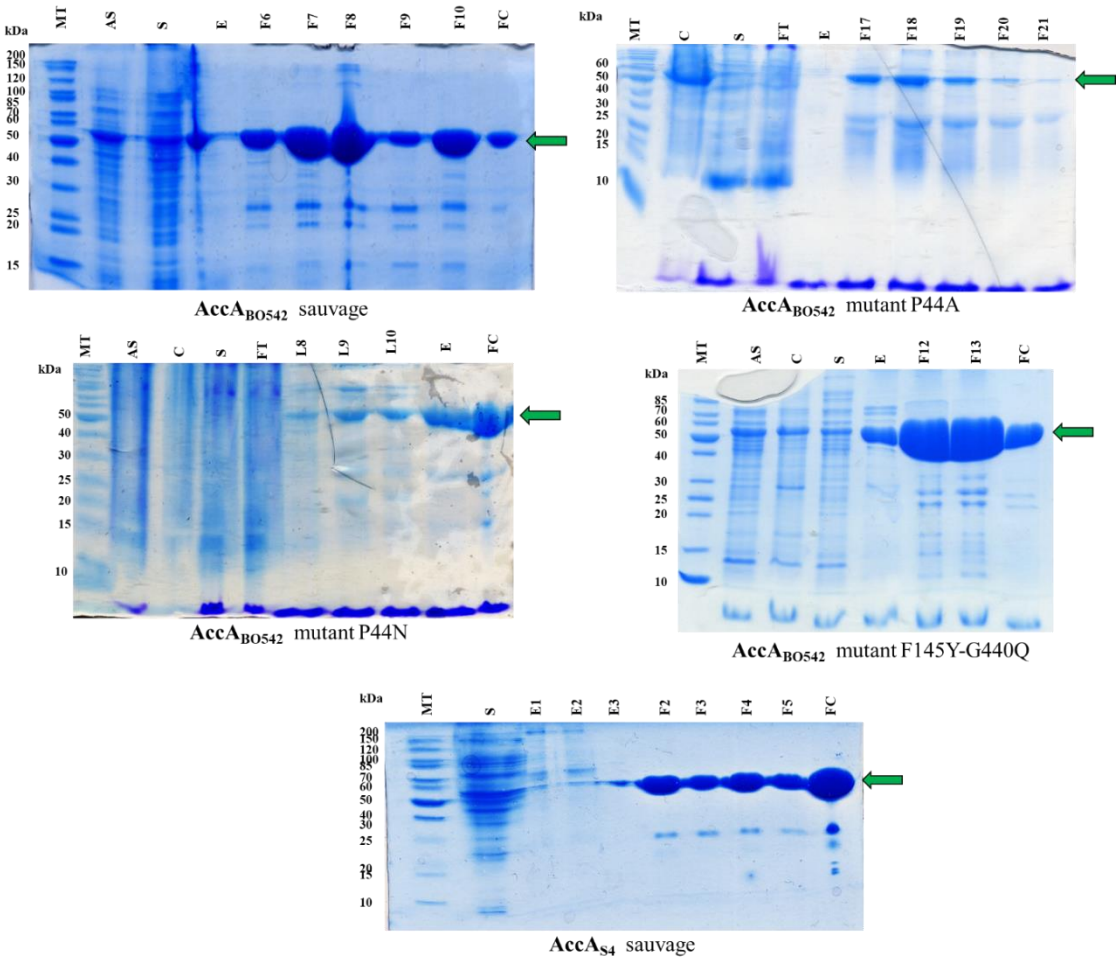
Une fois exprimées, les protéines sont purifiées comme décrit en Matériel et Méthodes : par chromatographie d'affinité sur une métal immobilisé (IMAC) où le métal est soit du nickel (**Ni**), soit du cobalt (**Co**), suivi par une chromatographie d'exclusion de taille (SEC). Le volume d'élution des protéines lors de cette 2eme étape est de 210 ml, ce qui correspond à un poids moléculaire d'environ 60 kDa. En effet, compte tenu de la calibration de la colonne, la forme relativement globulaire de la protéine et le volume d'élution observé m'ont permis de proposer que AccA est un monomère en solution (Figure 21). Le suivi de la purification des protéines sauvages ou mutées est



réalisé sur gel SDS-PAGE (Figure 22).

Figure 21 : Chromatogramme de l'étape de purification par chromatographie d'exclusion de taille de la protéine $AccA_{Bo542}$. La calibration de la colonne a été réalisée par l'équipe. Le volume d'élution des étalons utilisés: volume mort (V0) = 105 ml, volume colonne (VC) = 318 ml, $V_{Thyroglobulin}$ (670 kDa) =114 ml, $V_{aldolase}$ (158 kDa) = 173 ml, $V_{ovalbumine}$ (44 kDa) = 216 ml, $V_{Myoblobulin}$(17 kDa) = 247 ml.

On remarque la présence de bandes surexprimées aux alentours de 50-60 kDa, correspondant aux protéines d'intérêt (Figure 22). Les protéines sauvages et les mutantes, n'ayant pas de différences significatives au niveau du poids moléculaire, se retrouvent au même endroit sur le gel SDS. Cependant, malgré des conditions d'expression et de purification similaires, on observe une grande variabilité de la solubilité du mutant P44A en comparaison des autres protéines. La protéine, qui a un taux de production plus bas que celui des autres protéines, est de plus peu soluble et une grande partie est observable sur le gel dans la fraction correspondant au culot de lyse.



Protéine	AccA$_{Bo542}$				AccA$_{C58}$	AccA$_{S4}$
Mutation	Sauvage	P44A	P44A	F145Y/G440Q	E195A	H197A
mg$_{protéine\ purifiée}$/L$_{culture}$	59,6	7,2	3,6	6,6	-	30,1
Bactérie	Bl21 (DE3)	Bl21 (DE3)	Bl21 (DE3)	Bl21 (DE3)	Bl21 (DE3)	Rosetta (DE3) pLysS
Température d'expression	37°C	37°C	37°C	37°C	37°C	16°C

Figure 22 : Suivi de la purification des différentes protéines AccA$_{Bo542}$ et AccA$_{S4}$ sur gel SDS-PAGE et tableau récapitulatif de la quantité de protéine AccA sauvage et mutantes obtenues. Les protéines d'intérêt sont indiquées par une flèche verte (**MT** : Marqueur de Taille, **AS** : Avant sonication, **FT** : Fraction totale, **L** : Lavage, **C** : Culot, **S** : Surnageant, **E** : Elution Colonne Nickel, **Fx** : Fraction S200, **FC** : Fraction concentrée finale). Les culots des bactéries transformées ayant surexprimé AccA sont décongelés puis soniqués, la fraction soluble est déposée sur une résine de Nickel (Qiagen) ; après lavage, la protéine est éluée par ajout en gradient d'un tampon 300 mM NaCl, 50 mM Tris-HCl pH 8, 300 mM d'imidazole. Les fractions éluées sont regroupées et concentrées ou ajustées à 5 ou 10 mL, puis injectées sur une colonne de chromatographie d'exclusion de taille S200 (GE Healthcare). Les fractions éluées contenant les protéines d'intérêt sont regroupées et concentrées.

ii. Synthèse de molécules signal et Transfert conjugatif.

Les agrocinopines une fois transformées en A2P/G2P et en sucre, peuvent servir de source de carbone mais également comme molécules permettant d'engendrer la synthèse de molécules signal indispensables pour le quorum sensing (Voir Intro – Processus d'utilisation des agrocinopines). Ces molécules auto-inductrices (acyl homo-sérine lactone) déclenchent, lorsqu'elles atteignent une

certaine concentration, le transfert du plasmide de virulence, phénomène appelé transfert conjugatif. Comme nous l'avons vu dans l'introduction, en présence d'agrocinopines et plus particulièrement d'A2P ou de G2P, la répression de TraR par AccR est levée, grâce à la fixation du complexe dimérique TraR-AHSL (acyl homo-sérine lactones) sur une région ADN promotrice (Figure 15 page 39). Cette fixation permet l'expression de l'opéron TraCDG, ce qui engendre une augmentation de la production des signaux de quorum sensing, et déclenche le transfert conjugatif du plasmide Ti.

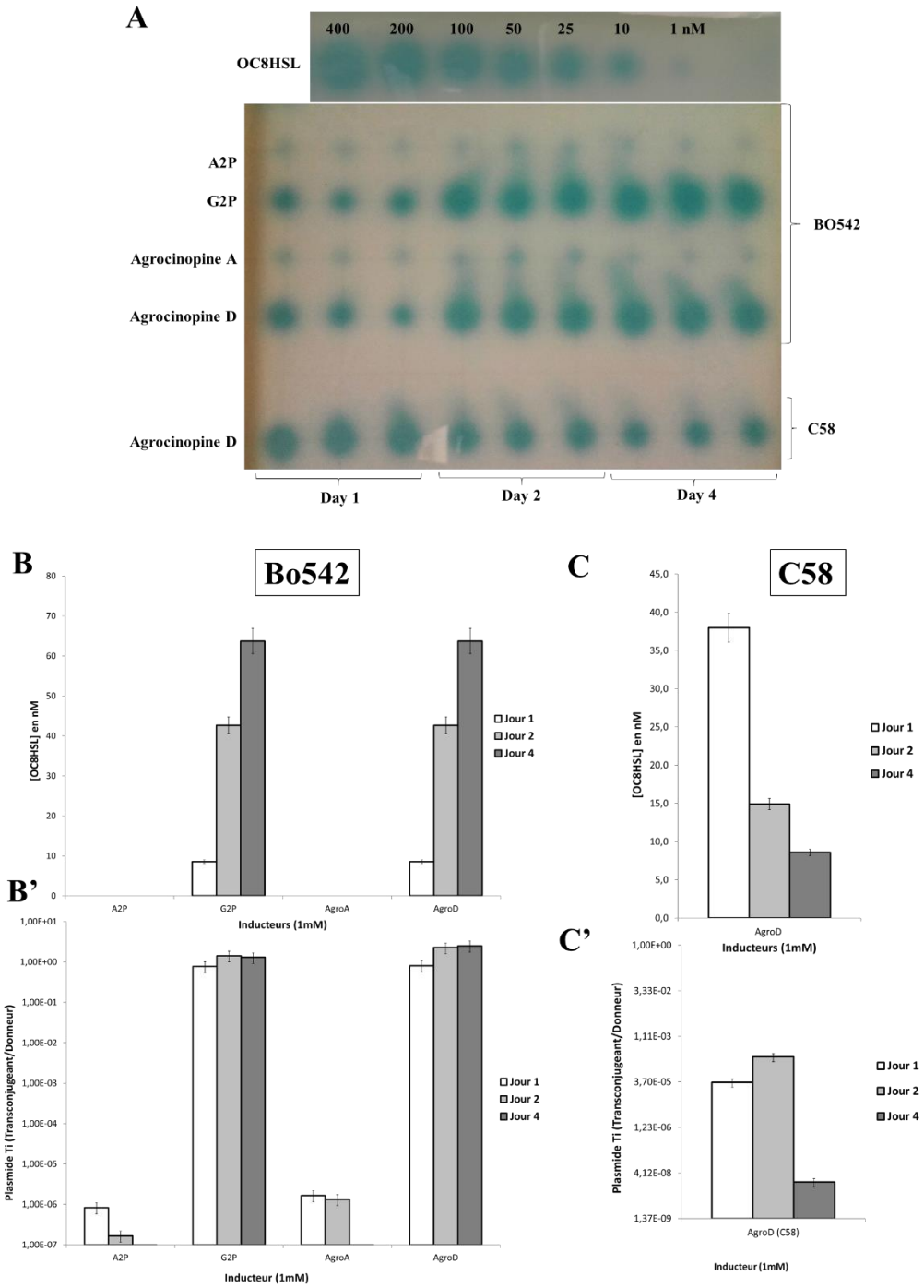
Dans de précédents travaux au laboratoire, il a été montré que la souche $AccA_{C58}$ produit des molécules signal en présence d'A2P, de G2P et d'agrocinopine A et que le transfert conjugatif est assuré (El Sahili et al., 2015).

Dans un premier temps j'ai d'abord vérifié que la souche C58 d'*A. tumefaciens* était également capable d'utiliser l'agrocinopine D (Figure 23). Ensuite, j'ai réalisé des expériences me permettant de vérifier s'il y avait production d'acyl homo-sérine lactone et si un transfert conjugatif avait lieu en présence des différentes agrocinopines et dérivés chez la souche Bo542 (Agrocinopine A, D et A2P/G2P) (Figure 23).

Le protocole utilisé afin de mettre en évidence les molécules d'AHSL dans des cultures des différentes souches en présence de différentes agrocinopines ou dérivés est décrit dans le Matériel et Méthodes (Cha, Gao, Chen, Shaw, & Farrand, 1998). La détection des résultats se fait sur des plaques de chromatographie en couche mince (M&M) recouvertes d'un milieu minimum (AB mannitol agar) mélangé à une solution contenant du Xgal ainsi que la souche de bactérie NTLR4. Cette souche a la particularité de contenir un plasmide possédant les gènes *traR* formant un complexe avec les AHSL, l'opéron TraCDG ainsi que son promoteur, suivi du gène *lacZ*. S'il y a production d'AHSL, ils seront reconnus pour la protéine TraR pour former un complexe TraR-AHSL et permettre l'expression de l'opéron TraCDG mais également celui de la β-galactosidase. Cette dernière transforme alors le Xgal en galactose et 5-bromo-4-chloro-3-hydroxyindole, dont la dimérisation donne une couleur bleue visible que nous allons observer pour l'interprétation des résultats. Des halos sont alors observables et la concentration en AHSL est calculée selon le diamètre du halo en comparaison avec une gamme de concentrations d'AHSL connues.

Pour valider la réalisation d'un transfert conjugatif, j'ai combiné deux souches : une souche receveuse contenant un gène de résistance à la rifampicine et dépourvue de plasmide, et une souche donneuse possédant le plasmide Ti. Les deux souches sont mises à DO_{600nm} 1 et mélangées puis

cultivées à 28°C durant 4 jours. Des aliquots sont ensuite étalés sur 3 milieux sélectifs afin de quantifier les bactéries totales, les bactéries receveuses ou les bactéries transformées (voir Matériel et Méthodes) et de calculer l'efficacité de transformation. La présence de halos bleus (Figure 23 A) nous confirme la présence d'AHSL. J'ai réalisé ce travail avec Armelle Vigouroux et en



collaboration avec l'équipe de Denis Faure et plus particulièrement avec Almudena Gonzàlez-Mula (I2BC, département Microbiologie, Intéractions Plantes-Bactéries).

Figure 23 : Résultats des expériences de transfert conjugatif et de détection de signaux de quorum sensing.
A : Production d'acyl homo-sérine lactone pour C58 et BO542: Photo de la plaque révélant la présence de molécules

signal en présence de différents inducteurs à différents temps. Différents échantillons ont été prélevés en triplicata et à différents temps (1, 2, 4 jours). La présence d'OC8HSL engendre la formation de halos bleus correspondant au X-gal dégradé. Les différents halos correspondent donc indirectement aux molécules signal (OC8HSL) et sont ensuite comparés à ceux de la gamme dont les concentrations sont connues.
B/C : Production d'acyl homo-sérine lactone pour BO542 et C58: Diagramme récapitulatif des résultats obtenus quant aux molécules signal détectées pour la souche BO542 (B) et la souche C58 (C). On remarque qu'il n'y a production d'OC8HSL qu'en présence de G2P et d'agrocinopine D et que la quantité de molécules « signal » croit en fonction des jours d'incubation. On note en revanche l'absence de tâches de diamètre de taille significative en présence d'A2P et d'agrocinopine A. La présence de molécules « signal » de la souche C58 en présence d'agrocinopine D est différente est une décroissance de la concentration de molécules « signal » est observée.
B'/C': Transfert conjugatif pour les souches BO542 et C58: L'étalement des souches de bactérie sur milieu sélectif à différents temps (1, 2 et 4 jours) m'a permis de réaliser un diagramme représentant l'efficacité de transformation, correspondant au ratio de bactéries (transformées/donneuses) ayant reçu le plasmide de virulence par rapport au nombre de bactéries pouvant le donner.

Les résultats obtenus nous montrent que la souche C58 produit des molécules « signal » en présence d'agrocinopine D, mais que la concentration décroit au court du temps. En effet le halo bleu correspond à une concentration d'environ 14,9 et 8,6 nM de OC8HSL au jour 2 et 4 respectivement, alors que pour le jour 1 elle est de 38 nM (Figure 23 A/C). De plus le transfert conjugatif en présence d'agrocinopine D montre une efficacité de transfert de l'ordre de 1 bactérie transformée pour 10 bactéries donneuses au jour 1-2 alors qu'elle est extrêmement basse et de l'ordre de 4 pour 100.000.000 au jour 4 (Figure 23 C'). S'il est possible que la souche C58 utilise l'agrocinopine D, il n'y a aucune raison que l'efficacité de transformation décroisse en fonction du temps. La seule raison pour laquelle l'efficacité de transformation diminue est un problème au niveau de la croissance bactérienne, qui semble être ralentie. Pour être sûr des résultats obtenus, il serait intéressant de reproduire une nouvelle fois cette expérience.

Pour la souche Bo542, on remarque que seuls l'agrocinopine D et le G2P permettent la synthèse d'acyl homo-sérine lactone. On remarque l'augmentation du diamètre des halos bleus correspondant à l'augmentation de la concentration d'OC8HSL en présence de G2P et d'agrocinopine D (Figure 23 A). La concentration passe de 10 nM au jour 1 à 65 nM au jour 4 (Figure 23 A/B). Cependant, en présence d'A2P ou d'agrocinopine A, l'absence d'évolution de la tache bleue de très petit diamètre nous laisse penser qu'il s'agit de fuite de synthèse et non pas d'une expression spécifique. De plus, l'efficacité de transformation obtenue nous montre que le transfert conjugatif n'est efficace qu'en présence d'agrocinopine D et de G2P, avec une efficacité de transfert de l'ordre de 1 pour 1 (Figure 23 B'). En revanche, les valeurs obtenues en présence d'Agrocinopine A et d'A2P sont très faibles et non significatives aux vues de l'efficacité de transformation de 1 pour 1.000.000.

Pour conclure, ces résultats confirment que l'agrocinopine D peut être utilisée par la souche C58, en plus de l'A2P, du G2P et de l'agrocinopine A comme montré précédemment au laboratoire,

et que la souche Bo542 a une spécificité de reconnaissance pour les agrocinopines C et D. Afin de comprendre pourquoi Bo542 est spécifique de ses propres agrocinopines, j'ai décidé d'étudier les interactions d'AccA avec ces différents ligands.

iii. Etude des interactions protéine : ligands des différentes PBP AccA.

Dans cette étude je me suis intéressé à la spécificité d'interaction des protéines AccA du genre *Agrobacterium*. J'ai poursuivi les travaux réalisés avant mon arrivée au laboratoire sur la souche C58, et élargi le sujet par l'étude des souches Bo542 d'*A. tumefaciens* et S4 d'*A. vitis*. Les tests d'interaction ont été réalisés par le biais de trois techniques : Fluorescence, Tycho et Calorimétrie différentielle à balayage ou DSC (differencial scattering calortimetry). Le principe de ces techniques est donné dans la partie Matériel et Méthodes.

1. PBP $AccA_{C58}$

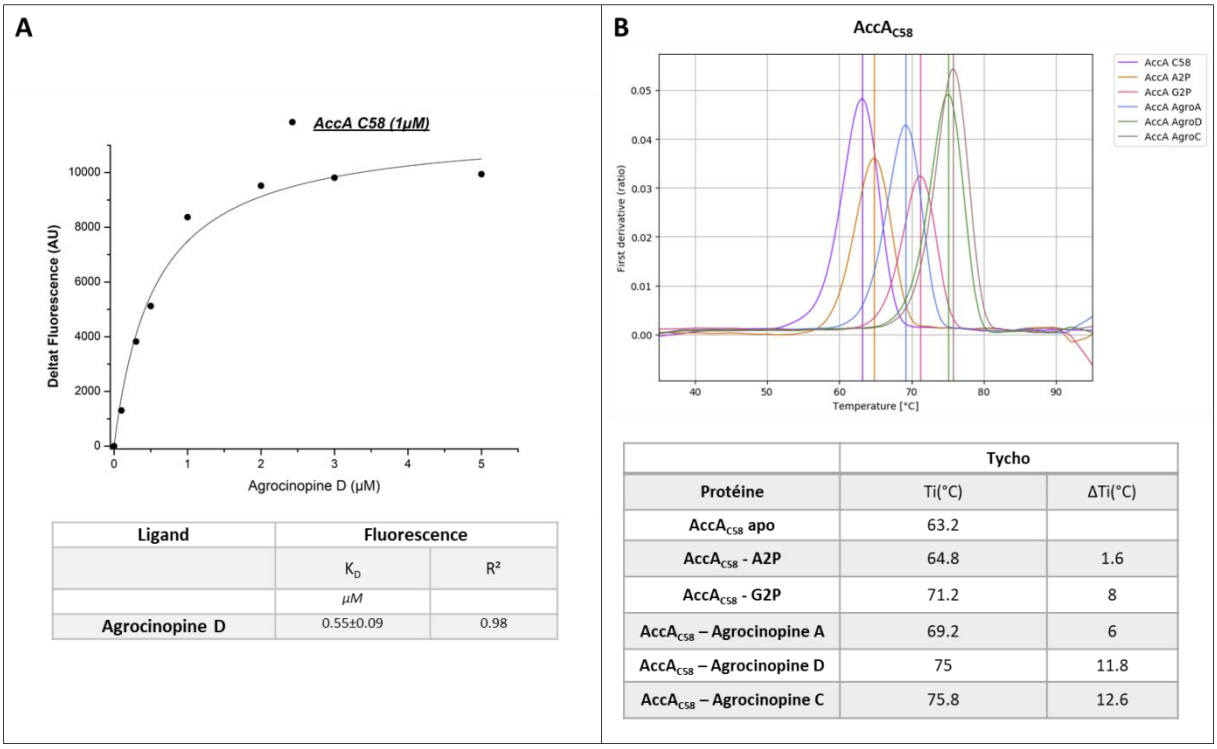
D'après les études précédentes, la protéine AccA de la souche C58 ne semble pas être spécifique et elle permettrait l'utilisation de toutes les agrocinopines. De plus, selon des résultats basés sur des études d'interaction réalisées par l'équipe avec des agrocinopines ou dérivés (arabinose-2-phosphate, glucose-2-phosphate, agrocinopine A), il semble que de manière générale, toutes les molécules possédant un motif pyranose-2-phosphate peuvent être utilisées par $AccA_{C58}$ (El Sahili et al., 2015).

Dans un premier temps, j'ai caractérisé l'interaction de la PBP $AccA_{C58}$ avec l'agrocinopine D. Des expériences de fluorescence et de Tycho ont montré respectivement une modification du signal des tryptophanes et une modification de la température de fusion en présence d'agrocinopine D, démontrant une interaction avec la protéine (Figure 24).

Grâce aux tests de titration par fluorescence (Figure 24 A), on observe la variation de l'émission de fluorescence des tryptophanes de la protéine en fonction de la variation de la concentration en ligand. L'agrocinopine D est reconnue par $AccA_{C58}$ et le coefficient de corrélation est de 0.98. La forte affinité de la protéine $AccA_{C58}$ pour l'agrocinopine D est représentée par une constante de dissociation de l'ordre du micromolaire.

Les résultats obtenus avec la technique du Tycho (Figure 24 B) nous montrent une augmentation de la température d'inflexion (Ti) en présence des différents ligands (A2P, G2P, Agrocinopine A, D et $C_{modifiée}$), et donc une stabilisation de la protéine lorsqu'elle est liée à l'un de

ces ligands. Les valeurs obtenues sont significatives à partir d'un ΔTi de 1°C et on peut observer une augmentation de toutes les températures d'inflexion de 1.6 à 12.6 °C. L'augmentation de la stabilité de la protéine en présence de ces différents ligands est due à la formation d'un complexe.



Ligand	Fluorescence	
	K_D	R^2
	µM	
Agrocinopine D	0.55±0.09	0.98

	Tycho	
Protéine	Ti(°C)	ΔTi(°C)
AccA$_{C58}$ apo	63.2	
AccA$_{C58}$ - A2P	64.8	1.6
AccA$_{C58}$ - G2P	71.2	8
AccA$_{C58}$ – Agrocinopine A	69.2	6
AccA$_{C58}$ – Agrocinopine D	75	11.8
AccA$_{C58}$ – Agrocinopine C	75.8	12.6

Figure 24 : Résultats des tests d'interaction entre AccA$_{C58}$ et l'agrocinopine D. **A :** Graphique représentant la variation du signal de fluorescence des tryptophanes par rapport à la protéine seule, en fonction de la concentration en agrocinopine D ajoutée. **B**: Courbes représentant la dérivée du ratio 350nm/330nm en fonction de la température (°C), laissant apparaitre la température d'inflexion (Ti) correspondant au point de fusion de la protéine seule ou liée à l'agrocinopine D. Les différentes valeurs de Ti et de variation de température d'inflexion entre la protéine seule et la protéine liée aux différents ligands sont résumées dans le tableau.

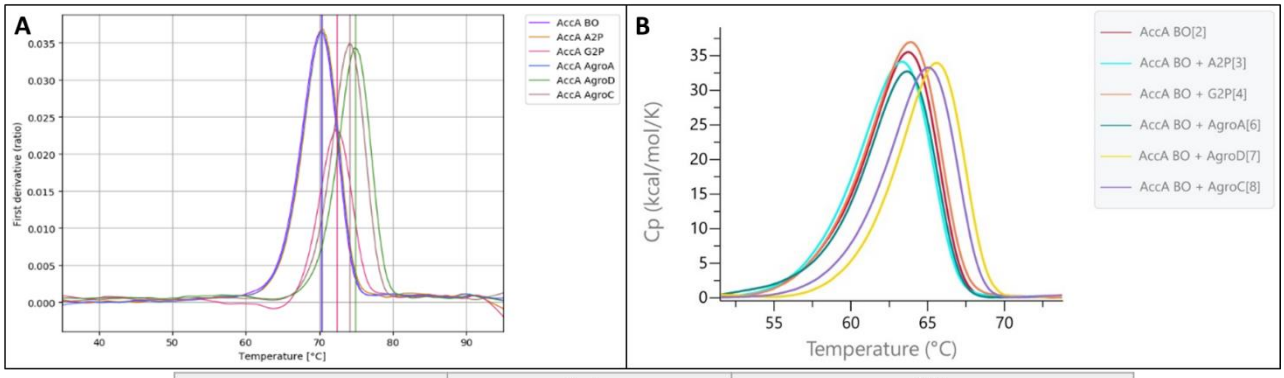
On peut conclure de ces expériences que la protéine AccA$_{C58}$ interagit bien avec les agrocinopines A, B et leurs dérivés (A2P, G2P) (El Sahili et al., 2015) mais également avec l'agrocinopine C* modifiée et D. Elle ne présente donc pas de spécificité.

La reconnaissance possible de l'agrocinopine C native reste par contre hypothétique étant donné que les résultats de l'agrocinopine C* modifiée ne lui sont pas transposables aux vues de la différence structurale entre C* modifiée et C, comme évoqué précédemment.

2. *PBP AccA$_{Bo542}$*

Grâce au gène *asc* (agrocinopine synthase) présent dans l'ADN-t du plasmide de virulence, la souche Bo542 d'*A. tumefaciens* fait produire à la plante les agrocinopines C et D, et il est donc normal que la protéine AccA$_{Bo542}$ reconnaisse et internalise ces molécules. Dans cette partie je me suis intéressé à la spécificité de la protéine AccA$_{Bo542}$ vis-à-vis des agrocinopines et plus particulièrement à savoir si la protéine permettait également l'import des agrocinopines A et B. De

la même façon que la protéine $AccA_{C58}$, l'interaction entre $AccA_{Bo542}$ (sauvage et mutants P44N, P44A, et F145Y-G440Q) et les différentes agrocinopines et dérivés a été mesurée par Tycho, DSC et tests de fluorescence (Figure 25).

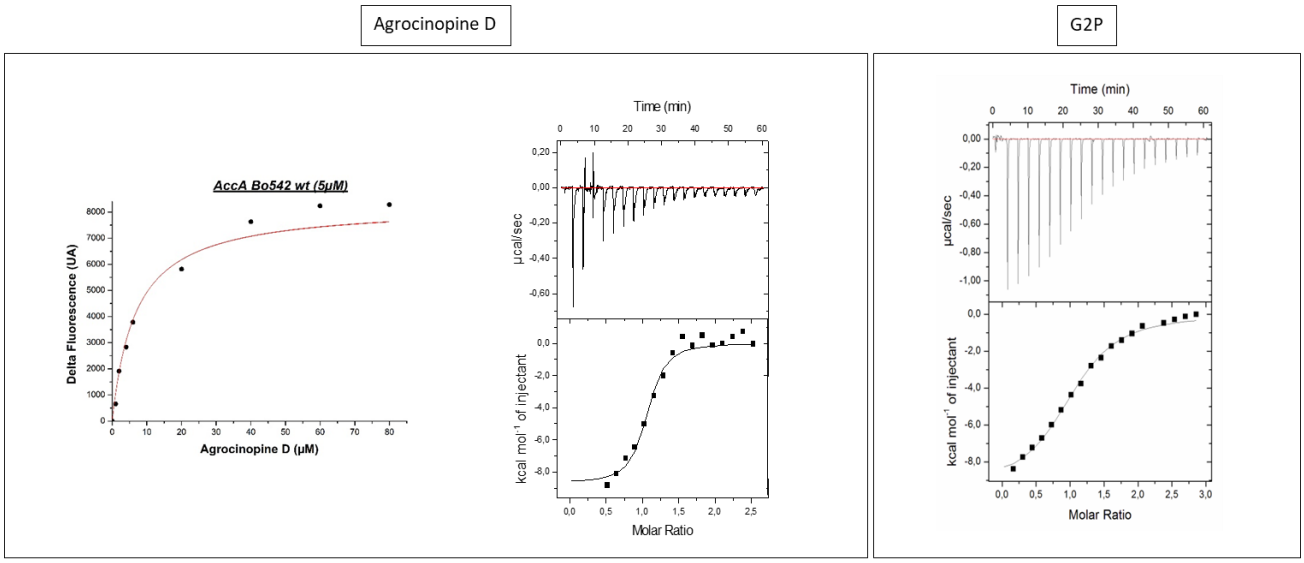


	Tycho		DSC		
Protéine	Ti(°C)	ΔTi(°C)	Tm(°C)	ΔTm(°C)	ΔH (kcal/mol)
$AccA_{Bo542}$ vide	70.2		63.1		209
$AccA_{Bo542}$ - A2P	70.4	0.2	62.7	-0.4	207
$AccA_{Bo542}$ - G2P	72.3	2.1	63.3	0.2	225
$AccA_{Bo542}$ – Agrocinopine A	70.3	0.1	63.1	0	190
$AccA_{Bo542}$ – Agrocinopine D	74.9	4.7	65.1	2	185
$AccA_{Bo542}$ – Agrocinopine C	74.1	3.9	64.5	1.4	190

Figure 25 : Résultats des tests d'interactions par DSC et Tycho d'$AccA_{Bo452}$ en présence des différents ligands. A: Courbes représentant la dérivée du ratio 350nm/330nm laissant apparaitre la température d'inflexion (Ti). **B**: Courbes représentent la capacité thermique en fonction de la température laissant apparaitre la température de fusion représentative de la température de dénaturation de la protéine. Le tableau résume les différentes températures d'inflexion/fusion obtenues par tycho/DSC.

Les résultats de l'expérience de Tycho (Figure 25 A) montrent une augmentation de la température d'inflexion lorsque la protéine est en complexe avec le glucose-2-phosphate (2,1°C), l'agrocinopine D (4,7°C) et l'agrocinopine C* modifiée (3,9°C). La protéine interagit donc bien avec ces différents ligands. En revanche l'absence de différence de température d'inflexion entre la protéine seule et en présence d'A2P ou agrocinopine A permet de conclure qu'il n'y a pas stabilisation et donc d'interaction dans ces 2 cas-là. Si on compare ces résultats aux résultats obtenus avec $AccA_{C58}$, on remarque que le Ti de la forme *apo* est plus élevée pour la souche Bo542 contrairement à la souche C58. Cependant, celui de la forme complexée est presque identique avec une valeur avoisinant les 74-75°C. On peut donc en conclure que $AccA_{Bo542}$ *apo* est plus stable que AccC58 *apo* mais que la reconnaissance et la fixation de l'agrocinopine D engendre une stabilité des protéines équivalente.

Des différences de Tm significatives ont également été observées par la technique de DSC (Figure 25 B). Ces résultats nous montrent que la protéine $AccA_{Bo542}$ possède un Tm de 63,1°C seul alors qu'il est de 65,1 et 64,5°C en présence d'agrocinopine D et d'agrocinopine C* modifiée. Ceci indique que la protéine est plus stable en présence de ces ligands et interagit avec eux.



Ligand	Fluorescence		ITC					
	K_D	R^2	K_D	N	Enthalpie (ΔH)	Entropie (ΔS)	Contribution entropique (-TΔS)	Energie libre (ΔG)
	µM		µM		cal/mol	cal/mol/deg	cal/mol	cal/mol
G2P			0.6	1.05	-9354	-7.71	2260	-7094
Agrocinopine D	1.36±0.80	0.98	0.3	1.03	-8696	0.186	-54	-8750

Figure 26 : Résultats des tests d'interactions par titration de fluorescence et de calorimétrie isotherme d'$AccA_{Bo542}$ en présence d'agrocinopine D et de G2P. Les mesures ont permis l'obtention de la constante de dissociation (K_D) ainsi que des paramètres thermodynamiques. Le tableau regroupe les données obtenues grâce aux expériences effectuées à savoir : la constante de dissociation (K_D), l'enthalpie (ΔH), l'entropie (ΔS), la contribution entropique (-TΔS) et l'énergie libre (ΔG).

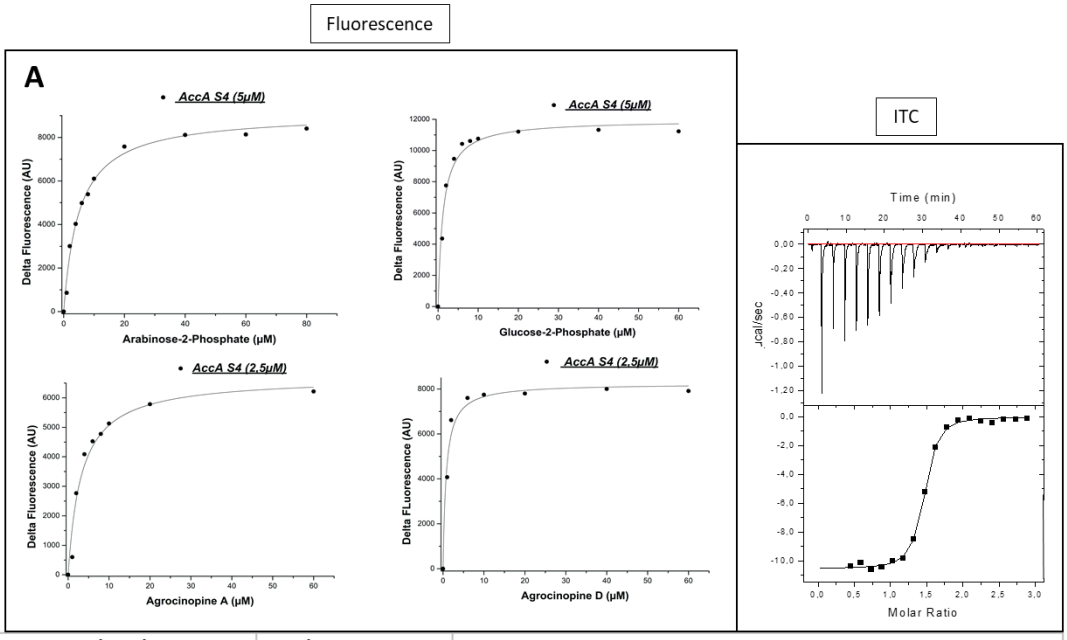
Ensuite, les mesures de constante de dissociation (K_D) par titration de la fluorescence montrent une évolution du signal de fluorescence en présence de G2P et d'agrocinopine D. La constante de dissociation obtenue est de l'ordre du micromolaire pour les deux ligands et ces résultats indiquent donc que la protéine interagit avec le G2P et l'agrocinopine D (Figure 26).

Enfin, il en est de même pour la détermination du K_D par titration de calorimétrie isotherme. L'enthalpie (ΔH) et l'enthalpie libre (ΔG) négative témoignent d'une réaction exothermique et exergonique. On remarque également que la contribution enthalpique est plus élevée que la contribution entropique, ce qui indique que la reconnaissance est spécifique. Enfin, l'affinité de la protéine pour les ligands est élevée avec une constante de dissociation de l'ordre du micromolaire. Si on compare la reconnaissance des deux ligands, on remarque que l'affinité est plus élevée lors de la reconnaissance de l'agrocinopine D, ce qui est cohérent étant donné que l'agrocinopine D est plus grande et qu'elle établit plus d'interactions avec la protéine comme observé lors de l'analyse structurale des complexes. (Cf Tableau 5 page 73)

Les résultats des différentes mesures biophysiques mènent vers la conclusion que la protéine $AccA_{Bo542}$ est spécifique et n'interagit qu'avec le G2P et les agrocinopines C et D, avec une affinité de l'ordre du micromolaire.

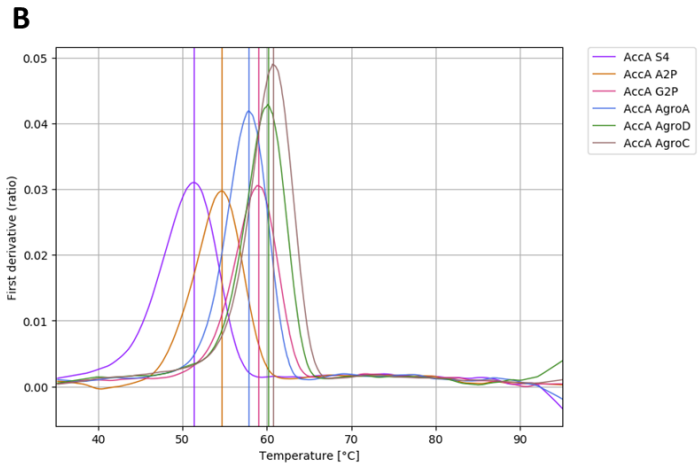
3. *PBP $AccA_{S4}$*

Le plasmide de la souche S4 *A. vitis* ne possède pas le gène *acs* (agrocinopine synthase) permettant de faire produire à la plante les agrocinopines. Nous ne savons donc pas si la PBP AccA peut interagir *in vitro* avec les agrocinopines. Le pourcentage d'identité de séquence pourtant très élevée avec les autres AccA laisse à penser qu'elle pourrait elle aussi interagir (82% id de séquence avec $AccA_{C58}$ et 74 % avec $AccA_{Bo542}$). Les résultats obtenus nous permettent de confirmer que la protéine peut interagir avec toutes les agrocinopines et n'est pas spécifique, tout comme dans le cas



Ligand	Fluorescence		ITC					
	K_D	R^2	K_D	N	Enthalpy (ΔH)	Entropy (ΔS)	Entropic contribution (-TΔS)	Free energy (ΔG)
	µM		*µM*		*cal/mol*	*cal/mol/deg*	*cal/mol*	*cal/mol*
$AccA_{S4}$ – G2P	0.25±0.02	0.97						
$AccA_{S4}$ – A2P	0.80±0.07	0.98						
$AccA_{S4}$ – Agrocinopine A	1.04±0.13	0.96						
$AccA_{S4}$ – Agrocinopine D	0.24±0.04	0.97	0.19	1.4	-10560	-5.26	1541	-9019

de $AccA_{C58}$ (Figure 27).



	Tycho	
Protéine	Ti(°C)	ΔTi(°C)
$AccA_{S4}$ apo	51.3	
$AccA_{S4}$ – A2P	54.7	3.4
$AccA_{S4}$ – G2P	59	7.7
$AccA_{S4}$ – Agrocinopine A	57.8	6.5
$AccA_{S4}$ – Agrocinopine D	60.1	8.8
$AccA_{S4}$ –Agrocinopine C*	60.8	9.5

Figure 27 : Résultats des tests d'interaction entre la protéine $AccA_{S4}$ et l'agrocinopine D. A : Graphiques représentant les deltas de fluorescence des tryptophanes d'$AccA_{S4}$ en fonction de la concentration des différents ligands. Contrairement aux graphiques, les tableaux affichent des résultats normalisés pour 1µM de protéine. On observe également un titrage de la calorimétrie isotherme en présence d'agrocinopine D. **B**: Courbes représentant la dérivée du ratio 350nm/330nm laissant apparaitre la température d'inflexion (Ti). Le tableau regroupe les données obtenues grâce aux expériences effectuées à savoir : la constante de dissociation (K_D), l'enthalpie (ΔH), l'entropie (ΔS), la contribution entropique (-TΔS) et l'énergie libre (ΔG).

La fluorescence de la protéine $AccA_{S4}$ augmente significativement en présence de chaque ligand. Ceci suggère un changement de l'environnement physicochimique des tryptophanes présents dans le site de fixation de la protéine et donc la présence du ligand dans celui-ci. Ainsi la protéine est capable de fixer tous les ligands (Agrocinopine A, D, A2P, G2P) avec une affinité de l'ordre du micromolaire.

La titration par calorimétrie isotherme permet d'indiquer que tout comme $AccA_{Bo542}$, $AccA_{S4}$ fixe l'agrocinopine D avec une constante de dissociation de l'ordre du micromolaire.

Le Tycho a permis de confirmer l'établissement d'interaction par une augmentation de la température d'inflexion en présence de tous les ligands. Comme on peut le remarquer (Figure 27), les températures d'inflexion augmentent de 3.5 à 9.5 degrés en fonction du ligand et semblent augmenter la stabilité de la protéine. La protéine $AccA_{S4}$ est donc capable de reconnaitre les différentes agrocinopines.

Pour conclure, $AccA_{S4}$ possède une large spécificité et peut interagir avec toutes les agrocinopines et dérivés testés.

4. *Conclusion*

Les précédents travaux réalisés avaient montré que la PBP AccAC58 n'était pas spécifique et interagissait avec de nombreuses agrocinopines et dérivés (A2, G2P, Agrocinopine A). Grâce aux différentes mesures d'interactions réalisées, j'ai pu confirmer la non spécificité de cette protéine, qui interagit également avec l'agrocinopine D. La PBP $AccA_{S4}$ reconnait également toutes les agrocinopines et dérivés et n'est pas spécifique (A2P, G2P, Agrocinopine A et Agrocinopine D).

Par contre la PBP AccA de la souche Bo542 d'*A. tumefaciens* ne semble pas reconnaitre l'agrocinopine A et l'A2P. Ces résultats concordent avec la détection des molécules « signal » du quorum sensing et le transfert conjugatif (Figure 23 A, B et B'). Ceci indique que la souche Bo542 est spécifique et ne peut utiliser que les agrocinopines C et D contrairement aux souches C58 et S4.

Pour essayer d'expliquer cette spécificité au niveau moléculaire, j'ai entrepris de déterminer la structure des 3 protéines AccA (C58, Bo542 et S4).

iv. Cristallogenèse

Avant mon arrivée au laboratoire, les AccA des souches Bo542 et S4 avaient déjà commencé à être étudiées et des cristaux avaient été obtenus. Toutes les protéines que j'ai exprimées et concentrées ont donc été cristallisées en utilisant les conditions de cristallisation déjà mises au point par l'équipe. J'ai repris ces conditions pour obtenir des cristaux des protéines en complexe avec les différents ligands, par co-cristallisation (voir Matériel et méthodes). Les différents cristaux obtenus sont montrés dans la Figure 28.

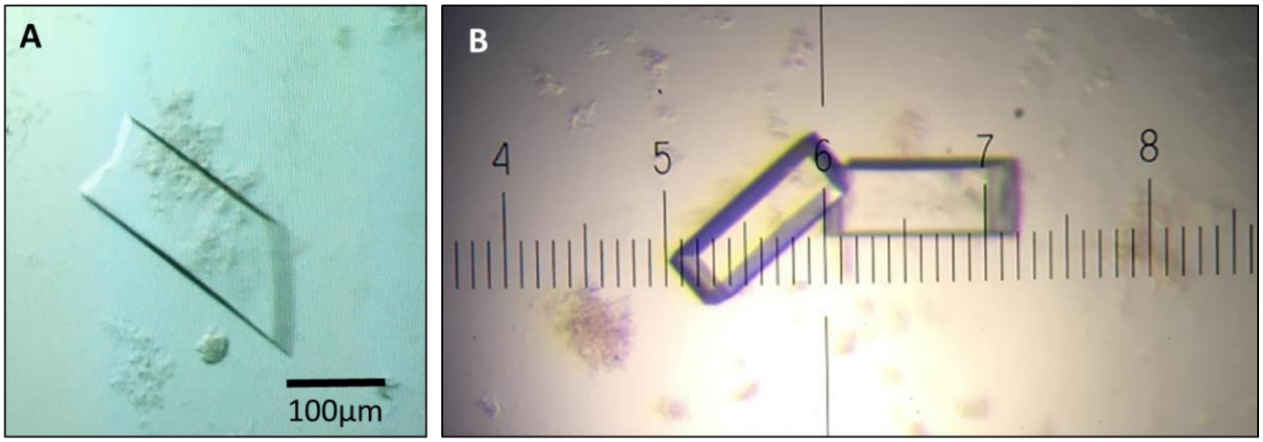


Figure 28 : Exemples caractéristiques de cristaux d'$AccA_{S4}$ obtenus. Les cristaux ont été obtenus par la technique de goutte suspendue. La taille des cristaux varie entre 100 et 300 µm comme on peut le voir sur les images B et A respectivement.

v. Diffraction aux rayons X et résolution des structures

Les cristaux congelés et stockés dans l'azote liquide ont été placés dans le faisceau de rayon X des lignes Proxima 1 ou 2 du Synchrotron SOLEIL (Saclay). Les données collectées ont été traitées avec le programme XDS (Kabsch, 2010). Suite au traitement des données, un groupe d'espace a été déterminé pour chaque complexe. Afin de résoudre les structures, le phasage par remplacement moléculaire a été réalisé. Une fois la structure résolue, la suite de logiciel BUSTER (Smart et al., 2012) a permis de réaliser plusieurs cycles d'affinement pour améliorer les coordonnées des atomes des structures ainsi que les cartes de densité électronique. Entre chaque cycle d'affinement, les structures ont été manuellement modifiées avec le logiciel COOT afin de corriger les erreurs entre les modèles et les cartes de densité électronique (Emsley, Lohkamp, Scott, & Cowtan, 2010) (Matériel et Méthodes).

Le modèle utilisé pour tous les calculs de remplacements moléculaires est la protéine $AccA_{C58}$ en complexe avec l'agrocinopine A (code PDP : 4ZEB) (El Sahili et al., 2015). Suite à l'obtention de la première structure de la protéine $AccA_{S4}$, celle-ci a été utilisée comme modèle pour la résolution de sa structure en complexe avec des ligands. Les tables cristallographiques sont présentées en Annexe page 147.

vi. Analyse et comparaison structurale de la protéine AccA des différentes souches et en complexe avec divers ligands.

1. *Les structures obtenues*

La PBP AccA des souches du genre *Agrobacterium* est une protéine qui reconnait les agrocinopines présentes dans le périplasme. Ces protéines font parties de la classe C des PBP précédemment décrites (Cf Figure 18 page 44) et présentent entre elles une grande identité de séquence (Tableau 2). Les structures obtenues ont été résolues par remplacement moléculaire en utilisant comme modèle la PBP AccA de la souche C58 (El Sahili et al., 2015) (Code PDB : 4ZEB).

Souche	**C58**	**Bo542**	**S4**
Identité de séquence (%)	-	74,9	81,7

Tableau 2 : Pourcentage d'identité de séquence des différentes PBP AccA du genre *Agrobacterium* par rapport à la souche C58.

Des études dans l'équipe avaient déjà permis d'obtenir quelques cristaux et des jeux de données, ainsi que de résoudre certaines structures des différentes PBP AccA des différentes souches. Le tableau ci-dessous regroupe la totalité des complexes structuraux obtenus, leur seuil de résolution ainsi que la contribution que j'ai apportée dans chaque étude (Tableau 3).

Souche	C58		Bo542	S4			
Protéine	AccA Sauvage						
Ligand	AgroA	AgroD	AgroD	A2P	G2P	AgroA	AgroD
Résolution (Å)	1,67	1,8	1,75	1,4	1,2	1,5	1,68
RMSD (Å)	-	0,44	0,85	1,2	1,05	1,03	1,01
Contribution	Traitement de données Affinement	Traitement de données Affinement	Affinement	Cristaux Collecte Traitement de données Affinement	Cristaux Collecte Traitement de données Affinement	Affinement	Cristaux Collecte Traitement de données Affinement

Tableau 3 : Récapitulatif des structures de la protéine $AccA_{Bo542/C58/S4}$, à savoir, leur résolution, le RMSD de chaque complexe AccA par rapport à $AccA_{C58}$ fixée à AgroA, et ma contribution apportée à chaque structure.

2. *Description structurale des PBP AccA.*

La description des structures cristallographiques de la PBP AccA de la souche C58 en complexe avec l'agrocinopine A, l'A2P et le G2P a été réalisée avant mon arrivée au laboratoire (El Sahili et al., 2015).

a. Structure globale des PBP $AccA_{Bo542/S4}$

Pour ma part, j'ai commencé par affiner la structure de la PBP $AccA_{Bo542}$ en complexe avec l'agrocinopine D (Figure 29 A). La protéine est organisée en deux domaines (lobe 1 et lobe 2), reliés par une région charnière composée de deux petites chaines polypeptidiques sans structure secondaire. Le premier domaine est composé des résidus 31 à 280 et 494 à 521 alors que le deuxième domaine s'étend des résidus 285 à 489. Enfin les deux linkers correspondant à la région charnière sont composés des résidus 281 à 284 et 490 à 493 (Figure 29).

Les cristaux d'$AccA_{S4}$ obtenus en complexe avec l'agrocinopine D m'ont permis de montrer que cette protéine possédait le même repliement (RMSD = 0.88Å avec $AccA_{Bo542}$, 477 résidus alignés). Les 2 domaines sont identiques à ceux de $AccA_{Bo542}$, et leur position relative est la même (Figure 29B).

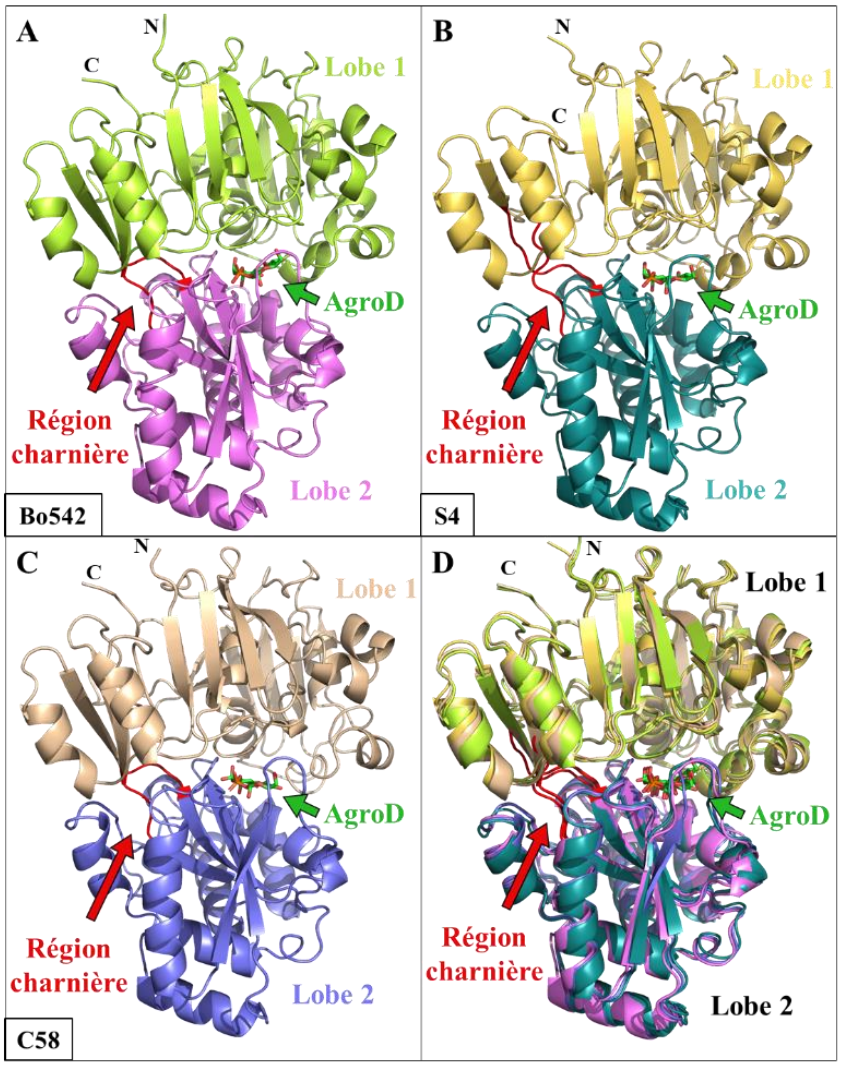


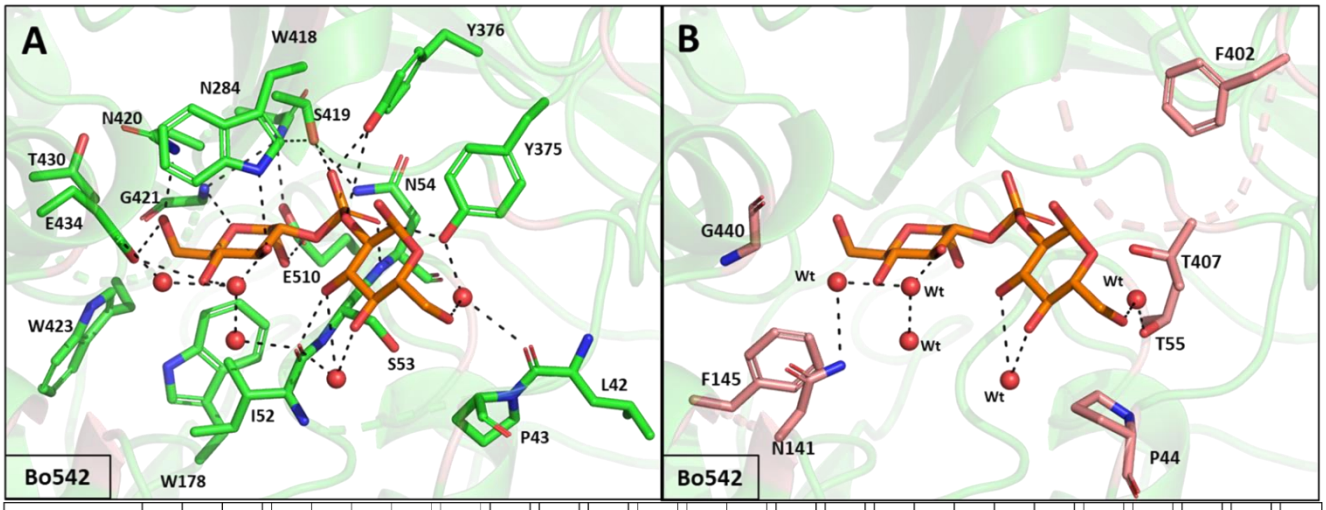
Figure 29 : Représentation des différentes PBP AccA en ruban et leur superposition. Les protéines sont toutes en complexes avec l'agrocinopine D en stick vert. Les deux domaines de chaque protéine sont d'une couleur différente. Les régions charnière sont représentées en rouge. **A :** PBP $AccA_{Bo542}$ en complexe avec l'agrocinopine D. Le domaine 1 est en vert et le domaine 2 en fuchsia. **B :** PBP $AccA_{S4}$ en complexe avec l'agrocinopine D. Les lobes 1 et 2 sont colorés en jaune et vert respectivement. Augmentation de la taille de la région charnière dû à la diminution de la taille du brin béta **C :** PBP $AccA_{C58}$ en complexe avec l'agrocinopine D. Le domaine 1 est en beige et en bleu pour le domaine 2. **D :** Superposition des trois structures en complexe avec l'agrocinopine D.

b. Site de fixation des PBP $AccA_{C58/Bo542/S4}$

Les protéines AccA des souches Bo542, C58 et S4 sont des protéines capables de fixer les agrocinopines et leurs dérivés. Les tests précédemment réalisés montrent que les protéines AccA sauvage de souches C58 et S4 peuvent reconnaitre toutes les agrocinopines et dérivés alors que celle de la souche Bo542 fixe uniquement les agrocinopines C* modifié et D ainsi que le G2P. Pour comprendre les bases moléculaires de cette spécificité, j'ai réalisé différentes approches afin d'identifier des résidus pouvant être potentiellement impliqués. Dans un premier temps j'ai fait une analyse des résidus de la poche de fixation d'AccA des 3 souches en présence d'agrocinopine D et des souches C58 et S4 en présence d'A2P. Ensuite j'ai répertorié tous les résidus des AccA qui sont directement en interaction avec les différents ligands, chez les 3 souches. Enfin j'ai réalisé une modélisation de l'agrocinopine A dans le site de fixation d'$AccA_{Bo542}$.

i. Analyse des poches de fixation d'$AccA_{C58/Bo542/S4}$

Les structures des 3 AccA possèdent entre elles un ΔRMSD d'environ 1Å (Figure 29 D). Elles sont donc extrêmement proches et l'étude globale de leur structure ne m'a pas permis d'obtenir des informations pouvant m'aider à expliquer la spécificité de reconnaissance de $AccA_{Bo542}$. J'ai donc orienté mes recherches vers les poches de fixation des ligands des 3 protéines. Les surfaces de contact protéine-ligand de chaque souche en présence d'agrocinopine D sont très similaires : 428 Å² pour $AccA_{Bo542}$, 466 Å² pour $AccA_{C58}$ et 433 Å² pour $AccA_{S4}$. L'agrocinopine D est en interaction avec de nombreux résidus et de nombreuses molécules d'eau (Figure 30 A). Le tableau en corrélation avec la figure liste les résidus de la poche de fixation des 3 AccA et nous permet de déduire le pourcentage d'identité de séquence entre les 3 sites de liaison du substrat, qui est de 73% et 63% entre Bo542/C58 et Bo542/S4 respectivement (Figure 30 C). Au niveau des similitudes et différences des sites de fixation en présence d'agrocinopine D, 7 résidus se démarquent des 26 autres, du fait qu'ils sont variables entre $AccA_{Bo542}$ d'une part et $AccA_{C58/S4}$ d'autre part: positions 44, 55, 141, 145, 402, 407, 440 (Figure 30 B).



C	42	43	44	52	53	54	55	141	145	178	284	372	375	376	402	407	417	418	419	420	421	423	430	434	440	510
Bo542	L	P	P	I	S	N	T	N	F	W	N	M	Y	Y	F	T	N	W	S	N	G	W	T	E	G	E
C58	L	P	N	I	S	N	V	L	Y	W	N	M	Y	Y	Y	F	N	W	S	N	G	W	T	E	Q	E
S4	I	P	V	I	S	N	V	L	Y	W	N	M	Y	Y	Y	A	N	W	S	N	G	W	T	E	Q	
Résidus Similaire/Différents																										
Zone ligand impactée	G2	G2	G2	G1/G2	G1/G2	G1	G2	G1	G1	G1	G1	G2	P	P/G2	G2	G2	G1	P	P	G1	G1	G1	G1	G1	G1	G1

Figure 30: Représentation du site de fixation de la PBP $AccA_{Bo542}$ en complexe avec l'agrocinopine D affichée en orange. A : En verts sont représentés les résidus similaires à au moins une des deux autres souches et en rouge les molécules d'eau. **B :** Les résidus différents de AccA entre la souche Bo542 et les souches C58/S4 sont de couleur rose. **C :** Tableau récapitulatif des résidus impliqués dans la fixation de l'agrocinopine D. Les résidus sur fond bleu sont impliqués dans la fixation du ligand par la protéine de la souche concernée alors que ceux sur fond blanc sont des résidus utilisés par les autres souches. Les similitudes sont validées par les cases vertes alors que les différences sont représentées par une case rouge. Les 3 zones impactées par le ligand sont **G1** : Glucose 1 ; **G2** : Glucose 2 ; **P** : Phosphate. Le glucose 1 représente la partie correspondant à l'arabinose des agrocinopines A et B plutôt enfouie. Alors que le glucose 2 représente la partie sucrose des agrocinopines tournée vers l'extérieur du site de fixation.

J'ai renouvelé cette analyse en considérant les structures en complexe avec l'A2P (Tableau 4). Les résidus qui se démarquent cette fois sont les résidus 141, 145 et 440. Les résidus 44, 55, 141, 145, 402, 407, 440 sont donc de bons candidats pour être impliqués dans la spécificité de reconnaissance des agrocinopines. On peut voir que les résidus 141, 145 et 440 sont localisés au niveau du glucose 1 (partie pyranose-phosphate des agrocinopines) dans le cas de la reconnaissance de l'agrocinopine D et au niveau de l'arabinose dans le cas de l'A2P. Ces résidus semblent donc importants pour la fixation de la région pyranose-phosphate des agrocinopines. A l'inverse, les 4 autres résidus (44, 55, 402 et 407) sont plutôt proches de la partie sucrose des agrocinopines.

	52	53	54	141	145	178	284	375	376	418	419	420	421	434	440	510
Bo542	I	S	N	N	F	W	N	Y	Y	W	S	N	G	E	G	E
C58	I	S	N	L	Y	W	N	Y	Y	W	S	N	G	E	Q	E
S4	I	S	N	L	Y	W	N	Y	Y	W	S	N	G	E	Q	
Similitude																
Zone ligand impactée	A	A	A	A	A	A	A	P	P	P	P	A	A	A	A	A

Tableau 4: Récapitulatif des résidus impliqués dans la fixation de l'A2P. Les résidus sur fond bleu sont impliqués dans la fixation du ligand par la protéine de la souche concernée alors que ceux sur fond blanc sont des résidus utilisés par les autres souches. Les similitudes sont validées par les cases vertes alors que les différences sont représentées par une case rouge. Les 2 zones du ligand impactées sont **A** : Arabinose ; **P** : Phosphate.

ii. Analyse des interactions protéine : ligands

Les résidus impliqués dans les interactions protéine-ligand ont été analysés utilisant le serveur PISA (https://www.ebi.ac.uk/pdbe/pisa/) en se concentrant sur les liaisons hydrogènes et les interactions hydrophobes (Tableau 5).

	Protéine	$AccA_{C58}$	$AccA_{S4}$	$AccA_{C58}$	$AccA_{S4}$	1 - $AccA_{C58}$	2- $AccA_{C58}$	2 - $AccA_{S4}$	2 - $AccA_{S4}$	$AccA_{Bo542}$	$AccA_{C58}$	$AccA_{S4}$
Type d'interaction	Ligand	A2P		Agrocinopine A		G2P				Agrocinopine D		
	Résolution (Å)	2,3	1,4	1,67	1,5	1,75		1,2		1,75	1,8	1,68
	Résidus											
Hydrogène	N44 (C58)	-	-	3,4/2,9	-	-	-	-	-	-	-	-
	I52	-	-	3	-	-	-	-	-	3,4	-	-
	N54	3,3/3	3,4/2,9/3,5/2,8	2,9/2,8	2,7/2,9	2,9	3,0/3,0	3,6/2,8/3,5	3,6/2,8/3,5/2,9	2,8	2,29/3,4	2,9
	Y145 (C58/S4)	-	-	-	-	3	3,1	2,8	2,8	-	2,7	2,7
	N284	3,3	3,2	-	3,1	-	3,1	-	3,1	-	-	-
	Y375	2,7	2,6	2,6	2,6	2,7	2,4	2,6	2,6	2,7	2,7	2,7
	Y376	2,6	2,7	2,8/2,7/3,1	2,7/2,9	2,6	2,5	2,6	2,6	2,7/3,5	2,7/3,2/2,7	2,8/3,2/3
	Y402 (C58/S4)	-	-	2,9	-	-	-	-	-	-	-	-
	W418	3	3,2	2,8	3,1/3,3	3	2,9	3	3	3,2	3	3/3,4
	S419	3,6/3,2/2,7/3,6	3,5/3/2,7/3,6	3,4/3/2,5	3,5/2,9/2,7/3,4	2,6/3,3	2,9/2,9	2,7	2,7/2,8	2,7	2,6/3,3	2,7
	N420	3,1	3	2,8	3	-	-	3,2	3,2	3,4	3,1	3,1
	G421	3,4/3	2,9	2,9	2,8	3	2,8/3,4	-	3,3	2,8	2,9	2,8
	E434	2,8	2,7	2,7	2,7	2,7/2,9	2,8/2,8	2,7/3	2,7/3	2,9/2,8	2,8/2,8	2,7/2,7
	E510 (Bo542/C58)	2,7	-	2,7/3,5	-	2,8	2,6	-	-	2,6	2,6	-
	WT	1	4	-	12	-	-	-	-	7	8	7
Hydrophobe	V44 (S4)	-	-	-	4/4,3	-	-	-	-	-	-	-
	P44 (Bo542)	-	-	-		-	-	-	-	3,4	-	-
	I52	3,8	4,1	4,1	4,1	4,4	4,4	4	4	4,2	4,1	4,1
	F145 (C58/S4)	-	-	-	-	-	-	-	-	3,5		
	W178	3,4	3,7	3,4	3,3/3,5	3,2/3,7	3,2/3,7	3,3/3,7	3,3/3,7	3,4/3,7	3,2/3,7	3,4/3,7
	M372	-	-	3,4	3,8/3,6	-	-	-	-	-	3,7	4,4
	G405	-	-	3,7	-	-	-	-	-	-	-	-
Nombre interaction Protéine:Ligand		18	17	23	22	14	18	14	18	17	19	17

Tableau 5: Récapitulatif des résidus et des interactions hydrogènes et hydrophobes impliqués dans la reconnaissance des différents ligands par la protéine AccA chez les différentes souches. Les chiffres du tableau représentent les distances (en Å) des différentes interactions hydrogènes et hydrophobes. Les slashs permettent de différencier plusieurs interactions réalisées. Les deux colonnes $AccA_{C58}$-G2P et $AccA_{S4}$-G2P correspondent aux deux conformations possible du G2P. En rouge sont repérés les résidus utilisés par $AccA_{C58/S4}$ contrairement à $AccA_{Bo542}$.

Les données individuelles du tableau ne sont pas forcément exploitables, mais l'observation globale de celui-ci m'a tout de même permis de mettre en avant plusieurs résidus intéressants. En effet, les protéines AccA des souches C58 et S4 nécessitent la présence de certains résidus pour la reconnaissance des ligands (A2P/ G2P/ Agrocinopine A/ Agrocinopine D), alors que ce n'est pas le cas pour la souche Bo542. Il s'agit des résidus 44, 145, 284, 402, 372 et 405. Parmi ces résidus seuls les résidus 44, 145 et 402 varient d'une souche à l'autre. Le résidus 44 est une proline pour $AccA_{Bo542}$ alors qu'il s'agit d'une asparagine et d'une valine pour la souche C58 et S4 respectivement. Les phénylalanines 145 et 402 de la souche Bo542 sont des tyrosines pour C58 et S4.

Pour conclure, si l'on recoupe les informations obtenues par l'analyse des résidus de la poche de fixation, on note que les seuls résidus qui diffèrent entre AccABo542 et AccAC58/S4 sont les résidus 44, 55, 141, 145, 402, 407 et 440. Ces résidus sont donc des potentiels candidats pour expliquer la spécificité de reconnaissance des agrocinopines (Figure 31).

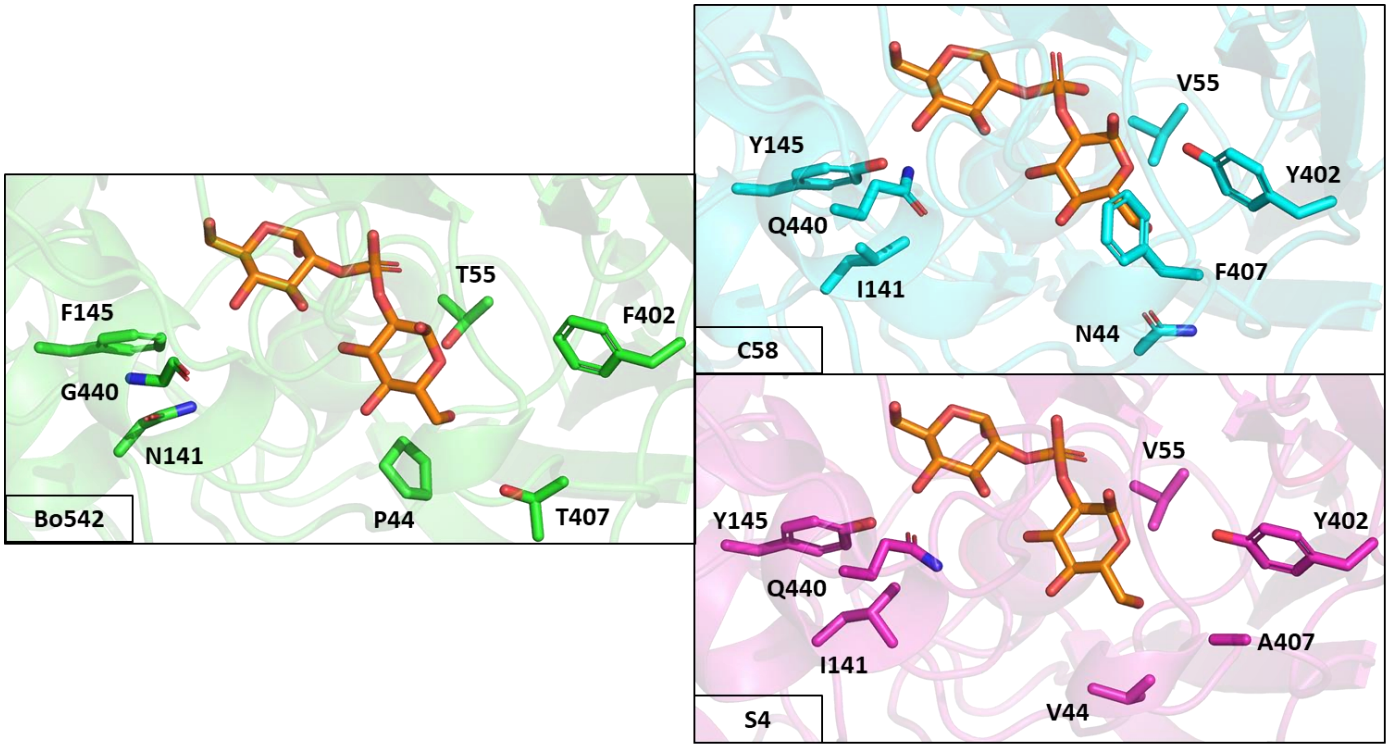


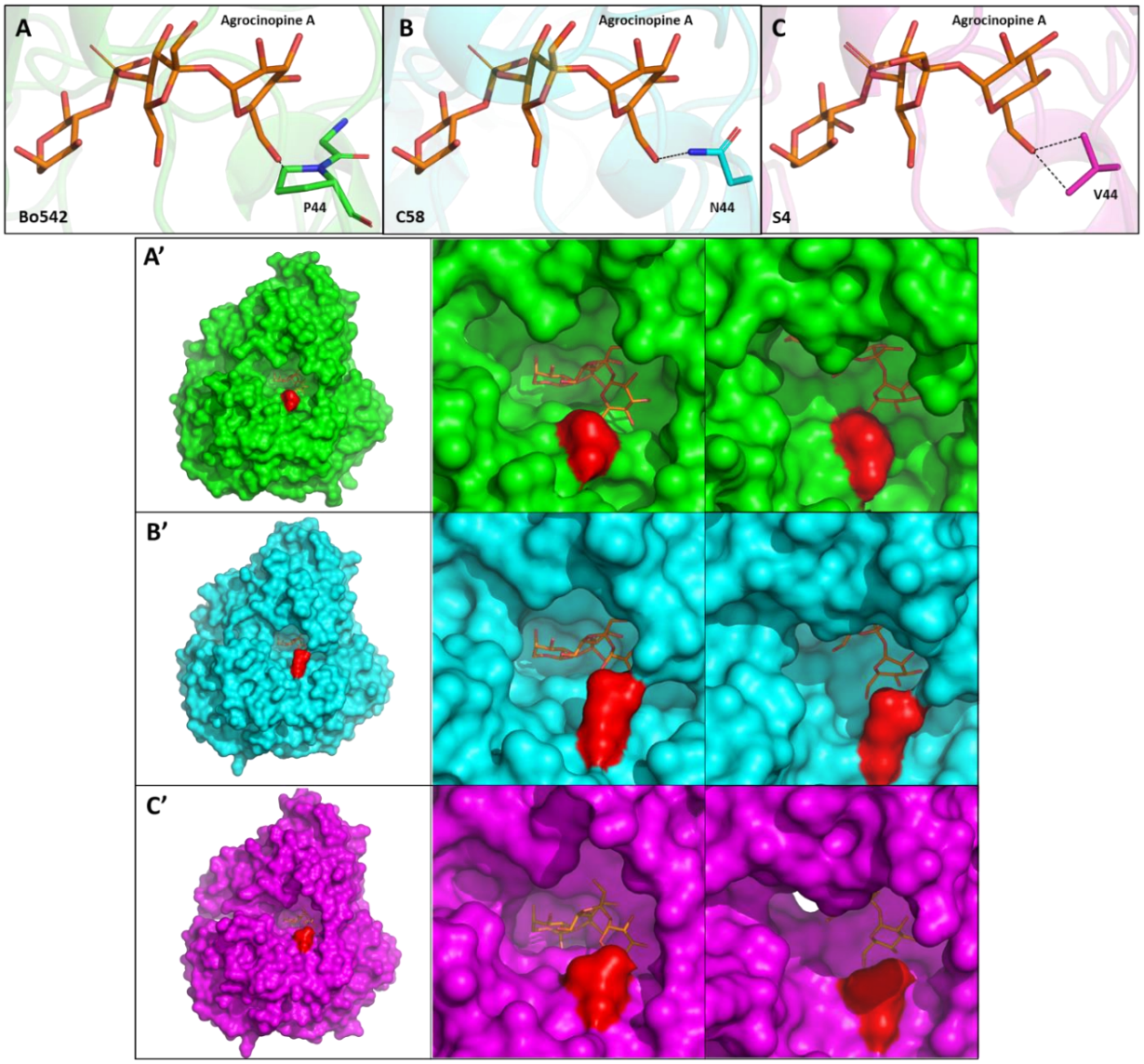
Figure 31: Représentation en bâton, des résidus potentiellement impliqués dans la spécificité de reconnaissance des agrocinopines. Les protéines AccABo542, AccAC58 et AccAS4 sont représentées en vert, bleu et rose respectivement. L'agrocinopine D est représentée en orange. Ces images sont issues des structures dont les tables sont présentes en annexe.

iii. Modélisation de l'Agrocinopine A dans le site de fixation de la PBP $AccA_{Bo542}$

Il a été montré précédemment qu'$AccA_{Bo542}$ ne reconnaissait pas l'agrocinopine A. Afin de pouvoir

poursuivre l'analyse comparative entre les 3 protéines dans le but d'expliquer pourquoi AccA de la souche Bo542 était spécifique, j'ai réalisé une modélisation manuelle de l'agrocinopine A dans le site actif de $AccA_{Bo542}$, par superposition de structure.

J'ai ensuite comparé le modèle obtenu aux structures $AccA_{C58}$: Agrocinopine A et $AccA_{S4}$: Agrocinopine A. De façon très intéressante, la modélisation permet de mettre en évidence un clash stérique entre la proline 44 d'$AccA_{Bo542}$ et l'agrocinopine A, contrairement à l'asparagine et la



valine présentes chez $AccA_{C58}$ et $AccA_{S4}$ respectivement (Figure 32).

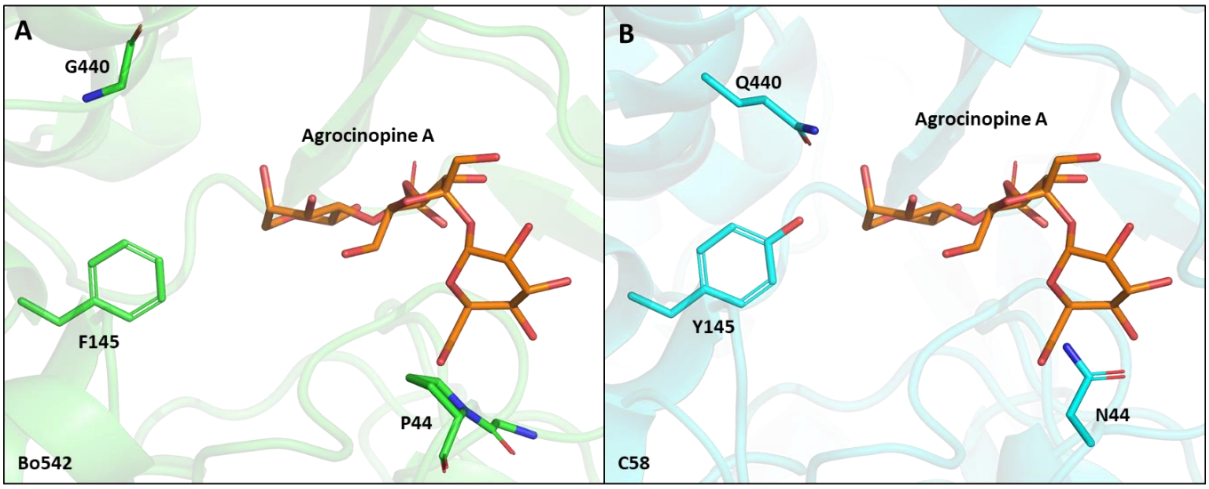
Figure 32: A/B/C Représentation du résidu 44 des souches Bo542, C58 et S4 en présence l'agrocinopine A. Le ligand est représenté en orange sur chacune des images **A :** Modélisation de l'agrocinopine A dans le site de fixation d'$AccA_{Bo542}$. La proline 44 est représentée en vert et est à une distance de 0.8Å. **B :** Représentation de l'agrocinopine A en présence d'asparagine 44 d'$AccA_{C58}$. Le résidu est à une distance de 3.4Å (code PBP : 4ZEB). **C :** Représentation de l'agrocinopine A en présence de la valine 44 d'$AccA_{S4}$ à une de distance de 3.4 Å.
A'/B'/C': Représentation du positionnement de l'agrocinopine A dans le site de fixation d'AccA des différentes

souches. Les 3 images correspondent à 3 prises de vues différentes permettant un aperçu du clash stérique produit lors de la modélisation de l'agrocinopine A dans le site actif d'$AccA_{Bo542}$. **A :** Représentation de la poche de fixation de la modélisation de l'agrocinopine A dans le site de fixation d'$AccA_{Bo542}$. **B :** Représentation du positionnement de l'agrocinopine A dans la protéine $AccA_{C58}$ (code PBP : 4ZEB). **C :** Représentation du positionnement de l'agrocinopine A d'$AccA_{S4}$.

La proline 44 est donc beaucoup plus proche de l'agrocinopine A que les résidus 44 d'$AccA_{C58/S4}$. Sans un réarrangement de l'agrocinopine A, il semblerait donc que la proline 44 d'$AccA_{Bo542}$ pose problème quant au positionnement de l'agrocinopine A dans le site actif d'AccABo542, ce qui pourrait expliquer qu'elle ne reconnaisse pas ce substrat.

iv. Choix de mutants

Les analyses réalisées ont permis de mettre en avant les résidus 44, 55, 141, 145, 402, 407, 440 de la protéine AccA. Le choix des mutations à réaliser afin de valider les hypothèses de spécificité vis-à-vis des ligands s'est finalement porté sur les résidus 44 et 145/440 (Figure 33). Ces choix ont été motivés dans un premier temps par le fait que ces résidus sont différents entre les protéines $AccA_{Bo542}$ et $AccA_{C58/S4}$. Dans un second temps, AccABo542 ne reconnaissant ni l'A2P, ni l'agrocinopine A et étant donné que les résidus 145 et 440 sont proches de la région pyranose-phosphate commune aux deux ligands, j'ai choisi de faire une double mutation. Enfin le clash stérique identifié au niveau de la P44 d'$AccA_{Bo542}$ lors de la modélisation de l'agrocinopine A dans le site actif de la $AccA_{Bo542}$ nous a permis de proposer que ce soit une des raisons de la spécificité



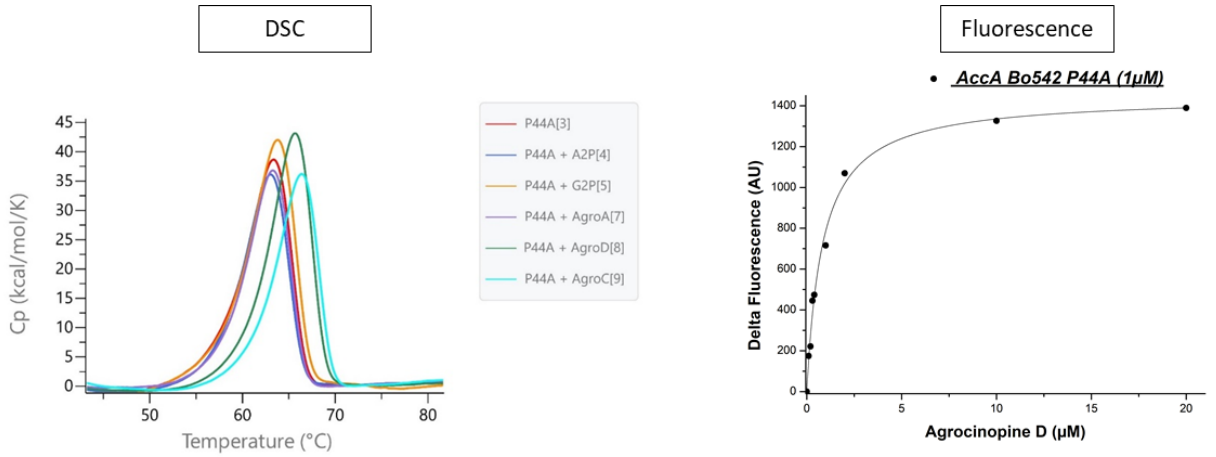
de cette protéine entre les agrocinopines C et D.

Figure 33: Représentation des résidus (44, 145 et 440) choisis comme étant potentiellement impliqués dans la spécificité de reconnaissance des agrocinopines. A : Modélisation de l'agrocinopine A dans le site de fixation d'$AccA_{Bo542}$. **B :** Représentation de l'agrocinopine A d'$AccA_{C58}$.

vii. Test des mutants identifiés

Ainsi les mutants P44N, P44A et F145Y/G440Q de $AccA_{Bo542}$ ont été construits, produits et purifiés (Matériel et Méthodes) et leur capacité à reconnaitre l'agrocinopine A et l'A2P a été testée. Pour ce faire, j'ai réalisé les mêmes expériences (Figure 34) que celles effectuées sur la protéine sauvage $AccA_{Bo542}$ (Figure 26 page 64).

1. *Analyse $AccA_{BO542}$ mutant P44A*



	DSC			Fluorescence		
Protéine	Tm(°C)	ΔTm(°C)	ΔH (kcal/mol)	K_D (µM)		R^2
$AccA_{Bo542}$ P44A apo	62,6		238			
$AccA_{Bo542}$ P44A – A2P	62,36	-0,24	217			
$AccA_{Bo542}$ P44A – G2P	63,02	0,42	263			
$AccA_{Bo542}$ P44A – Agrocinopine A	62,61	0,01	215			
$AccA_{Bo542}$ P44A – Agrocinopine D	65	2,4	247	0.84	± 0.05	0.99
$AccA_{Bo542}$ P44A –Agrocinopine C*	65,69	3,09	205			

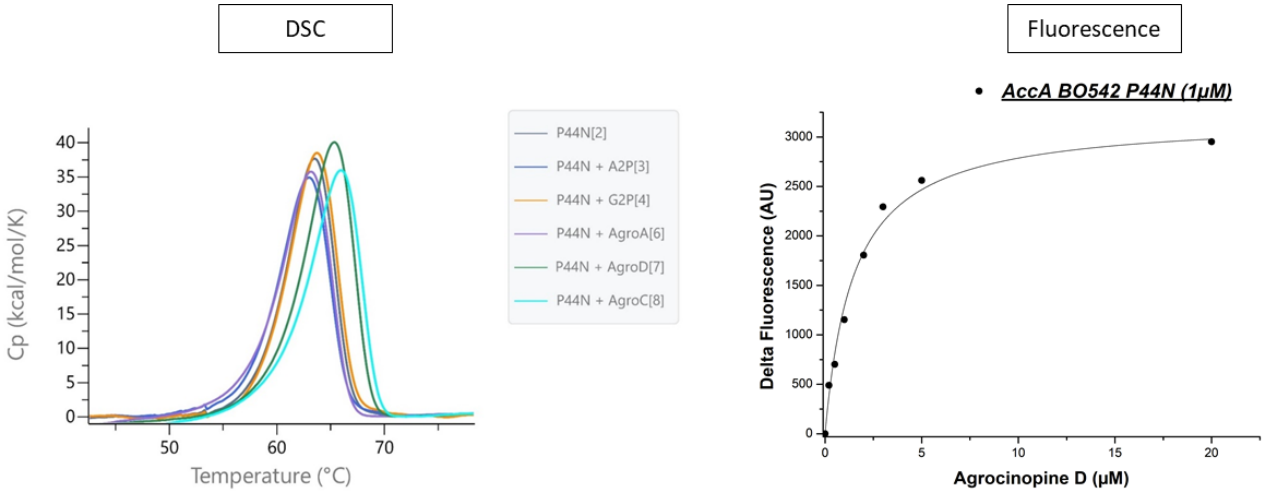
Figure 34 : Résultats des tests d'interactions par DSC et titrage par fluorescence d'$AccA_{Bo452}$ P44A en présence des différents ligands. Les courbes obtenues par DSC représentent la capacité thermique en fonction de la température laissant apparaitre la température de fusion représentative d'une variation de la température dû à la dénaturation de la protéine. Enfin le titrage par fluorescence permet d'obtenir une information concernant l'affinité pour son substrat via la constante de dissociation.

Si l'on compare les résultats de la protéine sauvage à ceux du mutant P44A, on remarque que la température de fusion (Tm) est plus faible pour le mutant sous sa forme *apo*. On peut donc déjà observer que la protéine mutée est moins stable que la protéine sauvage.

De plus, les résultats obtenus montrent une augmentation de la température de fusion en présence d'agrocinopine D et C*. Le très faible changement de fluorescence observé en présence de G2P ne permet pas d'affirmer une reconnaissance du G2P. Des experiences complémentaires seraient nécessaires pour valider la reconnaissance du G2P par la protéine mutée $AccA_{Bo542}$ P44A.

Enfin, l'enthalpie positive observée témoigne d'une absorbtion de chaleur globale dû aux ruptures de liaisons stabilisant la protéine et témoignant de sa dénaturation.

En ce qui concerne l'affinité du mutant P44A pour l'agrocinopine D, on remarque que la constante de dissociation est toujours de l'ordre du micromolaire mais qu'elle est tout de même plus elevée que pour la protéine sauvage. L'affinité du mutant pour l'agrocinopine D est donc plus faible.



	DSC			Fluorescence		
Protéine	Tm(°C)	ΔTm(°C)	ΔH (kcal/mol)	K_D (µM)		R^2
$AccA_{Bo542}$ P44N apo	62,97		210			
$AccA_{Bo542}$ P44N – A2P	62,4	-0,57	209			
$AccA_{Bo542}$ P44N – G2P	63,19	0,22	219			
$AccA_{Bo542}$ P44N – Agrocinopine A	62,5	-0,47	214			
$AccA_{Bo542}$ P44N – Agrocinopine D	64,66	1,69	234	1.48	± 0.09	0.99
$AccA_{Bo542}$ P44N –Agrocinopine C*	65,21	2,24	218			

2. Analyse $AccA_{BO542}$ mutant P44N

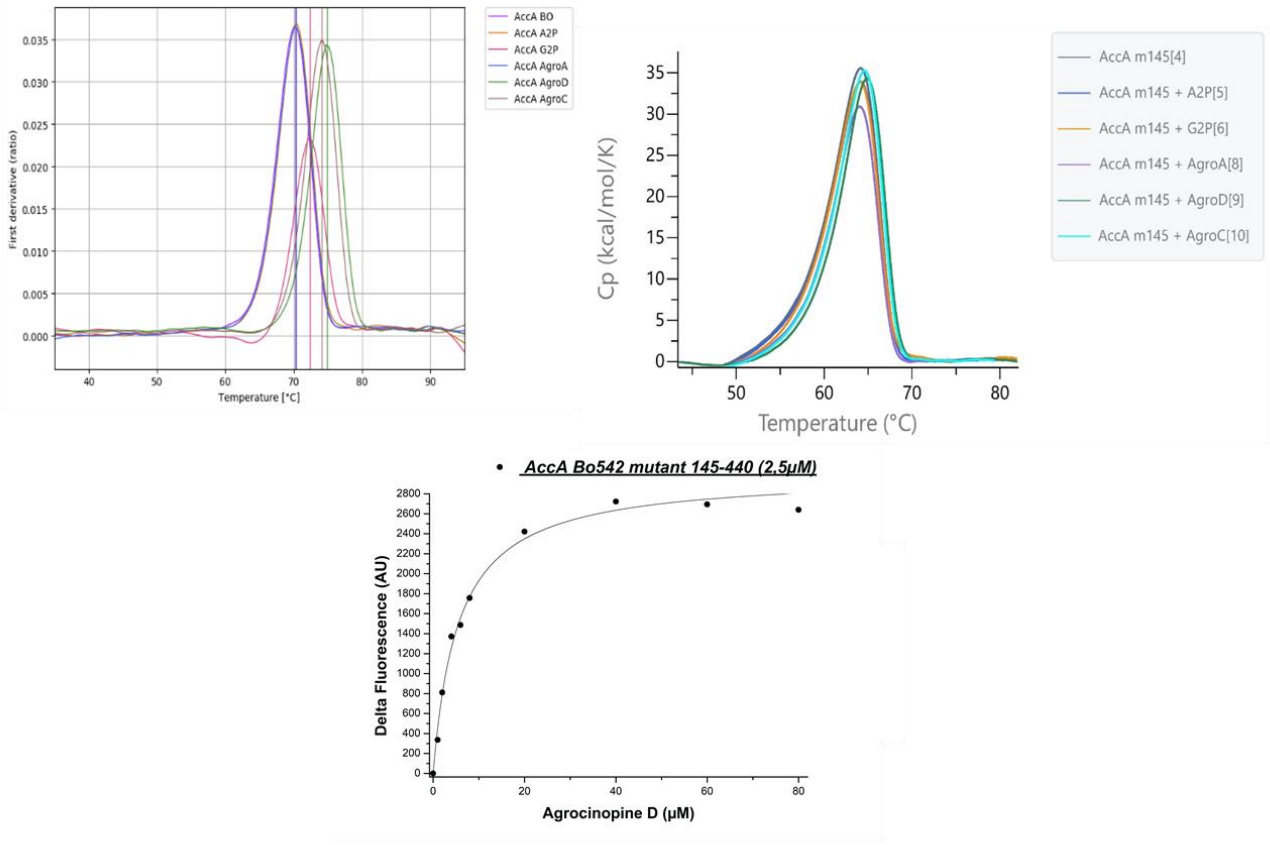
Figure 35 : Résultats des tests d'interactions par DSC et titrage par fluorescence d'$AccA_{Bo452}$ P44N en présence des différents ligands. Les courbes obtenus par DSC représentent la capacité thermique en fonction de la température laissant apparaitre la température de fusion représentative d'une variation de la température dû à la dénaturation de la protéine. Enfin le titrage par fluorescence permet d'obtenir une information concernant l'affinité pour son substrat via la constante de dissociation.

Lorsque l'on observe les résultats du mutant P44N (Figure 35), on aboutit aux mêmes conclusions que pour le mutant P44A, c'est-à-dire, une baisse de stabilité de la protéine mutée par rapport à la sauvage, et une augmentation de sa stabilité en présence d'agrocinopine D et C*. On ne remarque pas non plus de stabilisation en présence de G2P ou du moins très faible et non significative, ce qui ne permet pas de montrer une reconnaissance du G2P par la protéine mutée (significative à partir

d'un degré).

Enfin, la constante de dissociation de la protéine vis-à-vis de l'agrocinopine D est une nouvelle fois de l'ordre du micromolaire mais est un peu plus élevée que celle du mutant P44A.

3. *Analyse $AccA_{BO542}$ mutant F145Y/G440Q*



	Tycho		DSC			Fluorescence		
Protéine	Tm (°C)	ΔTm(°C)	Tm(°C)	ΔTm (°C)	ΔH (kcal/mol)	K_D(µM)		R^2
$AccA_{Bo542}$ F145Y/G440Q apo	71		62,46		255			
$AccA_{Bo542}$ F145Y/G440Q - A2P	71.1	0.1	62,54	0,08	242			
$AccA_{Bo542}$ F145Y/G440Q - G2P	72	1	62,61	0,15	238			
$AccA_{Bo542}$ F145Y/G440Q -Agrocinopine A	71.1	0.1	61,06	-1,4	210			
$AccA_{Bo542}$ F145Y/G440Q - Agrocinopine D	73.9	2.9	63,49	1,03	226	2,21	± 0,13	0,99
$AccA_{Bo542}$ F145Y/G440Q – Agrocinopine C*	73.2	2.2	63,16	0,7	237			

Figure 36 : Résultats des tests d'interactions par DSC, Tycho et titrage par fluorescence d'$AccA_{Bo452}$ F145Y/G440Q en présence des différents ligands.

Pour finir, les résultats du mutant F145Y/G440Q (Figure 36) permettent d'arriver aux mêmes conclusions mis à part ce qui concerne la reconnaissance du G2P. Les valeurs obtenues par la technique Tycho affirment que le mutant reconnait le G2P, l'agrocinopine D et C*. En effet, l'augmentation de la température d'inflexion entre le témoin et la protéine en présence du ligand est

significative à partir d'un degré.

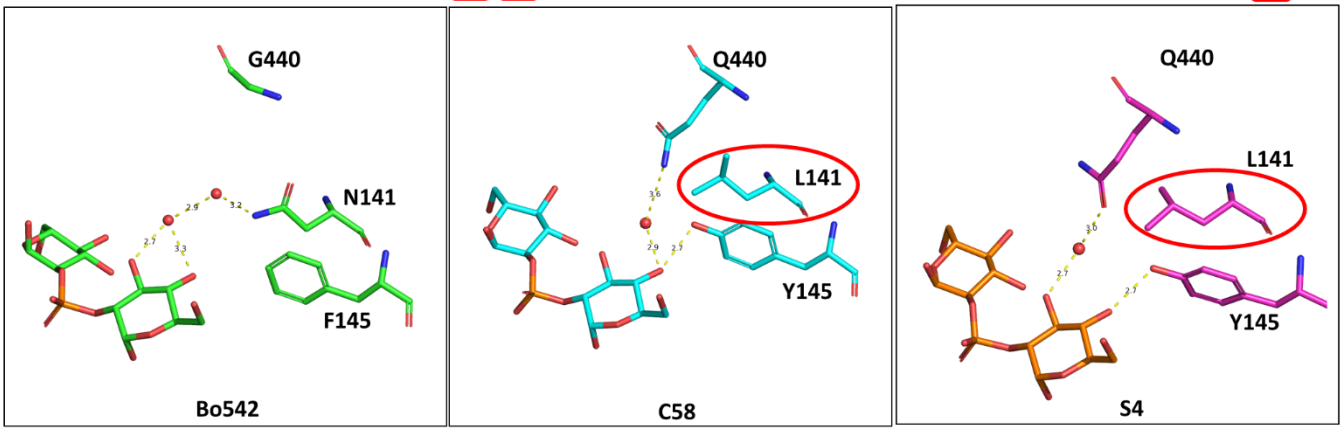
On peut donc en conclure que le mutant F145Y/G440Q est capable tout comme la protéine sauvage de reconnaitre le G2P et l'agrocinopine D et C* et que les mutants utilisés n'ont aucun impact sur la spécificité de reconnaissance puisque les protéines mutantes gardent le même pattern d'interaction que la protéine sauvage. On note tout de même que les températures de fusion ou d'inflexion diminuent légèrement pour les mutants, ce qui témoigne d'une légère déstabilisation des protéines recombinantes à cause des mutations. De plus, la variation de la constante de dissociation indique que les mutants sont légèrement impactés pour la reconnaissance de l'agrocinopine D. L'interaction entre les résidus et le ligand n'est pas directe et laisse à penser que c'est le réseau de molécules d'eau interagissant avec le ligand qui doit être impacté.

viii. Hypothèses concernant la spécificité de la protéine $AccA_{Bo542}$

1. *Révision des choix de mutation d'$AccA_{Bo542/C58/S4}$*

Aux vues de l'absence de reconnaissance de l'A2P et l'agrocinopine A par $AccA_{Bo542}$, mes recherches s'étaient dans un premier temps orientées vers la partie de la protéine fixant la région glucose/arabinose des agrocinopines. J'avais donc choisi de muter les résidus 145 et 440. Mon choix ne s'était pas porté sur le résidu 141 aux vues de sa distance avec le ligand (environ 5Å)

	42	43	44	52	53	54	55	141	145	178	284	372	375	376	402	407	417	418	419	420	421	423	430	434	440	510
Bo542	L	P	P	I	S	N	T	N	F	W	N	M	Y	Y	F	T	N	W	S	N	G	W	T	E	G	E
C58	L	P	N	I	S	N	V	L	Y	W	N	M	Y	Y	Y	F	N	W	S	N	G	W	T	E	Q	E
S4	I	P	V	I	S	N	V	L	Y	W	N	M	Y	Y	Y	A	N	W	S	N	G	W	T	E	Q	
Similitude																										
Zone ligand impactée	G2	G2	G2	G1/G2	G1/G2	G1	G2	G1	G1	G1	G1	G2	P	P/G2	G2	G2	G1	P	P	G1	G1	G1	G1	G1	G1	G1



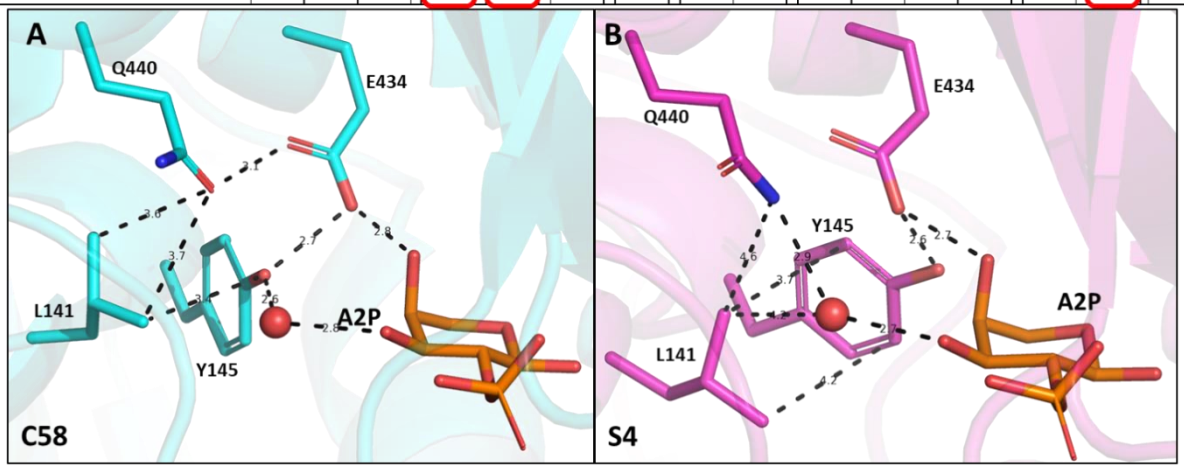
(Figure 37).

Figure 37: Récapitulatif des résidus impliqués dans la fixation de l'agrocinopine D. Les résidus sur fond bleu sont impliqués dans la fixation du ligand par la protéine de la souche concernée alors que ceux sur fond blanc sont des résidus utilisés par les autres souches. Les similitudes sont validées par les cases vertes alors que les différences sont

représentées par une case rouge. Les 3 zones du ligand impactées sont **G1** : Glucose 1 ; **G2** : Glucose 2 ; **P** : Phosphate. Le glucose 1 représente la partie correspondant à l'arabinose des agrocinopines A et B plutôt enfouie. Alors que le glucose 2 représente la partie sucrose des agrocinopines tournée vers l'extérieur du site de fixation. L'image des structures met ensuite en avant les résidus analysés car potentiellement impliqués dans la spécificité d'$AccA_{Bo542}$. La N141 d'$AccA_{Bo542}$ est à grande distance du substrat, mais entre peut-être en interaction avec à travers 2 molécules d'eau. Les G440 et F145 ne partagent pas d'interaction avec le ligand. La L141 des $AccA_{C58/S4}$ pourrait permettre de positionner correctement les chaines latérales de Q440 et Y145 pour qu'elles puissent interagir avec le ligand, grâce à la création d'un environnement hydrophobe.

Cependant, je n'avais pas remarqué l'importance indirecte de ce résidu. En effet la leucine141 des AccA de C58 et S4 est localisée derrière les résidus Y145 et Q440 et permettrait, grâce à la création d'un environnement hydrophobe, un bon positionnement de leurs chaines latérales par rapport au ligand afin de leur permettre de créer les liaisons chimiques nécessaires à son interaction. Cette situation est observée en présence d'agrocinopine D (Figure 37), mais

	52	53	54	141	145	178	284	375	376	418	419	420	421	434	440	510
Bo542	I	S	N	N	F	W	N	Y	Y	W	S	N	G	E	G	E
C58	I	S	N	L	Y	W	N	Y	Y	W	S	N	G	E	Q	E
S4	I	S	N	L	Y	W	N	Y	Y	W	S	N	G	E	Q	
Similitude																
Zone ligand impactée	A	A	A	A	A	A	A	P	P	P	P	A	A	A	A	A



également en présence d'A2P (Figure 38).

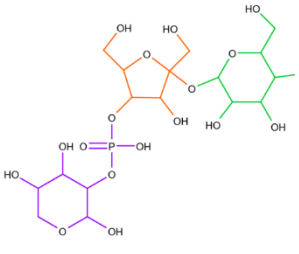
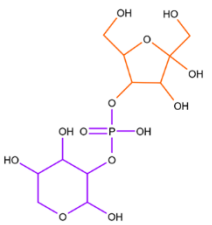
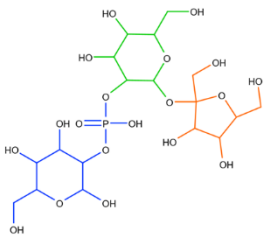
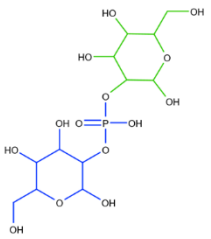
Figure 38 : Récapitulatif des résidus impliqués dans la fixation de l'A2P. Les résidus sur fond bleu sont impliqués dans la fixation du ligand par la protéine de la souche concernée alors que ceux sur fond blanc sont des résidus utilisés par les autres souches. Les similitudes sont validées par les cases vertes alors que les différences sont représentées par une case rouge. Les 2 zones du ligand impactées sont A : Arabinose ; P : Phosphate. **A :** Représentation en bâton de l'agrocinopine A.

Par contre nous voyons que ces 3 résidus sont différents entre AccA de la souche Bo542 et C58/S4. Sur les images on remarque que la leucine 141 est à une distance comprise entre 3.4 et 3.7 pour la souche C58 (Figure 38 A) et comprise entre 3.4 et 4.6 pour la souche S4 (Figure 38 B). Il

semblerait donc qu'un triple mutant (141/145/440) soit finalement nécessaire au rétablissement de la reconnaissance des agrocinopines A par $AccA_{Bo542}$.

c. ***Conclusion***

Les PBP AccA sont des protéines capables de reconnaitre des substrats dans le but de les adresser à un transporteur dans le but de leur internalisation. La reconnaissance de ces substrats est spécifique et met en jeu un grand nombre d'interactions. Les protéines étudiées reconnaissent un substrat spécifique nommé agrocinopine. Dans cette famille de substrat, il existe deux types d'agrocinopines, à savoir les agrocinopines A/B et C/D. La différence entre ces deux types se trouve au niveau de la nature des glucides utilisés. En effet, les agrocinopines A et B sont constituées d'un arabinose lié à un phosphate suivi d'un sucrose, alors que les agrocinopines C et D

Souche	**C58**		**Bo542**	
Agrocinopine synthétisée				
Molécule « précurseur »	**Agrocinopine A** Arabinose-2-phosphate-fructose-glucose	**Agrocinopine B** Arabinose-2-phosphate-fructose	**Agrocinopine C** Glucose-2-phosphate-glucose-fructose	**Agrocinopine D** Glucose-2-phosphate-glucose

sont constituées d'un glucose lié à un phosphate suivit d'un sucrose (Figure 39).

Figure 39 : Représentation des structures chimiques des agrocinopines A, B, C et D pouvant être synthétisés grâce aux gènes *acs* des souches C58 et Bo542 du genre *Agrobacterium*. En violet est représenté le motif arabinose-2-phosphate, en orange le fructose, en vert le glucose et en bleu de glucose-2-phosphate.

De précédentes études ont montré que la souche C58, dont les agrocinopines spécifiques sont les agrocinopines A et B, pouvait malgré tout utiliser *in vitro* et *in vivo* le glucose-2-phosphate grâce à une reconnaissance de ce substrat par la PBP AccA. Je me suis donc intéressé à la souche Bo542, dont les agrocinopines spécifiques sont les agrocinopines C et D, pour savoir si elle pouvait également utiliser les agrocinopines A et B.

Les résultats obtenus ont montré qu'$AccA_{Bo542}$ était spécifique à ses propres agrocinopines et ne permettait pas la reconnaissance d'autres agrocinopines. Au vu des identités de séquence et de structure très élevées sur l'ensemble de la protéine, j'ai concentré mes recherches sur le site de reconnaissance où j'ai identifié quelques résidus différents entre les souches comme étant potentiellement responsables de la spécificité de la protéine. J'ai réalisé des tests en utilisant les mutants P44N, P44A et F145Y/G440Q pour vérifier s'il s'agissait bien des résidus engendrant cette spécificité. Cependant les résultats ont été négatifs et ne m'ont pas permis d'observer une reconnaissance des agrocinopines A et B.

J'ai donc émis plusieurs hypothèses pour tenter d'expliquer cette spécificité. Étant donné que la PBP $AccA_{Bo542}$ ne reconnait ni l'A2P, ni l'agrocinopine A, la présence de résidus impliqués dans la spécificité se trouve soit au niveau du motif pyranose-2-phosphate, soit au niveau du pyranose-2-phosphate et de la partie sucrose.

Dans un premier temps, et aux vues des résultats obtenus, il serait maintenant judicieux de réaliser les même tests en utilisant le triple mutant N141L/F145Y/G440Q. En effet le choix du double mutant réalisé dans cette étude perd un peu de sa crédibilité en absence de la mutation N141L aux vues de son importance dans le maintien des deux autres résidus.

Dans un second temps, si la reconnaissance de l'A2P seul est observée, la reconnaissance de l'agrocinopine A pourrait être permise par un mutant possédant la triple mutation N141L/F145Y/G440Q ainsi que la mutation P44N localisée sur la partie sucrose de l'agrocinopine A. Il serait donc intéressant de tester ces différents mutants et d'observer leur impact sur la reconnaissance des agrocinopines.

Dans le cadre de cette étude, le but était de savoir comment cette spécificité était obtenue aux vues de la forte similitude des protéines des différentes souches. Dans un second temps, il serait intéressant de savoir quel est l'impact de cette spécificité *in planta*. En effet, les différentes souches d'agrocinopines peuvent être amenées à infecter les mêmes plantes. Il serait donc intéressant de savoir s'il y a un avantage ou pas à reconnaitre spécifiquement ses propres agrocinopines avec une affinité un peu plus élevée ou si au contraire il est préférable de reconnaitre plusieurs types d'agrocinopines avec une affinité un peu moindre.

Chapitre 2 :

Etude structure-fonction de la phosphodiesterase AccF

a. Introduction

Une fois internalisées, les agrocinopines ne sont pas utilisables directement par la bactérie : une étape de maturation est nécessaire pour qu'elles puissent servir de nutriments et/ou de molécules précurseurs au mécanisme de communication bactérienne « quorum sensing ». Cette étape de maturation est réalisée par l'enzyme AccF dont le gène est localisé dans l'opéron *acc*, et qui est essentielle pour l'utilisation des agrocinopines (Figure 40) (H. Kim & Farrand, 1998).

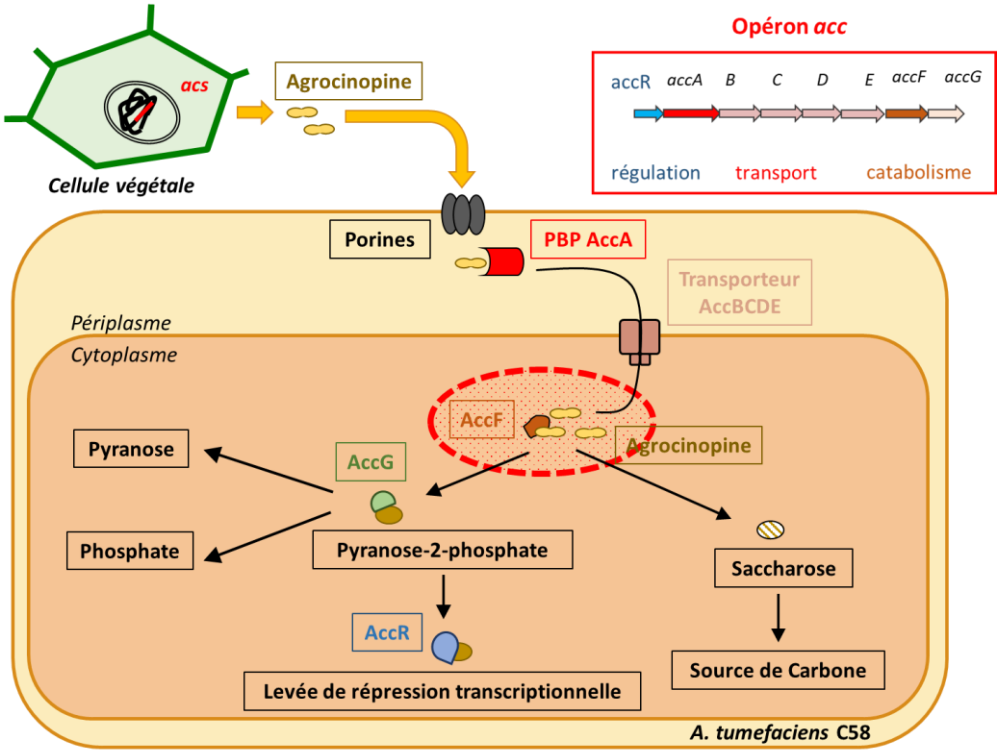


Figure 40 : Représentation schématique de la voie métabolique de l'opéron *acc* ainsi que de la localisation de la protéine AccF dans le cytoplasme d'une bactérie d'*Agrobacterium tumefaciens* (Cercle en pointillé rouge).

Ce projet est la poursuite de travaux préliminaires réalisés par un ancien membre de l'équipe, en collaboration avec une équipe de l'université de Caroline du Nord à Chapel Hill aux Etats Unis, dirigée par John Reader.

Les résultats préalablement obtenus ont permis de résoudre la structure de la protéine AccF de la souche C58 d'*Agrobacterium tumefaciens* et ont confirmé que cette protéine possède un repliement caractéristique de type triosephosphate isomerase. Cette enzyme a besoin de la présence d'un ion métallique divalent pour être active et son site catalytique est semblable à celui de certaines glycérophosphodiestérases. Les travaux réalisés n'ayant pas abouti à l'obtention de la structure de la protéine en complexe avec différents ligands, j'ai décidé d'étudier cette protéine chez la souche Bo542 d'*A. tumefaciens*. Le choix de cette souche a été motivé par le fait qu'elle utilise les agrocinopines de façon comparable à la souche C58, et qu'il existe une identité de séquence

élevée entre les deux protéines AccF (68% d'identité de séquence). Ces deux souches partagent le même biovar (n°1) utilisé lors de la classification des souches, mais diffère au niveau du groupe génomique (G8 pour la souche C58 et G1 pour la souche Bo542).

J'ai étudié la structure et la fonction de cette protéine dans le but de comprendre et d'expliquer son mécanisme réactionnel. J'ai tout d'abord résolu sa structure cristallographique par remplacement moléculaire. De nombreux essais de co-cristallisation et d'infiltration dans des cristaux m'ont permis également d'obtenir la structure de la protéine en complexe avec différents produits et un intermédiaire réactionnel. Cependant je n'ai pas obtenu de structure en complexe avec le substrat. Pour observer le substrat dans le site actif je me suis orienté vers la modélisation.

J'ai d'autre part mis en place un protocole permettant de tester l'activité enzymatique afin de caractériser l'enzyme et de mettre en évidence l'implication de certains résidus dans l'activité catalytique et/ou dans la fixation au substrat. J'ai pour cela sélectionné cinq mutants mettant en jeu des résidus identifiés (grâce aux structures obtenues) comme étant potentiellement importants pour le mécanisme réactionnel ou la fixation du substrat. Mes résultats ont permis d'émettre des hypothèses sur le mécanisme réactionnel de l'enzyme.

b. Résultats

i. Clonage, expression et purification

Le gène codant pour la protéine AccF de la souche Bo542 (sauvage) a été synthétisé avec une étiquette codant pour 6 histidines et cloné dans un vecteur de surexpression bactérien par l'entreprise Genescript (Matériel et Méthodes, Tableau 12 et 13 page 135), ainsi que ceux des mutants H16A, H60A, H174A, E195A, H197A. Le choix de ces mutants a été motivé par la littérature en ce qui concerne les résidus H16A et H60A (Shi et al., 2008), et après une analyse structurale personnelle suite à l'obtention de la structure, pour les mutants H174A, E195A, H197A.

L'expression des protéines $AccF_{Bo542}$ (type sauvage/mutants) a été mise au point pour déterminer les meilleures conditions d'expression (Matériel et Méthodes). La souche B21(DE3) permet une surexpression optimale de la protéine $AccF_{Bo542}$ type sauvage et mutant H174A, alors que c'est la souche Rosetta la plus adaptée pour l'expression des mutants H16A, H60A, E195A et H197A. La température optimale pour l'expression de chaque protéine a été référencée dans le tableau ci-dessous (Tableau 6). Les étapes nécessaires à l'expression et la purification des protéines sont décrites dans le Matériel et Méthodes. La purification des protéines a été possible grâce à

l'utilisation d'une chromatographie d'affinité sur un métal immobilisé (IMAC) ainsi qu'une chromatographie d'exclusion de taille (SEC). A la suite de cette étape de purification, du calcium à une concentration finale de 1mM a été ajouté aux fractions récupérées. L'ajout de calcium vise à stabiliser la protéine qui a tendance à s'agréger en absence d'ion métallique. Toutes les protéines (sauvage/mutées) ont été produites mais on remarque que la quantité de protéine purifiée en fonction du litre de culture diminue fortement pour les protéines mutées.

Protéine	AccF sauvage	AccF mutant				
Mutation	-	H16A	H60A	H174A	E195A	H197A
$mg_{protéine\ purifiée}/L_{culture}$	68,4	5	4	1,89	9,21	6,03
Bactérie	Bl21 (DE3)	Rosetta	Rosetta	Bl21 (DE3)	Rosetta	Rosetta
Température d'expression	16°C	16°C	16°C	28°C	16°C	16°C

Tableau 6 : Récapitulatif de la quantité des protéines AccF sauvage et mutantes obtenues

La colonne utilisée lors de la chromatographie d'exclusion de taille est une Hiprep Superdex 200 26/60, permettant la séparation de protéine allant de 10 à 600 kDa. Le volume d'élution de la protéine lors de cette étape de purification est de 210 ml, ce qui correspond à un poids moléculaire d'environ 60 kDa. La forme relativement globulaire de la protéine et le volume d'élution observé m'ont permis de proposer que AccF est un dimère en solution (Figure 41). La forme monomérique de la protéine à 30 kDa est ensuite retrouvée sur gel SDS-PAGE, qui permet de suivre le niveau de pureté des protéines au fil des étapes (Figure 41). Des 5 mutants utilisés, on remarque que c'est le mutant H174A qui est le moins bien exprimé. Le résidus H174 se situe au niveau de l'interface dimérique et interagit avec le résidus H174 du second monomère. Sa mutation engendrerait donc une déstabilisation du dimère protéique impactant la stabilité de la protéine. Pour chacune des protéines, la fraction finale concentrée contient 3 bandes dont 1 très majoritaire aux alentours de 30 kDa (Figure 41). Lors d'expériences de cristallogenèse ou d'enzymologie, l'ion métallique coordonné par la protéine a été modifié par ajout d'EDTA 1mM, concentration puis remise en tampon en présence d'un ion métallique pertinent (voir Matériel et méthodes).

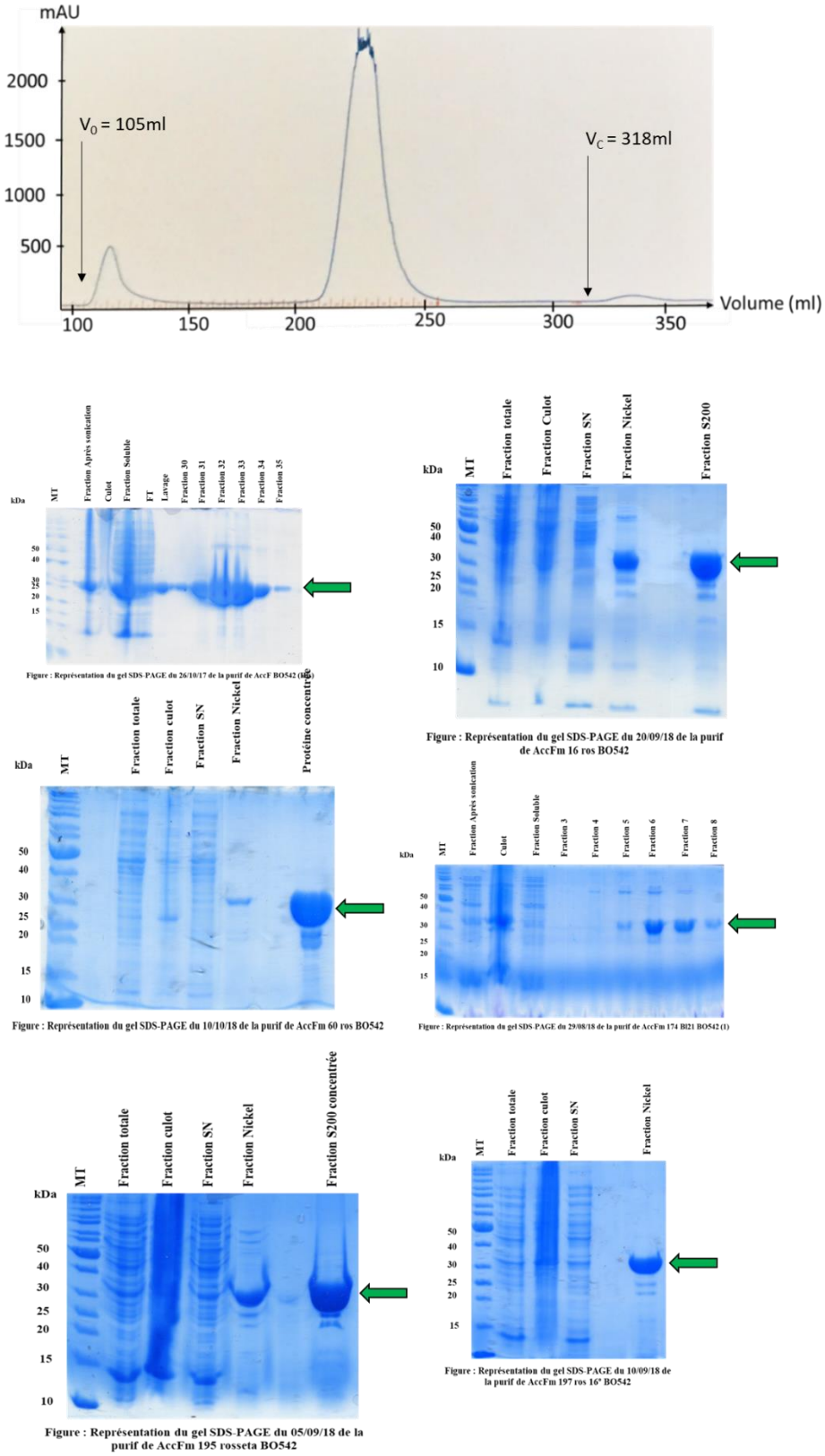


Figure 41 : Purification des AccF sauvage et mutants. Haut : Chromatogramme de l'étape de purification par chromatographie d'exclusion de taille de la protéine $AccA_{Bo542.}$ La calibration de la colonne a été réalisée par l'équipe. Le volume d'élution des marqueurs utilisés sont volume mort (V0) = 105 ml, volume colonne (VC) = 318 ml, $V_{Thyroglobulin}$ (670 kDa) =114 ml, $V_{aldolase}$ (158 kDa) = 173 ml, $V_{ovalbumine}$(44 kDa) = 216 ml, $V_{Myoblobulin}$(17 kDa)=247 ml.
Bas : Gels SDS-PAGE de suivi de la purification des différentes protéines $AccF_{Bo542}$ (sauvage et mutantes). (**MT** : Marqueur de Taille, **FT** : Fraction totale, **L** : Lavage, **C** : Culot, **S** : Surnageant, **E** : Elution Colonne Nickel, **Fx** : Fraction S200, **FC** : Fraction concentrée finale).

ii. Cristallogenèse

La cristallisation d'AccF a été réalisée comme décrit en Matériel et Méthodes. Au bout de quelques jours, de nombreux cristaux ont été obtenus à la suite de la réalisation de ces cribles de cristallisation (Figure 42). Certaines conditions ont été reproduites manuellement afin d'optimiser la taille des cristaux, ce qui a permis d'obtenir, après la diffraction du faisceau de rayons X de la ligne Proxima 1 du synchrotron SOLEIL, les structures de la protéine en complexe avec différents ligands.

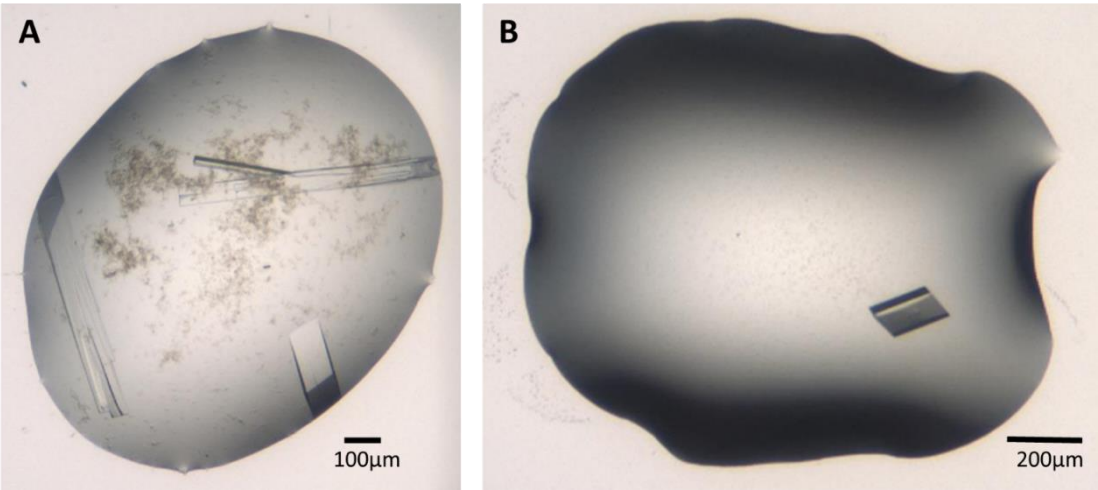


Figure 42 : Exemples de cristaux obtenus grâce aux kits de criblages Qiagen A : 2M Sulfate d'ammonium, 0.1M Tris HCL pH 8.5 – **B :** 30% Peg4000, 0.1M Tris HCL pH8.5, 0.2M Acétate de sodium

Durant cette étape d'optimisation, il a été nécessaire de faire varier le pH, les concentrations en précipitant, les concentrations en protéine et les volumes (ratio $V_{protéine}/V_{puit}$). Au bout de quelques jours, de nombreux cristaux sont apparus dans les gouttes de cristallisation réalisées. Pour obtenir la forme complexée de l'enzyme avec le substrat et/ou le produit, j'ai réalisé de la co-cristallisation (mélange protéine et ligand avant cristallisation) ou du trempage (cristal de la protéine placé dans une solution avec le ligand) comme décrit dans la partie Matériel et méthodes.

Une fois obtenus, les cristaux ont été trempés dans la solution de cristallisation additionnée d'un cryoprotectant pour être congelés et stockés (Matériel et méthodes). Selon les conditions de cristallisation, la nature du cryoprotectant peut varier en suivant l'article de McFerrin de 2002 (McFerrin & Snell, 2002).

iii. Diffraction aux rayons X et résolution des structures

Les cristaux congelés et stockés dans l'azote liquide ont été placés dans le faisceau de rayons X des lignes Proxima 1 ou 2 (Synchrotron SOLEIL, Gif-sur-Yvette). Les données collectées ont été traitées avec le programme XDS (Kabsch, 2010) (Matériel et méthodes). Suite au traitement des données, les groupes d'espace ont été déterminés. Afin de résoudre les structures, le phasage par remplacement moléculaire a été réalisé. Le modèle utilisé pour le remplacement moléculaire a été la structure de la protéine AccF de la souche C58 précédemment résolue au laboratoire. Une fois les structures résolues, la suite de logiciel BUSTER a été utilisée afin de réaliser plusieurs cycles d'affinement pour améliorer les coordonnées des atomes ainsi que les cartes de densité électronique. Entre chaque cycle d'affinement, les structures ont été manuellement modifiées avec le logiciel COOT afin de corriger les erreurs entre les modèles et les cartes de densité électronique. Pour toutes les autres structures obtenues c'est la $AccF_{Bo542}$ qui a été utilisée comme modèle pour le remplacement moléculaire.

Les tables cristallographiques sont présentées en annexe.

iv. Analyse et comparaison structurale de la protéine AccF.

1. *Les structures obtenues*

AccF est une phosphodiesterase capable de cliver certaines molécules spécifiques par reconnaissance d'une liaison phosphodiester. La première structure de cette phosphodiesterase capable de reconnaitre et de cliver les agrocinopines a été obtenue par un ancien étudiant de l'équipe. Il a utilisé comme modèle de remplacement moléculaire deux structures de glycerophosphodiesterases possédant un pourcentage d'identité de séquence relativement faible de 31 et 34% (Tableau 7) (Shi et al., 2008) (Rao 2006). Malgré ce faible pourcentage, l'utilisation de ces structures a permis de résoudre celle de la protéine AccF de la souche C58 d'*A. tumefaciens*. Malheureusement, il n'a réussi à obtenir que la structure de la protéine sauvage sous la forme *apo*.

Organisme	***Thermoanaerobacter tencongensis***	***Agrobacterium tumefaciens***
Protéine	GDPQ	GPDQ
Code PDB	2PZO	1ZCC

Identité de séquence (%)	34	31

Tableau 7 : Structures utilisées pour la résolution de la phase par remplacement moléculaire de la protéine $AccF_{C58}$.

Dans le but d'obtenir des structures d'une AccF en complexe avec des substrats et/ou des produits, j'ai décidé de travailler avec un homologue d'une autre souche. Mon choix s'est porté sur AccF de la souche Bo542 d'*A. tumefaciens*, avec laquelle j'ai pu obtenir de nombreuses structures (une centaine de structure dont plusieurs identiques), à des résolutions allant de 1,31 Å à 1,94 Å (Tableau 8). La première structure a été résolue par remplacement moléculaire en utilisant la phosphodiesterase AccF de la souche C58 comme modèle, et elle a ensuite servi pour résoudre les autres structures. J'ai pu obtenir différentes structures de la protéine sauvage sous forme *apo* ou complexée/*holo* (Sulfate, Glucose, Glucose-2-phosphate (G2P), Intermédiaire réactionnel), ainsi que des structures des 5 mutants sous la forme *apo* (mutants H16A, H60A, H174A, E195A, H197A). Le choix de ces mutations est expliqué et discuté par la suite page 104

Protéine	**AccF Sauvage**								
Mutation	-								
Ligand	-		Sulfate		Glucose		G2P		Intermediaire réactionnel
Ion	Ca^{2+}	Mn^{2+}	Ca^{2+}	Mn^{2+}	Ca^{2+}	Mn^{2+}	Ca^{2+}	Mn^{2+}	Mn^{2+}
Résolution (Å)	1.45	1.79	1.70	1.31	1.33	1.48	1.30	1.95	1.59
RMSD (Å)	-	0.13	0.30	0.42	0.05	0.28	0.06	0.26	0.38

Protéine	**AccF Mutants**				
Mutation	H16A	H60A	H174A	E195A	H197A
Ligand	-	-	-	-	-
Ion	Ca^{2+}	-	Ca^{2+}	-	-
Résolution (Å)	1.70	1.40	1.50	1.40	1.40
RMSD (Å)	0.30	0.41	0.22	0.38	0.37

Tableau 8 : Récapitulatif des structures de la protéine $AccF_{Bo542}$ obtenues (*apo* et complexées). Les différents détails concernant les jeux de données et conditions de cristallisation de chaque complexe sont regroupés dans les tables cristallographiques page 150.

2. *Description structurale de la protéine.*

La protéine cristallisée est sous forme dimérique stabilisée par 26 interactions hydrophobes, 17 interactions hydrogènes mettant en jeu 11 résidus (D65, N71, N76, F152, D153, F154, L172, H174, S175, R176, Y230) dont 2 ponts salins (D153 et R176) (Figure 43).

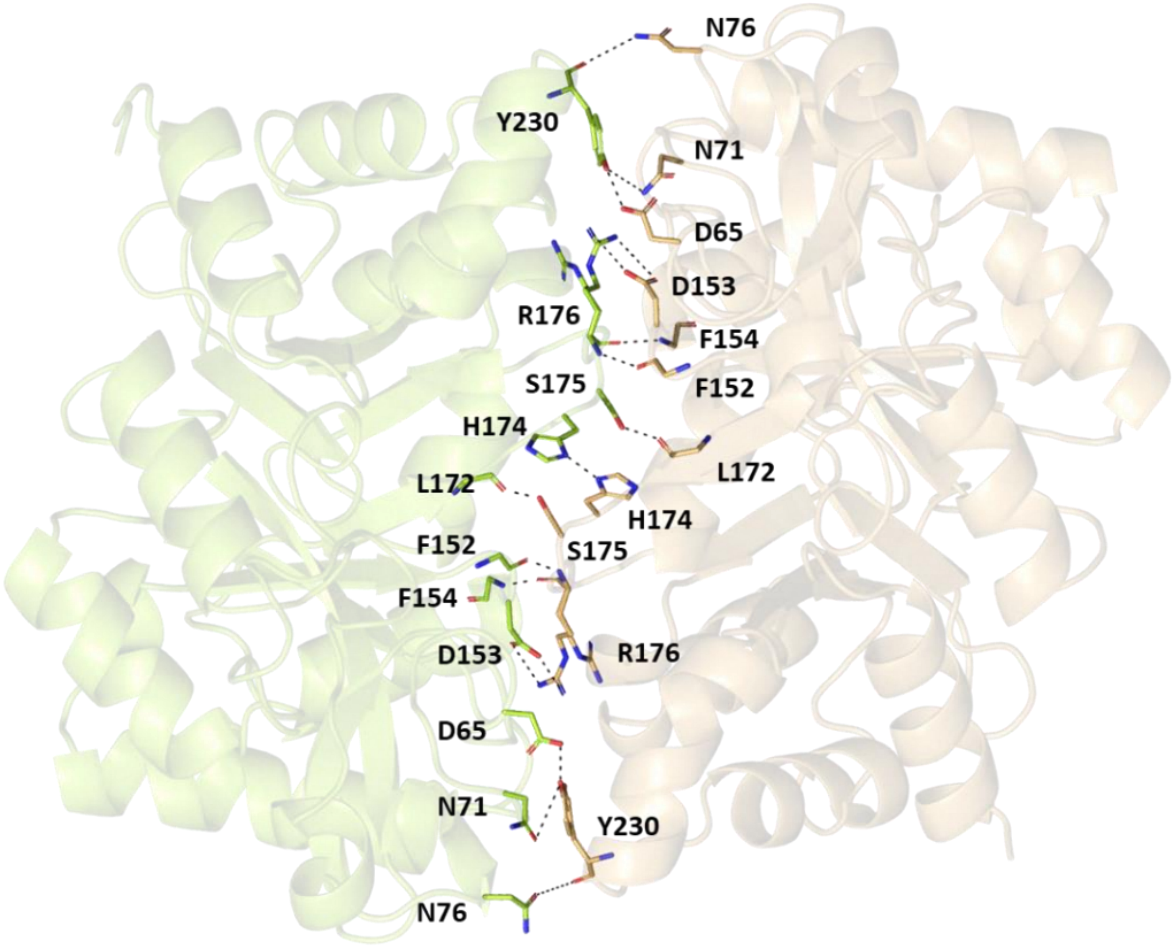


Figure 43: Représentation des interactions hydrogènes et ponts salins stabilisant le dimère AccF, avec en vert et orange les monomères 1 et 2.

Chaque monomère est composé de deux domaines. Quelle que soit sa forme (*apo* ou *holo*), la structure est conservée comme le montre les RMSD de 0.27 Å en moyenne entre les différentes structures. Le premier domaine, composé des résidus 1 à 47 et 102 à 267 possède un repliement de type TIM barrel (tonneau) (pour Triosephosphate IsoMerase), composé de huit brins bêta centraux entourés de huit hélices alpha et de deux hélices 3_{10} (Figure 44/vert). Le second domaine allant des résidus 48 à 102 est composé de deux brins bêta et de deux hélices 3_{10} reliés par de longues boucles (Figure 44/ fuchsia).

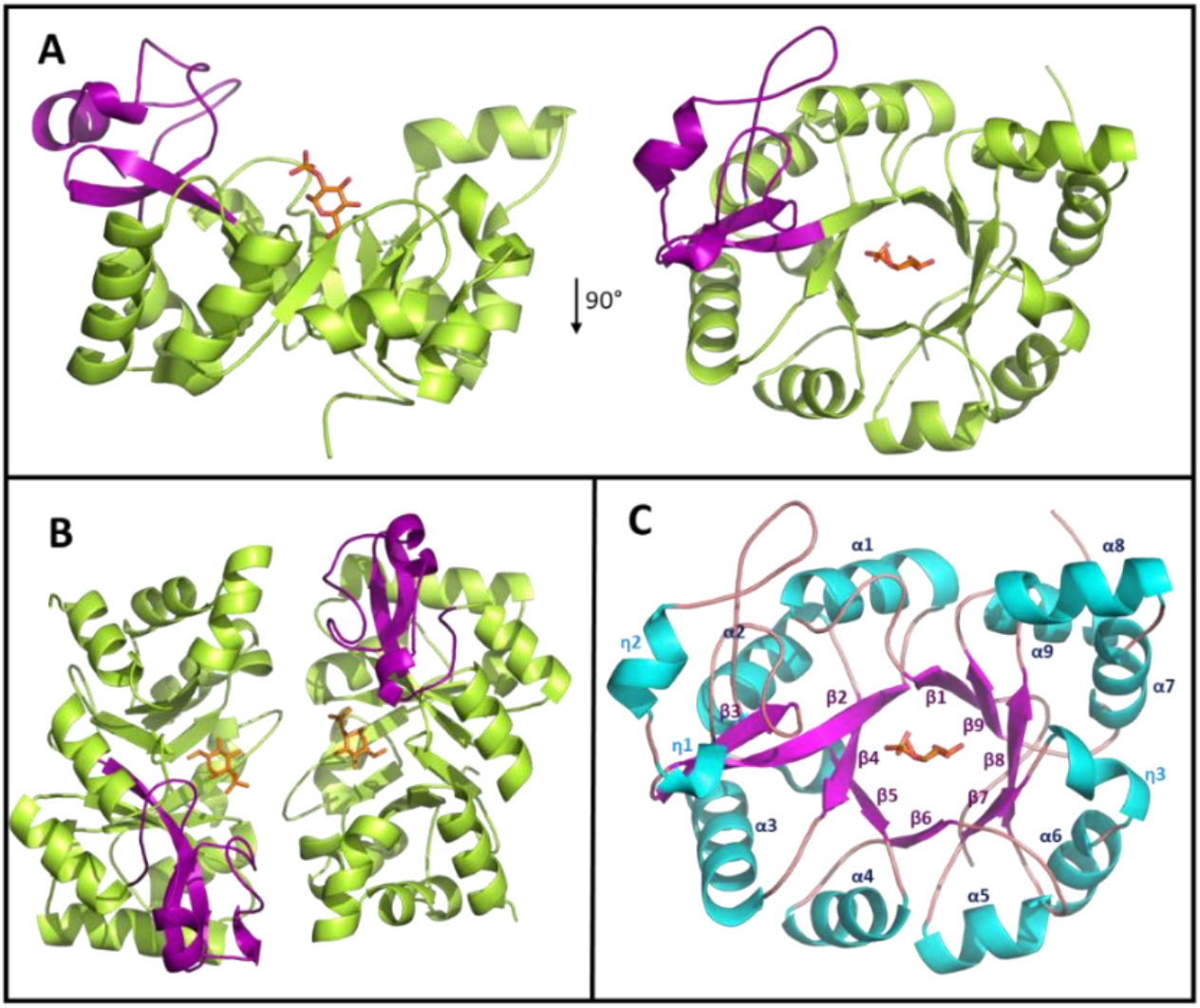


Figure 44 : Représentation en ruban de la structure de la protéine AccF de la souche Bo542. A et B : Représentation du monomère et du dimère de $AccF_{Bo542}$ en complexe avec le G2P. Les deux domaines sont en vert et fuchsia et le G2P en sticks orange. **C** : Eléments de structures secondaires du monomère. Les hélices alpha 3_{10}, les feuillets beta et les boucles sont représentés en cyan, violet et en beige respectivement. Le G2P est quant à lui représenté en orange.

La superposition structurale de la protéine AccF de Bo542, C58 et S4 avec des glycerophosphodiesterases de la même famille met en évidence une très forte conservation du repliement protéique malgré un faible pourcentage d'identité de séquence, et donc une potentielle similitude fonctionnelle entre les protéines (Tableau 9).

Protéine (souche/code PDB)	$AccF_{C58}$	2OOG	2PZ0	3KS6	3L12	3QVQ	4R7O
% Identité de séquence	68	27.3	32	20.4	21.7	24	23.4
ΔRMSD (Å)	0.61	1.8	1.58	2	1.97	2.04	1.77

Tableau 9: Comparaison du RMSD et des identités de séquence entre $AccF_{Bo542}$ et différentes phosphodiesterases. Ces comparaisons se sont faites par rapport à la protéine AccF de la souche Bo542. (2OOG : *Staphylococcus* aureus ; 2PZ0 : *Thermoanaerobacter* tengcongensis ; 3KS6 : *Agrobacterium* tumefacien C58; 3L12:

Silicibacter pomeroyi; 3QVQ: Oleispira antarctica ; 4R7O : Bacillus anthraci)

L'alignement des séquences (Figure 45 A) et la superposition structurale de ces enzymes mettent en évidence la conservation et donc l'importance de certains résidus. De plus, les positions spatiales de ces résidus ne varient pas entre les différentes structures et restent conservées en présence ou en absence de différents ligands et ions métalliques (Figure 45 B). Enfin, on peut donc en déduire une certaine importance des deux histidines de AccF (H16 et H60) (Figure 45 C), et l'implication dans la coordination d'un ion métallique des trois autres résidus (E45, D47 et E120)(Figure 45 C) (Shi et al., 2008).

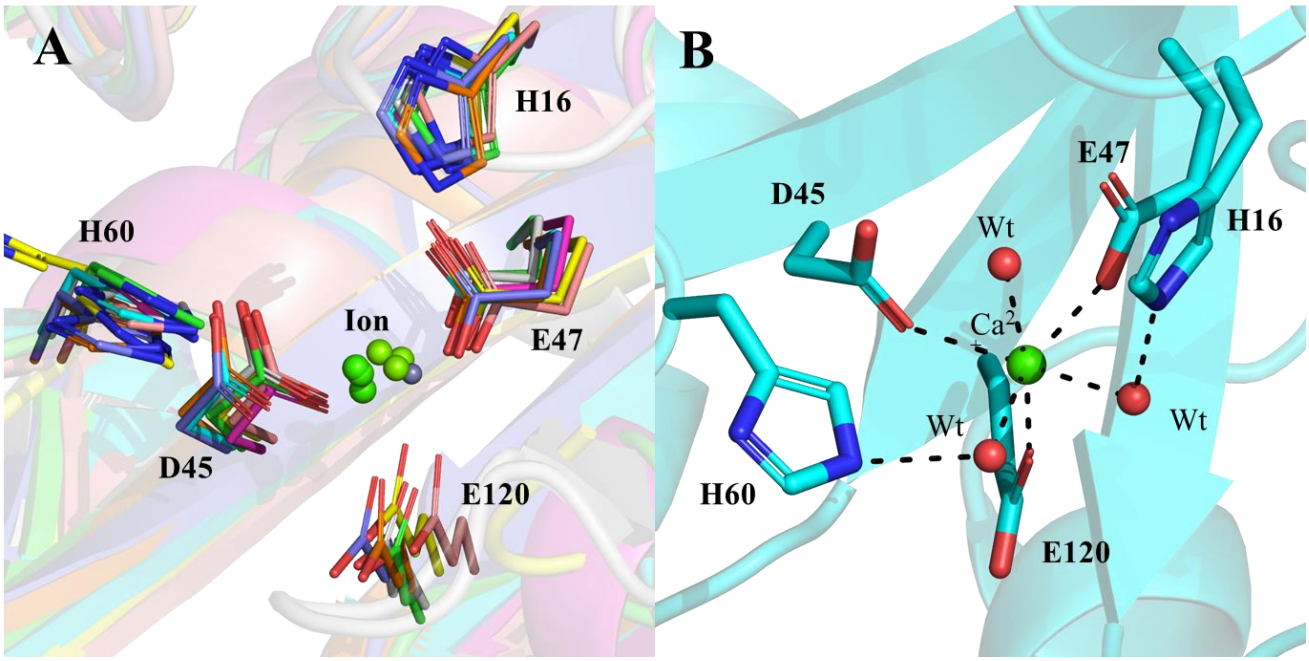
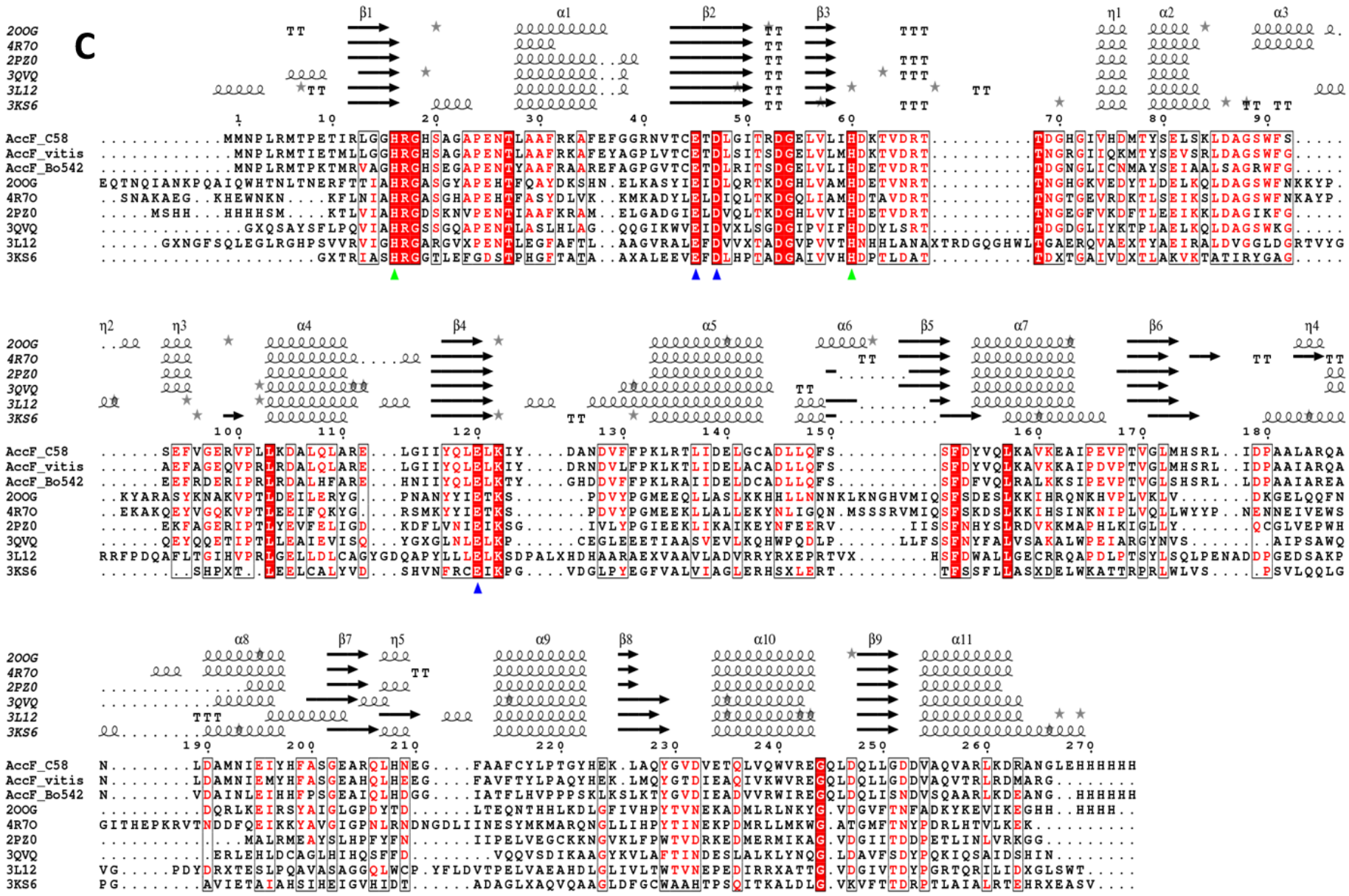


Figure 45 : Conservation du centre catalytique de AccF avec ses homologues structuraux. A: Superposition des sites catalytiques des protéines utilisées précédemment pour l'alignement de séquences. **B:** Site catalytique de la protéine $AccF_{Bo542}$ stabilisée par un ion métallique de calcium en vert et par 3 molécules d'eau (Wt) en rouge. **C:** Ci-après, alignement de séquences des protéines AccF des souches C58, Bo542 et S4 d'*Agrobacterium* avec les glycerophosphodiesterases les plus proches en terme d'identité de séquence (Codes PDB : 2OOG, 4R7O, 2PZ0, 3QVQ, 3L12 et 3KS6). Au-dessus de l'alignement de séquences sont représentées les structures secondaires des protéines connues et soumises à la PDB. Les cônes bleus représentent les résidus impliqués dans la coordination du métal, et ceux en vert les résidus impliqués dans la réaction catalytique.


C
200G
4R7O
2PZ0
3QVQ
3L12
3KS6
AccF_C58
AccF_vitis
AccF_Bo542

3. *Description de l'interaction protéine-ligand*

Les différents ligands utilisés pour obtenir les structures complexées de la protéine $AccF_{Bo542}$ sont représentés dans la Figure 46. Le choix de ces ligands a été fait grâce à la ressemblance de chacun d'eux avec le substrat principal : l'agrocinopine D. Ainsi les co-cristallisations de $AccF_{Bo542}$ ont été faites avec du glucose, du glucose-2-phosphate et l'intermédiaire réactionnel (issu de la protéolyse spontanée de l'agrocinopine D dans le cristal, voir page 103).

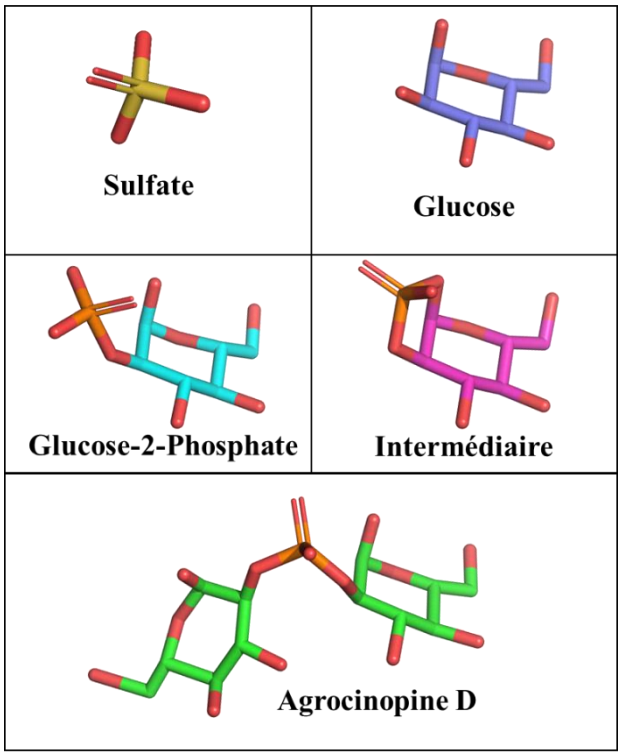


Figure 46 : Représentation des structures de l'agrocinopine D et des 4 différents ligands obtenus en complexe avec $AccF_{Bo542}$. L'atome de sulfate est représenté en jaune, en rouge les atomes d'oxygène, en orange l'atome de phosphore et enfin de couleurs différentes les carbones (bleu, cyan, fuchsia, vert). L'agrocinopine D est composée de deux glucoses, tous deux liés en position 2 à un phosphate (Glucose-2-phosphate-2-glucose).

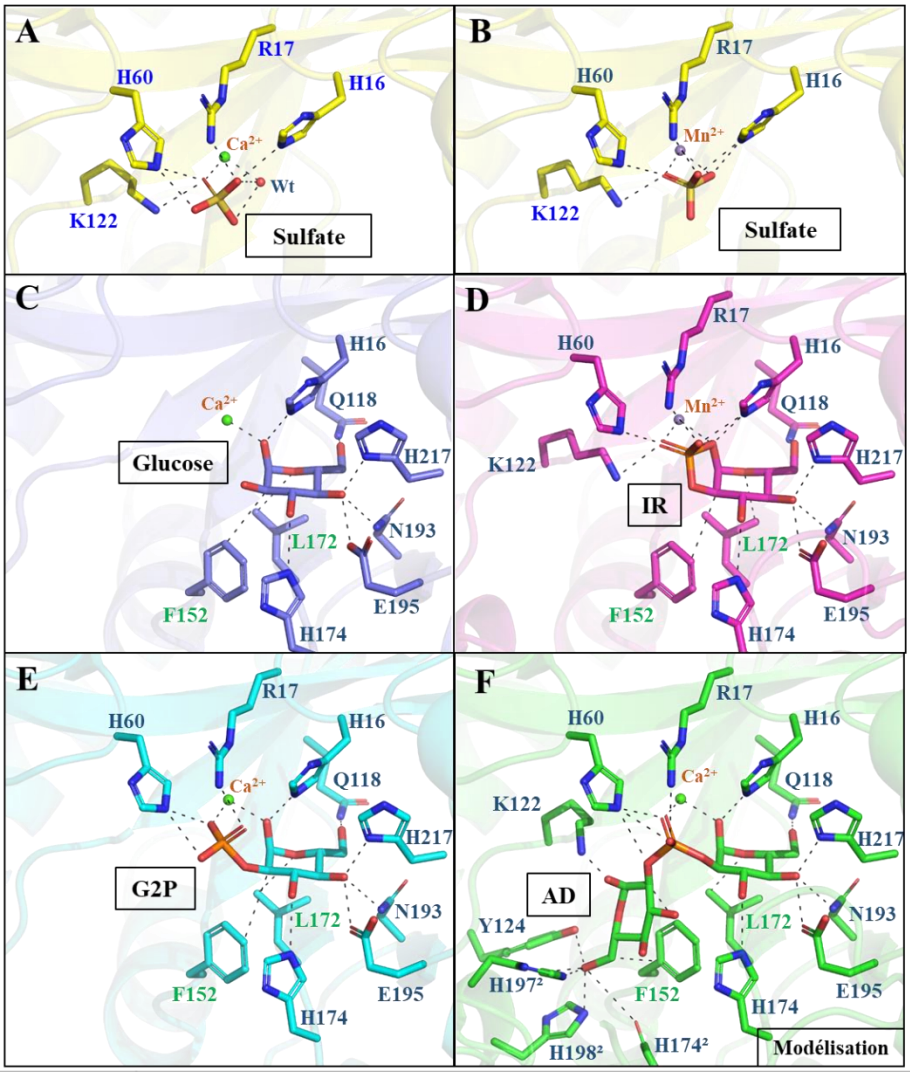
Lors de l'analyse plus approfondie concernant ces différentes structures, j'ai pu identifier plusieurs résidus conservés comme étant impliqués dans la fixation du sulfate, du glucose, du glucose-2-phosphate et de l'intermédiaire réactionnel (Figures 47). Un important réseau de molécules d'eau est impliqué dans la fixation des différents ligands mais n'a pas été représenté.

Le sulfate complexé à $AccF_{Bo542}$ peut occuper deux positions en fonction de l'ion métallique coordonné. En effet, comme on peut le remarquer sur la Figure 47 A, fixé au calcium (Ca^{2+}), le sulfate est stabilisé par des liaisons hydrogènes avec les résidus H16, R17, H60 et K122 d'$AccF_{Bo542}$ ainsi que par une molécule d'eau. De plus, le sulfate établit une liaison de coordination avec le calcium. Lorsque que le sulfate est en liaison avec un ion manganèse (Mn^{2+}) (Figure 47 B), les résidus d'AccF intervenant dans sa fixation sont les mêmes, mais le nombre de liaisons hydrogènes avec H16 et le nombre de liaisons de coordination avec l'ion métallique passent d'un à deux. Et les liaisons hydrogènes avec H60 passent de deux à une. De plus la molécule d'eau n'est

plus présente à proximité de l'ion métallique. Ce changement de position pourrait être dû à la différence de nature de l'ion coordonné. En effet, le calcium est alcalino-terreux alors que le manganèse est un métal de transition ce qui pourrait expliquer qu'ils n'interagissent pas de la même façon avec le sulfate. Enfin, j'ai pu remarquer que la position du sulfate mimait la position du phosphate du glucose-2-phophate ou de celui de l'intermédiaire réactionnel en complexe avec $AccF_{Bo542}$ (Figures 47 D et E).

Le glucose, le glucose-2-phosphate et l'intermédiaire réactionnel établissent de nombreuses liaisons hydrogènes ainsi que des liaisons de coordination avec AccF (Figures 47 C-E). Cependant ces ligands ont la particularité d'établir également des liaisons hydrophobes avec les résidus F152 et L172. Ces résidus communs chez les souches Bo542 et C58 ainsi que chez différentes glycerophosphodiesterases pourraient jouer un rôle de maintien du substrat, à distance des résidus catalytiques. Malgré de nombreux essais de cristallogenèse faisant varier les différents PEGS, les pH, ainsi que les ratios des volumes protéine/solution de cristallisation, je n'ai pas réussi à obtenir de cristaux de la protéine $AccF_{Bo542}$ en complexe avec son substrat agrocinopine D. J'ai donc decidé de réaliser sa modélisation manuelle à l'aide du logiciel COOT, à partir de la structure du complexe $AccF_{Bo542}$-glucose-2-phosphate. Le G2P étant positionné dans le site catalytique, j'ai seulement ajouté un glucose à la suite de celui-ci afin de modéliser l'agrocinopine D (Figure 47 F). De nombreux résidus dans la cavité du site de liaison sont disponibles pour le stabiliser. Les résidus impliqués pourraient donc être K122, Y124, H197 et H198, à travers l'établissement des liaisons hydrogènes (Figure 47 F). De plus, une interaction hydrophobe avec le résidu F152 permettrait également la stabilisation du second glucose. Cependant, en observant les distances entre le second glucose de l'agrocinopine D et K122 (2.1 Å – interaction hydrogène) et F152 (2.4 Å et 3.1 Å – interactions hydrophobes), un réarrangement de la protéine, ou un faible décalage du glucose de l'agrocinopine D sont sans doute nécessaire.

Suite à l'analyse de ces différentes structures complexées, on peut remarquer que de nombreuses liaisons hydrogènes, liaisons hydrophobes et liaisons de coordinations permettent la reconnaissance des différents ligands. Cependant ces résidus ne sont pas conservés chez toutes les glycerophosphodiesterases (Figure 45 A). Ceci indique que ces résidus sont impliqués dans la reconnaissance du substrat mais pas forcément dans le mécanisme réactionnel. L'alignement de séquences de différentes glycerophosphodiesterases avec AccF ainsi que plusieurs études ont montré l'importance des résidus H16 et H60. Ils seraient directement impliqués dans l'activité enzymatique, alors que les résidus E45, D47 et E120 seraient eux impliqués dans la coordination de l'ion métallique.



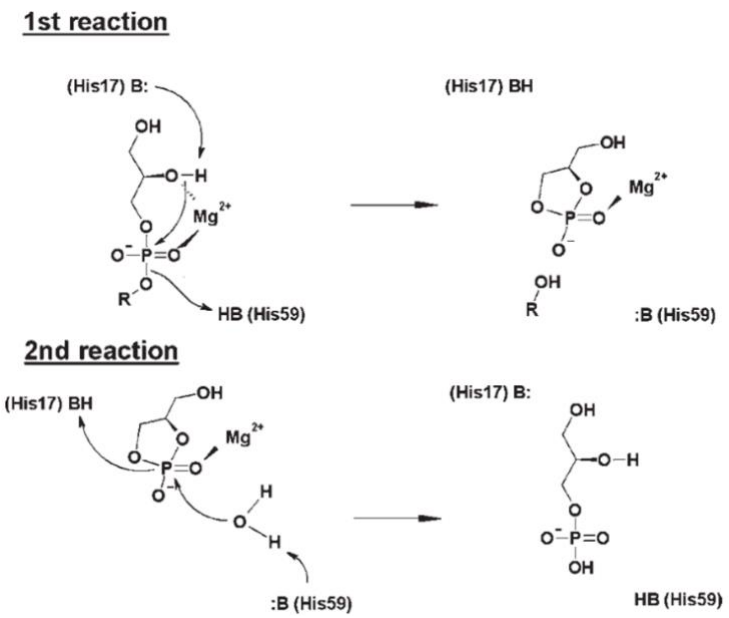
	Protéine	AccF$_{Bo542}$					
	Ligand-ion	Sulfate-Ca^{2+}	Sulfate-Mn^{2+}	Glucose	G2P	Intermédiaire	Agrocinopine D
	Résidus	Distance Å					
Liaisons hydrogènes	H16	2,8	3/2,8	2,8	2,7	2,8/2,8	2,8
	R17	2,9	2,8	-	2,8	2,9	3
	H60	2,8/3,1	3,1	-	2,9/3	2,9	3/2,8/3,1
	Q118	-	-	3,2	3,2	3,4	3,2
	K122	2,9	2,8	-	-	2,9	2,1
	Y124	-	-	-	-	-	3
	H174	-	-	2,9	2,9	2,9	2,9
	H174²	-	-	-	-	-	3,6
	N193	-	-	3,3	3,4	3,2	3,4
	E195	-	-	2,7	2,6	2,7	2,7
	H197²	-	-	-	-	-	2,5
	H198²	-	-	-	-	-	2,5
	H217	-	-	2,7	-	2,9	2,7
	Wt	2,7/3,2	-	-	-	-	-
Liaisons hydrophobes	F152	-	-	3,5	3,8	3,5	2,4
	L172	-	-	3,6	3,6	3,7	3,5
Liaison de coordination	Ca^{2+}	2,3	-	2,4	2,3/2,4	2,3/2,3	2,3/2,4
	Mn^{2+}	-	2,/2,3	-	-	-	-

Figure 47 : Représentation des sites actifs de la protéine AccF$_{Bo542}$ en complexe avec différents ligands. Toutes les

interactions entre la protéine et les ligands sont référencées dans le tableau. Un important réseau de molécules d'eau est également présent mais n'a pas été représenté. Le symbole « ² » représente les résidus issus de la seconde sous unité. **A/B:** $AccF_{Bo542}$ en complexe avec du sulfate et un ion calcium (Ca^{2+}) pour la figure **A** et un ion manganèse (Mn^{2+}) pour la figure **B**. La position du sulfate en présence d'ion manganèse est la même que celle du phosphate de l'intermédiaire réactionnel (**D**). Alors qu'en présence d'ion calcium, la position du sulfate est la même que le phosphate du G2P (**E**). **C:** $AccF_{Bo542}$ en complexe avec du glucose et un ion manganèse. La structure du complexe $AccF_{Bo542}$-glucose en présence d'ion manganèse ou calcium est identique. **D:** $AccF_{Bo542}$ en complexe avec l'intermédiaire réactionnel (IR) et un ion manganèse. **E:** $AccF_{Bo542}$ en complexe avec du glucose-2-phosphate et un ion calcium. La structure du complexe $AccF_{Bo542}$-G2P en présence d'ion manganèse ou calcium est identique. **F:** Représentation de l'agrocinopine D modélisée dans le site catalytique de $AccF_{Bo542}$. La modélisation ayant été réalisée à partir de la structure en complexe avec le G2P, les résidus interagissant avec le ligand sur la partie « G2P » sont identiques. Au niveau du deuxième glucose, les résidus potentiellement impliqués dans l'interaction sont K122, Y124, H174, H197 et H198.

v. Le mécanisme enzymatique des glycerophosphodiesterases

Le mécanisme enzymatique proposé pour la famille des glycerophosphodiesterases met en jeu les résidus conservés évoqués précédemment, à savoir H16, H60, E45, D47 et E120 chez AccF de la souche Bo542. Les deux histidines catalytiques H16 et H60 agiraient comme acide et base alors que trois autres résidus (E45, D47 et E120, Figure 45 A) sont impliqués dans la coordination d'un ion métallique qui agirait comme un électrophile stabilisant un intermédiaire réactionnel. La réaction serait donc réalisée en deux étapes et génèrerait un état de transition (Figure 48) (Shi et al., 2008). Ce mecanisme réactionnel reste toutefois une supposition dans l'état puisqu'aucune structure n'a



actuellement été resolue, ce qui permettrait de prouver l'existence de cet intermediaire réactionnel.

Figure 48 : Mécanisme réactionnel proposé pour la glycerophosphodiesterase *Thermoanaerobacter tengcongensis* (Shi et al., 2008).

En prenant en compte la conservation de différents paramètres (le repliement protéique, la conservation des résidus identifiés comme impliqués dans le mecanisme réactionnel, la

ressemblance des sites actifs et le positionnement des produits dans le site actif), on peut supposer que le mécanisme réactionnel proposé des glycerophosphodiesterases puisse être appliqué à l'enzyme $AccF_{Bo542}$ avec son substrat naturel, l'agrocinopine D (Figure 46). Aux vues du positionnement identique des histidines identifiées (Figure 45 A) comme étant impliquées dans la réaction catalytique, on peut émettre l'hypothèse que ces mécanismes réactionnels sont similaires. Comme vu précedement, durant ces travaux, j'ai réussi à obtenir la structure de $AccF_{Bo542}$ seule ou en complexe avec différents ligands et notamment l'intermediaire réactionnel (Figure 47 D) et le produit de la réaction G2P (Figure 47 E). Ces structures m'ont permis de vérifier partiellement le mécanisme enzymatique proposé (Figure 49). Ce mécansime propose également l'intervention de deux histidines dans la clivage des agrocinopines. Celles-ci vont jouer le rôle d'acide et de base comme accepteur et donneur de protons. L'hydroxyde de l'agrocinopine D serait ainsi déprotonné par l'histidine 16, suivi d'une attaque du phosphate par l'oxygène instable chargé négativement. Le phosphore possède 5 électrons sur sa derniere couche de valence et ne peut donc accepter une liaison supplémentaire. Pour stabiliser le phosphore, l'oxygène (qui suit l'atome de phosphore) arrache un proton à l'histide 60 (Figure 49 A). Ainsi, la liaison entre le phosphore et l'oxygène est brisée, engendrant la formation de l'intermédiaire réactionnel (Figure 49 B).

Cependant, la géométrie du phosphate lié au glucose n'est pas la plus stable. Lors de la libération du glucose produit, une molécule d'eau viendrait se mettre entre le phosphate et l'histidine 60. C'est alors le mécanisme inverse qui devrait se produire. La molécule d'eau est déprotonée par l'histidine 60 et l'ion hydroxyde (HO^-) formé attaque le phosphate (Figure 49 C). L'oxygène de l'intermédiaire réactionnel O- produit déprotonne ensuite l'histidine 16 afin de se stabiliser et de laisser place au glucose-2-phosphate (Figure 49D).

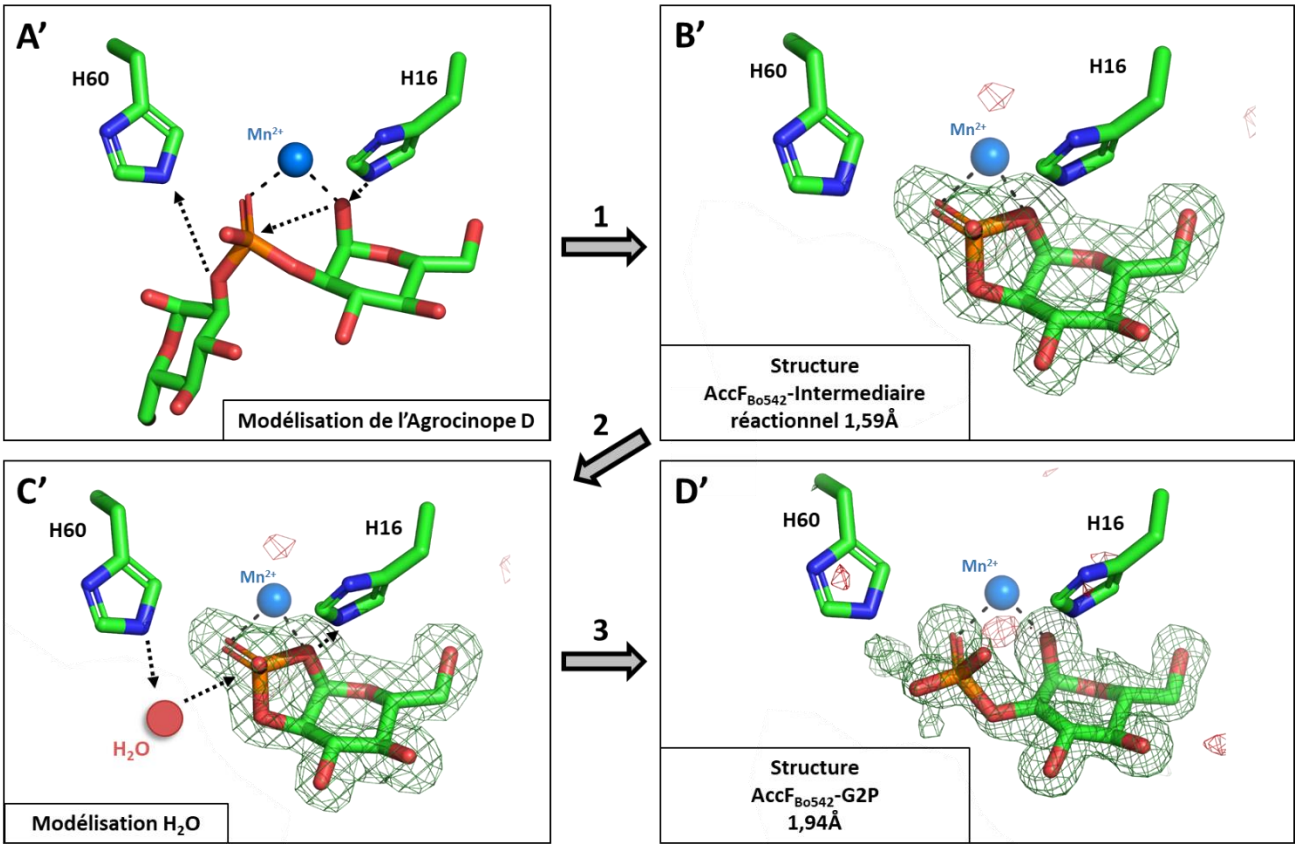


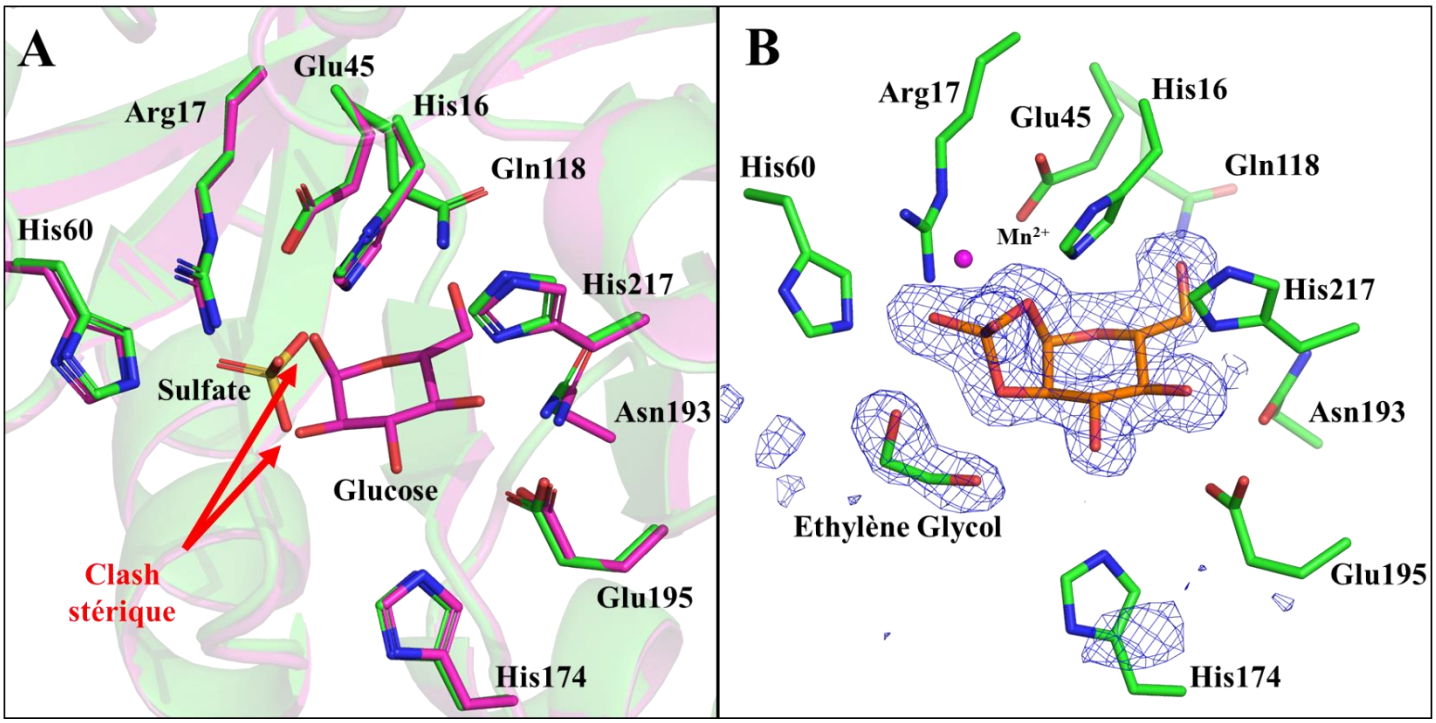
Figure 49 : A'/B'/C'/D' : Représentation en bâton du mécanisme réactionnel proposé pour la transformation de l'agrocinopine D en glucose-2-phosphate via un intermédiaire réactionnel, et en glucose, selon l'article (Shi et al., 2008). En vert les deux histidines catalytiques (H16 et H60) et le substrat (**A'**), l'intermédiaire réactionnel (**B'/C'**) et le produit réactionnel (**D'**). **A' :** L'agrocinopine est composée de glucose-2-phosphate (en bleu) et de glucose (en vert). L'histidine 16 de la phosphodiesterase $AccF_{Bo542}$, agit comme une base et déprotone l'hydroxyle de l'agrocinopine D. L'oxygène déprotoné établit ensuite une liaison avec le phosphate pour se stabiliser. Cette 6ème interaction n'étant pas possible, l'histidine 60 libère un proton et permet la libération d'un oxygène lié au phosphate. Cette libération engendre le clivage de l'agrocinopine D en glucose et un intermédiaire réactionnel (**B'**). **C':** Ensuite l'intermédiaire est transformé en glucose-2-phosphate. La molécule d'eau est déprotonée par l'histidine 60. L'hydroxyde attaque le phosphate et l'oxygène de l'intermédiaire réactionnel à portée de l'histidine 16 la déprotone. Enfin du glucose-2-phosphate est produit **(D')**.

La structure de la protéine en complexe avec l'intermédiaire réactionnel n'a été obtenue qu'une seule fois et c'est à une concentration de 16mg/ml et en présence de 1,8M de sulfate d'ammonium, de 100mM HEPES pH7.5, de manganèse et d'agrocinopine D que la protéine a cristallisé. Par la suite le cristal a été plongé dans une solution cryoprotectante contenant les conditions de cristallisation ainsi que du Peg400. La structure obtenue a ensuite permis d'identifier un ligand ressemblant à un intermédiaire réactionnel.

Suite à la possible présence d'un intermédiaire réactionnel, de nombreuses hypothèses ont été émises concernant la pertinence de cet intermédiaire réactionnel. En effet, les produits de réaction de la catalyse de l'agrocinopine D sont du glucose-2-phosphate et du glucose. Il est donc probable que le glucose produit puisse être présent dans le site catalytique après réaction (Cf Figure 47 C page 100). Cependant la cohabitation du sulfate et du glucose n'est pas possible au vu des

clashs stériques rencontrés entre l'hydroxyle du carbone 1 et 2 et le sulfate (Figure 50 A). De plus, le taux d'occupation observé pour les atomes correspondant à l'intermédiaire réactionnel sont tous de 70 à 80%. Il n'y a donc pas de taux d'occupation partagé permettant d'expliquer une cohabitation des deux molécules. Enfin la présence de cet intermédiaire réactionnel pourrait être expliqué par la présence d'une molécule ressemblant à de l'éthylène glycol à la place de la molécule permettant ainsi de bloqué la réaction à l'état d'intermédiaire réactionnel (Figure 50 B).

Figure 50 : Représentation du site catalytique d'AccF$_{Bo542}$. **A :** Superposition de la structure d'AccF$_{Bo542}$ en présence de sulfate et de glucose. On observe l'impossibilité de cohabitation dû aux clashs stériques engendrés. **B :** Représentation



du site catalytique en présence du substrat bloqué à l'étape d'intermédiaire réactionnel dû à la présence d'un éthylène glycol.

vi. Choix des mutants

J'ai obtenu une structure cristallographique en présence d'un intermediaire réactionnel, ce qui me permet d'affirmer qu'il existe un état transitoire entre l'agrocinopine D et le G2P (Figure 49 D). Pour prouver l'importance de certains résidus impliqués directement ou indirectement dans l'activité catalytique, j'ai décidé d'utiliser des protéines mutées au niveau de certains résidus. Tout les mutants utilisés ont subi une mutation en alanine. J'ai choisi de muter H16 et H60 mises en avant pour leur implication dans la réaction catalytique. Après l'observation des structures, j'ai ensuite arbitrairement selectionné les résidus H174, E195 et H197. Le résidu H174 a été choisi car il est présent au niveau de l'interface du dimère de façon symétrique et qu'il intéragit également avec le produit de réaction (G2P), et donc potentiellement avec le substrat. Selon la modélisation que j'ai réalisée du complexe AccF-agrocinopine D, E195 pourrait d'autre part intéragir avec le

ligand au nivau de glucose-2-phosphate, et H197 intéragir potentiellement au niveau du deuxième glucose (Figure 51). Ces résidus ont été mutés en alanine.

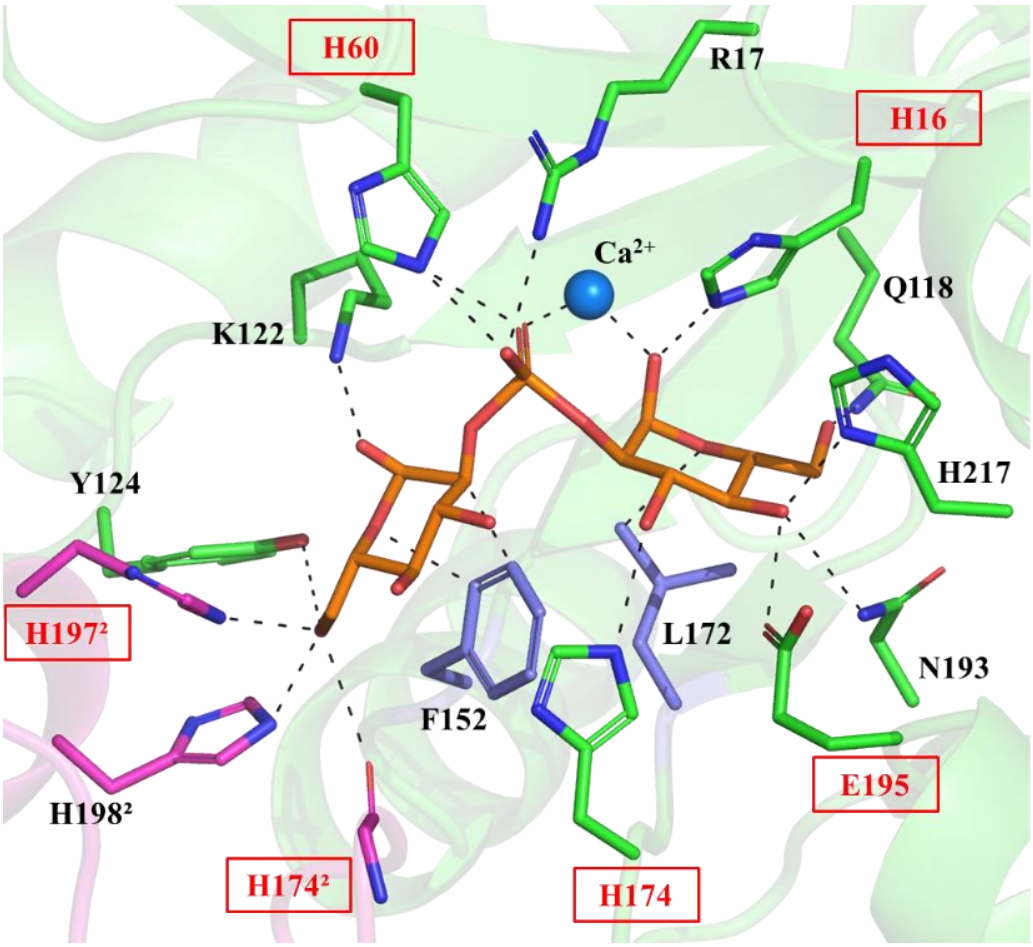


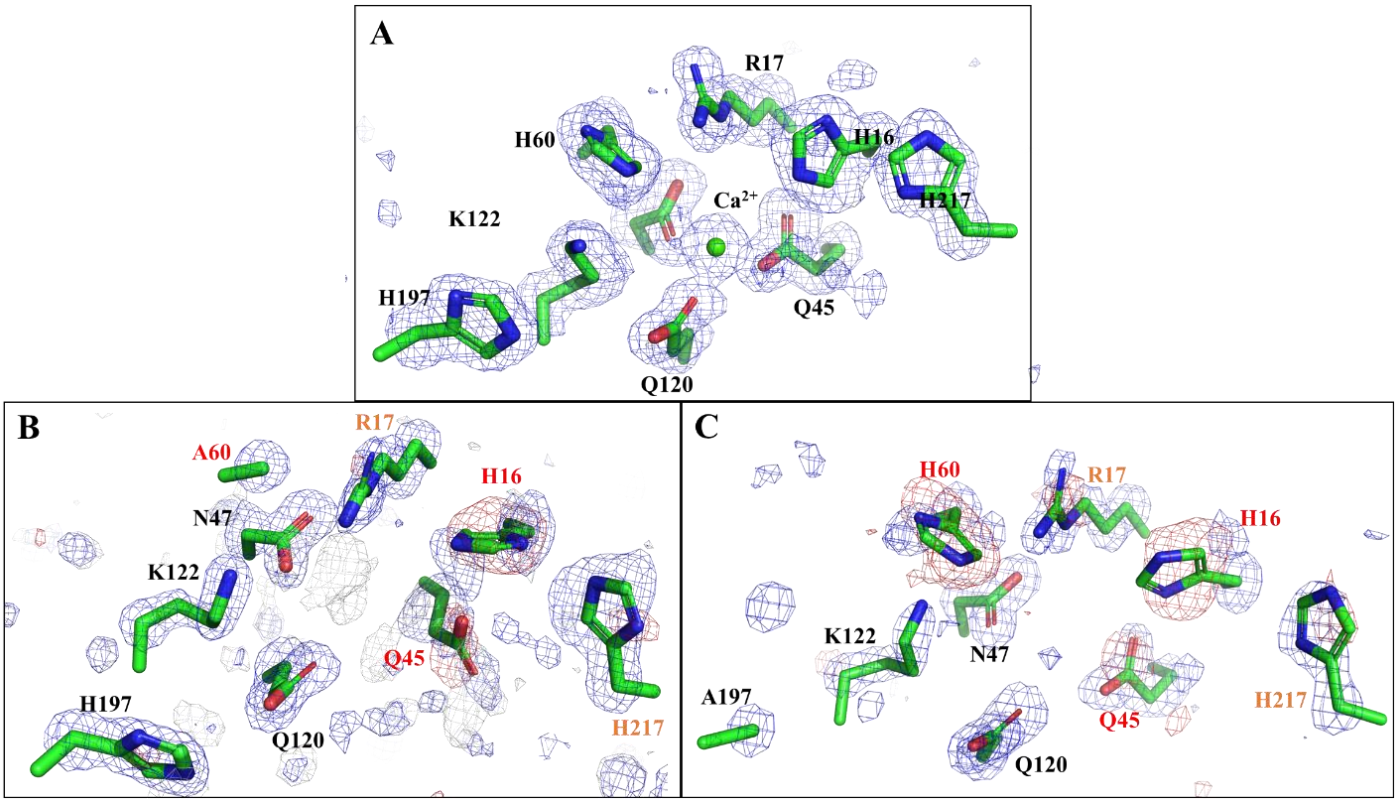
Figure 51 : Modélisation du site catalytique de l'enzyme AccF en complexe avec l'agrocinopine D (en orange) et un ion calcium. Cette figure représente les résidus impliqués dans l'activité catalytique (H16 et H60) et la potentielle implication de certains résidus dans la fixation du substrat (H16, R17, H60, Q118, K122, Y124, F152, L172, H174, H174², E195, H197, H198 et H217). Le signe « ² » représente les résidus de la deuxième sous unité. Les résidus verts et violets représentent les résidus de la premiere sous unité alors que les résidus roses font partie dans la seconde sous unité. Enfin les résidus violets (F152 et L172) représentent les résidus impliqués dans la création d'un environnement hydrophobe.

Après de nombreux essais infructueux de co-cristallisation et de trempage des differents mutants avec l'Agrocinopine D, seules les structures de la forme *apo* de ces protéines mutées ont été obtenues. Lorsque j'ai affiné et observé les structure des différents mutants, j'ai pu constater que les structures étaient globalement identiques, avec un RMSD différence moyen allant de 0.2 à 0.4Å (Tableau 8 page 93).

Cependant, j'ai également pu constater quelques différences et l'obtention de ces structures

a permis de mettre en évidence des points importants quant à la stabilité du site catalytique de la protéine (Figure 52). En effet, les mutations H16A, H174A et H195A n'engendrent pas la perte de fixation de l'ion métallique alors que c'est le cas pour les résidus H60A et H197A. L'impact de l'absence des résidus H60 et H197 chez les protéines mutées est une déstabiliation des résidus R17, H16 et E45. Cette hypothèse a été proposée à cause de la perte de densité éléctronique du résidu due à sa plus grande mobilité, ainsi qu'une perte de la fixation de l'ion métallique (Figure 52 B et C). La mutation E195 déstabilise très legèrement les résidus H217 et H174. En plus des résidus déstabilisés, le réseau de molécules d'eau l'est également.

Afin de vérifier l'impact des mutations sur l'activité de l'enzyme, j'ai mis au point un

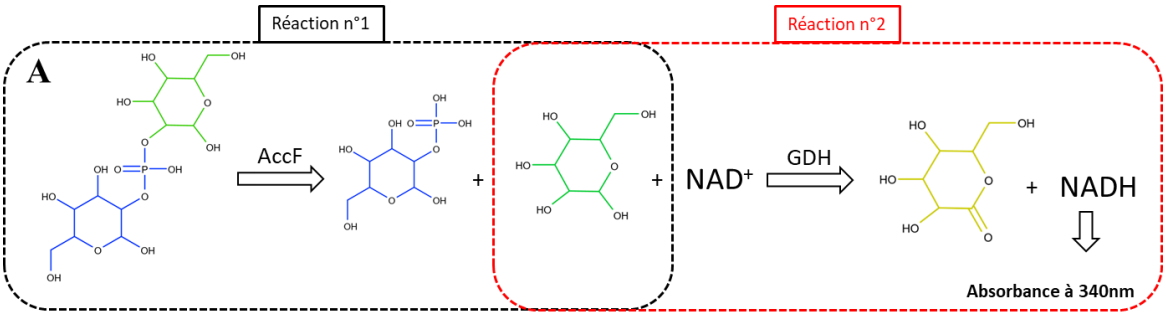


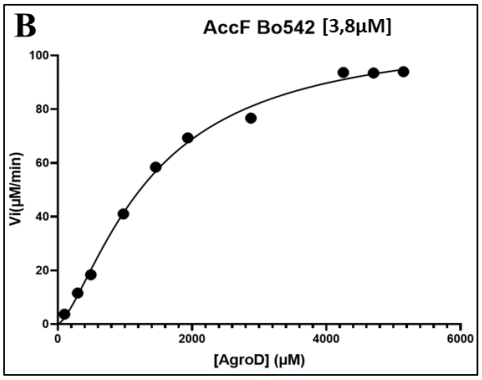
protocole enzymatique permettant de déterminer l'activité d'AccF sauvage et mutants (Figure 53).

Figure 52 : Représentation des sites catalytiques des structures de la protéine sauvage, du mutant H60A et du mutant H197A. La carte de densité 2Fo-Fc est représentée en bleu, alors que la carte Fo-Fc est représentée en rouge lorsque celle-ci est négative alors qu'elle est représentée en blanc lorsqu'elle est positive. Le but de cette figure étant de montrer la flexibilité de certains résidus par l'absence de densité électronique, la carte Fo-Fc>0 (résiduelle) a volontairement était mise en blanc afin de ne pas gêner la visualisation. Les résidus notés en noir sont fixes alors que ceux notés en orange et rouge sont peu ou très flexibles. **A** : Représentation du site catalytique de la protéine $AccF_{Bo542}$ vide en présence de calcium. Structure obtenue à une résolution de 1.45 Å. **B** : Représentation du site catalytique du mutant $AccF_{Bo542}$ H60A. Structure obtenue à une résolution de 1.40 Å. **C** : Représentation du site catalytique du mutant $AccF_{Bo542}$ H197A. Structure obtenue à une résolution de 1.40 Å.

vii. Etude de l'activité enzymatique de la protéine $AccF_{C58/Bo542}$ sauvage et de ses mutants.

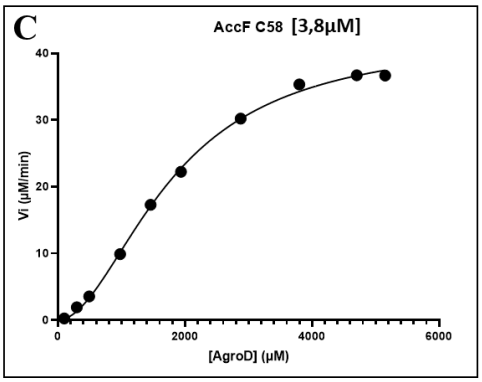
La protéine AccF est capable de catalyser le clivage des agrocinopines par reconnaissance de la liaison phosphodiester. Sur les quatre agrocinopines existantes, j'ai utilisé l'agrocinopine D qui après clivage, produit du glucose facile à détecter. Le test est basé sur le dosage du glucose, un des produits d'AccF, via sa déhydrogénation par la GDH (Glucose-Dehydrogenase) en présence de NAD+, et la mesure d'absorbance du NADH produit, à 340 nm. La mise au point de ce protocole, qui met donc en jeu deux réactions (Figure 53 A), a permis de montrer que les conditions optimales d'activité pour la protéine $AccF_{Bo542}$ sont pH 7.5 et une température de 20°C. J'ai également mesuré





Activité enzymatique	
AccF C58 Type sauvage	
Vmax	108,9
h	1,441
Khalf (µM)	1372
R^2	0,9967

Activité mutants (%)	
En présence de l'ion calcium	
Bo542	
Sauvage	100
H16A	0
H60A	0
H174A	55
E195A	6,9
H197A	0



Activité enzymatique	
AccF C58 Type sauvage	
Vmax	42,51
h	1,905
Khalf (µM)	1797
R^2	0,9987

Activité mutants (%)	
En présence de l'ion calcium	
C58	
Sauvage	100
H18A	0
H62A	0

l'activité enzymatique de AccF de la souche C58.

Figure 53 : A : Réactions enzymatiques du protocole mis en place pour déterminer l'activité enzymatique de l'enzyme AccF en présence du substrat agrocinopine D. **B/C** : Cinétiques enzymatiques de l'enzyme $AccF_{Bo542}$(**B**) /$AccF_{C58}$(**C**) en présence d'agrocinopine D à 25°C. Les paramètres de celles-ci sont représentés en haut à droite ainsi que l'activité observée chez les mutants en bas à droit. Le Khalf représente la concentration de substrat nécessaire pour atteindre la

moitié de la vitesse maximale de réaction. L'activité des mutants a été déterminé à une concentration de substrat permettant d'atteindre une vitesse initiale maximale et le pourcentage correspond au ratio de cette valeur.

Les résultats obtenus montrent que les courbes pour les deux AccF correspondent à une activité enzymatique de type allostérique, avec un coefficient de corrélation de 0.99 (Figure 53 B). Aux vues des concentrations en enzyme utilisées et des résultats obtenus, la glucose dehydrogenase [0,16µM] a une activité spécifique de 510 UI/mg, alors que l'enzyme AccF [3,8µM] a une activité spécifique de 0.3 UI/mg. La glucose dehydrogenase possède donc une activité spécifique environ 1500 fois supérieure à celle de l'enzyme AccF. AccF est donc limitante et c'est bien elle qui est détectée. Les résultats ont ensuite permis de fitter les courbes selon l'équation Y=Vmax*X^h/(Khalf^h + X^h). Les courbes et les résultats obtenus nous montrent et nous permettent de proposer que l'enzyme AccF est une enzyme allostérique dont le Khalf (agrocinopine D) des enzymes allosteriques est de l'ordre du millimolaire (mM). Le Khalf représente la concentration de substrat nécessaire pour atteindre la moitié de la vitesse maximale de réaction. Cette valeur élevée de Khalf représente une faible affinité pour le substrat mais cette faible affinité pour le substrat est retrouvée chez certaines enzymes de la même famille (Denloye et al., 2012; Ohshima et al., 2008; Yanaka, 2007). Les réactions ont également été réalisées en présence de calcium et un autre ion aurait sans doute un impact sur la reconnaissance et l'activité de l'enzyme. L'affinité de l'enzyme $AccF_{Bo542}$ pour l'agrocinopine D est plus élevée que celle de l'enzyme $AccF_{C58}$ (Figure 53 C), ce qui est cohérant étant donné qu'il s'agit de son substrat « naturel ». Le coefficient de Hill obtenu étant supérieur à la valeur de 1, ceci peut témoigner d'une fixation coopérative. La coopérativité de l'enzyme $AccF_{C58}$ est plus élevée que celle de la souche Bo542 et ceci se remarque également par une allure sigmoïde plus prononcée.

L'absence d'activité observée pour les mutants H16, H60 (Souche Bo542) et H18, H62 (Souche C58) nous montre l'importance de ces résidus dans la catalyse et/ou la reconnaissance du substrat. De même, l'activité est perdue pour la protéine mutée en position 197, alors que l'enzyme reste partiellement active pour les mutants H174A et E195A. L'activité résiduelle des mutants H174A et E195A a donc été testée mais n'a pas pu être quantifiée lié à des problèmes de pureté de substrat. Tous ces résultats témoignent de l'importance de ces résidus dans l'activité enzymatique. De plus, comme précédemment évoqué, les résidus 16 et 60 de la protéine $AccF_{Bo542}$ jouent un rôle très important dans l'activité enzymatique puisqu'ils agiraient en tant que base et acide dans cette catalyse. La perte d'activité et les observations structurales précédemment réalisées se recoupent et montrent bien que les résidus 174 et 195 sont impliqués dans la stabilisation du dimère pour le résidu 174 et dans la fixation de la partie glucose-2-phosphate de l'agrocinopine D et les résidus

174 et 195. Enfin, la modélisation de l'agrocinopine D dans le site actif permet d'émettre l'hypothèse d'une implication du résidu 197 dans la fixation de l'agrocinopine D.

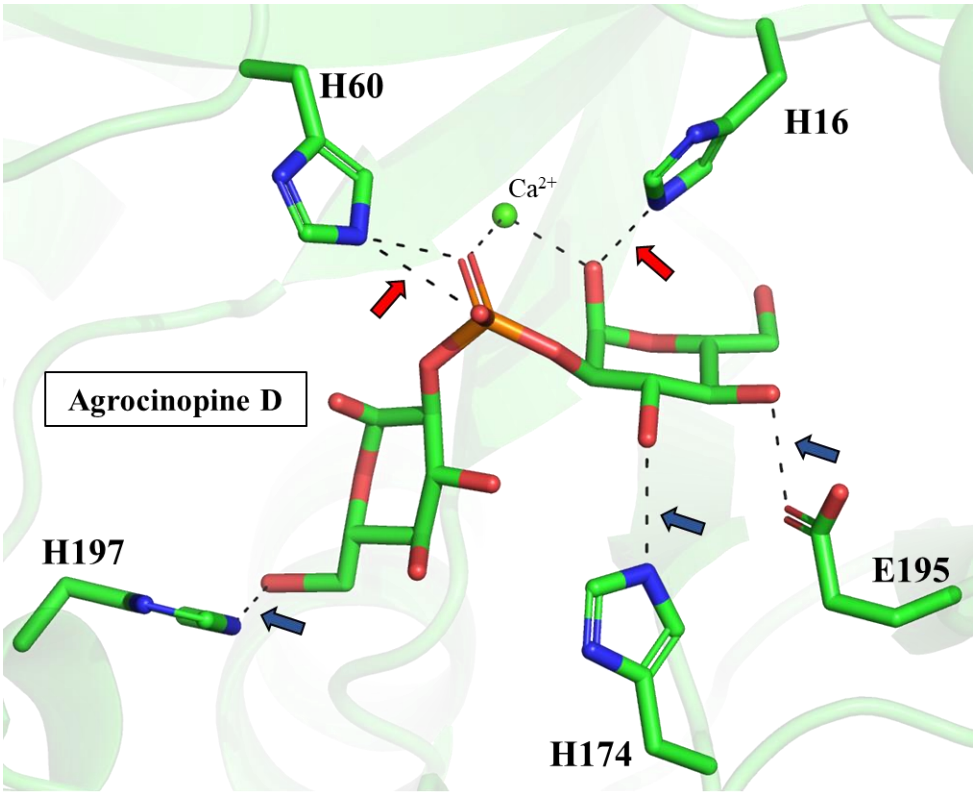


Figure 54 : Modélisation de l'agrocinopine D dans le site catalytique de l'enzyme $AccF_{Bo542}$ en présence de calcium. Représentation des résidus impliqués dans l'activité catalytique (H16 et H60) et dans la fixation du substrat (H16,H60, H174, E195 et H197). Les flèches indiquent les liaisons chimiques qui sont perdues avec les mutants utilisés.

La perte d'activité lors de la mutation de ce résidu semble confirmer son importance dans la stabilisation du site catalytique et dans la fixation du substrat (Figures 53 B et 54).

Les résultats précédemment obtenus concernant l'activité enzymatique montrent que les protéines $AccF_{Bo542}$ et $AccF_{C58}$ sont des enzymes allostériques avec une fixation coopérative. Je me suis donc intéressé aux différentes structures obtenues au cours de ma thèse afin d'expliquer au niveau structural cette coopérativité de fixation.

viii. Explication structurale de l'activité allostérique observée chez la protéine $AccF_{Bo542/C58}$ sauvage

1. *L'enzyme $AccF_{Bo542}$*

Une coopérativité enzymatique étant identifiée, j'ai voulu expliquer structuralement cette allostérie, en utilisant les structures *apo* et complexées d'AccF et en prenant en compte l'état dimérique de cette protéine. J'ai dans un premier temps remarqué la proximité des deux sites actifs. En effet,

contrairement aux glycerophosphodiesterases déjà étudiées (Figure 55 A et C), les deux sites catalytiques du dimère d'AccF sont beaucoup plus proches (Figure 55 B et D). Ces protéines partagent un repliement identique, mais l'organisation du dimère varie.

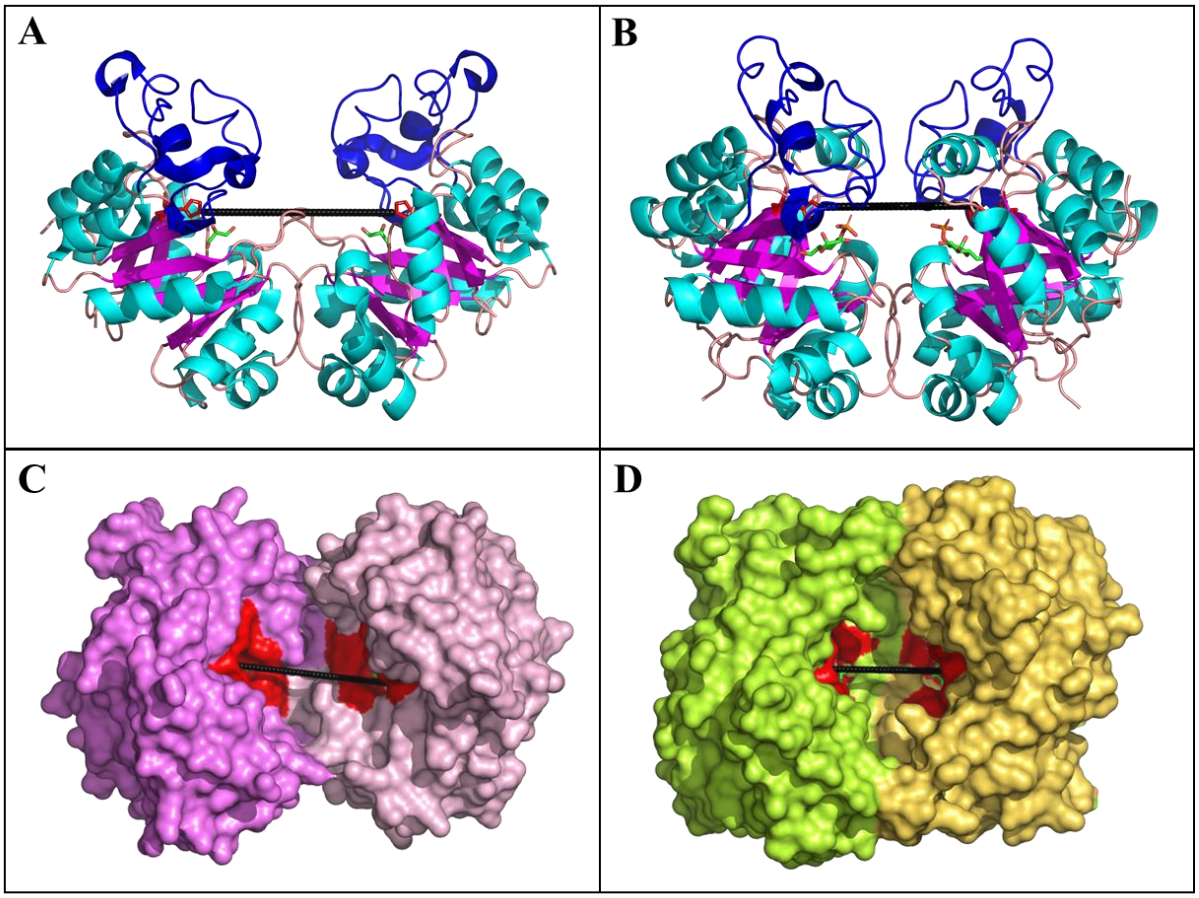


Figure 55 : Différentes représentations du dimère de la glycerophosphodiesterase (GPDQ) de *T. tengcongensis* (Code : 2PZ0)(**A/C**) et de la phosphodiesterase AccF d'*A. tumefaciens* Bo542 (**B/D**). La barre noire relie les deux résidus H16 des deux monomères et est représentative de la distance entre les deux sites actifs, et a été mesurée à 25.9 Å (**A/C**) et 19.6 Å (**B/D**). Le ligand est représenté en battons **verts**. **A/B** : Représentation en ruban de la $GPDQ_{T.tengcongensis}$ (**A**) et $AccF_{Bo542}$ (**B**), vus de côté. Les structures secondaires qui composent le domaine possédant le repliement caractéristique en TIM Barrel sont représentés de différentes couleurs (**Cyan** : hélice α ; **Rose** : Brin β ; **Beige** ; Boucle). Le domaine supplémentaire est représenté en **Bleu**. **C/D :** Surface des protéines $GPDQ_{T.tengcongensis}$ (**C**) et $AccF_{Bo542}$ (**D**), vus du dessus. La surface des dimères protéiques sont colorés en rose et en vert/jaune avec en rouge les résidus qui composent le site catalytique.

Comme on peut le remarquer sur la Figure 55, les deux monomères sont orientés de manière différente et engendrent ainsi une diminution des surfaces d'interaction pour la GPDQ. Cependant, même si la surface d'interaction est diminuée, le nombre d'interactions reste conservé. En effet, l'interface du dimère de GPDQ est stabilisée par un pont disulfure, 12 liaisons hydrogène et 2 ponts salins. Il en est de même pour l'interface dimérique d'$AccF_{Bo542}$ (12 liaisons hydrogène et 2 ponts salins) mais le pont disulfure n'existe pas. La distance rapprochée des deux sites catalytiques des monomères d'$AccF_{Bo542}$, contrairement à d'autres glycerophosphodiesterases, engendre un rapprochement des résidus impliqués dans la liaison au substrat.

De plus, j'ai pu remarquer durant l'analyse des structures, que les deux résidus E195 et H197 pouvaient adopter deux positions différentes (positionnement différent dans deux structures *apo* différentes) (Figure 56 A et B). Suite à ces observations, j'ai émis une hypothèse quant à la reconnaissance et la fixation coopérative de l'agrocinopine D. La reconnaissance du premier substrat dans le site de l'un des 2 monomères permettrait de contraindre les premiers résidus 195 et 197 dans une position qui leur permet d'interagir avec le substrat (Figure 56 C). Ces interactions obligeraient ensuite les deuxièmes résidus 195 et 197 à se retrouver dans la même position (Figure 56 D). Le fait de contraindre les deux autres résidus 195 et 197 favoriserait la reconnaissance du substrat pour la deuxième sous-unité (Figure 56 E et F). Ainsi la fixation d'un premier substrat faciliterait la fixation du deuxième et la coopérativité pourrait ainsi être expliquée.

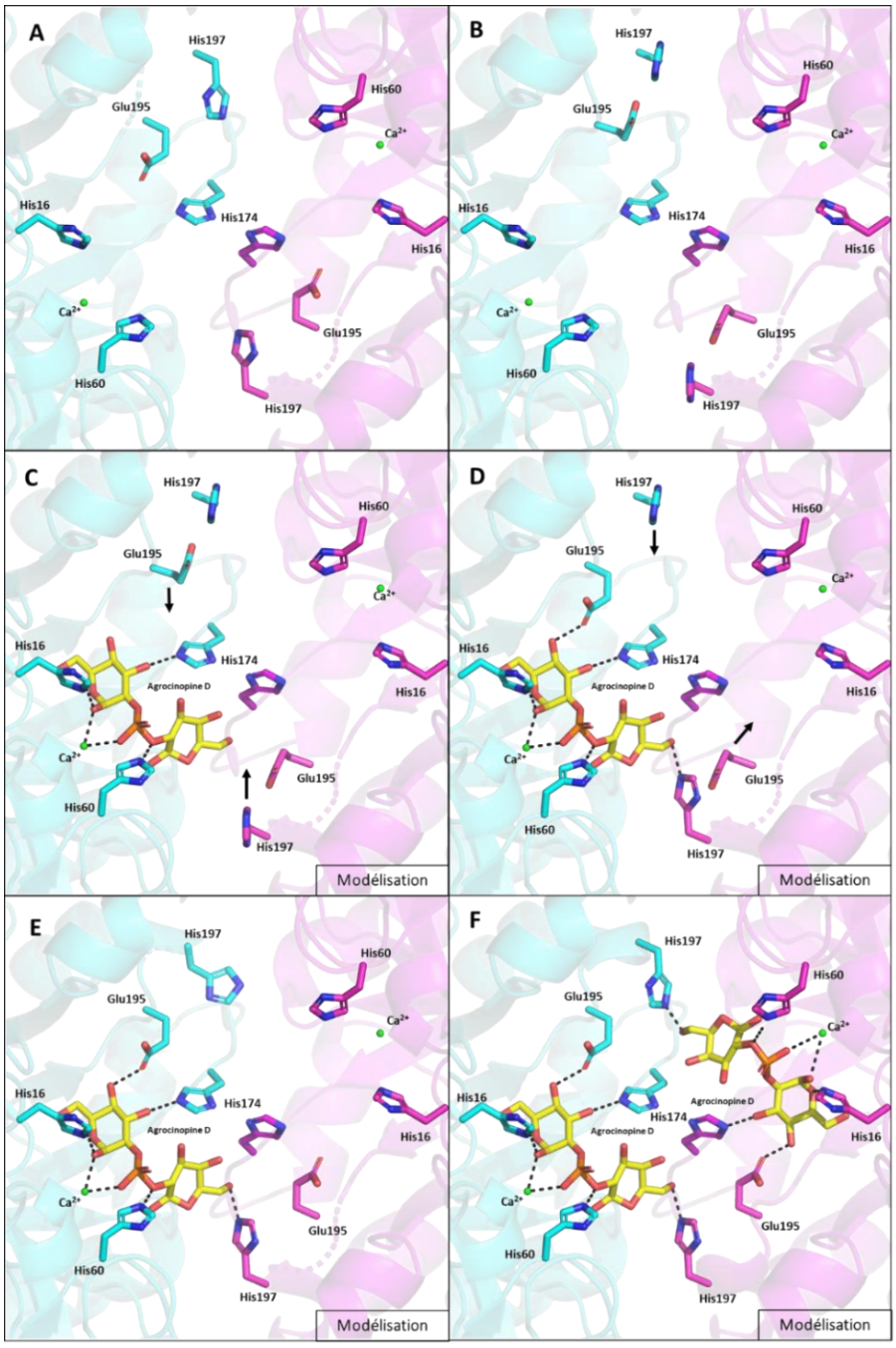
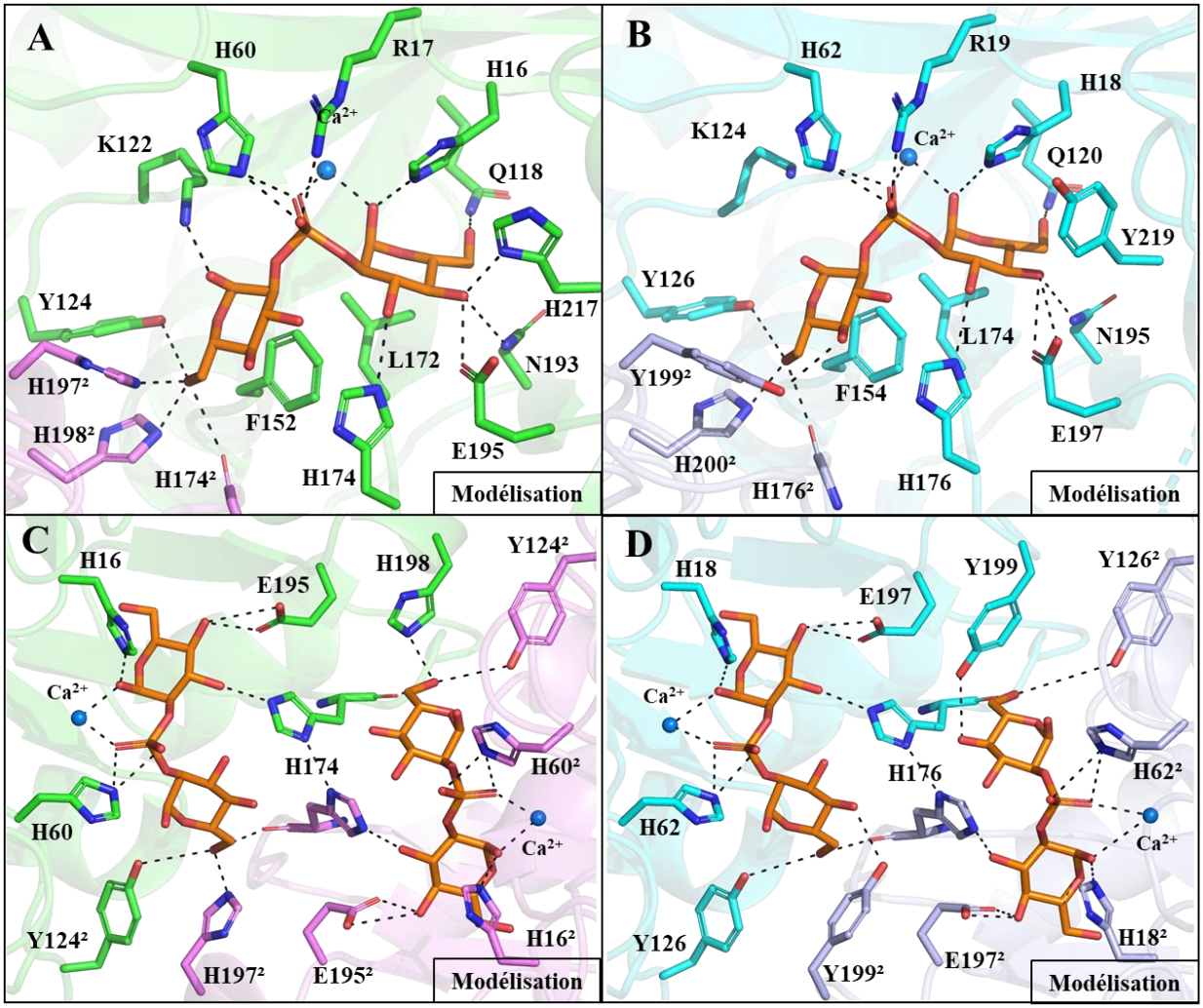


Figure 56 : Représentation du positionnement différent des deux résidus E195 et H197 ainsi que la modélisation de la fixation coopérative de l'agrocinopine D chez l'enzyme $AccF_{Bo542}$. Dans cette représentation, certaines interactions n'ont pas été affichées pour ne pas surcharger l'image et se concentrer sur le changement de position de His197 et E195. **A :** Représentation d'un premier positionnement possible des résidus E195 et H197. **B :** Représentation du second positionnement possible des résidus E195 et H197. **C :** Représentation et modélisation du déplacement des premiers résidus E195 et H197 lors de la fixation d'une première molécule d'agrocinopine D. Le changement de position a pour but d'augmenter le nombre de liaisons et de stabiliser le substrat. **D :** Représentation et modélisation du déplacement des seconds résidus E195 et H197 lors de la fixation d'une première molécule d'agrocinopine D. Au vu du changement de position des premiers résidus E195 et H197, les deux autres n'ont pas d'autre choix que de changer également de position, d'un point de vue stérique. **E-F :**

L'établissement de ces liaisons visant à stabiliser la première molécule d'agrocinopine D a contraint les résidus E195 et H197 dans une position précise qui facilite la fixation d'une seconde molécule d'agrocinopine D.

2. *L'enzyme $AccF_{C58}$*

Si on analyse la modélisation de la structure de la protéine $AccF_{C58}$ en complexe avec l'agrocinopine D, on observe seulement de faibles variations du site catalytique comparé à celui d'$AccF_{Bo542}$. Le résidu K122 semble être plus éloigné, les histidines H197 et H217 de $AccF_{Bo542}$ deviennent des tyrosines dans $AccF_{C58}$ (Figure 57). Cependant, la tyrosine 199 pourrait selon cette modélisation toujours interagir avec l'agrocinopine D. La plus faible affinité de la protéine $AccF_{C58}$ est sans doute due à la perte d'interaction des résidus K124 et Y219, contrairement aux K122 et



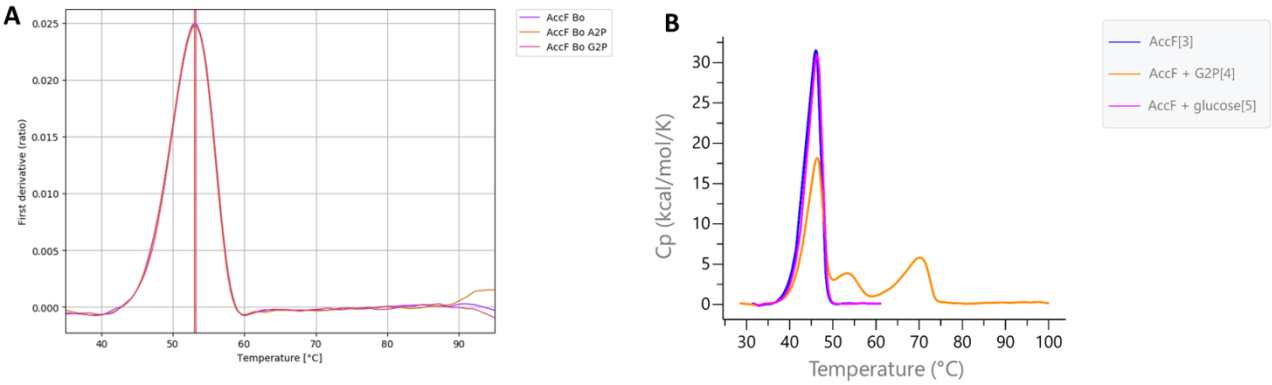
H217 d'$AccF_{Bo542}$ (Figure 57).

Figure 57 : Modélisation de l'agrocinopine D dans le site catalytique de l'enzyme $AccF_{Bo542}$ vide (A/C) et $AccF_{C58}$ vide (B/D) en présence de l'ion calcium. A: Représentation du site catalytique de la protéine $AccF_{Bo542}$. Les deux sous unités sont représentées en cyan et en rose. Tous les résidus représentés interagissent avec l'agrocinopine D. **B**: Représentation du site catalytique d'$AccF_{C58}$. Les deux sous unités sont représentées en vert et en gris. En comparaison avec la souche Bo542, les résidus Y219 et K124 ne semblent pas interagir avec l'agrocinopine D. Quant à la tyrosine 199, elle interagie avec un autre oxygène de l'agrocinopine D. Les deux sous unités sont représentées en vert et en gris. **C/D**: Représentation des sites actifs de l'interface dimérique des deux sous unités. On remarque que le positionnement de la tyrosine 199 chez la souche C58 est le même que l'histidine 197 de la souche Bo542.

La plus forte coopérativité de la protéine $AccF_{C58}$ peut être expliquée par la mutation du résidu H197 de $AccF_{Bo542}$ en Y199. En effet l'histidine étant potentiellement impliquée dans la coopérativité de fixation du substrat de la deuxième sous unité, son changement en tyrosine pourrait avoir un impact sur la coopérativité. Une hypothèse peut alors être formulée : sa flexibilité avant la reconnaissance du substrat pourrait gêner la fixation de celui-ci et lorsqu'il est dans le site catalytique, les résidus contraints (Y199 et E197) favoriseraient la fixation du deuxième substrat. La plus forte coopérativité serait due à une plus forte gêne de la flexibilité de la tyrosine lors de la reconnaissance (Figure 57 B et D). Nous ne disposons pas de structure montrant la présence de deux positions possibles de la tyrosine, cette hypothèse n'est donc pas très fiable. Cependant, les structures de la protéine $AccF_{Bo542}$ prouvant l'existence de deux positions pour ces résidus, elle n'en reste pas moins à considérer.

ix. Etude des interactions Enzyme : Ligands

Le but de cette partie a été de déterminer si une interaction entre l'enzyme et le produit était observable pour ensuite essayer d'observer une interaction entre un mutant inactif et le substrat. Pour se faire, j'ai réalisé des tests Tycho et DSC en mettant la protéine en présence de glucose, de glucose-2-phosphate et d'arabinose-2-phosphate (Figure 58).



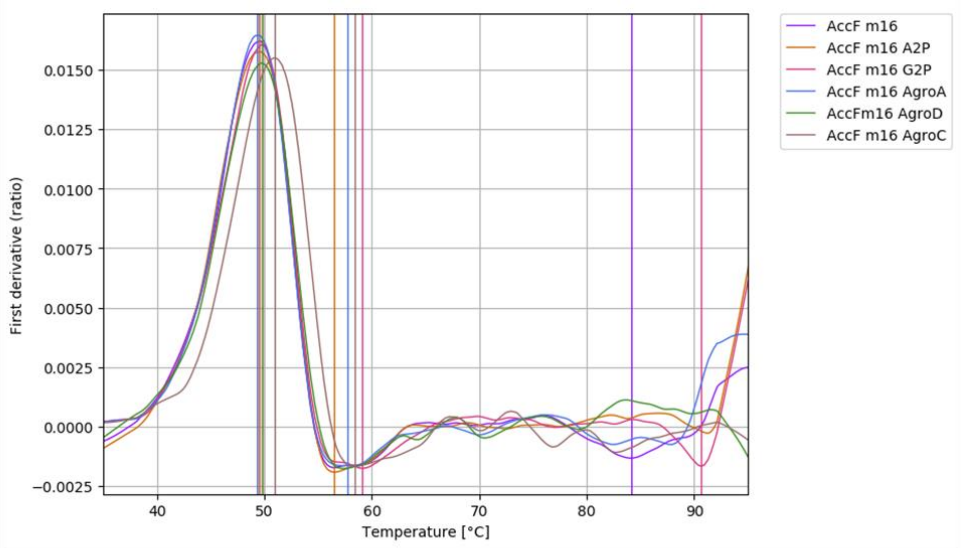
	Tycho		DSC						
Protéine	Ti(°C)	ΔTi(°C)	Tm_1(°C)	Tm_2(°C)	Tm_3(°C)	$ΔTm_1$(°C)	$ΔTm_2$(°C)	$ΔTm_3$(°C)	ΔH (kcal/mol)
$AccF_{Bo542}$ vide	53		44,9						144
$AccF_{Bo542}$ – G2P	53.2	0.2	45,6	53,8	68,9	0,7	9,9	24,8	162
$AccF_{Bo542}$ – Glucose			45,3			0,4			139
$AccF_{Bo542}$ – A2P	53	0							

Figure 58 : Résultats des tests d'interactions par DSC et Tycho d'$AccF_{Bo452}$ en présence de différents ligands. Les expériences ont été réalisé à pH 8 avec une température évoluant de 30 à 100°C. Les concentrations utilisées sont de 10µM pour la protéine et de 500µM pour le ligand.

Comme on peut le voir, les expériences de Tycho n'ont montré aucune variation significative des températures d'inflexion entre la protéine seule et la protéine en présence de l'A2P

ou G2P (Figure 58 A). Cependant, les expériences de DSC ont montré une variation de la température de fusion en présence de G2P (courbe orange Figure 58 B). En effet, il y a apparition de 3 populations différentes lors de l'augmentation de la température, correspondant à des populations de plus en plus stables. Lors de l'expérience, le ratio protéine : produit de réaction est de 1 pour 25 et les conditions expérimentales ne sont pas saturantes au vu des concentrations utilisées et du k_{half} de 1,3 mM pour le substrat. Si l'on observe les Tm des différents pics, on observe que le premier pic commun aux alentours de 45°C correspond à la protéine seule. L'observation de deux pics supplémentaires à 53,8 et 68.9°C en présence de G2P laisse à penser qu'il s'agit de la protéine dimérique en complexe avec un et deux G2P respectivement. Pour vérifier cette hypothèse, il serait intéressant de renouveler cette expérience en augmentant la concentration en ligand dans le but de déplacer l'équilibre entre le pic 2 et le pic 3 en présence de G2P ou voir apparaitre un second pic en présence de glucose ou d'A2P.

Suite aux résultats obtenus, j'ai tenté d'observer une variation du Ti du mutant AccF H16A, qui pourrait témoigner d'une stabilisation de la protéine en présence du ligand et donc d'une interaction entre les deux. Le mutant H16A a été choisi aux vues de sa faible implication dans la reconnaissance du substrat (1 interaction) et son implication dans l'activité catalytique de la



	Tycho	
Protéine	Tm(°C)	ΔTm(°C)
AccF Bo542 m16 vide	49.5	
$AccF_{Bo542}$ H16A - A2P	49.5	0
$AccF_{Bo542}$ H16A - G2P	49.8	0.3
$AccF_{Bo542}$ H16A - Agrocinopine A	49.3	0.2
$AccF_{Bo542}$ H16A -Agrocinopine D	49.8	0.3
$AccF_{Bo542}$ H16A - Agrocinopine C modifiée	50.9	1.4

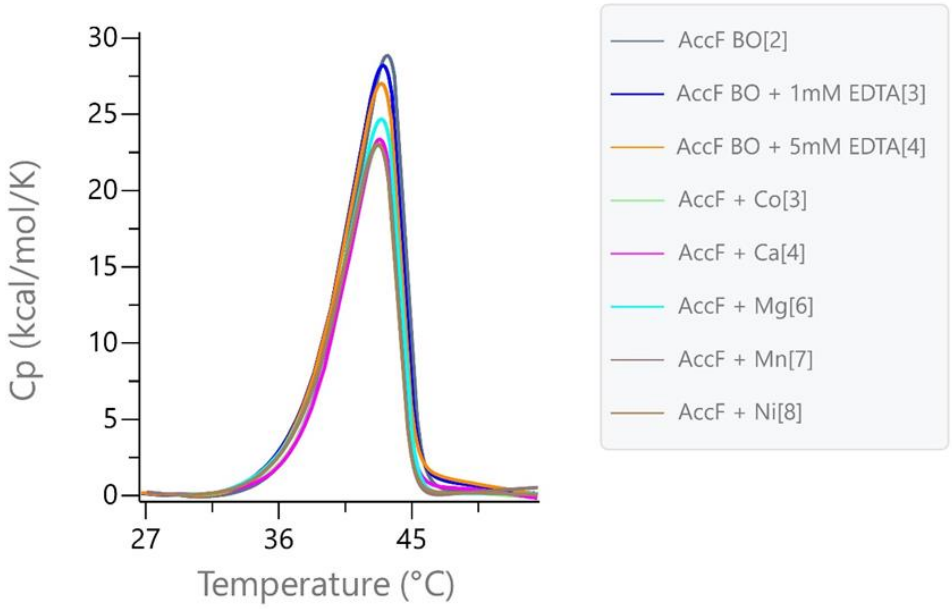
protéine.

Figure 59 : Résultats des tests d'interactions par Tycho du mutant $AccF_{Bo452}$ H16A en présence des différents ligands. Les courbes des expériences via l'utilisation du Tycho représentent la dérivée du ratio 350nm/330nm laissant apparaitre la température d'inflexion (Ti).

Grâce aux tests Tycho réalisés (Figure 59), j'ai pu mettre en évidence une légère augmentation de la température d'inflexion (ΔTm = 1.4 °C) en présence d'agrocinopine C* modifiée. Cette information n'est malheureusement pas utilisable et significative seule étant donné la modification de l'agrocinopine C qui la différencie du substrat naturel. Cependant, ceci m'a permis de m'orienter vers une possible stabilité protéine : ligand pouvant éventuellement permettre d'obtenir le complexe enzyme-substrat. La co-cristallisation de $AccF_{Bo542}$ n'a cependant pas encore abouti.

x. Impact de l'ion sur la stabilité de la protéine

La phosphodiesterase AccF est une enzyme qui nécessite la présente d'un cation pour stabiliser le substrat ainsi que pour permettre son activité catalytique. Afin d'approfondir les connaissances et la



Protéine	Tm(°C)
$AccF_{Bo542}$	41.13
$AccF_{Bo542}$ + 1mM EDTA	40.76
$AccF_{Bo542}$ + 5mM EDTA	40.75
$AccF_{Bo542}$ + Cobalt (Co^{2+})	40.64
$AccF_{Bo542}$ + Calcium (Ca^{2+})	40.75
$AccF_{Bo542}$ + Magnésium (Mg^{2+})	40.52
$AccF_{Bo542}$ + Manganèse (Mn^{2+})	40.39
$AccF_{Bo542}$ + Nickel (Ni^{2+})	40.36

compréhension du mécanisme enzymatique, j'ai voulu regarder si la nature de l'ion avait un impact sur la stabilité de la protéine. Pour cela j'ai analysé par DSC la protéine AccF en présence de différents ions (Figure 60)

Figure 60 : Résultats des tests d'interaction par DSC d'AccF$_{Bo452}$ en présence de différents ions. Les courbes obtenues par DSC représentent la capacité thermique en fonction de la température, laissant apparaitre la température de fusion représentative de la température de dénaturation de la protéine.

Les différents échantillons représentent :

- La protéine purifiée et stockée avec un ion calcium ajouté lors de la purification (AccF$_{Bo542}$)
- La protéine dépourvue d'ion grâce à l'ajout d'EDTA 30 fois et 150 fois supérieur à la concentration protéique. (AccF$_{Bo542}$ + EDTA)
- La protéine préalablement dépourvue d'ion divalent puis complémentée avec un nouvel ion par dialyse afin d'éliminer l'EDTA.

Comme on l'observe sur cette figure, la température de fusion ne change absolument pas si la protéine est seule, en présence ou en absence de différents ions. Ceci indique que la température de fusion n'est pas impactée par le changement d'ion et donc que l'absence ou la présence d'ion ne change en aucun cas la stabilité thermique de la protéine. Cependant, lors des premières purifications de cette protéine, à forte concentration, la présence d'ion est indispensable puisqu'en son absence l'agrégation est favorisée. Pour s'assurer du changement effectif de l'ion, aucune vérification n'a été réalisée. Néanmoins, la même manipulation a été réalisée pour l'obtention des structures et le changement d'ion a cette fois été observé visuellement dans les cartes de densité électronique.

La stabilité de la protéine n'étant pas impactée, il serait intéressant désormais de réaliser des tests d'activité en changeant les ions captés par la protéine, ce qui nous permettrait de voir si la nature des ions liés dans le centre actif de la protéine a un impact sur son activité enzymatique. Tout ceci ayant pour but de caractériser biochimiquement l'enzyme et de trouver un ion capable d'inactiver l'enzyme sans la déstabiliser, et ainsi permettre d'obtenir le complexe enzyme : substrat pour une étude structurale.

c. Conclusion

La phosphodiesterase AccF des souches Bo542 et C58 d'*A. tumefaciens* est capable de reconnaitre et de cliver les agrocinopines. Cette protéine est dimérique et possède un repliement de type Triophosphate isomerase (TIM). Des alignements de séquences et de structures ont montré que cette

protéine était proche des glycerophosphodiesterases. En effet le repliement est conservé, ainsi que les résidus identifiés comme étant impliqués dans la catalyse enzymatique (H16, H60 chez $AccF_{Bo542}$).

Lors de la reconnaissance et la stabilisation du substrat, certains résidus sont impliqués dans l'établissement de liaisons hydrogènes. Contrairement aux autres phosphodiesterases, un bon nombre de résidus ne sont pas conservés et portent la spécificité de reconnaissance de chacune des protéines vis-à-vis de son substrat respectif. Cependant, deux résidus hydrophobes localisés sous le ligand sont conservés et pourraient avoir pour rôle de maintenir le substrat à une distance convenable/optimale des résidus catalytiques (F154 et L174).

Des tests enzymatiques ont ensuite montré une faible affinité (de l'ordre du millimolaire) de la protéine $AccF_{Bo542/C58}$ pour son substrat, qui peut peut-être être expliquée par la flexibilité de certains résidus du site actif mais également la nature de l'ion utilisé. Ces tests ont également mis en évidence la présence d'une coopérativité des sites catalytiques. Cette allostérie a été démontrée par le positionnement des résidus E195 et H197 favorable à l'interaction d'un second substrat après fixation d'un premier.

Les différentes analyses de DSC et Tycho permettant d'obtenir des informations sur une stabilisation de la protéine grâce à la mesure des Tm et Ti (température de fusion et d'inflexion) n'ont permis d'observer qu'une seule variation, en présence d'agrocinopine C* modifiée, et la variation de Tm, une reconnaissance du G2P. Malheureusement, les optimisations des petits cristaux obtenus par criblage automatisé, réalisées pour obtenir la structure du complexe AccF-Agrocinopine C* modifiée, n'ont pour l'instant pas encore abouti à l'obtention de nouveaux cristaux.

Discussion et perspectives

Les bactéries du genre *Agrobacterium* sont majoritairement saprophytes mais peuvent être pathogènes lorsqu'elles possèdent un plasmide de virulence nommé pTi (Tumor Inducing). Ces pathogènes ont la capacité d'infecter une plante blessée pour ensuite transférer et intégrer un fragment d'ADN-t (ADN de transfert) dans le chromosome de la plante. Cet ADN-t présent chez la plante a pour impact le détournement de certaines voies métaboliques au profit de la bactérie. Le plasmide possède des gènes codant pour des enzymes qui favorisent la formation de phytohormones. Leur forte production engendre par la suite un déséquilibre hormonal qui conduit à une multiplication anarchique des cellules et finit par former des tumeurs servant de niches écologiques aux bactéries. Cet ADN-t possède également des gènes codant pour des enzymes qui permettent la synthèse de molécules nommées opines. Il en existe 8 familles, mais celles qui ont été utilisées dans cette étude sont les agrocinopines. Ces opines sont des sucres phosphorylés produits par la plante grâce au gène *acs* (agrocinopine synthase).

L'utilisation de ces molécules par les bactéries n'est possible que pour celles possédant le plasmide de virulence pTi et plus particulièrement l'opéron *acc*. En effet, cet opéron possède les gènes permettant la reconnaissance (*accA)* et la transformation (*accF*) des agrocinopines afin de les utiliser. Ces molécules spécifiques vont jouer un rôle dans le développement de la cellule bactérienne mais également dans la pérennité de cet avantage trophique. En effet, ces sucres phosphorylés une fois transformés sont une source de carbone favorisant le développement de la bactérie dans les tumeurs. Cependant, la partie pyranose-phosphate produite après transformation des agrocinopines est utilisée comme molécule initiatrice de la communication cellulaire (Quorum-sensing). Cette communication va avoir un impact sur la conjugaison bactérienne et plus particulièrement le transfert horizontal du plasmide de virulence (pTi). Ce mécanisme permet donc le développement de la bactérie pathogène et la pérennité de cet avantage trophique.

Il existe plusieurs souches de bactéries du genre *Agrobacterium* possédant l'opéron *acc* et pouvant utiliser ces agrocinopines et notamment les souches *A. tumefaciens* C58, *A. tumefaciens* Bo542 et *A. vitis* S4. Chacune de ces souches ont la particularité de faire produire à la plante des types agrocinopines différentes avec par exemple les agrocinopines A et B pour la souche C58 et C et D pour la souche Bo542. L'architecture de ces agrocinopines est semblable mais se différentie par la nature des sucres phosphorylés. Les agrocinopines A et B sont constituées d'arabinose-phosphate lié à un sucrose ou un fructose, respectivement. Alors que les agrocinopines C et D sont composées d'un glucose-phosphate lié à un fructose ou à un glucose, respectivement (Figure 61).

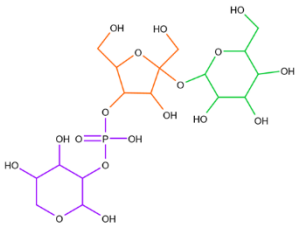
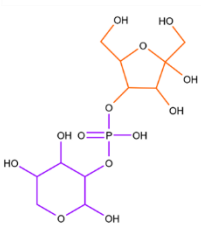
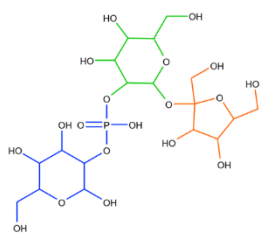
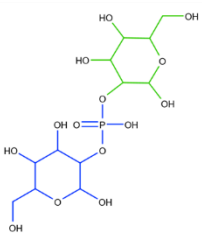
Souche	C58		Bo542	
Agrocinopine synthétisée				
Molécule « précurseur »	**Agrocinopine A** Arabinose-2-phosphate-fructose-glucose	**Agrocinopine B** Arabinose-2-phosphate-fructose	**Agrocinopine C** Glucose-2-phosphate-glucose-fructose	**Agrocinopine D** Glucose-2-phosphate-glucose

Figure 61 : Représentation des structures chimiques des agrocinopines A, B, C et D pouvant être synthétisés grâce aux gènes *acs* des souches C58 et Bo542 du genre *Agrobacterium*. En violet est représenté le motif arabinose-2-phosphate, en orange le fructose, en vert le glucose et en bleu de glucose-2-phosphate.

De précédents travaux réalisés concernant la souche modèle C58 ont montré qu'elle était capable de reconnaitre et d'utiliser tous les types d'agrocinopines comme source de carbone et comme molécules initiatrice du quorum sensing (El Sahili et al., 2015). Dans le but d'augmenter les connaissances concernant la reconnaissance et la transformation des agrocinopines par les bactéries du genre *Agrobacterium*, je me suis intéressé aux protéines AccA et AccF de la souche *A. tumefaciens* Bo542.

Dans un premier temps avec l'étude de la spécificité de la PBP AccA permettant la reconnaissance des agrocinopines pour leur adressage à un transporteur ABC. Puis dans un second temps avec l'étude structurale et fonctionnelle de la phosphodiesterase AccF, qui permet de transformer les agrocinopines en pyranose-phosphate (A2P/G2P) et en sucrose.

a. Spécificité des PBP $AccA_{Bo542}$ et $AccA_{S4}$

Comme évoqué ci-dessus, de précédentes recherches ont permis de caractériser biochimiquement et structuralement la protéine $AccA_{C58}$. En effet, il a été montré que la protéine présentait une affinité élevée pour l'agrocinopine A et les dérivés d'agrocinopine (A2P et G2P), avec des constantes de dissociation de l'ordre du µM. De plus les structures obtenues ont montré une reconnaissance plus élevée de la partie pyranose-phosphate comparé à la partie sucrose. En suivant ce même modèle, j'ai voulu faire cette même analyse avec les protéines AccA des souches Bo542 et S4.

i. L'apport de connaissance des travaux effectués

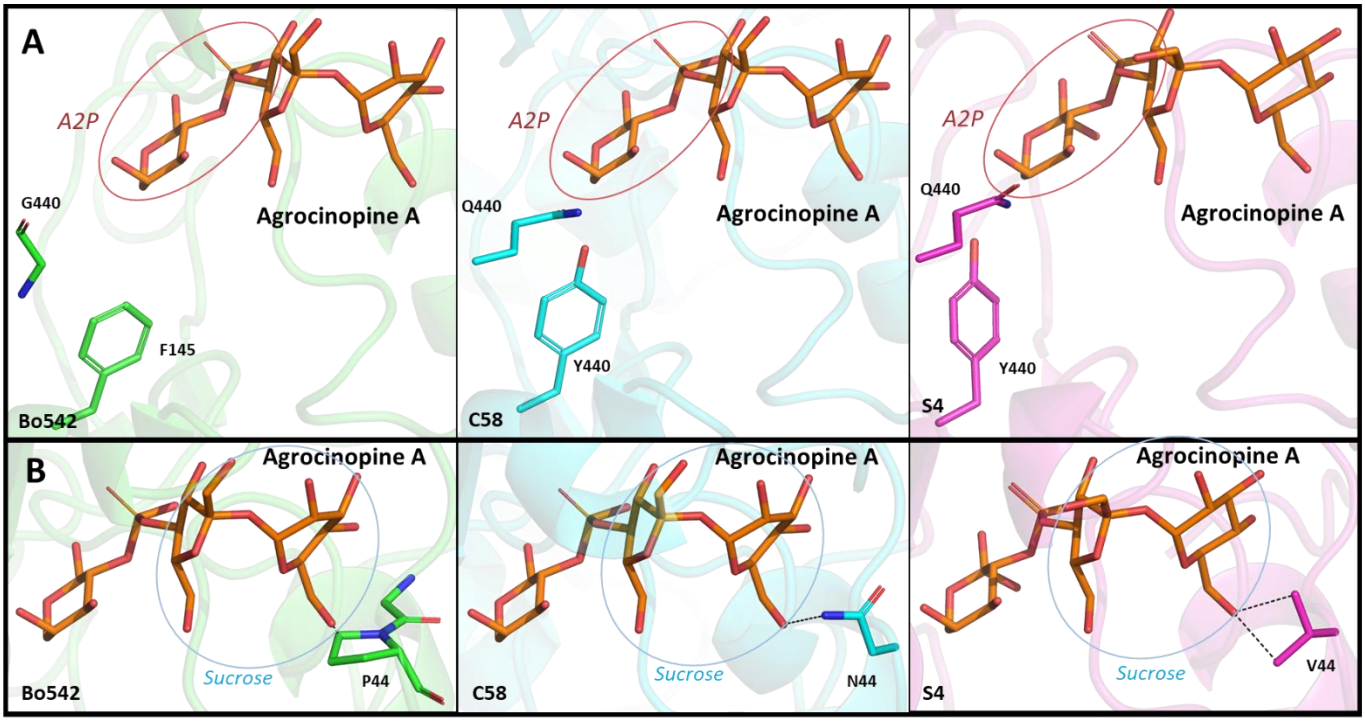
Dans un premier temps j'ai complété les travaux réalisés par l'ancien étudiant concernant la souche C58. En effet il avait démontré que la PBP AccA de la souche C58 pouvait reconnaitre de façon générale tous types de molécules possédant le motif pyranose-phosphate et en avait déduit la possible utilisation des agrocinopines C et D. Suite à la synthèse de l'agrocinopine D réalisée par nos collaborateurs (Yves QUENEAU, Mohammed AHMAR, INSA de Lyon – Equipe COB) j'ai pu caractériser biochimiquement l'interaction entre $AccA_{C58}$ et l'agrocinopine D. Sa forte affinité pour l'agrocinopine D se caractérise par une constante de dissociation de l'ordre du µM. Suite à l'obtention de cristaux et de jeux de données du complexe $AccA_{C58}$-Agrocinopine, j'ai pu réaliser l'analyse structurale et montrer que le ligand était stabilisé par 15 interactions hydrogènes et 4 interactions hydrophobes.

Dans un second temps, j'ai également caractérisé biochimiquement et structuralement $AccA_{S4}$. La protéine, possédant une identité de séquence de 81% avec $AccA_{C58}$, est donc capable de reconnaitre les agrocinopines A et D ainsi que leurs dérivés. Elle se comporte donc de la même façon qu'$AccA_{C58}$ avec une bonne affinité pour les agrocinopines et une constante de dissociation de l'ordre du µM. La protéine établit entre 14 et 22 liaisons avec les ligands lors de la reconnaissance des agrocinopines et dérivés (Cf Tableau 5 page 73).

Enfin, la caractérisation biochimique d'$AccA_{Bo542}$ a montré que cette protéine réagissait différemment en ne reconnaissant que l'agrocinopine D et le G2P. La protéine est donc spécifique de ses propres agrocinopines avec une affinité forte et une constante de dissociation de l'ordre du µM.

Afin de comprendre la spécificité de $AccA_{Bo542}$, j'ai réalisé l'étude structurale de cette protéine en complexe avec l'agrocinopine D. La protéine établit donc 17 interactions donc 5 hydrophobes avec l'agrocinopine D. De plus, j'ai comparé la structure de cette dernière avec celles des souches C58 et S4 afin d'observer des différences pouvant expliquer cette spécificité. J'ai également réalisé une superposition des sites de fixation me permettant d'avoir une idée du positionnement de l'agrocinopine A dans le site de fixation d'$AccA_{Bo542}$. Ceci m'a donc permis, malgré un RMSD de 1Å et une identité de séquence de 74 et 81 % avec les protéines des souches C58 et S4 respectivement, de mettre en évidence plusieurs résidus potentiellement impliqués dans la spécificité de reconnaissance des agrocinopines : résidus 44, 141, 145, 402 et 440. Suite à ces observations, j'ai décidé d'utiliser des mutants dans le but de permettre à $AccA_{Bo542}$ (normalement

spécifique de ses propres agrocinopines) de reconnaitre l'agrocinopine A et l'A2P. Durant cette étude, j'ai donc utilisé des mutants simples P44A et P44N ainsi que le double mutant F145Y/G440Q. Le choix de ces mutants a été motivé par leur variabilité entre la souche Bo542 et C58/S4, leur localisation à proximité de la région arabinose-2-phopshate pour le double mutant



F145Y/G440Q (Figure 62 A) et le potentiel clash stérique entre la proline 44 d'AccABo542 avec l'agrocinopine A (Figure 62 B).

Figure 62 : Représentation des résidus mutés dans le but de permettre une reconnaissance de l'agrocinopine A par $AccA_{Bo542}$. A : Représentation des résidus 145 et 440 mutés dans le double mutant $AccA_{Bo542}$ F145Y/G440Q. B : Représentation du résidu 44 muté dans les simples mutants P44A et P44N.

Aucune structure de ces mutants n'a malheureusement pu être obtenue et la caractérisation des interactions biochimiques a montré une affinité vis-à-vis de l'agrocinopine D légèrement plus faible avec une constante de dissociation toujours de l'ordre du micromolaire. Cependant aucune reconnaissance de l'agrocinopine A ou de l'arabinose-2-phosphate n'a été observée. Les mutants simples ou double utilisés n'ont donc pas permis d'expliquer la spécificité d'$AccA_{Bo542}$ et il semblerait donc que d'autres résidus soient impliqués dans la spécificité de reconnaissance de l'agrocinopine A.

ii. Perspectives

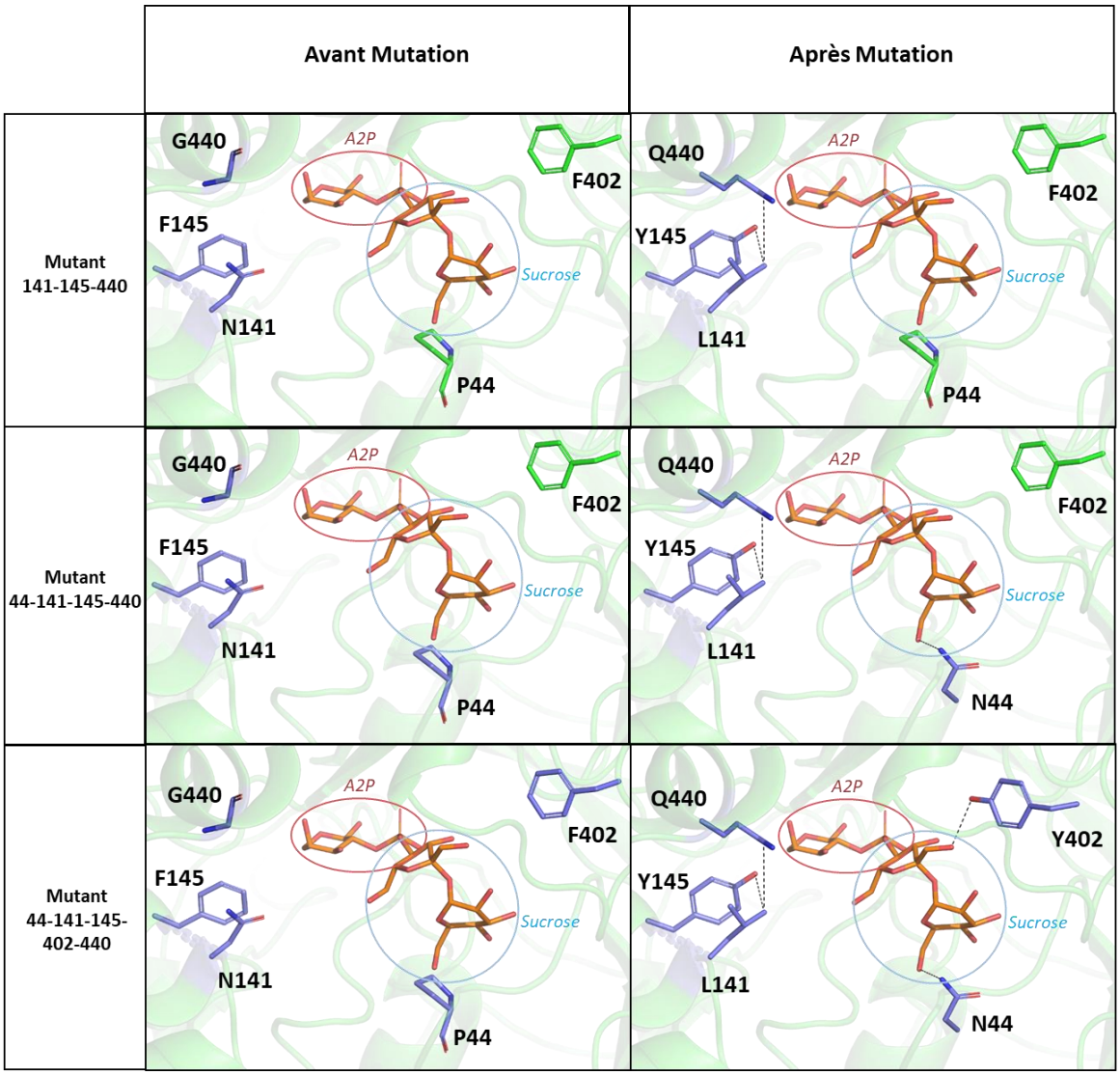
1. *Comprendre la spécificité de reconnaissance*

Les découvertes engendrées par ces résultats ont ouvert de nouvelles pistes à explorer et a

également soulevé de nombreuses questions auxquelles il serait intéressant de répondre. En effet, il serait intéressant d'identifier tous les résidus impliqués dans la spécificité de la protéine. Les résidus responsables de l'absence de reconnaissance de l'A2P et de l'agrocinopine par d'$AccA_{Bo542}$ seraient localisés vers la partie pyranose-phosphate des agrocinopines. Cependant, il ne faut pas écarter une potentielle implication des résidus situés près de la partie sucrose des agrocinopines pouvant avoir un impact sur la reconnaissance de l'agrocinopine A indépendamment de l'A2P. Les travaux réalisés ont permis de mettre en avant différents résidus potentiellement importants dans la reconnaissance de l'A2P avec les résidus L141, Y145 et Q440. Alors que les résidus P44N et Y402 seraient quant à eux plutôt impliqués dans la reconnaissance de l'agrocinopine A. Il serait intéressant de réaliser ces mutants, étape par étape, à savoir :

- N141L, F145Y et G440Q
- N141L, F145Y, G440Q et P44N
- N141L, F145Y, G440Q, P44N et F402Y

Le premier mutant permettrait de vérifier une possible reconnaissance de l'A2P. Si celle-ci est



reconnue contrairement à l'agrocinopine A, alors les autres mutants seraient à tester (Figure 63).

Figure 63 : Représentation des résidus d'AccA$_{Bo542}$ potentiellement impliqués dans la reconnaissance de l'A2P et l'agrocinopine A. En violet les résidus potentiellement impliqués et modifiés et en vert les résidus potentiellement impliqués mais non modifiés.

2. *Implications in vivo de la spécificité*

Les bactéries du genre *Agrobacterium* ont la capacité d'infecter divers types de plantes et d'arbres, pour lesquels cette infection engendrerait le développement d'une galle. Certaines souches ayant le même mode d'infection, elles pourraient être amenées à infecter la même plante si elles sont localisées au même endroit. La difficulté, le manque de clarté des écrits et sans doute le manque de connaissances concernant la localisation des différentes souches dans le monde ne permettent pas de citer d'exemples concrets. Cependant, d'un point de vue plus général, il serait intéressant d'analyser un potentiel désavantage trophique de la spécificité de reconnaissance des agrocinopines par la souche Bo542 par rapport à la souche C58. En effet, dans le cas de co-infection d'un même plant par la souche C58 et Bo542, y aurait-il un désavantage à ne reconnaitre qu'un type d'agrocinopine ? Pour essayer de répondre à cette question, des tests préliminaires de co-infection *in planta* avec les souches C58(pTiC58) et C58(pTiBo542) sur plants de tomates ont été réalisés. Cependant, les résultats obtenus et plus particulièrement les dénombrements bactériens réalisés sur les différents réplicas n'étaient pas concordants et n'ont pas permis de répondre clairement à cette question. La réalisation de travaux *in planta* est quelque chose de complexe où l'entièreté des paramètres n'est pas contrôlée. La discordance des résultats obtenus est sans doute dû à la variation d'un de ces paramètres non contrôlés. Afin de pouvoir partiellement répondre à cette question, il aurait été intéressant de réaliser des tests complémentaires tels que des études *in vitro* afin de s'affranchir des paramètres de croissance végétale pouvant impacter les résultats obtenus.

b. *Phosphodiesterase AccF d'Agrobacterium C58 et Bo542*

Contrairement à AccA, la phosphodiesterase AccF est une protéine qui a été très peu étudiée et dont les connaissances sont très limitées. Un ancien doctorant de l'équipe avait commencé des travaux sur AccF$_{C58}$ au cours desquels il avait réussi à résoudre la structure de la protéine seule par remplacement moléculaire mais n'avait pas entrepris d'autres expériences. Mes travaux se sont inspirés de ses expériences mais avec la souche Bo542 d'*A. tumefaciens*.

i. L'apport de connaissance des travaux effectués

Dès le début de ma thèse, j'ai exprimé et purifié la protéine $AccF_{Bo542}$. La protéine s'est avérée être un dimère en solution et un ajout de calcium a été nécessaire pour stabiliser la protéine. Ensuite, des cristaux de la protéine seule ou en complexe m'ont permis de résoudre la structure de la protéine par remplacement moléculaire. J'ai réussi à obtenir la structure de cette protéine en présence de calcium et manganèse sous sa forme *apo* et sous sa forme complexée avec le produit de réaction comme le glucose-2-phosphate, le glucose et un intermédiaire réactionnel. La caractérisation structurale a montré que le dimère était stabilisé par 26 interactions hydrophobes, 17 interactions hydrogènes mettant en jeu 11 résidus (D65, N71, N76, F152, D153, F154, L172, H174, S175, R176, Y230) dont 2 ponts salins (D153 et R176) (Figure 64).

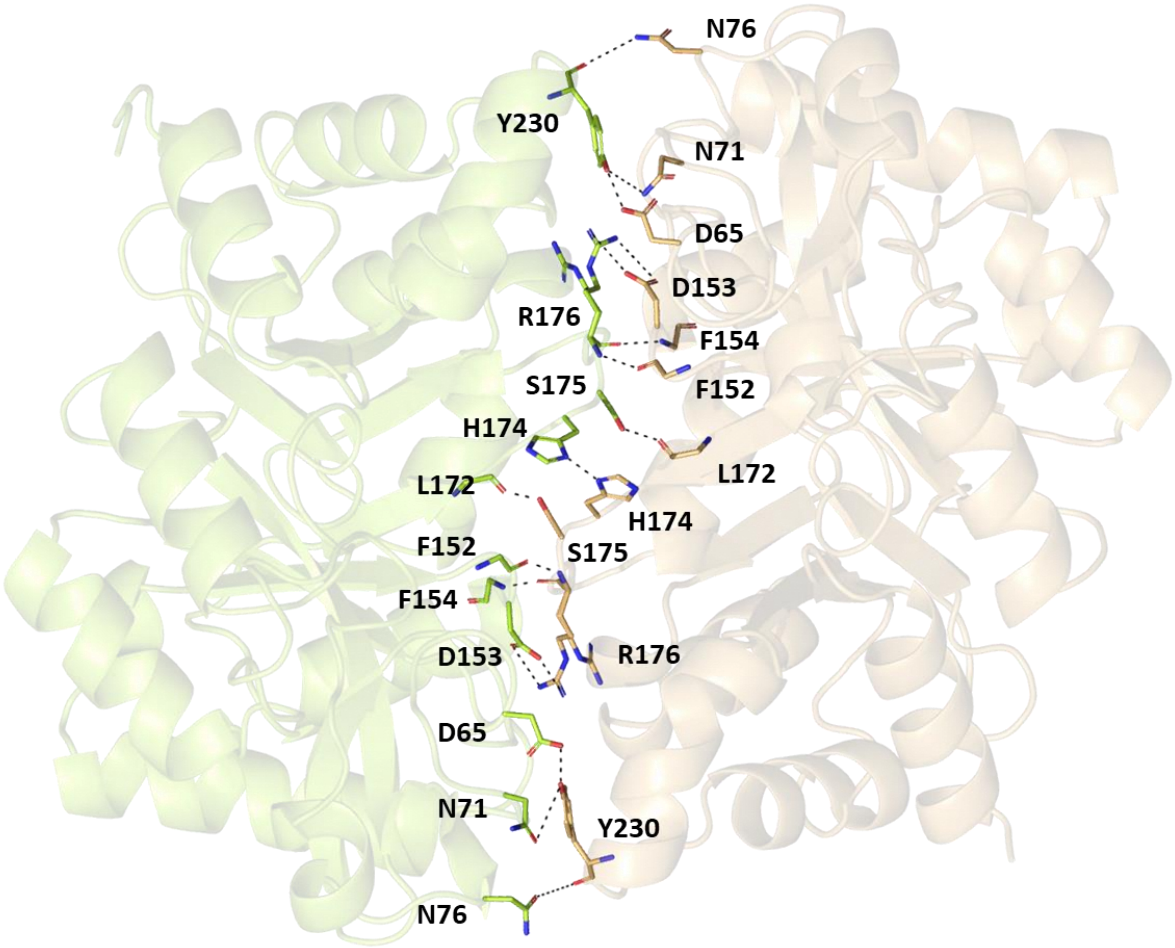
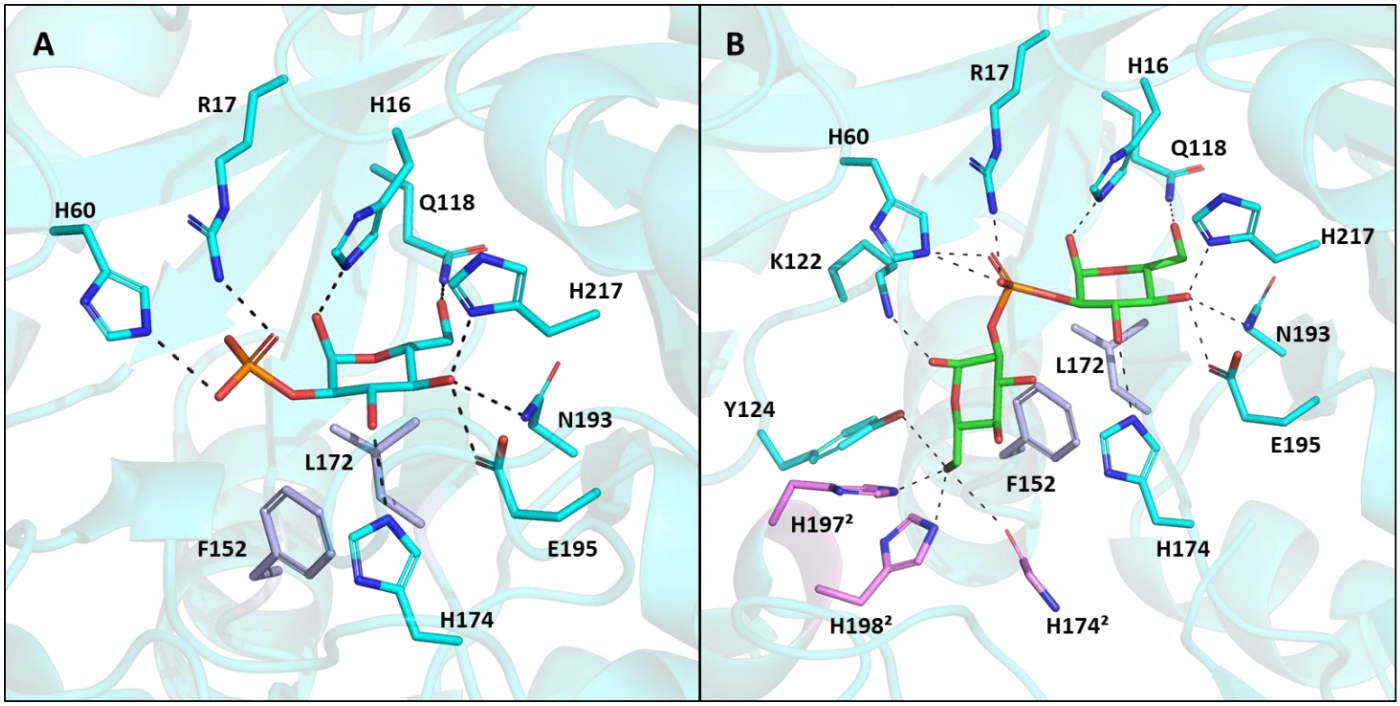


Figure 64: Représentation des interactions hydrogènes et ponts salins stabilisant le dimère AccF avec en vert et orange les monomères 1 et 2.

De plus, les différentes structures d'AccF *apo* ou *holo* ont été obtenues en présence de calcium ou de manganèse. Le choix des ions divalents utilisés fut motivé par des travaux réalisés par un ancien collaborateur John Reader de l'université de Caroline du nord à Chapel Hill (résultats non publiés). De plus, la bibliographie concernant des enzymes dont le site actif partage de nombreuses similitudes avec AccF ainsi que l'analyse structurale m'ont ensuite permis de mettre en

avant différents résidus comme étant impliqués dans le mécanisme réactionnel (H16 et H60) et impliqués dans la fixation de la partie glucose-2-phosphate de l'agrocinopine D (H16, R17, H60, Q118, H174, N193, E195, H217, F152 et L172) (Figure 65 A). De plus, la réalisation d'une superposition d'une molécule d'agrocinopine D sur le glucose-2-phosphate du complexe AccFBo542-G2P m'a permis d'émettre des hypothèses sur des résidus potentiellement impliqués dans la reconnaissance et la stabilisation du substrat (modélisation agrocinopine D : H16, R17, H60, Q118, K122, Y124, H174, H174², N193, E195, H197², H198², H217, F152 et L172) (Figure 65 B). J'ai observé également une possible position alternée des résidus 195 et 198 et d'après la modélisation de l'agrocinopine D dans le site de fixation, les deux résidus seraient potentiellement à



proximité du ligand.

Figure 65: Représentation de la protéine AccF en complexe avec le glucose-2-phophate en bleu et l'agrocinopine D (modélisation) en vert. Les résidus du monomère 1 sont représentés en bleu alors que ceux de domaine 2 sont représentés en rose. Les résidus en gris représentent les résidus du monomère 1 impliqués dans des interactions hydrophobes. **A :** Représentation en bâton des résidus impliqués dans la fixation du glucose-2-phosphate. **B :** Représentation en bâton des résidus potentiellement impliqués dans la reconnaissance de l'agrocinopine D

Dans le but d'obtenir la structure du complexe AccF-Agrocinopine D, je me suis orienté vers l'utilisation de mutants inactifs. L'analyse structurale m'a donc permis d'identifier de bons candidats me permettant d'obtenir une protéine stable capable de reconnaitre le substrat tout en étant inactif. Je me suis donc orienté vers une mutation des résidus H16, H60, H174, E195 et H197 en alanine (Figure 66).

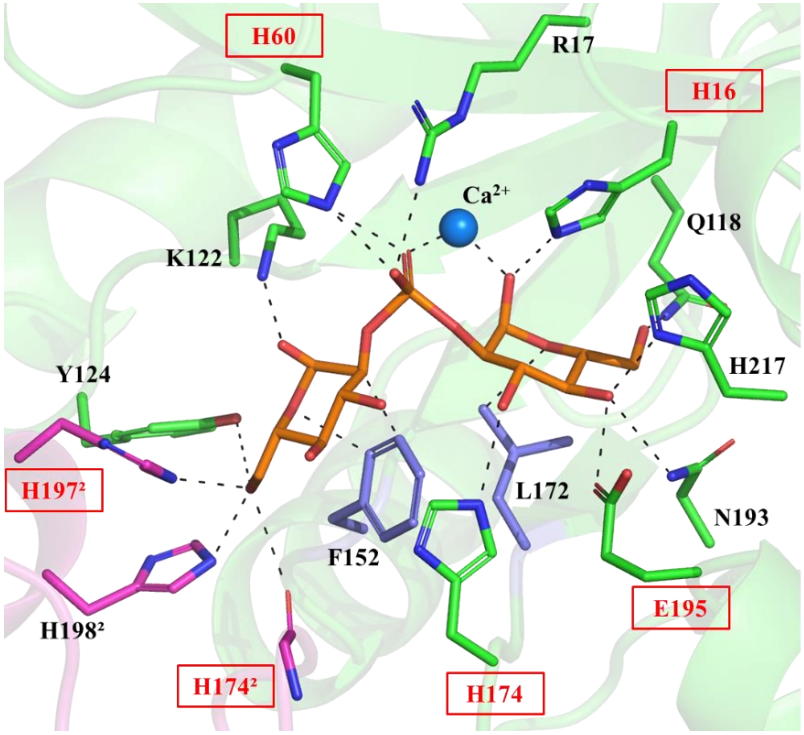


Figure 66 : Modélisation du site catalytique de l'enzyme AccF en complexe avec l'agrocinopine D (en orange) et un ion calcium. Cette figure représente les résidus impliqués dans l'activité catalytique (H16 et H60) et la potentielle implication de certains résidus dans la fixation du substrat (H16, R17, H60, Q118, K122, Y124, F152, L172, H174, H174², E195, H197, H198 et H217). Le signe « ² » représente les résidus de la deuxième sous unité. Les résidus verts et violets représentent les résidus de la premiere sous unité alors que les résidus roses font partie dans la seconde sous unité. Enfin les résidus violets (F152 et L172) représentent les résidus impliqués dans la création d'un environnement hydrophobe.

J'ai ensuite mis au point un test me permettant de mesurer l'activité catalytique de protéine. Les produits de la réaction étant du glucose-2-phosphate et du glucose, j'ai mis au point un protocole me permettant de doser le glucose en utilisant la glucose déshydrogénase et du NAD^+. Grâce à ces tests, j'ai pu mettre en évidence que l'enzyme était allostérique et que les mutants H16A, H60A et H197A étaient inactifs alors que les mutants H174A et E195A possédaient une activité résiduelle. Par la suite, j'ai essayé d'obtenir la structure de ces mutants en complexe avec le substrat, mais je n'ai à ce jour obtenu que la forme *apo* des mutants utilisés.

Toujours dans le but de diminuer l'activité enzymatique de la protéine afin d'obtenir la structure du complexe enzyme : substrat, je me suis orienté vers la nature de l'ion catalytique. En effet des tests préliminaires de notre collaborateur mettaient en avant la perte d'activité de la protéine en fonction de l'ion métallique. J'ai donc réalisé des tests préliminaires visant à modifier l'ion métallique tout en m'assurant de la stabilité de la protéine. Ces tests obtenus m'ont permis de montrer que la protéine était stable en présence de calcium, nickel, cobalt, magnésium et manganèse.

Pour finir, toutes ces études ont tout de même pu apporter et confirmer des informations importantes concernant le mécanisme enzymatique d'AccF et notamment l'importance des résidus catalytique H16 et H60 en tant qu'acide et base dans cette catalyse.

ii. Perspectives

Suite aux travaux réalisés et aux résultats obtenus, il serait intéressant de valider les tests d'activité afin de confirmer les informations obtenues. Dans un premier temps, les travaux ont été effectués en présence de calcium et de manganèse. Les résultats de l'observation de l'impact de différents métaux sur l'activité de l'enzyme pourraient être utilisés afin d'obtenir la structure enzyme-substrat. Des tests préliminaires ont permis de valider la stabilité de la protéine en présence des différents métaux. Il serait désormais intéressant d'analyser l'activité de la protéine en leur présence. L'utilisation d'ions divalents inactivant ou diminuant l'activité enzymatique de la protéine permettrait donc d'obtenir la structure enzyme-substrat.

Dans un second temps, les tests de DSC ont permis de montrer une fixation du G2P par la protéine sauvage. Il serait intéressant de continuer dans cette direction et de regarder si les protéines mutées peuvent cette fois-ci fixer le substrat sans risque d'artefact lié à la réalisation de la réaction. En effet, l'augmentation de la température d'une protéine en présence ou en absence de son ligand peut nous permettre d'observer une augmentation de la stabilité et ainsi témoigner de la présence d'une interaction. Cependant, lorsqu'il s'agit d'une enzyme, l'interaction n'est pas stable et des artéfacts peuvent être observés lors de l'expérience sauf si l'enzyme est inactive. Etant donné que la reconnaissance du produit de réaction (G2P) est possible par l'enzyme sauvage, il serait donc intéressant de voir si la reconnaissance du substrat (agrocinopine D ou A) est observable par les mutants inactifs d'AccF (H16A, H60A, H197A). Enfin, si il est possible d'observer une interaction stable du substrat par les mutants, il serait intéressant de poursuivre les tests de cristallisation enzyme(mutée) : substrat.

Dans un troisième temps, j'ai pu montrer que les protéines $AccF_{Bo542}$ et $AccF_{C58}$ sont toutes les deux capables de reconnaitre et de transformer l'agrocinopine D en G2P et glucose. $AccF_{C58}$ n'est donc pas spécifique de ses agrocinopines et peut utiliser, internaliser (cf : Partie II) et transformer les agrocinopines. Tout en sachant que la souche Bo542 ne peut pas internaliser l'agrocinopine A, il serait intéressant d'un point de vue structural de savoir si $AccF_{Bo542}$ est capable de reconnaitre et d'hydrolyser l'agrocinopine A.

Dans un quatrième temps, l'utilisation d'$AccF_{Bo542}$ a permis l'obtention des structures AccF-G2P et AccF-intermédiaire réactionnel contrairement aux étude réalisées avec la souche C58. Il serait intéressant d'étudier AccF de la souche S4 pour voir si elle ne permettrait pas d'obtenir des

informations complémentaires.

Enfin, toutes les recherches évoquées précédemment, s'inscrivent dans la compréhension du mécanisme catalytique et de la phosphodiesterase AccF du genre *Agrobacterium* qui est une enzyme très peu connue.

Matériels et Méthodes

a. Synthèse et purification des agrocinopines et dérivés.

La synthèse et la purification des ligands utilisés tout au long de cette thèse ont été réalisés par des collaborateurs chimistes (Yves QUENEAU, Mohammed AHMAR, INSA de Lyon – Equipe COB). La synthèse de l'agrocinopine A, du L-arabinose-2-phosphate et du D-glucose-2-phosphate) a été précédemment décrite dans l'article (El Sahili et al., 2015).

b. Bilan des clonages et mutagénèses

Les gènes codants pour les protéines d'intérêt sont soit synthétisés (Genescript) soit amplifiés par réaction de polymérisation en chaîne (PCR), à partir d'ADN génomique purifié (Tableau 11). Le produit de synthèse ou de PCR est ensuite inséré au niveau du site de clonage multiple d'un plasmide d'expression bactérien. Les plasmides de routine au laboratoire sont des vecteurs commerciaux de type pET20 et pET9 (Novagen). Une étiquette 6His est ajoutée en 5' des gènes, afin de faciliter la purification des protéines recombinantes. Les mutants ont également été obtenus et synthétisés par la société Genescript ou par le travail de nos collaborateurs.

	A. tumefaciens	*A. tumefaciens*	*A. vitis*
	Bo542	**C58**	**S4**
AccA		El Sahili et al., 2015	
Clonage	Synthèse de gène (Genescript, piscataway, NJ)	PCR	Synthèse de gène (Genescript, piscataway, NJ)
Mutations	P44A, P44N et F140Y/G440Q		
Plasmide	pET-9a	pET-9aSN1	pET-20b
Antibiotique	Kanamycine	Kanamycine	Ampicilline
AccF		Collaboration John READER	
Clonage	Synthèse de gène (Genescript,		
Mutations	H16A, H60A, H174A, E195A, H197A	H18A et H62A	
Plasmide	pET-9b (wt) et pET29a	pET21b	
Antibiotique	Kanamycine	Kanamycine	

Tableau 11 : Récapitulatif des méthodes d'obtention des plasmides contenant les gènes codants pour les protéines d'intérêts utilisées

c. ***Expression des protéines***

Le plasmide de surexpression, une fois cloné et vérifié par séquençage (Beckman Coulter Genomics), est introduit dans des bactéries compétentes de la souche *E. coli* afin de permettre la surproduction recombinante des protéines étudiées.

Souches et plasmides utilisés

Souches	Caractéristiques	Source & référence
Xl1-blue	*endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+*	Stratagen
Rosetta (DE3)pLysS	F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm* (DE3) pLysSRARE (Cam^R)	Novagen
BL21-Gold (DE3)	F - *ompTgal[dcm][lon] hsdSB(rB-mB-)* (DE3) endA Hte	Novagen
C41 (DE3)	F – *ompT hsdSB (rB- mB-) gal dcm* (DE3)	Novagen
NTLR4		(Cha et al., 1998) Obtenu grâce à l'équipe de Denis FAURE, I2BC.

Tableau 12 : Listes des souches

Plasmides	Caractéristiques	Sources
pET-20	*pT7, ori pBR322,* Amp^R	Novagen
pET-9	*pT7, ori pBR322,* Kan^R	Novagen
pET-9aSN1	*pT7, ori pBR322,* Kan^R	Equipe FAAM

Tableau 13 : Listes des plasmides

i. Transformation par choc thermique

La transformation des bactéries par les plasmides contenant les gènes codant pour les protéines étudiées a été tout au long de cette thèse réalisée par choc thermique et le protocole utilisé ne varie pas en fonction des souches et des plasmides mis à part l'antibiotique de sélection des transformants.

Les bactéries compétentes stockées congelées à -80°C sont tout d'abord décongelées dans la glace puis 1-2µl de plasmide (≈100/200 ng) sont ajoutés aux 50µl de bactéries puis laissé 30 min dans la glace. Les tubes eppendorf sont ensuite plongés 40 secondes dans un bain-marie préalablement chauffé à 42°C. Les tubes sont ensuite remis dans la glace pendant 2-5 min avant ajout d'1 ml de milieu LB et incubés 1h à 37°C. 100µl de la solution bactérienne sont ensuite étalés

sur boite LB-agar-antibiotique à 37°C pendant la nuit.

ii. Tests d'expression

Dans le but de déterminer les conditions optimales d'expression de chacune des protéines étudiées, un test d'expression est réalisé. Trois bactéries, Bl21 (DE3), C41 (DE3) et Rosetta (DE3) pLysS (Novagen) sont transformées par choc thermique avec les plasmides contenant le gène de la protéine d'intérêt puis isolées sur milieu sélectif. Trois cultures de 10 ml de chacun de ces transformants sont réalisées et inoculées à DO 0.1 par une préculture réalisée sur le nuit. Elles sont incubées à 37°C pendant 1 heure puis induites avec 0.5 mM d'IPTG. Les trois cultures de chaque transformant sont ensuite placées à 16, 28 et 37°C durant 20, 4 et 3 heures respectivement. Une fois terminé, les cultures sont centrifugées 10 min à 8000g. Le surnageant est éliminé et le culot est repris dans 400 µl de tampon A (50 mM Tris-HCl pH8, 300 mM NaCl et 20 mM Imidazole) puis congelé à -20°C. Une fois tous les culots récupérés et repris en tampon, 5 cycles de congélation (azote liquide) /décongélation (37°C) sont réalisés. Les tubes sont centrifugés à 13.000 rpm durant 30 min et les surnageants sont ensuite isolés et mis en présence de 15 µl de billes de nickel NiNTA (Sigma-Aldrich) (1:1) préalablement lavées en tampon A. Après une incubation de 10 min à 4°C sur une roue permettant l'agitation, la réalisation de 5 lavages des billes sont effectués. Les lavages se font en 3 étapes : centrifugation des tubes à 13.000 rpm pendant 1 min, élimination du surnageant et ajout de 500 µl de tampon A.

Au cours du test d'expression, 10 µl d'échantillons sont prélevés à différentes étapes afin de suivre l'évolution du test. Enfin, un tampon de chargement est ajouté aux différents échantillons et ils sont ensuite déposés sur gel SDS PAGE, afin de juger des niveaux d'expression et de solubilité des protéines recombinantes.

iii. Expressions à grande échelle

Les expressions à grande échelle des protéines étudiées ont toutes été réalisées de la même manière, mis à part la température d'expression, le temps d'incubation et les antibiotiques de sélection utilisés. Les différences sont répertoriées dans le tableau 14 pour chaque souche.

Pour débuter l'expression de la protéine d'intérêt, 5 ml de préculture + antibiotique sont ajoutés à 1L de LB pour ensuite être mis en culture à 27°C o/n. Les bactéries ont ensuite été culotées et resuspendues dans 1L de milieu LB frais et mis en culture à 37°C pendant 1 heure. La surexpression protéique est déclenchée par ajout de d'isopropyl β-D-1-thiogalactopyranoside

(IPTG) 0.5 mM et est réalisée durant 4 heures. Les cultures sont ensuite centrifugées à 8000g pendant 20 min à 4°C et stocké à -20°C.

	A. tumefaciens	*A. tumefaciens*	*A. vitis*
	Bo542	**C58**	**S4**
AccA		El Sahili et al., 2015	
Mutations	P44A, P44N et F140Y/G440Q		
Antibiotique	Kanamycine		Ampicilline
Souche d'expression	BL21 (DE3)	BL21 (DE3)	Rosetta(DE3)pLysS
Induction	0,5mM IPTG	0,5mM IPTG	0,5mM IPTG
Expression	37°C 4h	37°C 3h	16°C o/n
Etapes de purification	HiTrap, Superdex200	HiTrap, Superdex200	HiTrap, Superdex200
AccF		Collaboration John READER	
Mutations	H16A, H60A, H174A, E195A, H197A	H18A et H62A	
Antibiotique	Kanamycine	Kanamycine	
Souche d'expression	BL21 (wt et H16A) et Rosetta (autres mutants)	HSM174	
Induction	0,5mM IPTG	0,5mM IPTG	
Expression	16°C o/n et 28°C 4h (mutant H174A)	16°C o/n	

Tableau 14: Tableau récapitulatif des méthodes d'expression de chacune des protéines d'intérêts.

d. Purification des protéines

Toutes les protéines étudiées durant cette thèse ont été purifiées de la même façon. Les cellules sont tout d'abord suspendues dans 50 mM Tris-HCL pH8, 300 mM NaCl, 20 mM imidazole (tampon A) et lysées par sonication. Après centrifugation à 25000g durant 45 min, le filtrat du surnageant est injecté sur une colonne d'affinité de nickel (HiTrap 5 ml, Healthcare). La protéine est éluée avec du tampon A supplémenté avec 300 mM d'imidazole (tampon B) et est ensuite injectée sur une colonne de gel filtration Superdex 200 26/60 (GE Healthcare) en utilisant un tampon composé de 50 mM Tris-HCL pH 8 et 150 mM NaCl. Pour la protéine AccF, du calcium est ajouté aux fractions récupérées avant concentration pour stabiliser la protéine et éviter sa précipitation. Les différentes étapes de purification sont suivies sur SDS-PAGE. Pour finir, les protéines purifiées sont concentrées (Vivaspin, GE Healthcare), aliquotées, congelées dans l'azote liquide et stockée à -80°C. Leur concentration est mesurée par absorbance à 280 nm.

	A. tumefaciens	*A. tumefaciens*	*A. vitis*
	Bo542	**C58**	**S4**
AccA		El Sahili et al., 2015	
Mutations	P44A, P44N et F145Y/G440Q		
Etapes de purification	HiTrap, Superdex200	HiTrap, Superdex200	HiTrap, Superdex200
Tampon	50mM Tris-HCL pH 8 et 150mM NaCl	50mM Tris-HCL pH 8 et 150mM NaCl	50mM Tris-HCL pH 8 et 150mM NaCl
Poids moléculaire (kDa)	56,2	57,1	59,6
Forme Oligomerique	Monomère	Monomère	Monomère
Volume Elution	220	220	220
Calibre vivaspin	30	30	30
AccF		Collaboration John READER	
Mutations	H16A, H60A, H174A, E195A, H197A	H18A et H62A	
Etapes de purification	HiTrap, Superdex200	HiTrap, Superdex200	
Tampon	50mM Tris-HCL pH 8 et 150mM NaCl	50mM Tris-HCL pH 8 et 150mM NaCl	
Poids moléculaire (kDa)	30,6	30,9	
Forme Oligomerique	Dimère	Dimère	
Volume Elution	210	210	
Calibre vivaspin	10	10	

Tableau 15: Récapitulatif des méthodes de purification et des paramètres biochimiques des protéines d'intérêt utilisées

e. Signaux de quorum sensing et transfert conjugatif

La reconnaissance de l'agrocinopine A, de l'agrocinopine D, de l'A2P ou du G2P peut engendrer la production de molécules signal OC8HSL chez les souche C58 et Bo542 d'*A. tumefaciens* et induire le transfert conjugatif du plasmide de virulence. Afin de vérifier si les souches bactériennes reconnaissent et utilisent les agrocinopines et dérivés, des mesures de la production des molécules « signal » et de l'efficacité de transfert conjugatif ont été réalisées.

La première étape est la construction des souches à utiliser. Les deux expériences ont été réalisées en parallèle et ont nécessité l'utilisation de 3 souches :

- **Receveuse :** C58:00 (résistante à la rifampicine)
- **Donneuse :** C58 (pTiC58)
- **Donneuse :** Bo542 (pTiBo542)

Ces souches ont été obtenues grâce à un don de l'équipe Denis FAURE.

i. Détermination de l'efficacité du transfert conjugatif

Dans un premier temps, une culture de chacune des 3 souches a été réalisée à 28°C sur la nuit. La densité optique à 600 nm a ensuite été mesurée afin d'inoculer à une DO de départ de 1 (bactéries receveuses et donneuses) dans un milieu de culture composé de TY et d'un inducteur (Agrocinopine A, Agrocinopine D, A2P, G2P). 150 µl de ces mélanges ont été déposés dans une plaque 96 puits en triplicate puis laissés à incuber 24/48/96 heures à 28°C. Des aliquots de 10 µl ont ensuite été congelés afin d'être utilisés pour les mesures de molécules signal OC8HSL et d'autres aliquots ont été étalés sur 3 milieux sélectifs afin de permettre la quantification des bactéries totales, des bactéries receveuses, et des bactéries transconjugantes.

- **Bactéries totales** : TY agar
- **Bactéries receveuses** : TY agar + Rif 100
- **Bactérie transconjugantes C58** : AB minimum agar + Rif 100 + Nopaline 2 mM
- **Bactérie transconjugantes Bo542** : AB minimum agar + Rif 100 + Mannopine 2 mM

L'utilisation de nopaline et de mannopine est propre aux souches C58 et Bo542, respectivement, et leur utilisation est possible grâce à la présence du plasmide Ti (C58/Bo542).

Enfin un comptage a permis de poser le ratio entre nombre de bactéries transconjugantes/donneuses, permettant ainsi de calculer l'efficacité de transformation en présence des différents inducteurs.

TY: (Bacto tryptone, 5 g/L; yeast extract, 3 g/L; 15 g/L)

AB (Agrobacterium broth) minimum: (K_2HPO_4,3g/L; NaH_2PO_4,1g/L; $MgSO_4.7H_2O$, 0.3 g/L; KCl, 0.15 g/L; $CaCl_2$, 0.01 g/L; $FeSO_4.7H_2O$, 2.5 mg/L; pH7)

ii. Mesure de la production de molécules signal

La mesure des molécules signal en présence d'inducteur (Agrocinopine A, agrocinopine D, A2P et G2P) a été réalisée dans un but qualitatif et quantitatif. Dans un premier temps, 5 µl d'échantillons bactériens congelés, obtenus comme décrit précédemment, ont été déposés sur des plaques de chromatographie en couche mince (plaques TLC). De la même façon, une gamme de molécules signal NAHLs allant de 1 à 400 nM a été déposée dans le but de quantifier les signaux. Un milieu de culture composé d'AB mannitol est ensuite ajouté sur la plaque TLC, ainsi qu'une culture

bactérienne NTLR4 (Tableau 12 page 135) et de l'Xgal[40], puis le tout est incubé toute la nuit à 28°C. Les plaques sont ensuite récupérées et le diamètre des halos bleus formés, dû à la dégradation du Xgal, est comparé à la gamme de NAHLs. Ainsi une mesure qualitative et quantitative est réalisée.

f. Cristallisation et collecte de données

Afin de déterminer les conditions de cristallisation des différentes protéines étudiées sous leur forme *apo* ou holo, plusieurs kits commerciaux de 96 conditions (Qiagen, Valencia, CA, USA), les robots de pipetage Cartesian (Genomic Solutions) et Mosquito (TTP LabTech) ont été utilisés selon la technique de la diffusion de vapeur, en goutte assise. Cette technique repose sur un phénomène d'échange de vapeur dans un milieu hermétiquement clos, entre un grand volume de solution concentrée en agent précipitant contenu dans un puit et une goutte plus faiblement concentrée en agent précipitant car mélangée avec la solution protéique. Ces échanges induisent une dessiccation progressive de la goutte, causant une augmentation de la concentration en protéine et en agent précipitant, suffisante pour permettre la nucléation et la formation progressive du cristal (Figure 67).

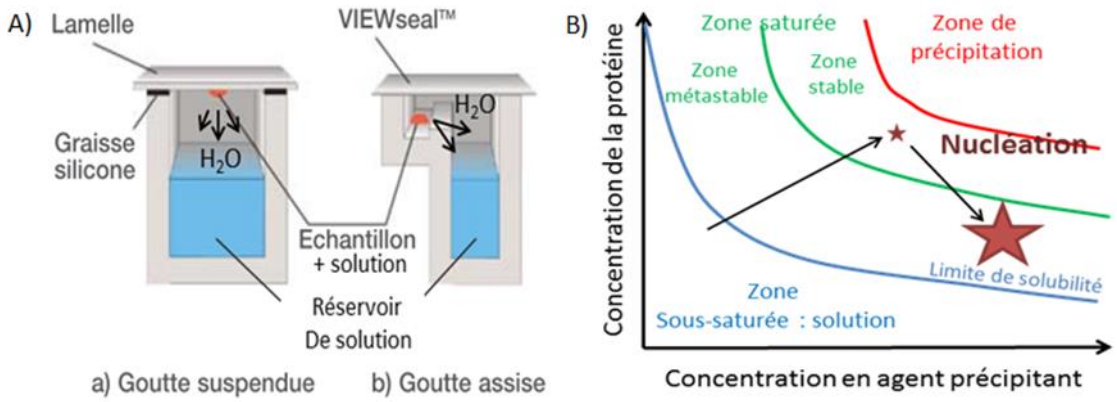


Figure 67: A) Principe de la diffusion en phase vapeur. **B)** Diagramme de phase du processus de cristallisation.

Grâce à la méthode de la goutte suspendue, les cristaux obtenus ont été ensuite manuellement reproduits et optimisés, en mélangeant les mêmes quantités de solution de protéines pures et d'agent de cristallisation, puis transférés dans une solution cryoprotectante mentionnée dans les tables cristallographiques présentes en Annexe page 147. Les cristaux ont enfin été vitrifiés dans l'azote liquide et transportés au synchrotron.

Les données obtenues par diffraction aux rayons X de ces cristaux ont été collectées à 100K sur les lignes de lumière Proxima 1/2 (synchrotron SOLEIL, Saint-Aubin, France). Les données ont

été traitées grâce à l'utilisation de la suite logiciel XDS (Kabsch, 2010) (Tables cristallographiques en Annexe page 147).

i. Cristaux de la protéine AccA

Avant mon arrivée au laboratoire, les études sur les protéines sauvages AccA des souches Bo542 et S4 avaient déjà commencé. Toutes les protéines que j'ai purifiées ont donc été cristallisées en utilisant le criblage comme expliqué précédemment ou en utilisant les conditions de cristallisation déjà mises au point par l'équipe. Les conditions ayant permis l'obtention de cristaux sont les suivantes (**Condition 1** : PEG 4000 35% ; **Condition 2** : PEG 8000 20%, 0.1M Hepes pH7 ; **Condition 3** : PEG 8000 20%, 0.1M MES pH6.5). Pour obtenir la forme complexée de la protéine avec le substrat, j'ai réalisé des co-cristallisations qui consistent à mélanger la protéine et le ligand en solution, selon un ratio (protéine/ligand) de 1/10 et un ratio (mix/solution de cristallisation) de 1/1, 1/2, 2/1. Les cristaux ont ensuite été trempés dans la solution de cristallisation additionnée d'un cryoprotectant pour être congelés et stockés. Selon les conditions de cristallisation, j'ai fait varier la nature et la concentration du cryoprotectant (22% PEG 400 pour les conditions précédemment citées) en suivant l'article de McFerrin de 2002 (McFerrin & Snell, 2002).

ii. Cristaux de la protéine AccF

En ce qui concerne la protéine AccF, deux conditions ont été retenues et ont permis la cristallisation de la protéine $AccF_{Bo542}$ en boites manuelles. (**Condition 1** : PEG 4000 15-30%, 50mM Tris pH8; **Condition 2** : Sulfate d'ammonium 1,8 à 2,1M, Hepes pH7.5). Pour obtenir la forme complexée de la protéine avec le substrat, j'ai réalisé des co-cristallisations qui consistent à mélanger la protéine et le ligand en solution, selon un ratio (protéine/ligand) de 1/10 et un ratio (mix/solution de cristallisation) de 1/1, 1/2, 2/1. J'ai également testé des infiltrations durant différents temps visant à permettre l'insertion du ligand dans le site de fixation après cristallisation de la protéine. Les cristaux ont ensuite été trempés dans la solution de cristallisation additionnée d'un cryoprotectant comme évoqué précédemment (22% PEG 400 pour la conditions 1 et 25% de glycerol pour la conditions 2).

g. Détermination des structures et affinements

i. Collecte des données et affinement des différentes structures

Après collecte des données de diffraction sur les ligne Proxima 1 ou 2 du Synchrotron SOLEIL

(Saclay), j'ai utilisé la suite de programme XDS (Kabsch, 2010) pour traiter tous mes jeux de données. A cette étape, j'ai sélectionné la résolution maximale à laquelle couper mes jeux de données. Pour cela, je me suis basé sur plusieurs critères. J'ai essayé d'obtenir des données complètes à au moins 95%. Le rapport signal sur bruit (I/σ) doit être supérieur à 1 et le coefficient de corrélation (CC) doit être supérieur à 0.5.

Pour le phasage par remplacement moléculaire, j'ai utilisé les programmes Phaser (McCoy, 2006) et Molrep (Vagin & Teplyakov, 2010). Ces deux logiciels fonctionnent par comparaison des cartes de Patterson générées par transformée de Fourier des modules des facteurs de structure, et qui représentent, par des pics de positions et d'intensités variables, tous les vecteurs intramoléculaires et intermoléculaires de la molécule. Le programme recherche le maximum de corrélation entre la carte de Patterson du modèle et du cristal en faisant varier l'orientation et la position du modèle. On détermine ainsi la meilleure matrice de superposition (rotation + translation).

Pour la réalisation de l'affinement, le logiciel BUSTER-2.10 a été utilisé avec les restrictions des symétries non cristallographiques lorsque les unités asymétriques étaient composées de plus de deux molécules (Blanc et al. 2004). L'utilisation du logiciel COOT a permis l'inspection des densités électroniques et la reconstruction manuelle des structures tridimensionnelles (Emsley ans cowtan 2004). Les modèles des structures tridimensionnelles des ligands utilisés ont été générés avec le serveur internet ProDRG (Schüttelkopf & Van Aalten, 2004). Les données caractéristiques de chaque structure sont présentées dans la Table cristallographique en Annexe page 147 et leurs représentations graphiques ont été obtenues grâce au logiciel PyMOL (http://www.pymol.org).

ii. Structure de la protéine AccA

Les structures de la protéine des souches Bo542 et S4 ont été obtenues par remplacement moléculaire. Le modèle utilisé est la protéine AccA de la souche C58 (Code PDB : 4ZEB) possédant 75 et 82% d'identité de séquence avec $AccA_{Bo542}$ et $AccA_{S4}$ respectivement (El Sahili et al., 2015). L'analyse du contenu en solvant a permis de mettre en évidence la présence d'une molécule (condition PEG 35%) par unité asymétrique (AU) pour $AccA_{Bo542}$. Il en est de même pour AccAS4 avec une molécule par AU (**Condition 1** : 20% PEG 8K, 0.1 M Hepes pH7 ; **Condition 2** : 20% PEG 8K, 0.1 MES pH 6.5).

iii. Structure de la protéine AccF

La structure cristallographique de la protéine $AccF_{C58}$ a été résolue par remplacement moléculaire à

une résolution de 2Å par un doctorant précédent. Le modèle utilisé pour ce remplacement moléculaire est la combinaison de deux structures de glycerophosphodiesterase (code PDB : 2PZ0 et 1ZCC) et a été obtenu grâce à l'utilisation du serveur internet @tome2 (Pons & Labesse, 2009; Shi et al., 2008).

En ce qui concerne la protéine $AccF_{Bo542}$, j'ai réussi à résoudre la structure tridimensionnelle par remplacement moléculaire en utilisant comme modèle celle de la souche C58. L'analyse du contenu du solvant a permis de mettre en évidence la présence de 2 (condition NH_4SO_4) et 1 ou 4 (condition PEG) monomères par unité asymétrique (AU).

h. Détermination de l'activité enzymatique d'AccF

La phosphodiesterase AccF transforme l'agrocinopine D en glucose-2-phosphate et glucose. L'activité enzymatique a été étudiée à 25°C en tampon Tris pH7,5 en mesurant la vitesse initiale de production de glucose issu de l'agrocinopine D par la glucose déshydrogénase de Pseudomonas sp (Sigma). La solution de réaction utilisée contenait 70mM de tampon Tris, 10mM de NAD (Simga-Aldrich), 0.16µM de glucose déshydrogénase et 3.8µM de protéine AccF. La réaction a ensuite été déclenché par ajout de différentes concentrations d'agrocinopine D. La production de NADH, H+ a été suivie à une longueur d'onde de 340nm grâce au spectrophotomètre UVIKON xs. Le K_{half} (concentration nécessaire pour atteindre la vitesse maximale divisée par 2) de la réaction a ensuite été obtenu en utilisant GraphPad Prism 5.0 (www.graphpad.com) par fitting de la courbe obtenue avec l'équation Y=Vmax*X^h/(Khalf^h + X^h). Le choix de cette équation a été basé sur l'allure sigmoïde de la courbe correspondant à une enzyme allostérique, ainsi que l'obtention d'un coefficient de corrélation élevé.

i. Caractérisation des interactions protéines : ligands et de la stabilité des protéines

i. Mesure de la constante de dissociation par titration de fluorescence

L'établissement de l'interaction entre les protéines et les différents ligands a été déterminé par observation de la fluorescence émise à 330 nm par l'échantillon protéique (principalement les tryptophanes), après excitation à 270 nm. La différence de fluorescence entre la protéine seule et en présence de ligand est enregistrée avec un TECAN infinite M1000 PRO microplate reader (Tecan Group SA). Toutes les expériences ont été enregistrées à 20°C en triplicata en tampon Tris-HCL

pH8 et 150 mM de NaCl. Les concentrations utilisées sont comprises entre 1 et 5 µM pour les protéines et une gamme de concentration croissante variant de 0,1 à 100µM est utilisée pour les ligands. 100 µl de protéine sont dans un premier temps déposés dans les puits et une première mesure à 330 nm est enregistrée après excitation à 270 nm. Ensuite, 20 µl de différentes concentrations de ligands sont ajoutés. Une dernière mesure est enregistrée après excitation. En absence de ligand, l'environnement des tryptophanes reste inchangé et la valeur de fluorescence ne change pas. A l'inverse, en présence de ligand proche des tryptophanes, leur environnement est modifié et la valeur du signal de fluorescence change. Les données enregistrées ont été analysées grâce au logiciel Origin 7 software (ORIGIN®) et « fittées » à l'équation $f=\Delta Fluorescence_{max}*abs(x)/(Kd+abs(x))$.

ii. Détermination des températures d'inflexion (Tycho)

Les analyses Tycho concernent la stabilité des protéines via l'analyse de l'état replié ou déplié. En effet la protéine possède un signal de fluorescence intrinsèque à 330 et 350 nm grâce aux tryptophanes et aux tyrosines. L'augmentation de la température au court du temps va provoquer la dénaturation de la protéine et ainsi impacter sa fluorescence intrinsèque, ce qui nous renseigne sur la transition de l'état replié vers l'état déplié. La température pour laquelle la moitié des protéines sont dépliées est appelée la température d'inflexion. Si la protéine est en complexe avec son ligand, elle établit de nouvelles liaisons qui peuvent stabiliser la protéine et faire augmenter sa température d'inflexion.

En ce qui concerne la réalisation de la mesure, 10 µl de protéine seule ou de complexe protéine-ligand (10µM de protéine et 500µM ligand) sont introduits dans le capillaire par capillarité, et le faisceau de lumière est ensuite positionné sur le capillaire pour réaliser la mesure. Le signal est enregistré en même temps qu'un gradient de température est exercé sur le capillaire dans le but de dénaturer la protéine.

iii. Analyse et détermination des températures de fusion (DSC)

Le principe de cette technique consiste à faire subir à l'échantillon un gradient de température qui va progressivement déplacer l'équilibre de la forme native vers la forme dénaturée de la protéine. Lorsqu'une protéine se dénature, elle absorbe ou dégage de la chaleur. La différence de température (ΔT) entre deux cellules, l'une contenant uniquement le tampon et l'autre contenant la protéine, est mesurée et la puissance utilisée pour ramener ΔT à 0 permet de tracer la courbe de la capacité

calorimétrique Cp (kcal/mole/°C) en fonction de la température. La température de transition entre les deux états (natif et dénaturé (Tm)) est déterminée au sommet du pic dont l'aire correspond à l'enthalpie (ΔH) qui reflète les changements de conformation de la protéine. Plus la protéine est stable et plus son Tm va être élevé.

Les mesures ont été réalisées sur un système automatisé MicroCal PEAQ-DSC (Malvern Panalytical). Les échantillons (350 µL) (1mg/ml protéine et un ratio protéine:ligand de 1:3) sont stockés sur plaques de type Deepwell 96 puits thermostatées à 15°C. Chaque échantillon et le tampon correspondant sont prélevés pour analyse dans des cuves progressivement chauffées de 20 à 100°C à une vitesse de 1°C par minute. Les données ont été analysées par le logiciel MicroCal PEAQ-DSC.

Les mesures de Tm ont une variabilité allant de ± 0.05 à 0.005 °C. La valeur extrêmement faible de cette variabilité est donc négligeable pour l'interprétation des résultats obtenus.

iv. Mesure de la constante de dissociation par titrage de calorimétrie isotherme

J'ai utilisé l'ITC pour mesurer la constante de dissociation K_D caractérisant les complexes protéine/ligand étudiés. Le principe de cette technique est basé sur le fait que, lorsqu'une interaction entre biomolécules a lieu, de la chaleur est absorbée ou libérée. L'appareil est constitué de deux cellules, l'une contenant le tampon de référence et l'autre la protéine. Le ligand est progressivement injecté dans la cellule d'échantillon à l'aide d'une seringue et à chaque injection on mesure l'énergie nécessaire (en µcal/sec) pour compenser la petite différence de température avec la cellule de référence. La variation de température décroît progressivement jusqu'à saturation de la protéine par le ligand. La superficie de chaque pic est intégrée en fonction du rapport de concentration ligand/protéine, ce qui permet de tracer un isotherme qui va permettre de calculer les différents paramètres d'interaction (Figure 71).

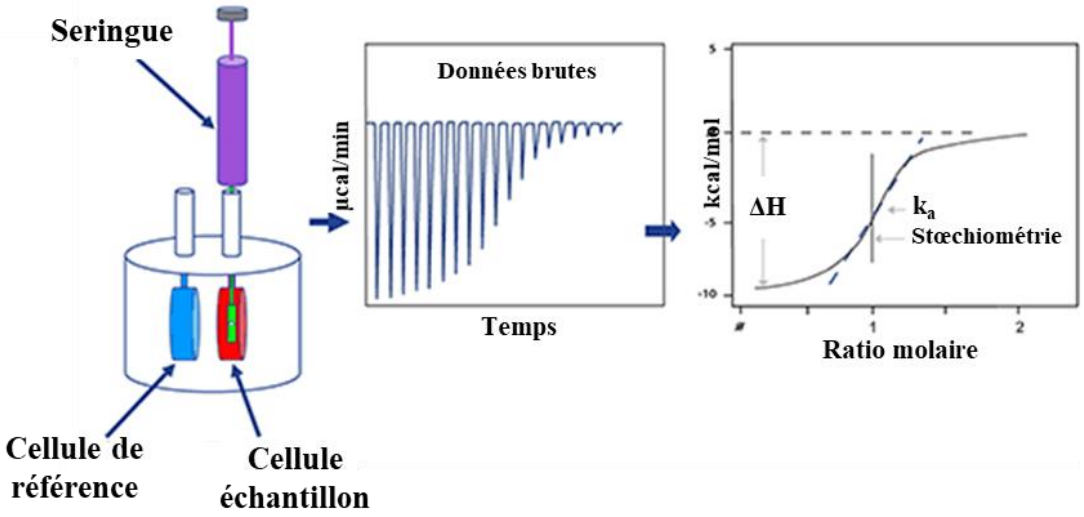


Figure 71: Principe de l'analyse ITC avec comme exemple une réaction exothermique

La pente correspond à la constante d'association à l'équilibre, K_A, inverse de la constante de dissociation K_D, qui renseigne sur l'affinité. Le rapport molaire correspondant au centre de l'isotherme renseigne sur la stœchiométrie de l'interaction, c'est-à-dire sur le nombre de sites n de fixation du ligand sur la protéine. Enfin, l'isotherme permet aussi de mesurer l'enthalpie (ΔH) de la réaction, et donc de calculer les autres paramètres thermodynamiques de l'interaction, ΔG et ΔS. Ces mesures sont très sensibles aux différences de tampon entre les deux cellules.

Les mesures ont été réalisées à une température constante de 20°C sur un appareil MicroCal iTC200 (Malvern Panalytical), constitué de deux cellules de 200 µL. La concentration en protéine dans la cellule d'échantillon était de 100 à 500 µM et la seringue contenait les ligands à une concentration environ 10 fois supérieure à celle de la protéine, ajustée en fonction de l'affinité du complexe afin de voir toute la courbe de saturation.

Annexes

Tables cristallographiques $AccF_{Bo542}$

	AccF apo Ca^{2+} LRE 3-1	AccF apo Mn^{2+} Eqsm5-5	AccF Sulfate Ca^{2+} Eqsm4-16	AccF Sulfate Mn^{2+} Eqsm4-8	AccF Glucose Ca^{2+} LRE3-5
Date	13/02/18	07/12/17	13/02/18	09/02/18	13/02/18
Conditions de Cristallisation	28% Peg 4000 0,1M Tris pH8 CaCl2	18% Peg4000 0,1M Tris pH8 MnCl2	Sulfate d'ammonium 2.1M 0,1MHepes pH7.5	Sulfate d'ammonium 1.9M 0,1 M Hepes pH7.5	28% Peg4000 0,1M Tris pH8
Nombre de molécules dans l'unité asymétrique	4	4	2	2	4
Groupe d'espace **Paramètre de la cellule : a, b, c (Å) et b (°)**	$P2_1$ 60 80 103 103	$P2_1$ 60 80 103 103	$P2_1 2_1 2_1$ 70 76 120	$P4_1 2_1 2$ 67 67 241	$P2_1$ 65 80 102 103
Résolution (Å)	48.02-1.45 (1.54-1.45)	49.52-1.7 (1.90-1.79)	47.07-1.70 (1.80-1.70)	18.81-1.31 (1.39-1.31)	49.54-1.33 (1.41-1.33)
Nombre de réflexion observées	624684 (97758)	538043 (78561)	947919 (148103)	324922 (383124)	877229 (132036)
Nombre de réflexion uniques	176866 (27916)	94474 (14222)	71408 (11068)	134133 (20983)	229829 (36280)
R_{sym} (%)	6.4% (73.4%)	9.1% (91.4%)	13% (8.4%)	8.8% (226.3%)	5% (107.2%)
R_{pim}	4.0 (45.5)	4.1 (41.5)	3.7 (53.9)	1.8 (54.1)	2.9 (63.0)
Complétude (%)	97.6% (95.6%)	98.7% (92.4%)	99.4% (96.4%)	99.7% (97.9%)	99.4% (97.2%)
I/σ	10.12 (1.38)	9.99 (1.31)	10.78 (1.11)	19.01 (1.05)	11.63 (1.04)
$CC_{1/2}$	99.9 (71.8)	99.9 (79.9)	99.7 (68.0)	100 (77.1)	99.9 (59.4)
R_{cryst} (%)	0.1764	0.2088	0.1819	0.1683	0.1847
R_{free} (%)	0.2023	0.2304	0.2046	0.1917	0.2072
RMSD longueurs liaisons (Å)	0.010	0.009	0.010	0.010	0.010
RMDS angles liaisons (°)	0.98	0.97	0.94	1.02	0.98
Moyenne B ($Å^2$) **Protéine A/B/C/D** **Métal** **Solvant** **Ligand**	19.1/21.9/19.7/23.9 13.9 33.9 -	32.3/36.8/38.0/36.8 24.7 40.9 -	34.5/33.1 26.0 44.7 73.2	24.1/22.7 16.3 38.0 16.2	19.83/22.5/20.6/23.9 17.5 36.11 24.25
Clashscore[a] **MolProbity score[a]**	1 0.8	1.6 1.08	1.28 0.85	2.3 1.01	0.59 0.70
Ramachandran plot (%) **Favorable** **Accepté** **Aberrant**	98.78% 1.22% 0%	98.49% 1.51% 0%	98.49% 1.51% 0%	99.06% 0.94% 0%	98.68% 1.32% 0%
Contribution	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement

	AccF Glucose Mn^{2+} Eqsm5-13	AccF G2P Ca^{2+} Eqsm3-8	AccF G2P Mn^{2+} Eqsm5-2	AccF intermédiaire Mn^{2+} Eqsm5-16
Date	07/12/17	13/02/18	07/12/17	07/12/17
Conditions de Cristallisation	Sulfate d'ammonium 1.9M Hepes pH7.5	PEG4K 28% Tris pH8	PEG4K 14% Tris pH8	Sulfate d'ammonium 1.8M Hepes pH7.5
Nombre de molécules dans l'unité asymétrique	2	4	1	2
Groupe d'espace **Paramètre de la cellule : a, b, c (Å) et b (°)**	$P2_1\ 2_1\ 2_1$ 70 76 126	$P2_1$ 65 80 102 103	C_2 54 80 65 112	$P4_1\ 2_1\ 2$ 68 68 242
Résolution (Å)	21.65-1.48 (1.57-1.48)	44.63-1.30 (2.02-1.91)	41.83-1.95 (2.06-1.94)	46.94-1.59 (1.69-1.59)
Nombre de réflexion observées	1317659 (150519)	1306201 (196927)	111968 (16896)	1344692 (204783)
Nombre de réflexion uniques	107784 (16522)	245114 (38741)	19161 (2974)	75739 (11645)
R_{sym} (%)	7.2% (136%)	6.4% (116%)	8.5% (107.4%)	8.03% (133.8%)
R_{pim}	2.1 (46.1)	4.3 (59.2)	3.9 (48.7)	2.0 (32.0)
Complétude (%)	99.2% (95.1%)	99.5% (97.4%)	99.1% (90.4%)	99.4% (96.1%)
I/σ	19.02 (0.99)	11.58 (1.13)	20.54 (1.23)	20.29 (1.57)
$CC_{1/2}$	99.9 (89.9)	99.9 (63.3)	99.7 (78.7)	100 (83)
R_{cryst} (%)	0.1860	0.1708	0.1897	0.1735
R_{free} (%)	0.2121	0.1930	0.2403	0.1990
RMSD longueurs liaisons (Å)	0.010	0.010	0.010	0.010
RMDS angles liaisons (°)	0.95	1.03	1.07	0.96
Moyenne B ($Å^2$) **Protéine A/B/C/D** **Métal** **Solvant** **Ligand**	26.4/27.2 19.6 41.9 21.4	18.9/21.2/19.3/23.2 13.2 34.5 17.2	60.103 38.2 56.2 39.0	30.1/30.3 18.3 41.8 23.8
Clashscore[a] **MolProbity score[a]**	0.8 0.84	1.35 0.86	1.41 0.87	0.58 0.69
Ramachandran plot (%) **Favorable** **Accepté** **Aberrant**	98.88% 1.12% 0%	98.87% 1.13% 0%	98.11% 1.89% 0%	98.69% 1.31% 0%
Contribution	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement

	AccF H16A Ca^{2+} Eqsm2-10	AccF H60A Eqsm6-7	AccF H174A Ca^{2+} Eqsm2-4	AccF E195A Eqsm2-10	AccF H197A Eqsm3-2
Date	16 nov 18	25 jan 19	22 sep 18	22 sep 18	22 sep 18
Conditions de Cristallisation	30% Peg4000 0,1M Tris pH8	30% Peg4000 0,1M Tris pH8	30% Peg4000 0,1M Tris pH8	30% Peg4000 0,1M Tris pH8	30% Peg4000 0,1M Tris pH8
Nombre de molécules dans l'unité asymétrique	1	4	4	4	4
Groupe d'espace **Paramètre de la cellule : a, b, c (Å) et b (°)**	C2 55 80 65 113	$P2_1$ 65 80 102 103	$P2_1$ 65 80 102 103	$P2_1$ 65 80 102 103	$P2_1$ 65 80 102 103
Résolution (Å)	42.89-1.70 (1.80-1.70)	48.84-1.40 (1.48-1.40)	48.85-1.50 (1.59-1.50)	43.95-1.40 (1.48-1.40)	44.16-1.40 (1.48-1.40)
Nombre de réflexion observées	198329 (31071)	1409645 (224164)	1124151 (171109)	924409 (144593)	1411311 (208919)
Nombre de réflexion uniques	28924 (4664)	209714 (32256)	166105 (25735)	206284 (32806)	206920 (32052)
R_{sym} (%)	5.2% (51.8%)	7.6% (124.5%)	9.9% (87.3%)	6.3% (60.2%)	7.1% (75.9%)
R_{pim}	2.1 (21.5)	31.1 (49.8)	4.1 (36.0)	3.3 (31.8)	2.9 (31.5)
Complétude (%)	99.7% (99.3%)	99.1% (97.3%)	99.2% (95.4%)	99.6% (98.4%)	99.2% (95.5%)
I/σ	17.33 (2.76)	10.91 (1.26)	9.97 (2.21)	10.57 (1.90)	12.90 (1.79)
$CC_{1/2}$	99.9 (88.9)	99.9 (65.4)	99.5 (80.3)	99.7 (87.2)	99.8 (83.5)
R_{cryst} (%)	0.1915	0.1850	0.1897	0.2334	0.1916
R_{free} (%)	0.2123	0.2049	0.2163	0.2601	0.2119
RMSD longueurs liaisons (Å)	0.010	0.010	0.010	0.010	0.010
RMDS angles liaisons (°)	0.98	1.00	0.98	0.98	0.98
Moyenne B ($Å^2$) **Protéine A/B/C/D** **Métal** **Solvant** **Ligand**	40.3 38.8 45.9	22.5/21.9/22.7/23.9 - 36.6	23.5/24.4/23.5/25.0 16.6 36.4	25.2/22.8/22.6/22.2 15.15 36.0	33.0/21.0/19.8/20.0 - 35.7
Clashscore[a] **MolProbity score[a]**	0.95 0.78	1.36 0.87	0.89 0.77	2.25 1.01	2.58 1.04
Ramachandran plot (%) **Favorable** **Accepté** **Aberrant**	98.49% 1.51% 0%	98.59% 1.41% 0%	98.59% 1.41% 0%	98.59% 1.41% 0%	98.68% 1.32% 0%
Contribution	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement

Tables cristallographiques $AccA_{C58/Bo542/S4}$

	AccA Bo542 Agrocinopine D Col2-helical
Date	18/12/15
Conditions de Cristallisation	35% Peg4000
Nombre de molécules dans l'unité asymétrique	1
Groupe d'espace **Paramètre de la cellule : a, b, c (Å) et b (°)**	107 40 103 114
Résolution (Å)	33.4-1.75 (1.86-1.75)
Nombre de réflexion observées	158068 (23114)
Nombre de réflexion uniques	40840 (6420)
R_{sym} (%)	12% (95.7%)
R_{pim}	6.7% (57.6%)
Complétude (%)	99.6% (52.5%)
I/σ	8.18 (1)
$CC_{1/2}$	99.6 (52.5)
R_{cryst} (%)	0.1867
R_{free} (%)	0.2224
RMSD longueurs liaisons (Å)	0.009
RMDS angles liaisons (°)	1.04
Moyenne B (Å²) **Protéine A/B/C/D** **Métal** **Solvant** **Ligand**	38.2 - 42 32.7
Clashscore[a] **MolProbity score[a]**	1.54 1.20
Ramachandran plot (%) **Favorable** **Accepté** **Aberrant**	96.32% 3.68% 0%
Contribution	DEICSICS Grégory Affinement

	AccA C58 – Agrocinopine A	AccA C58 Agrocinopine D
Date	17 mars 2013	18 dec 2015
Conditions de Cristallisation	20% Peg 4000 0,1M Acetate NH4 0,2M Citrate Na pH 5.6	20% Peg 4000 0,1M Acetate NH4 0,2M Citrate Na pH 5.6
Nombre de molécules dans l'unité asymétrique	1	4
Groupe d'espace **Paramètre de la cellule : a, b, c (Å) et b (°)**	C2221 78.1 114.9 113.3	P1 78.1 87.5 87.53 76.4 63.5 63.46
Résolution (Å)	23.16-1.67 (1.77-1.67)	42.21-1.80 (1.90-1.80)
Nombre de réflexion observées	428865 (67433)	302562 (44676)
Nombre de réflexion uniques	59578 (9375)	159929 (24869)
R_{sym} (%)	8.2% (164.9%)	6.1% (56.5%)
R_{pim}	3.3% (64.7%)	8.3% (53%)
Complétude (%)	99.7% (98.1%)	92.1% (88.4%)
I/σ	13.61 (1.06)	9.42 (1.31)
$CC_{1/2}$	99.9 (63.7)	99.6 (52.0)
R_{cryst} (%)	0.1974	0.1787
R_{free} (%)	0.2210	0.2059
RMSD longueurs liaisons (Å)	0.010	0.010
RMDS angles liaisons (°)	0.98	1.02
Moyenne B (Å²) **Protéine A/B/C/D** **Métal** **Solvant** **Ligand**	 32.1 39.8 37.9	 31.4/31.25/31.43/31.23 38.1 28.8
Clashscore[a] **MolProbity score[a]**	0.25 0.84	1.54 1.11
Ramachandran plot (%) **Favorable** **Accepté** **Aberrant**	 96.31 3.69 0	 96.62 3.38 0
Contribution	DEICSICS Grégory Traitement des données Affinement	DEICSICS Grégory Traitement des données Affinement

	AccA S4 Arabinose-2-P Eqsm05-1	AccA S4 Glucose-2-P Eqsm03-4-2	AccA S4 Agrocinopine A Col03-01	AccA S4 Agrocinopine D Eqsm03-9-3
Date	**22/06/19**	**18/05/17**	**29/06/16**	**18/05/17**
Conditions de Cristallisation	20% Peg8000 0.1 M Hepes pH7	20% Peg8000 0.1 MES pH 6.5	20% Peg8000 0.1 M Hepes pH7	20% Peg8000 0.1 M Hespes pH7
Nombre de molécules dans l'unité asymétrique	1	1	1	1
Groupe d'espace **Paramètre de la cellule : a, b, c (Å) et b (°)**	$P2_1\ 2_1\ 2_1$ 47 61 152	$P2_1\ 2_1\ 2_1$ 47 61 151	$P2_1\ 2_1\ 2_1$ 47 61 151	$P2_1\ 2_1\ 2_1$ 47 61 151
Résolution (Å)	50-1.40 Å (1.49-1.40)	47.66-1.20 Å (1.27-1.2)	44.97-1.5 Å (1.59-1.50)	33.39-1.68 Å (1.78-1.68)
Nombre de réflexion observées	1137987 (182224)	1645347 (167102)	833423 (125060)	609872 (76638)
Nombre de réflexion uniques	86791 (13746)	138637 (21841)	70531 (10968)	49148 (6588)
R_{sym} **(%)**	4.5% (41.2%)	7.2% (88.1%)	9% (130.3%)	11.1% (159.1%)
R_{pim}	1.3% (11.6%)	2.1 (33.1)	2.7 (39.5)	3.3 (46.5)
Complétude (%)	99.8% (99.1%)	99.7 (98.2)	99.5% (97%)	97.1 (81.9)
I/σ	30.57 (5.73)	16.48 (1.67)	14.87 (1.5)	12.84 (1.24)
$CC_{1/2}$	100 (96.4)	99.9 (70.7)	99.9 (97)	99.9 (77.8)
R_{cryst} **(%)**	0.1830	0.1835	0.1698	0.1783
R_{free} **(%)**	0.1951	0.1835	0.1964	0.2116
RMSD longueurs liaisons (Å)	0.009	0.010	0.010	0.010
RMDS angles liaisons (°)	1.02	1.06	1.02	1.01
Moyenne *B* (Å²) **Protéine A/B/C/D** **Métal** **Solvant** **Ligand**	27.01 35.03 21.33	24.9 18.2 36.5	28.25 23.6 36.7	40.9 38.8 42
Clashscore[a] **MolProbity score[a]**	2.67 1.15	2.30 1.13	1.29 0.91	1.71 1.20
Ramachandran plot (%) **Favorable** **Accepté** **Aberrant**	97.48% 2.25% 0%	97.27% 2.73% 0%	97.69% 2.31% 0%	96.64% 3.36% 0%
Contribution	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement	DEICSICS Grégory Affinement	DEICSICS Grégory Cristaux Collecte Traitement des données Affinement

Références

Aloni, R., Wolf, A., Feigenbaum, P., Avni, A., & Klee, H. J. (1998). The never ripe mutant provides evidence that tumor- induced ethylene controls the morphogenesis of Agrobacterium tumefaciens-induced crown galls on tomato stems. *Plant Physiology*, *117*, 841–849. https://doi.org/10.18574/nyu/9781479814336.001.0001

Beck von Bodman, S., Hayman, G. T., & Farrand, S. K. (1992). Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(2), 643–647. https://doi.org/10.1073/pnas.89.2.643

Berntsson, R. P. A., Smits, S. H. J., Schmitt, L., Slotboom, D. J., & Poolman, B. (2010). A structural classification of substrate-binding proteins. *FEBS Letters*, *584*(12), 2606–2617. https://doi.org/10.1016/j.febslet.2010.04.043

Breedveld, M. W., & Miller, K. J. (1994). Cyclic β-glucans of members of the family Rhizobiaceae. *Microbiological Reviews*, *58*(2), 145–161.

Brencic, A., Angert, E. R., & Winans, S. C. (2005). Unwounded plants elicit Agrobacterium vir gene induction and T-DNA transfer: Transformed plant cells produce opines yet are tumour free. *Molecular Microbiology*, *57*(6), 1522–1531. https://doi.org/10.1111/j.1365-2958.2005.04763.x

Brencic, A., Winans, S. C., & Colonization, P. (2005). Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiology and Molecular Biology Reviews*, *69*(1), 155–194. https://doi.org/10.1128/MMBR.69.1.155–194.2005

Cha, C., Gao, P., Chen, Y., Shaw, P. D., & Farrand, S. K. (1998). Production of acyl-homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria. *Molecular Plant-Microbe Interactions*, *11*(11), 1119–1129.

Chen, L., Chen, Y., Wood, D. W., & Nester, E. W. (2002). A new type IV secretion system promotes conjugal transfer in Agrobacterium tumefaciens. *Journal of Bacteriology*, *184*(17), 4838–4845. https://doi.org/10.1128/JB.184.17.4838-4845.2002

Cho, H., & Winans, S. C. (2007). TraA, TraC and TraD autorepress two divergent quorum-regulated promoters near the transfer origin of the Ti plasmid of Agrobacterium tumefaciens. *Molecular Microbiology*, *63*(6), 1769–1782. https://doi.org/10.1111/j.1365-2958.2007.05624.x

Christie, P. J., Whitaker, N., & González-Rivera, C. (2014). BBA review revised mechanism and structure of the bacterial type IV secretion systems. *Biochimica et Biophysica Acta (BBA)/Protein Structure and Molecular*, *46*(4), 564–574. https://doi.org/10.1016/j.cortex.2009.08.003.Predictive

Conn, H. J. (1942). Validity of the genus alcaligenes. *Journal of Bacteriology*, *44*(3), 353–360.

Corda, D., Mosca, M. G., Ohshima, N., Grauso, L., Yanaka, N., & Mariggiò, S. (2014). The emerging physiological roles of the glycerophosphodiesterase family. *FEBS Journal*, *281*(4), 998–1016. https://doi.org/10.1111/febs.12699

Deakin, W. J., Sanderson, J. L. C. I., Goswami, T., & Shaw, C. H. (1997). The Agrobacterium tumefaciens motor gene, motA, is in a linked cluster with the flagellar switch protein genes, fliG, fliM and fliN. *Gene*, *189*(1), 139–141. https://doi.org/10.1016/S0378-1119(96)00853-0

Denloye, T., Dalal, S., & Klemba, M. (2012). Characterization of a glycerophosphodiesterase with an unusual tripartite distribution and an important role in the asexual blood stages of

Plasmodium falciparum. *Molecular and Biochemical Parasitology*, *186*(1), 29–37. https://doi.org/10.1016/j.molbiopara.2012.09.004

Dessaux, Y., & Farrand, S. K. (1998). The Rhizobiaceae. *The Rhizobiaceae*, (January). https://doi.org/10.1007/978-94-011-5060-6

Dessaux, Y., & Faure, D. (2018). Quorum sensing and quorum quenching in agrobacterium: A "go/no go system"? *Genes*, *9*(4). https://doi.org/10.3390/genes9040210

Doty, S. L., Chang, M., & Nester, E. W. (1993). The chromosomal virulence gene, chvE, of Agrobacterium tumefaciens is regulated by a LysR family member. *Journal of Bacteriology*, *175*(24), 7880–7886. https://doi.org/10.1128/jb.175.24.7880-7886.1993

Efetova, M., Zeier, J., Riederer, M., Lee, C.-W., Stingl, N., Mueller, M., … Deeken, R. (2007). A central role of abscisic acid in drought stress protection of Agrobacterium-induced tumors on Arabidopsis. *Plant Physiology*, *145*(3), 853–862. https://doi.org/10.1104/pp.107.104851

El Sahili, A., Li, S. Z., Lang, J., Virus, C., Planamente, S., Ahmar, M., … Morera, S. (2015). A pyranose-2-phosphate motif is responsible for both antibiotic import and quorum-sensing regulation in Agrobacterium tumefaciens. *PLoS Pathogens*, *11*(8), 1–24. https://doi.org/10.1371/journal.ppat.1005071

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 486–501. https://doi.org/10.1107/S0907444910007493

Erika V. Soriano, Rajashankar Kanagalaghatta R., Hanes Jeremiah W., Shridhar Bale, Begley Tadhg P., & E.Ealick, S. (2008). Structural similarities between thiamin-binding protein and thiaminase-I suggest a common ancestor. *Biochemistry*, *47*(5), 1346–1357. https://doi.org/10.1021/bi7018282.

Fukami-Kobayashi, K., Tateno, Y., & Nishikawa, K. (1999). Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *Journal of Molecular Biology*, *286*(1), 279–290. https://doi.org/10.1006/jmbi.1998.2454

Fukamizo, T., Kitaoku, Y., & Suginta, W. (2019). Periplasmic solute-binding proteins: Structure classification and chitooligosaccharide recognition. *International Journal of Biological Macromolecules*, *128*, 985–993. https://doi.org/10.1016/j.ijbiomac.2019.02.064

Fuqua, W. C., Winans, S. C., & Greenberg, E. P. (1994). Quorum sensing in bacteria: The LuxR-LuxI family of cell density- responsive transcriptional regulators. *Journal of Bacteriology*, *176*(2), 269–275. https://doi.org/10.1128/jb.176.2.269-275.1994

Goldman, A. D., Beatty, J. T., & Landweber, L. F. (2016). The TIM barrel architecture facilitated the early evolution of protein-mediated metabolism. *Journal of Molecular Evolution*, *82*(1), 17–26. https://doi.org/10.1007/s00239-015-9722-8

Gonin, S., Arnoux, P., Pierru, B., Lavergne, J., Alonso, B., Sabaty, M., & Pignol, D. (2007). Crystal structures of an Extracytoplasmic Solute Receptor from a TRAP transporter in its open and closed forms reveal a helix-swapped dimer requiring a cation for α-keto acid binding. *BMC Structural Biology*, *7*, 1–14. https://doi.org/10.1186/1472-6807-7-11

Guo, M., Huang, Z., & Yang, J. (2017). Is there any crosstalk between the chemotaxis and virulence induction signaling in Agrobacterium tumefaciens? *Biotechnology Advances*, *35*(4), 505–511. https://doi.org/10.1016/j.biotechadv.2017.03.008

Halder, U., Banerjee, A., & Bandopadhyay, R. (2017). Structural and Functional Properties, Biosynthesis, and Patenting Trends of Bacterial Succinoglycan: A Review. *Indian Journal of Microbiology*, *57*(3), 278–284. https://doi.org/10.1007/s12088-017-0655-3

Harighi, B. (2008). Role of cheY1 and cheY2 in the chemotaxis of A. tumefaciens toward acetosyringone. *Current Microbiology*, *56*(6), 547–552. https://doi.org/10.1007/s00284-008-9120-1

Harighi, B. (2009). Genetic evidence for CheB- and CheR-dependent chemotaxis system in A. tumefaciens toward acetosyringone. *Microbiological Research*, *164*(6), 634–641. https://doi.org/10.1016/j.micres.2008.11.001

Haygood, M., & Nealson, K. H. (1985). Mechanisms of iron regulation of luminescence in Vibrio fischeri. *Journal of Bacteriology*, *162*(1), 209–216.

He, F., Nair, G. R., Soto, C. S., Chang, Y., Hsu, L., Ronzone, E., … Binns, A. N. (2009). Molecular basis of ChvE function in sugar binding, sugar utilization, and virulence in Agrobacterium tumefaciens. *Journal of Bacteriology*, *191*(18), 5802–5813. https://doi.org/10.1128/JB.00451-09

Higgins, M. A., Abbott, D. W., Boulanger, M. J., & Boraston, A. B. (2009). Blood Group Antigen Recognition by a Solute-Binding Protein from a Serotype 3 Strain of Streptococcus pneumoniae. *Journal of Molecular Biology*, *388*(2), 299–309. https://doi.org/10.1016/j.jmb.2009.03.012

Hu, X., Zhao, J., DeGrado, W. F., & Binns, A. N. (2013). Agrobacterium tumefaciens recognizes its host environment using ChvE to bind diverse plant sugars as virulence signals. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(2), 678–683. https://doi.org/10.1073/pnas.1215033110

Jackson, C. J., Carr, P. D., Liu, J., Watt, S. J., Beck, J. L., & Ollis, D. L. (2007). The structure and function of a novel glycerophosphodiesterase from Enterobacter aerogenes. *Journal of Molecular Biology*, *367*, 1047–1062. https://doi.org/10.1016/j.jmb.2007.01.032

Kabsch, W. (2010). XDS. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 125–132. https://doi.org/10.1107/S0907444909047337

Kado, C. I. (2014). Historical account on gaining insights on the mechanism of crown gall tumorigenesis induced by Agrobacterium tumefaciens. *Frontiers in Microbiology*, *5*(340). https://doi.org/10.3389/fmicb.2014.00340

Karpowich, N. K., Huang, H. H., Smith, P. C., & Hunt, J. F. (2003). Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. *Journal of Biological Chemistry*, *278*(10), 8429–8434. https://doi.org/10.1074/jbc.M212239200

Kemner, J. M., Liang, X., & Nester, E. W. (1997). The Agrobacterium tumefaciens virulence gene chvE is part of a putative ABC-type sugar transport operon. *Journal of Bacteriology*, *179*(7), 2452–2458. https://doi.org/10.1128/jb.179.7.2452-2458.1997

Kim, H., & Farrand, S. K. (1997). Characterization of the acc operon from the nopaline-type Ti plasmid pTiC58, which encodes utilization of agrocinopines A and B and susceptibility to agrocin 84. *Journal of Bacteriology*, *179*(23), 7559–7572. https://doi.org/10.1128/jb.179.23.7559-7572.1997

Kim, H., & Farrand, S. K. (1998). Opine catabolic loci from Agrobacterium plasmids confer

chemotaxis to their cognate substrates. *Molecular Plant-Microbe Interactions*, *11*(2), 131–143. https://doi.org/10.1094/MPMI.1998.11.2.131

Kim, S. I., Veena, & Gelvin, S. B. (2007). Genome-wide analysis of Agrobacterium T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. *Plant Journal*, *51*(5), 779–791. https://doi.org/10.1111/j.1365-313X.2007.03183.x

Kuzmanović, N., Gašić, K., Ivanović, M., Prokić, A., & Obradović, A. (2012). Identification of Agrobacterium vitis as a causal agent of grapevine crown gall in Serbia. *Archives of Biological Sciences*, *64*(4), 1487–1494. https://doi.org/10.2298/ABS1204487K

Lacroix, B., & Citovsky, V. (2013). The roles of bacterial and host plant factors in Agrobacterium-mediated genetic transformation. *International Journal of Developmental Biology*, *57*(6–8), 467–481. https://doi.org/10.1387/ijdb.130199bl

Lacroix, B., & Citovsky, V. (2019). Pathways of DNA transfer to plants from Agrobacterium tumefaciens and related bacterial species. *Phytopathology*, *176*(57), 231–251. https://doi.org/10.1016/j.physbeh.2017.03.040

Lassalle, F., Campillo, T., Vial, L., Baude, J., Costechareyre, D., Chapulliot, D., … Nesme, X. (2011). Genomic species are ecological species as revealed by comparative genomics in Agrobacterium tumefaciens. *Genome Biology and Evolution*, *3*(1), 762–781. https://doi.org/10.1093/gbe/evr070

Li, L., Jia, Y., Hou, Q., Charles, T. C., Nester, E. W., & Pan, S. Q. (2002). A global pH sensor: Agrobacterium sensor protein ChvG regulates acid-inducible genes on its two chromosomes and Ti plasmid. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(19), 12369–12374. https://doi.org/10.1073/pnas.192439499

Luo, Z. Q., Qin, Y., & Farrand, S. K. (2000). The antiactivator TraM interferes with the autoinducer-dependent binding of TraR to DNA by interacting with the C-terminal region of the quorum- sensing activator. *Journal of Biological Chemistry*, *275*(11), 7713–7722. https://doi.org/10.1074/jbc.275.11.7713

Matthysse, A. G., Jaeckel, P., & Jeter, C. (2008). attG and attC mutations of Agrobacterium tumefaciens are dominant negative mutations that block attachment and virulence. *Canadian Journal of Microbiology*, *54*(4), 241–247. https://doi.org/10.1139/W08-005

McCoy, A. J. (2006). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D: Biological Crystallography*, *63*(1), 32–41. https://doi.org/10.1107/S0907444906045975

McFerrin, M. B., & Snell, E. H. (2002). The development and application of a method to quantify the quality of cryoprotectant solutions using standard area-detector X-ray images. *Journal of Applied Crystallography*, *35*(5), 538–545. https://doi.org/10.1107/S0021889802009238

Melene A. Thompson, Onyeziri, M. C., & Fuqua, C. (**2018)**. Function and regulation of Agrobacterium tumefaciens cell surface structures that promote attachment. *Current Topics in Microbiology and Immunology*, *359*(March), 79–94. https://doi.org/10.1007/82

Mishima, Y., Momma, K., Hashimoto, W., Mikami, B., & Murata, K. (2003). Crystal structure of AlgQ2, a macromolecule (alginate)-binding protein of Sphingomonas sp. A1, complexed with an alginate tetrasaccharide at 1.6-Å resolution. *Journal of Biological Chemistry*, *278*(8), 6552–6559. https://doi.org/10.1074/jbc.M209932200

Momma, K., Mishima, Y., Hashimoto, W., Mikami, B., & Murata, K. (2005). Direct evidence for Sphingomonas sp. A1 periplasmic proteins as macromolecule-binding proteins associated with the ABC transporter: Molecular insights into alginate transport in the periplasm. *Biochemistry*, *44*(13), 5053–5064. https://doi.org/10.1021/bi047781r

Morton, E. R., Merritt, P. M., Bever, J. D., & Fuqua, C. (2013). Large deletions in the pAtC58 megaplasmid of Agrobacterium tumefaciens can confer reduced carriage cost and increased expression of virulence genes. *Genome Biology and Evolution*, *5*(7), 1453–1464. https://doi.org/10.1093/gbe/evt095

Nair, G. R., Liu, Z., & Binns, A. N. (2003). Reexamining the role of the accessory plasmid pAtC58 in the virulence of Agrobacterium tumefaciens strain C58. *Plant Physiology*, *133*(3), 989–999. https://doi.org/10.1104/pp.103.030262

Oh, B. H., Kang, C. H., De Bondt, H., Kim, S. H., Nikaido, K., Joshi, A. K., & Ames -, G. F. L. (1994). The bacterial periplasmic histidine-binding protein. Structure/function analysis of the ligand-binding site and comparison with related proteins. *Journal of Biological Chemistry*, *269*(6), 4135–4143. https://doi.org/10.2210/pdb1hpb/pdb

Ohshima, N., Yamashita, S., Takahashi, N., Kuroishi, C., Shiro, Y., & Takio, K. (2008). Escherichia coli cytosolic glycerophosphodiester phosphodiesterase (UgpQ) requires Mg2+,Co2+,orMn2+ for its enzyme activity. *Journal of Bacteriology*, *190*(4), 1219–1223. https://doi.org/10.1128/JB.01223-07

Păcurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, M. L., & Bellini, C. (2011). Agrobacterium tumefaciens: From crown gall tumors to genetic transformation. *Physiological and Molecular Plant Pathology*, *76*(2), 76–81. https://doi.org/10.1016/j.pmpp.2011.06.004

Pappas, K. M. (2008). Cell-cell signaling and the Agrobacterium tumefaciens Ti plasmid copy number fluctuations. *Plasmid*, *60*(2), 89–107. https://doi.org/10.1016/j.plasmid.2008.05.003

Peralta, E. G., & Ream, L. W. (1985). T-DNA border sequences required for crown gall tumorigenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(15), 5112–5116. https://doi.org/10.1073/pnas.82.15.5112

Perveen, S., Rashid, N., & Papageorgiou, A. C. (2016). Crystal structure of a phosphoribosyl anthranilate isomerase from the hyperthermophilic archaeon Thermococcus kodakaraensis. *Acta Crystallographica Section:F Structural Biology Communications*, *72*(11), 804–812. https://doi.org/10.1107/S2053230X16015223

Platt, T. G., Morton, E. R., Barton, I. S., Bever, J. D., & Fuqua, C. (2014). Ecological dynamics and complex interactions of Agrobacterium megaplasmids. *Frontiers in Plant Science*, *5*(635). https://doi.org/10.3389/fpls.2014.00635

Pons, J. L., & Labesse, G. (2009). @TOME-2: A new pipeline for comparative modeling of protein-ligand complexes. *Nucleic Acids Research*, *37*(SUPPL. 2), 485–491. https://doi.org/10.1093/nar/gkp368

Quiocho, F. A., & Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: Variation of common themes. *Molecular Microbiology*, *20*(1), 17–25. https://doi.org/10.1111/j.1365-2958.1996.tb02484.x

Ramírez-Bahena, M. H., Vial, L., Lassalle, F., Diel, B., Chapulliot, D., Daubin, V., … Muller, D. (2014). Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the Agrobacterium/Rhizobium supercluster characterized by the presence of a linear chromid. *Molecular Phylogenetics and Evolution*, *73*(1), 202–207. https://doi.org/10.1016/j.ympev.2014.01.005

Rao, K. N., Bonanno, J. B., Burley, S. K., & Swaminathan, S. (2006). Crystal structure of glycerophosphodiester phosphodiesterase from Agrobacterium tumefaciens by SAD with a large asymmetric unit. *PROTEINS: Structure, Function, and Bioinformatics*, *65*, 514–518. https://doi.org/DOI: 10.1002/prot.21079

Rucktooa, P., Antoine, R., Herrou, J., Huvent, I., Locht, C., Jacob-Dubuisson, F., … Bompard, C. (2007). Crystal structures of two Bordetella pertussis periplasmic receptors contribute to defining a novel pyroglutamic acid binding DctP subfamily. *Journal of Molecular Biology*, *370*(1), 93–106. https://doi.org/10.1016/j.jmb.2007.04.047

Scharf, B., & Schmitt, R. (2002). Sensory transduction to the flagellar motor of Sinorhizobium meliloti. *Journal of Molecular Microbiology and Biotechnology*, *4*(3), 183–186.

Scheepers, G. H., Lycklama a Nijeholt, J. A., & Poolman, B. (2016). An updated structural classification of substrate-binding proteins. *FEBS Letters*, *590*(23), 4393–4401. https://doi.org/10.1002/1873-3468.12445

Schüttelkopf, A. W., & Van Aalten, D. M. F. (2004). PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, *60*(8), 1355–1363. https://doi.org/10.1107/S0907444904011679

Shi, L., Liu, J. F., An, X. M., & Liang, D. C. (2008). Crystal structure of glycerophosphodiester phosphodiesterase (GDPD) from Thermoanaerobacter tengcongensis, a metal ion-dependent enzyme: Insight into the catalytic mechanism. *Proteins: Structure, Function and Genetics*, *72*(1), 280–288. https://doi.org/10.1002/prot.21921

Shouldice, S. R., Skene, R. J., Dougan, D. R., Snell, G., McRee, D. E., Schryvers, A. B., & Tari, L. W. (2004). Structural basis for iron binding and release by a novel class of periplasmic iron-binding proteins found in gram-negative pathogens. *Journal of Bacteriology*, *186*(12), 3903–3910. https://doi.org/10.1128/JB.186.12.3903

Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Sharff, A., & Vonrhein, C. (2012). Supplementary materials : and target structure restraints in BUSTER. *Acta Crystallographica Section D: Biological Crystallography*, 1–8.

Stachel, S. E., Nester, E. W., & Zambryski, P. C. (1986). A plant cell factor induces Agrobacterium tumefaciens vir gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(2), 379–383. https://doi.org/10.1073/pnas.83.2.379

Sutcliffe, I. C., & Russell, R. R. B. (1995). Lipoproteins of gram-positive bacteria. *Journal of Bacteriology*, *177*(5), 1123–1128. https://doi.org/10.1128/jb.177.5.1123-1128.1995

Trakhanov, S., Vyas, N. K., Luecke, H., Kristensen, D. M., Ma, J., & Quiocho, F. A. (2005). Ligand-free and -bound structures of the binding protein (LivJ) of the Escherichia coli ABC leucine/isoleucine/valine transport system: trajectory and dynamics of the interdomain rotation

and ligand specificity. *Biochemistry*, *44*, 6597–6608. https://doi.org/10.1021/bi047302o

Turovskiy, Y., Kashtanov, D., Paskhover, B., & Chikindas, M. L. (2007). Quorum sensing: Fact, fiction, and everything in between. *Advances in Applied Microbiology*, *62*, 191–234. https://doi.org/10.1016/S0065-2164(07)62007-3

Vagin, A., & Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallographica Section D: Biological Crystallography*, *66*(1), 22–25. https://doi.org/10.1107/S0907444909042589

Val, D. L., & Cronan, J. E. (1998). In vivo evidence that S-adenosylmethionine and fatty acid synthesis intermediates are the substrates for the LuxI family of autoinducer synthases. *Journal of Bacteriology*, *180*(10), 2644–2651.

Vergunst, A. C., van Lier, M. C. M., den Dulk-Ras, A., Stüve, T. A. G., Ouwehand, A., & Hooykaas, P. J. J. (2005). Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(3), 832–837. https://doi.org/10.1073/pnas.0406241102

Veselov, D., Langhans, M., Hartung, W., Aloni, R., Feussner, I., Götz, C., … Ullrich, C. I. (2003). Development of Agrobacterium tumefaciens C58-induced plant tumors and impact on host shoots are controlled by a cascade of jasmonic acid, auxin, cytokinin, ethylene and abscisic acid. *Planta*, *216*, 512–522. https://doi.org/10.1007/s00425-002-0883-5

Wächter, R., Langhans, M., Aloni, R., Götz, S., Weilmünster, A., Koops, A., … Ullrich, C. I. (2003). Vascularization, high-volume solution flow, and localized roles for enzymes of sucrose metabolism during tumorigenesis by Agrobacterium tumefaciens. *Plant Physiology*, *133*(3), 1024–1037. https://doi.org/10.1104/pp.103.028142

Welsh, M. A., & Blackwell, H. E. (2016). Chemical probes of quorum sensing: From compound development to biological discoverya. *FEMS Microbiology Reviews*, *40*(5), 774–794. https://doi.org/10.1093/femsre/fuw009

White, C. E., & Winans, S. C. (2007). Cell-cell communication in the plant pathogen Agrobacterium tumefaciens. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1483), 1135–1148. https://doi.org/10.1098/rstb.2007.2040

Williams, W. A., Zhang, R., Zhou, M., Joachimiak, G., Gornicki, P., Missiakas, D., & Joachimiak, A. (2004). The membrane-associated lipoprotein-9 GmpC from Staphylococcus aureus binds the dipeptide GlyMet via side chain interactions. *Biochemistry*, *43*(51), 16193–16202. https://doi.org/10.1021/bi048877o

Wright, E. L., Deakin, W. J., & Shaw, C. H. (1998). A chemotaxis cluster from Agrobacterium tumefaciens. *Gene*, *220*(1–2), 83–89. https://doi.org/10.1016/S0378-1119(98)00438-7

Wu, C. F., Lin, J. S., Shaw, G. C., & Lai, E. M. (2012). Acid-induced type VI secretion system is regulated by ExoR-ChvG/ChvI signaling cascade in Agrobacterium tumefaciens. *PLoS Pathogens*, *8*(9). https://doi.org/10.1371/journal.ppat.1002938

Yanaka, N. (2007). Mammalian glycerophosphodiester phosphodiesterases. *Bioscience, Biotechnology, and Biochemistry*, *71*(8), 1811–1818. https://doi.org/10.1271/bbb.70062

Young, J. M. (2008). Agrobacterium-taxonomy of plant-pathogenic rhizobium species. *Agrobacterium: From Biology to Biotechnology*, 183–220. https://doi.org/10.1007/978-0-387-72290-0_5

Zambryski, P., Tempe, J., & Schell, J. (1989). Transfer and function of T-DNA genes from Agrobacterium Ti and Ri plasmids in plants. *Cell*, *56*(2), 193–201. https://doi.org/10.1016/0092-8674(89)90892-1

Zhu, J., & Winans, S. C. (1998). Activity of the quorum-sensing regulator TraR of Agrobacterium tumefaciens is inhibited by a truncated, dominant defective TraR-like protein. *Molecular Microbiology*, *27*(2), 289–297. https://doi.org/10.1046/j.1365-2958.1998.00672.x

Zupan, J., Hackworth, C. A., Aguilar, J., Ward, D., & Zambryski, P. (2007). VirB1* promotes T-pilus formation in the vir-type IV secretion system of Agrobacterium tumefaciens. *Journal of Bacteriology*, *189*(18), 6551–6563. https://doi.org/10.1128/JB.00480-07

ÉCOLE DOCTORALE
Innovation thérapeutique du fondamental à l'appliqué (ITFA)

université PARIS-SACLAY

Titre : Spécificité d'import et de transformation des agrocinopines et dérivés chez les pathogènes du genre *Agrobacterium*

Mots clés : *Agrobacterium*, AccA, AccF, Phosphodiesterase, Agrocinopine, Spécificité

Résumé : ***Agrobacterium*** *tumefaciens* est une espèce bactérienne du sol qui infecte plus de 600 espèces de plantes. Elle est responsable du développement de la galle du collet. Lors de l'infection, la bactérie transfère au sein de la cellule végétale un ADN-T situé sur le plasmide de virulence Ti (Tumor inducing). L'ADN-T s'intègre à l'ADN génomique de la plante et détourne, au profit de la bactérie, certaines voies métaboliques comme la production de nutriments et d'hormones végétales. Les hormones végétales induisent le développement des tumeurs servant de niches écologiques aux bactéries. Les nutriments, appelés opines, qui sont des sucres phosphorylés dans le cas des agrocinopines, sont uniquement utilisables par la bactérie. Ces agrocinopines lui servent de source de carbone et initient la synthèse de signaux de quorum sensing. Les opines produites par la plante diffèrent en fonction des souches d'*A. tumefaciens* et sont déterminées par le gène bactérien *acs* (**agrocinopine** synthase). Par exemple, la souche modèle C58 transfère un gène *acs* ayant la particularité de faire produire à la plante les agrocinopines A et B, alors que pour la souche Bo542 il s'agit de la production des **agrocinopines** C et D. Une fois produites, les agrocinopines sont reconnues dans un premier temps par une protéine périplasmique (**PBP) AccA**, qui permet leur internalisation dans le cytoplasme de la bactérie, puis sont transformées par la **phosphodiesterase AccF**.

Durant cette thèse, par des approches biochimiques et structurales, je me suis intéressé à la spécificité de reconnaissance de la **PBP AccA** des souches *A. tumefaciens* Bo542 et S4 vis-à-vis des 4 agrocinopines et de leurs intermédiaires. J'ai pu établir par des techniques d'interaction protéine/ligand qu'**$AccA_{Bo542}$** reconnaissait spécifiquement les agrocinopines C et D, contrairement au souches C58 et S4 qui reconnaissant les 4 agrocinopines. Dans le but de comprendre les bases moléculaires de la spécificité d'**$AccA_{Bo542}$**, j'ai ensuite réalisé une étude structurale par cristallographie afin d'identifier des résidus potentiellement impliqués dans la spécificité de cette protéine. Je me suis également intéressé à la **phosphodiesterase AccF** d'un point de vue structural et fonctionnel afin de comprendre son mécanisme réactionnel. Les résultats obtenus combinées aux données scientifiques déjà publiées ont permis d'identifier les résidus impliqués dans le mécanisme réactionnel et les résidus potentiellement impliqués dans la reconnaissance du substrat.

L'ensemble de ces études ont permis d'avoir une meilleure compréhension de la reconnaissance et de la transformation des **agrocinopines** par les bactéries du genre ***Agrobacterium*** (souches *A. tumefaciens* C58/Bo542 et *A. vitis* S4).

Title : Import and processing specificity of agrocinopines and derivatives in the pathogen *Agrobacterium*

Keywords : *Agrobacterium*, AccA, AccF, Phosphodiesterase, Agrocinopine, Specificity

Abstract : ***Agrobacterium*** *tumefaciens* is a soil bacterial species that infects over 600 plant species. It is responsible for the development of crown gall. During infection, the bacterium transfers to the plant cell a T-DNA located on the virulence plasmid Ti (Tumor inducing). The T-DNA integrates with the plant's genomic DNA and diverts certain metabolic pathways, such as the production of plant nutrients and hormones, to the benefit of the bacterium. Plant hormones induce the development of tumours that serve as ecological niches for bacteria. The nutrients, called opines, which are phosphorylated sugars in the case of **agrocinopines**, are only usable by the bacterium. These agrocinopines serve as a carbon source and initiate the synthesis of quorum sensing signals. The opines produced by the plant differ according to the strains of A. tumefaciens and are determined by the bacterial gene acs (**agrocinopine** synthase). For example, the model strain C58 transfers the acs gene with the particularity of making the plant produce the A and B **agrocinopines**, while for the strain Bo542 it is the production of the C and D agrocinopines. Once produced, agrocinopines are initially recognized by a periplasmic binding protein **(PBP) AccA** , which allows their internalization in the cytoplasm of the bacterium, then are transformed by **phosphodiesterase AccF** .

During this thesis, through biochemical and structural approaches, I was interested in the specificity of recognition of the **PBP AccA** strains *A. tumefaciens* Bo542 and S4 vis-à-vis the 4 agrocinopines and their intermediaries. I was able to establish by techniques of protein/ligand interaction that **$AccA_{Bo542}$** specifically recognized the C and D agrocinopines, contrary to the C58 and S4 strains that recognize the 4 **agrocinopines**. In order to understand the molecular basis of the specificity of **$AccA_{Bo542}$**, I then carried out a structural crystallography study to identify residues potentially involved in the specificity of this protein. I was also interested in **phosphodiesterase AccF** from a structural and functional point of view in order to understand its reaction mechanism. The results obtained, combined with the previously published scientific data, identified the residues involved in the reaction mechanism and the residues potentially involved in the recognition of the substrate.

All these studies provided a better understanding of the recognition and transformation of **agrocinopine**s by bacteria of the genus ***Agrobacterium*** (strains A. tumefaciens C58/Bo542 and A. vitis S4).

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France